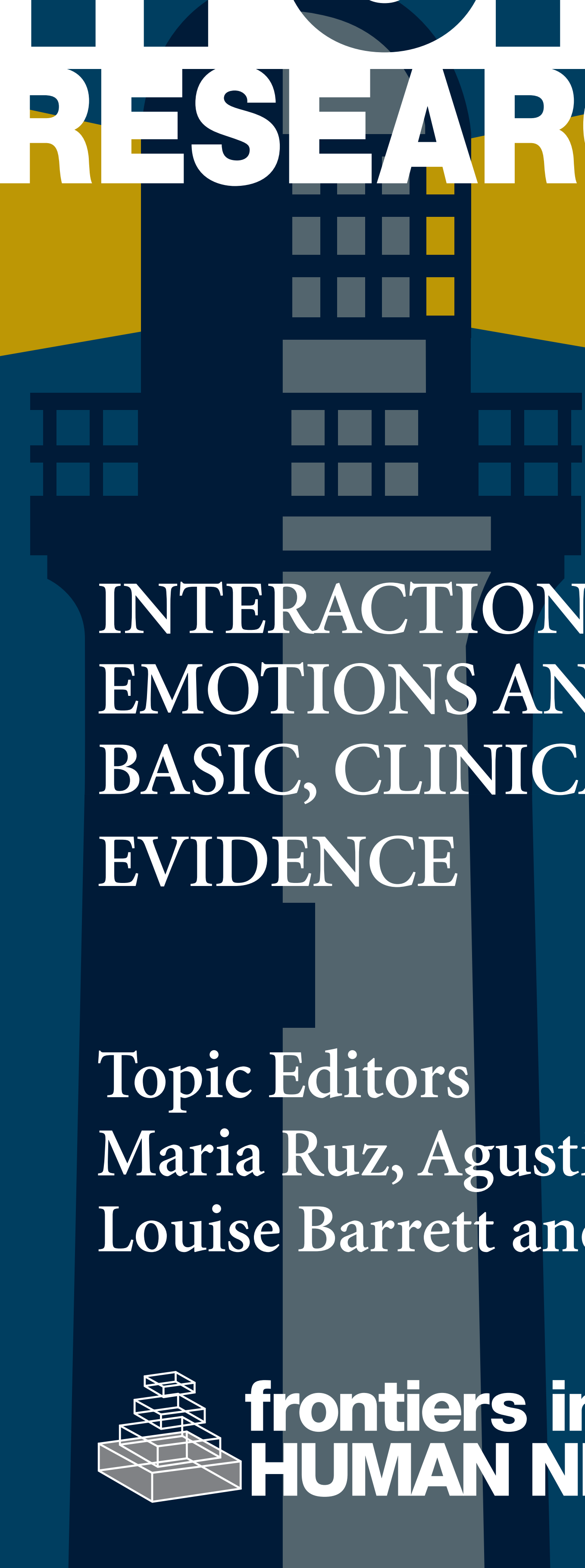
**frontiers**

**RESEARCH TOPICS**



# INTERACTIONS BETWEEN EMOTIONS AND SOCIAL CONTEXT: BASIC, CLINICAL AND NON-HUMAN EVIDENCE

Topic Editors
Maria Ruz, Agustin Ibanez, Sonja A. E. Kotz,
Louise Barrett and Jorge Moll

**frontiers in**
**HUMAN NEUROSCIENCE**

**FRONTIERS COPYRIGHT STATEMENT**

© Copyright 2007-2014
Frontiers Media SA.
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

Cover image provided by Ibbl sarl, Lausanne CH

**ISSN** 1664-8714
**ISBN** 978-2-88919-319-6
**DOI** 10.3389/978-2-88919-319-6

## ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

# INTERACTIONS BETWEEN EMOTIONS AND SOCIAL CONTEXT: BASIC, CLINICAL AND NON-HUMAN EVIDENCE

Topic Editors:
**Maria Ruz,** Universidad de Granada, Spain
**Agustin Ibanez,** UDP-INECO Foundation Core on Neuroscience (UIFCoN), Diego Portales University, Santiago, Chile; Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive Neurology (INECO), Favaloro University, Buenos Aires, Argentina; Universidad Autónoma del Caribe, Barranquilla, Colombia; National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina; Centre of Excellence in Cognition and its Disorders, Australian Research Council (ACR), Sydney, Australia
**Sonja A. E. Kotz,** Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; School of Psychological Sciences, University of Manchester, Manchester, UK
**Louise Barrett,** University of Lethbridge, Canada
**Jorge Moll,** D'Or Institute for Research and Education (IDOR), Brazil

The emotions that we feel and also those that we perceive in others are crucial to the social functioning of both humans and non-human animals. Although the role of context has been extensively studied in basic sensory processing, its relevance for social cognition and emotional processing is little understood.

In recent years, several lines of research at the behavioral and neural levels have highlighted the bidirectional interactions that take place between emotions and social context. Experienced emotions, even when incidental, bias decision-making. Remarkably, even basic emotions can be strongly influenced by situational contexts. In addition, both humans and non-human animals can use emotional expressions strategically as a means of influencing and managing the behavioral response of others in relation to specific environmental situations. Moreover, social emotions (e.g., engaged in moral judgment, empathic concern and social norms) seem to be context-dependent, which also questions a purely abstract account of emotion understanding and expression, as well as other social cognition domains.

The present Research Topic of Frontiers in Human Neuroscience highlights the need for a situated approach to emotion and social cognition. We presented theoretical and empirical work at the behavioral and neural levels that contribute to our understanding of emotion within a highly contextualized social realm, and vice-versa. Relevant contributions are presented from diverse fields, including ethology, neurology, biology, cognitive and social neuroscience, and as well as psychology and neuropsychiatry. This integrated approach that entails the interaction between emotion and social context provide important new insights into the growing field of social neuroscience.

# Table of Contents

frontiers in
HUMAN NEUROSCIENCE

EDITORIAL
published: 28 July 2014
doi: 10.3389/fnhum.2014.00547

# Situated affective and social neuroscience

**Agustin Ibanez[1,2,3,4,5]*, Sonja A. Kotz[6,7], Louise Barrett[8], Jorge Moll[9] and Maria Ruz[10]**

[1] Laboratory of Experimental Psychology and Neuroscience, Institute of Cognitive Neurology (INECO), Favaloro University, Buenos Aires, Argentina
[2] National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
[3] UDP-INECO Foundation Core on Neuroscience, Diego Portales University, Santiago, Chile
[4] Universidad Autónoma del Caribe, Barranquilla, Colombia
[5] Centre of Excellence in Cognition and its Disorders, Australian Research Council (ACR), Sydney, NSW, Australia
[6] Cognitive Neuroscience and Experimental Psychology Section, School of Psychological Sciences, University of Manchester, Manchester, UK
[7] Department of Neuropsychology, Max Planck Institute of Human Cognitive and Brain Sciences, Leipzig, Germany
[8] Department of Psychology, University of Lethbridge, Lethbridge, AB, Canada
[9] Cognitive and Behavioral Neuroscience Unit, D'Or Institute for Research and Education, Rio de Janeiro, Brazil
[10] Department of Experimental Psychology, Brain, Mind and Behavior Research Center, University of Granada, Granada, Spain
*Correspondence: aibanez@ineco.org.ar

**Edited and reviewed by:**
Hauke R. Heekeren, Freie Universität Berlin, Germany

**Keywords: social neuroscience, contextual social cognition, emotions, neuropsychiatry, embodied cognition, social behavior, emotion regulation, social decision making**

This Research Topic features several papers tapping the situated nature of emotion and social cognition processes. The volume covers a broad scope of methodologies [behavioral assessment, functional magnetic resonance imaging (fMRI), structural neuroimaging, event-related potentials (ERPs), brain connectivity, and peripheral measures], populations (non-human animals, neurotypical participants, developmental studies, and neuropsychiatric and pathological conditions), and article types (original research, review papers, and opinion articles). Through this wide-ranging proposal, we introduce a fresh approach to the study of contextual effects in emotion and social cognition domains.

We report four levels of evidence. First, we present studies examining how cognitive and neural functions are influenced by basic affective processes (interoception, motivation and reward, emotional impulsiveness, and appraisal of violent stimuli). A second set of behavioral and neuroscientific studies addresses how performance is modulated by different emotional variables (categorical and dimensional approaches to emotion, language-as-context for emotion, emotional suppression of the attentional blink, and reappraisal effects on the up-regulation of emotions). The studies in our third selection deal with different influences in social cognition (SC) domains (human and non-human comparative studies, long-term effects of social and physical stress, developmental theory of mind, neural bases of passionate love for others, social decision making in normal and psychopathic participants, and frontal lobe contributions to psychosocial adaptation models). Finally, the fourth set of papers investigates the blending of social and emotion-related processes (valence and social salience in amygdala networks, emotional contributions to identification of genuine and faked social expressions, emotional predispositions and social decision making bias, valence of fairness and social decisions, structural neuroimaging of emotional and social impairments in neurodegenerative diseases, and subjective reactivity to emotional stimuli and their association with moral cognition). A brief summary of all these studies is offered in the following sections.

## BASIC AFFECTIVE MODULATION

Body signals, especially of the interoceptive cardiac type, have been recently claimed to modulate emotion and decision-making processes. Leone et al. (2012) used chess decisions to analyze heart rate (HR) modulations in specific cognitive events. HR signals predicted the conception of a plan and the likelihood to blunder by fluctuations (e.g., performing random errors or bad moves). Such signals also reflected reactions, such as a blunder made by the opponent or fluctuations after a move. These data suggest that body signals are rich enough to reveal relevant episodes of inner decisions.

In another study, the affective motivational dimension of behavioral inhibition was assessed through the manipulation of reward magnitudes during a classical inhibitory task (Herrera et al., 2014). The effect of reward magnitude and context on behavioral inhibition in humans showed that dynamical behavioral inhibition depends on contextual parameters (reward magnitude modulation and initial reward history).

Torres et al. (2013) tested whether emotional and non-emotional dimensions of impulsiveness were differentially predictive of decision-making and addictive behavior in cocaine-dependent individuals (CDIs), pathological gamblers (PGs), and healthy controls. They used several instruments, including a Go/No-go paradigm assessed with ERPs and a delay-discounting task. Among the dimensions of trait impulsiveness, negative urgency was unique at independently covarying with gambling in PGs. Relative to these subjects, CDIs performed more poorly and showed ERP abnormalities. The effects of impulsiveness in negative emotion processing played a key role in decision-making and addiction.

Using fMRI, Porges and Decety (2013) evaluated the appraisal of violent stimuli and their relation with self-report measures of pleasure/displeasure. Participants watched video-clips depicting Mixed Martial Arts (MMA). Capoeira videos were used as a baseline. *Pleasurable* ratings of MMA predicted increased functional connectivity (FC) seeded in the nucleus accumbens, anterior cingulate cortex (ACC), and anterior insular cortex (AIC). These structures are related with positive/negative outcomes as well as feelings and somatic representations. Instead, *displeasure* ratings of MMA were related to increased FC among regions of the prefrontal cortex and superior parietal lobule (areas involved in cognitive control and executive attention). The results suggest that FC indexed the relationship between subjective feelings and anticipation of positive and negative outcomes.

## EMOTIONAL APPRAISAL

Matsuda et al. (2013) tested the hypothesis that separate neural loci might intrinsically encode categorical and dimensional facial emotion perception. Participants were scanned with fMRI while they passively viewed emotional faces and performed unrelated tasks. Activity in the right fusiform face area (FFA) was dependent on the categorical ambiguity of facial expressions. The amygdala, insula, and medial prefrontal cortex evidenced dimensional (linear) processing, which correlated with physical changes in expressions. The results suggest that distinct neural loci process the physical and psychological aspects of facial emotion perception in a region-specific and implicit manner.

Herbert et al. (2013) used words as contextual cues for emotion processing in two ERP experiments. They focused on self- vs. sender-related emotional pronoun-noun pairs (e.g., my fear vs. his fear) as cues for emotional face processing. Participants performed automatic (Experiment 1) and intentional (Experiment 2) affect labeling tasks. ERP patterns varied as a function of the label's reference (self vs. sender) and the intentionality of the labeling task (Experiment 1 vs. Experiment 2). Emotion decoding from facial expressions was not fully determined by sensory facial information, but proved sensitive to contextual factors and the perceiver's experience. These findings support a differentiated view of language-as-context for emotion processing.

The study conducted by Kanske et al. (2013) evaluated whether the attentional blink effect in rapid serial visual presentations is modulated by the emotionality of the stimuli (emotional and neutral images depicting social scenes as target). To this end, the authors used ERP recordings and offline self-reports of empathy. The results revealed enhanced performance for emotional stimuli and increased P3 amplitudes, which correlated with individual differences in empathy. These data suggest that empathy is associated with enhanced emotional processing in social contexts, even during unconscious target detection.

Peng et al. (2013) examined description-based reappraisal effects on the up-regulation of positive emotions. They measured ERP fluctuations as Chinese participants viewed erotic and neutral images shown after either a neutral or positive description. Further data was obtained through self-reported ratings. The results demonstrate that description-based reappraisal significantly modulated the emotional experience and ERP responses to erotic as well as neutral images.

## SOCIAL COGNITION

The review by Van Den Bos et al. (2013) considers animal and human studies tapping the influence of social context on decision-making. From a causal and functional perspective, the authors advance methodological considerations to improve the experimental assessment of social factors in decision-making.

In a study with rats, Chaby et al. (2013) investigated how exposure to social and physical stress during adolescence affects adult decision-making, coping response, cognitive bias, and exploratory behavior. Compared to control animals, rats exposed to chronic unpredictable stress (e.g., isolation, crowding, cage tilt) evinced long-term behavioral and cognitive changes, including negative cognitive bias, altered coping response, and accelerated decision-making. The results showed that stress during adolescence has a long-term impact on behavior and cognition. The most salient effects concern ambiguous stimulus interpretation, behavioral response to adverse events, and decision-making strategies.

Calero et al. (2013) propose a novel approach to quantifying the scaling property of theory of mind (ToM). Focusing on children between 6- and 8-years-old, they consider a scaling complexity of skills and their modulation by varied factors, such as gender, number of siblings, and personality traits.

The meta-analysis by Juan et al. (2013) considers a decade worth of fMRI studies to identify differential brain areas and cortical networks involved in (i) passionate love for others and (ii) understanding the intention of others' actions. Thus, this approach goes beyond classical experimental studies regarding individuals as strictly isolated entities. Both overlapping and distinct cortical and subcortical regions were identified for intention and love, respectively. By targeting these brain regions in future research, scientists and clinicians could promote breakthroughs in the neuroscience of pair-bonding.

Radke et al. (2013) investigated fairness considerations in psychopathic and non-psychopathic offenders as well as healthy controls. In a modified Ultimatum Game (UG) involving opposing intentionality constraints (intentional vs. unintentional), unfair offers were paired with different unselected alternatives, thereby establishing the context of a proposal. Psychopathic offenders resembled healthy controls in their rejection pattern—i.e., they took the unselected alternative into account. In contrast, non-psychopathic delinquents failed to adjust their decisions to an offer's alternatives, suggesting stronger impairments in social decision-making. Crucially, the mechanisms and processes underlying rejection decisions might differ in both groups, particularly in terms of cognitive vs. emotional competencies.

In an ERP experiment, Moser et al. (2014) investigated the levels of processing at which positive and negative descriptions of other people bias social decision-making. Participants played a game in which they had to accept or reject economic offers. Other variables manipulated were the fairness of the assets' distribution, the offers' advantageousness, and the game context's uncertainty. Negative description of the interaction partner enhanced medial frontal negativity (MFN) in an additive manner with fairness evaluations. The description of the partner interacted with personal benefit considerations, showing that this positive or negative information biased the evaluation of offers only when they did not favor the participant. P300 amplitudes were enhanced by advantageous offers, suggesting their heightened motivational significance at later stages of processing. In all phases of the study, processing of the offer was increased in the certain, as compared to the uncertain, contexts. These results provide new evidence that decision-making is influenced by interpersonal information and considerations of one's own interests relative to those of others.

Finally, Huepe and Salas (2013) set forth a new conceptualization of the prefrontal cortex for psychosocial adaptation. Their review of the evidence suggests that cognitive functions related to this lobule include fluid intelligence (FI), SC, and perspective changing abilities (PCA). These domains are crucial in adapting to social contexts and solving problems in new situations. Moreover, they appear to depend on contextual keys, thus requiring flexibility—yet another function associated with the frontal lobe. The model proposed integrates these components (FI, SC, and PCA) as indicators of psychosocial adaptation in contexts of social vulnerability or impoverished social/cultural conditions.

## CONTEXTUAL BLENDING OF SOCIAL AND EMOTION-RELATED PROCESSES

Vrtička et al. (2013) assessed whether the human amygdala preferentially responds to both emotionally and socially significant information, and whether these factors might display interactive encoding properties. Through an fMRI study, they demonstrated that amygdala activation is (1) greater for neutral social vs. non-social information, (2) similar for positive and negative social images, and (3) sensitive to a valence effect (negative vs. positive) for non-social images. The valence × social content interaction was also found in the right fusiform gyrus, right anterior superior temporal gyrus, and medial orbitofrontal cortex. Overall, these findings suggest that valence and social contents possess distinct kinds of relevance that interact within the human amygdala and throughout a more extensive cortical network.

The ability to discriminate between felt and faked expressions is a crucial social skill. Manera et al. (2013) investigated whether individual differences in smile authenticity recognition are explained by distinct predispositions to experience other people's emotions (susceptibility to emotional contagion). Susceptibility to emotional contagion for negative emotions increased smile authenticity detection. Instead, susceptibility to emotional contagion for positive emotions worsened detection performance, because it led to categorize most faked smiles as sincere. It follows that susceptibility to emotional contagion plays a key role in complex social emotions.

The study by Klapwijk (2013) examined the effects of three different emotional responses (anger, disappointment, and happiness) on social decision-making in adolescents. In a version of the Dictator Game, unfair offers by the participants received emotional responses from peers. Relative to angry and happy reactions, expressions of disappointment prompted more generous offers. Older adolescents were better than younger adolescents at differentiating among the three emotions. In addition, individual differences in social value orientation played a role in decisions after happy reactions to unfair offers. Thus, adolescents take into account the emotions of their peers when making social decisions and are affected by social value orientation and age.

Couto et al. (2013) report selective behavioral impairments of face recognition, emotion recognition, and ToM in patients with behavioral variant frontotemporal dementia (bvFTD) and progressive non-fluent aphasia (PNFA). Voxel-based morphometry revealed fronto-temporo-insular atrophy in both patient groups. SC deficits were differentially associated to fronto-insular-temporal atrophy in bvFTD and PNFA, respectively. While SC impairments were similar in both groups, they seem to reflect intrinsic ToM affectation in bvFTD and more basic deficits (face and emotion recognition) in PNFA.

Carmona-Perera et al. (2013) examined subjective reactivity to emotional stimuli and its possible association with moral decision-making. Healthy adult participants responded to a set of moral and non-moral dilemmas. The researchers focused on emotional experience in valence, arousal, and dominance dimensions in response to different types of pictures (neutral, pleasant, unpleasant non-moral, and unpleasant moral). Significant correlations emerged between less unpleasantness to negative stimuli, more pleasantness to positive stimuli, and a higher proportion of utilitarian choices. Also, a positive association was found between higher arousal ratings to negative moral laden pictures and more utilitarian choices. Low dominance was associated with greater perceived difficulty over moral judgment. These results evidenced a contextual role of emotional experience in moral choice.

## CONCLUSIONS

Despite the diversity of their topics, research questions, and methodologies, most of these studies highlight the contextual situatedness of emotional and social cognition processes (Garrido-Vasquez et al., 2011; Ruz and Tudela, 2011; Ibanez and Manes, 2012; Melloni et al., 2014). Moreover, they provide new evidence for the interaction among low and high-level cognition, emotion, and social domains (Moll and Schulkin, 2009; Pessoa, 2009; Alguacil et al., 2013; Ibanez et al., 2013, 2014; Ruz et al., 2013; Baez et al., 2014b). In the same vein, part of the evidence presented shows that our emotional arousal biases our decisions in the social world (Beauregard, 2007; Heatherton, 2011). More generally, this Research Topic indicates that a brain network approach to social and emotional processes (Moll et al., 2005, 2008; Kennedy and Adolphs, 2012; Baez et al., 2014a) seems more adequate than simple approximations ascribing such complex domains to a single region. This integrated approach to embedded emotional and social processes provides exciting new avenues into the growing field of social neuroscience.

## ACKNOWLEDGMENTS

Agustin Ibanez is supported by CONICET, CONICYT/FONDECYT Regular (1130920 and 1140114), FONCyT-PICT 2012-0412/2012-1309, and INECO Foundation. Maria Ruz is supported by the Spanish Ministry of Science and Innovation, "Ramón y Cajal" fellowship (RYC-2008-03008) and grant PSI2013-45567-P. Jorge Moll is supported by intramural grants, D'Or Institute for Research and Education, and FAPERJ (Rio de Janeiro State Foundation for Research). Sonja A. Kotz is supported by Canadian Institutes of Health Research (CIHR: 62867) and German Science Foundation (KO-2268/6-1).

## REFERENCES

Alguacil, S., Tudela, P., and Ruz, M. (2013). Cognitive and affective control in a flanker word task: common and dissociable brain mechanisms. *Neuropsychologia* 51, 1663–1672. doi: 10.1016/j.neuropsychologia.2013.05.020

Baez, S., Couto, B., Torralva, T., Sposato, L., Huepe, D., Montañes, P., et al. (2014a). Comparing the moral judgments of frontotemporal dementia and frontal stroke patients. *JAMA Neurol*. doi: 10.1001/jamaneurol.2014.347

Baez, S., Marengo, J., Perez, A., Huepe, D., Font, F. G., Rial, V., et al. (2014b). Theory of mind and its relationship with executive functions and emotion recognition in borderline personality disorder. *J. Neuropsychol*. doi: 10.1111/jnp.12046. [Epub ahead of print].

Beauregard, M. (2007). Mind does really matter: evidence from neuroimaging studies of emotional self-regulation, psychotherapy, and placebo effect. *Prog. Neurobiol.* 81, 218–236. doi: 10.1016/j.pneurobio.2007.01.005

Calero, C. I., Salles, A., Semelman, M., and Sigman, M. (2013). Age and gender dependent development of Theory of Mind in 6- to 8-years old children. *Front. Hum. Neurosci.* 7:281. doi: 10.3389/fnhum.2013.00281

Carmona-Perera, M., Martí-García, C., Pérez-García, M., and Verdejo-García, A. (2013). Valence of emotions and moral decision-making: increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults. *Front. Hum. Neurosci.* 7:626. doi: 10.3389/fnhum.2013.00626

Chaby, L. E., Cavigelli, S. A., White, A., Wang, K., and Braithwaite, V. A. (2013). Long-term changes in cognitive bias and coping response as a result of chronic unpredictable stress during adolescence. *Front. Hum. Neurosci.* 7:328. doi: 10.3389/fnhum.2013.00328

Couto, B., Manes, F., Montañés, P., Matallana, D., Reyes, P., Velasquez, M., et al. (2013). Structural neuroimaging of social cognition in progressive nonfluent aphasia and behavioral variant of frontotemporal dementia. *Front. Hum. Neurosci.* 7:467. doi: 10.3389/fnhum.2013.00467

Garrido-Vasquez, P., Jessen, S., and Kotz, S. A. (2011). Perception of emotion in psychiatric disorders: on the possible role of task, dynamics, and multimodality. *Soc. Neurosci.* 6, 515–536. doi: 10.1080/17470919.2011.620771

Heatherton, T. F. (2011). Neuroscience of self and self-regulation. *Annu. Rev. Psychol.* 62, 363–390. doi: 10.1146/annurev.psych.121208.131616

Herbert, C., Sfärlea, A., and Blumenthal, T. (2013). Your emotion or mine: labeling feelings alters emotional face perception—an ERP study on automatic and intentional affect labeling. *Front. Hum. Neurosci.* 7:378. doi: 10.3389/fnhum.2013.00378

Herrera, P. M., Speranza, M., Hampshire, A., and Bekinschtein, T. N. A. (2014). Monetary rewards modulate inhibitory control. *Front. Hum. Neurosci.* 8:257. doi: 10.3389/fnhum.2014.00257

Huepe, D., and Salas, N. (2013). Fluid intelligence, social cognition, and perspective changing abilities as pointers of psychosocial adaptation. *Front. Hum. Neurosci.* 7:287. doi: 10.3389/fnhum.2013.00287

Ibanez, A., Aguado, J., Baez, S., Huepe, D., Lopez, V., Ortega, R., et al. (2014). From neural signatures of emotional modulation to social cognition: individual differences in healthy volunteers and psychiatric participants. *Soc. Cogn. Affect. Neurosci.* 9, 939–950. doi: 10.1093/scan/nst067

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., and Toledo, M. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Pers. Individ. Dif.* 54, 616–621. doi: 10.1016/j.paid.2012.11.022

Ibanez, A., and Manes, F. (2012). Contextual social cognition and the behavioral variant of frontotemporal dementia. *Neurology* 78, 1354–1362. doi: 10.1212/WNL.0b013e3182518375

Juan, E., Frum, C., Bianchi-Demicheli, F., Wang, Y.-W., Lewis, J. W., and Cacioppo, S. (2013). Beyond human intentions and emotions. *Front. Hum. Neurosci.* 7:99. doi: 10.3389/fnhum.2013.00099

Kanske, P., Schönfelder, S., and Wessa, M. (2013). Emotional modulation of the attentional blink and the relation to interpersonal reactivity. *Front. Hum. Neurosci.* 7:641. doi: 10.3389/fnhum.2013.00641

Kennedy, D. P., and Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends Cogn. Sci.* 16, 559–572. doi: 10.1016/j.tics.2012.09.006

Klapwijk, E. (2013). Emotional reactions of peers influence decisions about fairness in adolescence. *Front. Hum. Neurosci.* 7:745. doi: 10.3389/fnhum.2013.00745

Leone, M. J., Petroni, A., Fernandez Slezak, D., and Sigman, M. (2012). The telltale heart: heart rate fluctuations index objective and subjective events during a game of chess. *Front. Hum. Neurosci.* 6:273. doi: 10.3389/fnhum.2012.00273

Manera, V., Grandi, E., and Colle, L. (2013). Susceptibility to emotional contagion for negative emotions improves detection of smile authenticity. *Front. Hum. Neurosci.* 7:6. doi: 10.3389/fnhum.2013.00006

Matsuda, Y.-T., Fujimura, T., Katahira, K., Okada, M., Ueno, K., Cheng, K., et al. (2013). The implicit processing of categorical and dimensional strategies: an fMRI study of facial emotion perception. *Front. Hum. Neurosci.* 7:551. doi: 10.3389/fnhum.2013.00551

Melloni, M., Lopez, V., and Ibanez, A. (2014). Empathy and contextual social cognition. *Cogn. Affect. Behav. Neurosci.* 14, 407–425. doi: 10.3758/s13415-013-0205-3

Moll, J., De Oliveira-Souza, R., and Zahn, R. (2008). The neural basis of moral cognition: sentiments, concepts, and values. *Ann. N.Y. Acad. Sci.* 1124, 161–180. doi: 10.1196/annals.1440.005

Moll, J., and Schulkin, J. (2009). Social attachment and aversion in human moral cognition. *Neurosci. Biobehav. Rev.* 33, 456–465. doi: 10.1016/j.neubiorev.2008.12.001

Moll, J., Zahn, R., De Oliveira-Souza, R., Krueger, F., and Grafman, J. (2005). Opinion: the neural basis of human moral cognition. *Nat. Rev. Neurosci.* 6, 799–809. doi: 10.1038/nrn1768

Moser, A., Gaertig, C., and Ruz, M. (2014). Social information and personal interests modulate neural activity during economic decision-making. *Front. Hum. Neurosci.* 8:31. doi: 10.3389/fnhum.2014.00031

Peng, J., Qu, C., Gu, R., and Luo, Y.-J. (2013). Description-based reappraisal regulate the emotion induced by erotic and neutral images in a Chinese population. *Front. Hum. Neurosci.* 6:355. doi: 10.3389/fnhum.2012.00355

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends Cogn. Sci.* 13, 160–166. doi: 10.1016/j.tics.2009.01.006

Porges, E. C., and Decety, J. (2013). Violence as a source of pleasure or displeasure is associated with specific functional connectivity with the nucleus accumbens. *Front. Hum. Neurosci.* 7:447. doi: 10.3389/fnhum.2013.00447

Radke, S., Brazil, I. A., Scheper, I., Bulten, B. H., and De Bruijn, E. R. A. (2013). Unfair offers, unfair offenders? Fairness considerations in incarcerated individuals with and without psychopathy. *Front. Hum. Neurosci.* 7:406. doi: 10.3389/fnhum.2013.00406

Ruz, M., Madrid, E., and Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Soc. Cogn. Affect. Neurosci.* 8, 838–844. doi: 10.1093/scan/nss080

Ruz, M., and Tudela, P. (2011). Emotional conflict in interpersonal interactions. *Neuroimage* 54, 1685–1691. doi: 10.1016/j.neuroimage.2010.08.039

Torres, A., Catena, A., Megías, A., Maldonado, A., Cándido, A., Verdejo-García, A., et al. (2013). Emotional and non-emotional pathways to impulsive behavior and addiction. *Front. Hum. Neurosci.* 7:43. doi: 10.3389/fnhum.2013.00043

Van Den Bos, R., Jolles, J. W., and Homberg, J. R. (2013). Social modulation of decision-making: a cross-species review. *Front. Hum. Neurosci.* 7:301. doi: 10.3389/fnhum.2013.00301

Vrtička, P., Sander, D., and Vuilleumier, P. (2013). Lateralized interactive social content and valence processing within the human amygdala. *Front. Hum. Neurosci.* 6:358. doi: 10.3389/fnhum.2012.00358

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 12 June 2014; accepted: 05 July 2014; published online: 28 July 2014.*
*Citation: Ibanez A, Kotz SA, Barrett L, Moll J and Ruz M (2014) Situated affective and social neuroscience. Front. Hum. Neurosci.* ***8****:547. doi: 10.3389/fnhum.2014.00547*
*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 08 October 2012
doi: 10.3389/fnhum.2012.00273

# The tell-tale heart: heart rate fluctuations index objective and subjective events during a game of chess

***María J. Leone*[1]*, Agustín Petroni*[1]*, Diego Fernandez Slezak*[2] *and Mariano Sigman*[1]**

[1] *Physics Department, School of Sciences, University of Buenos Aires, Buenos Aires, Argentina*
[2] *Computer Science Department, School of Sciences, University of Buenos Aires, Buenos Aires, Argentina*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*Sebastian J. Lipina, Unidad de Neurobiología Aplicada (UNA, CEMIC-CONICET), Argentina*
*Guillermo Solovey, Columbia University, USA*

***Correspondence:***
*María J. Leone, Laboratorio de Neurociencia Integrativa, Departamento de Física, FCEyN UBA and IFIBA, Conicet, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina.*
*e-mail: juli.leone@gmail.com*

During a decision-making process, the body changes. These somatic changes have been related to specific cognitive events and also have been postulated to assist decision-making indexing possible outcomes of different options. We used chess to analyze heart rate (HR) modulations on specific cognitive events. In a chess game, players have a limited time-budget to make about 40 moves (decisions) that can be objectively evaluated and retrospectively assigned to specific subjectively perceived events, such as setting a goal and the process to reach a known goal. We show that HR signals events: it predicts the conception of a plan, the concrete analysis of variations or the likelihood to blunder by fluctuations before to the move, and it reflects reactions, such as a blunder made by the opponent, by fluctuations subsequent to the move. Our data demonstrate that even if HR constitutes a relatively broad marker integrating a myriad of physiological variables, its dynamic is rich enough to reveal relevant episodes of inner thought.

**Keywords: decision-making, cognitive processes, problem solving, heart rate, chess, planning, calculation**

## INTRODUCTION

The decision-making process is accompanied by modification in heart rate (HR). In the period before making a decision, HR and skin conductance changes have been related to specific cognitive events and to the load of mental work (Bradley, 2009; Jennings et al., 2009; Taelman et al., 2011); they also have been postulated to act as signals of the possible future outcomes of a decision (Damasio, 1994). After the decision, body changes have been related to the outcome of the decision (Crone et al., 2004).

One specific theory, the somatic marker hypothesis (SMH), postulates that body signals guide decision making in uncertain situations (Damasio, 1994). According to this theory, the appearance of a specific body state associated with a specific outcome previously learnt is hypothesized to signal the expected value of a choice (Tranel et al., 1999). This provides physiological evidence for what has been popularly referred as “hunches” or “gut-feelings” which provide a rapid approximate evaluation of a complex problem (Bechara and Damasio, 2005; Dijksterhuis et al., 2006; Ariely, 2008). However, this theory has been contested and there is heated controversy about the specific relation and causality of body signals and decision making (Dunn et al., 2006).

One of the most widely used setups to investigate the interaction of emotions, decision making and body signals is the Iowa Gambling Task (IGT), a card game where subjects have to choose between four decks to maximize their money gain (Bechara et al., 1994). Changes in skin conductance and HR have been showed to predict performance on the IGT (Bechara et al., 1997; Crone et al., 2003; Drucaroff et al., 2011) even before this knowledge is expressed as conscious rational thought (Bechara et al., 1997). However, the necessity of conscious knowledge has been questioned (Maia and McClelland, 2004).

Compared to the IGT and other simple decision-making protocols used in laboratory setups, the game of chess constitutes a very rich and quantitative model of real life decisions, with a virtually infinite number of states and paths. First, players make successive decisions (around 40 each) on a finite time-budget and every decision (move) can be accurately evaluated with current algorithms. Second, player expertise level can be accurately assessed (Elo, 1978; Van Der Maas and Wagenmakers, 2005). Third, players can recognize specific events of their inner thinking during the game (like planning, calculation, and error moments), which explains why chess has been a goldmine for studies of introspection (De Groot, 1965). Last but not least, this game is played in a social setup in which the relation to the other opponent sets a regulatory focus which governs the type of play (Slezak and Sigman, 2011) and constitutes a highly motivating setup. Throughout the game, players undergo strong emotional fluctuations.

Here we have studied HR variation as a physiological correlate of decision making using rapid chess as a natural experimental setup. All subjects were expert chess players, and rating differences between players were minimized to avoid opponent level-related effects (Slezak and Sigman, 2011). This time budget (15 min per player) is set as a compromise to generate move durations which are fast enough to investigate transitions in HR but also sufficiently slow to allow a player to retrospectively recall relevant moments perceived and experimented during the game. Our aim was to investigate which aspects of HR index objective variables (the quality of a move, determined by the change in the objective evaluation of the position) and subjective reports such as the conception of a plan or a moment of calculus, as reported by the player in an after game recollection of its inner thought.

## MATERIALS AND METHODS

### PARTICIPANTS

Twenty-five games were played by nine different-subjects (one to five games each). Twenty-five independent games were played, 19 by men and 6 by women, mean age 35.6 ± 11.7 years old (age range: 21–58), mean international rating (Elo) 2111 ± 60.4 (Elo range: 2021–2216). Eight games were played in a special tournament with electronic chess boards and clocks (DGT), and the rest were played using a computer. For experimented players, there should be not differences between these modalities.

### EXPERIMENT DESIGN

In all games we recorded 2 min of rest before and after the games (except the tournament games for which we recorded rest only after the game). Each chess game lasted at most 30 min.

After the final rest, players were asked to complete a metacognitive questionnaire (Appendix). Players reported moments in which they were engaged in establishing a plan (planning) and moments in which they were engaged with depth search, examining and evaluating concrete tactical variations (calculation). Here we use the chess convention, where planning refers to the process of setting a goal, a strategic and general aim (De Groot, 1965; Kotov, 1971). In General Problem Solving, planning often refers to explicit examination of the process to reach a known goal, i.e., the evaluation of a tree of variations, which here, as in chess, is called calculation. They filled the form including the specific move in chess algebraic notation only on those fields that they could recognize and remember from the game, not from a retrospective evaluation of the position.

### DATA ACQUISITION AND PREPROCCESING

Electrocardiogram (ECG) activity was recorded using two external electrodes on a Biosemi Active-Two system (Biosemi, Amsterdam, Holland) with a sample rate of 256 Hz (electrode location: one on the left chest and the other on the sternum). ECG data were filtered between 1 and 50 Hz, and after a global visual inspection a threshold was set to detect peaks on the ECG signal (R peaks). Signal was then converted in instantaneous HR by interpolation and referred to the mean HR of each game.

Computer games were played using JinChess (http://www.jinchess.com/), an open-source chess client which connects to a server for playing chess through Internet (FICS, Free Internet Chess Server, http://www.freechess.org/). To control network lag, we used JinChess with timeseal, a program that act as a relay station and keeps track of transmission times. To synchronize the games with the ECG signal, we modified the JinChess code, to register and save all relevant tags of the game. This signal was sent to the Active-Two system through the parallel port, identifying each event with a different 16-bit code.

### CHESS DATA

#### *Time variables*

For each move we recorded player and opponent available times (AT), and the time it takes to make the move, defined for consistency with psychological experiments as response time (RT). We also defined the time after move (TAM) as the time between a move and the following opponent move. For all games, AT started in 900 s (15 min) and decreased during player's turn to play. In chess, each player has its own clock which stops during the other player's turn. If a player uses all his/her AT, the game is over (player lost by time). When AT gets close to a few seconds players have to play very fast, a situation referred as time trouble.

#### *Score (S)*

Score is a measure of the value of the position in pawns units. It can be seen as an estimate of the likelihood of the final result. We used the Rybka 4 engine to calculate chess moves score, using a 12 movements depth (Sigman et al., 2010). $S > 0$ indicates a white player advantage and $S < 0$, black player advantage. Score was saturated in +10 and −10. For simplicity and consistency of data presentation, we calculated a player corrected score whose sign indicates the goodness of the recorded player's position: positive values when he/she had advantage, independently on if he/she is playing with white or black pieces.

#### *Delta score (ΔS)*

The change in the position value (score, not player corrected) is a measure of the move goodness defined as $\Delta S = [S(i+1) - S(i)] \times C$, where $C$ is –1 for black moves and +1 for white ones. As with the score, $C$ is just a correction variable to measure $\Delta S$ relative to the player independently of piece colors. Close to or zero values of $\Delta S$ indicate that player made a good move. Significant negative departures of $\Delta S$ from zero indicate that the player move was far from the best. We defined moves with $\Delta S \leq -1$ as blunders (errors or bad moves). $\Delta S > 0$ values indicate that the player made a move that was better than all the ones conceived by the engine. Since we use an engine much stronger than all our players, this is very infrequent (see Sigman et al., 2010).

#### *Phases*

Chess games were classified in three conventional phases: Opening, Middle game, and Endgame by author (MJL) who is a Woman International Master (WIM). Phases were determined according to the piece distribution in the board. For instance, opening was finished when piece development was completed (not according to theorical knowledge). Although the precise transition between two phases (for instance the end of the opening and beginning of middle game) might be controversial, none of the analysis reported here is sensitive to slight changes in this criterion.

#### *Move statistics*

Two thousand and eighty-six moves were obtained from the 25 games (Opening: 565, Middle game: 1007, Endgame: 514). Eight hundred and sixteen of these moves had RT and TAM ≥5 s. Blunders ($\Delta S \leq -1$) were 153 (recorded players: 68; opponents: 85). Players identified 26 planning and 41 calculation moves across all games (8 moves were highlighted as both planning and calculation).

### HR DYNAMICS

We analyzed HR dynamics around moves in a 10 s time-window centered in the move. The baseline for each move was defined between 5 and 3 s before the move, and subtracted. We used a strict criterion to avoid wrapping artifacts and contamination by move overlapping, considering only those moves with RT and TAM of at least 5 s.

#### *Matching*

To analyze the effect of a move category (player blunder, opponent blunder, planning, and calculation) on HR, we matched other variables, to assure that the results were not accounted by covariations in the data. For instance, as the game proceeds, players have less AT, start playing faster and are more prone to make errors.

For every category with a small number of exemplar moves we found a matched category in the complementary group with other variables matched. For example, to investigate the effect of planning we first considered all moves where subjects reported a plan. This group of moves was much smaller than its complement which assured that in principle we could find sufficient non-planning moves with the same properties in other variables (AT, Score, etc.). If matching could not be made accurately, we only considered a subset of planning moves which could be adequately matched, through a random replacement procedure. Matching conditions were determined allowing a maximal difference between each exemplar move and its match in other variables. For player versus opponent moves, matched variables were as follows: player and opponent AT ($<30$ s), score ($<1$), and $\Delta S$ ($<0.5$). Planning and calculation moves were matched on player AT ($<30$ s), score ($<1$) and $\Delta S$ ($<0.5$). All these moves were also not blunder moves. For blunder versus non-blunder moves (both player and opponent) matching variables were player AT ($<30$ s) and score ($<1$).

The resulting number of moves for each category which could be matched for all other variables was as follows: player blunders $N = 24$, opponent blunder $N = 34$, planning $N = 15$, calculation $N = 17$.

### LINEAR CLASSIFIER ANALYSIS

We trained a support vector machine (SVM) algorithm (Cristianini and Shawe-Taylor, 2000) to test if HR could be used to classify a move as a target (a move defined by a category, as a planning or calculation move) or a non-target move (for each group target-matched moves) using a leave-four-out procedure. We used 300 independent iterations by randomly selecting the four exemplars not used in training and left for classification. For robustness of this procedure, we run the classifier five times with different matched moves for each target group.

### STATISTICAL ANALYSIS

Correlation analysis was assessed using Pearson correlation test. Analysis of HR dynamics was carried out using Wilcoxon rank sum test comparing two groups. For each pair of conditions, we considered significantly different if $p < 0.05$ and if the point is part of a cluster of 64 points (250 ms time window).

## RESULTS

### HR VARIATIONS THROUGHOUT THE GAME

The evolution of chess-variables during a game followed an expected path, AT decreased from its initial budget of 900 s first slowly (opening moves are played fast) and in the middle game with sharp transitions revealing long moments of thought (**Figure 1A**). Score begun equal and showed moderate fluctuation in the opening stage. As the game proceeded, the likelihood of making an error increases due to shortage of time and complexity of the position, revealing larger fluctuations in score (**Figure 1B**).

HR increased steadily throughout the game (**Figure 1C**, average slope of HR versus fraction of game played: $6.90e - 04 \pm 3.18e - 04$, One-sample $t$-test, $p < 0.00001$), with its categorical equivalent, through the three stages of the game [One-Way ANOVA, $F_{(4,\ 113)} = 27.37$, $p < 0.0001$] (**Figure 1D**) and with AT (**Figure 1E**, $r = -0.6347$, $p < 0.00001$). This effect was very robust, every single game of the 25 studied here showed a negative linear correlation with AT (game slopes: $-3.88e - 04 \pm 2.05e - 04$, One-sample $t$-test, $p < 0.00001$). HR also showed a positive correlation with absolute score indicating that HR increases as the game imbalances in favor of one side (One-sample $t$-test of the regression coefficients obtained from each game, $p < 0.005$).

On summary, HR increased throughout the phases of the game, when less time is available and when score became unbalanced. These three variables are correlated, as shown in **Figures 1A,B**, and our data could not distinguish how these strongly correlated factors differentially contribute to HR since a multiple regression to these factors was highly unstable. However, the non-stationary nature of HR throughout the game must be carefully taken into account for a robust analysis directed to our main goal: understanding how transient events of the game (occurrence of plans, calculation, blunders) relate to HR fluctuations.

### TRANSIENT MODIFICATIONS OF HR

We analyzed HR dynamics in a 10 s time-window centered in the execution of the move. We used two procedures to assure that this analysis was not biased by non-stationarities of the data reported in the previous section. First, each move was normalized to its baseline, hence compensating for linear global trends. Second, to further compensate non-linear global trends, we performed a matching procedure (see “Materials and Methods”).

#### *HR dynamics in blunders and correct moves*

First, we simply compared HR dynamics on player versus opponent moves, excluding all blunders (**Figures 2A,B**). This comparison showed a significant difference between player and opponent moves from $-0.5$ to 5 s after the move ($p < 0.05$, see “Materials and Methods”).

HR responses locked to an opponent move were virtually flat until about 500 ms before the move. This anticipated response is not unexpected since a player can predict the timing of an opponent estimating the Hazard rate (Janssen and Shadlen, 2005) and also from gestures by the opponent. After the opponent move there is a change in HR with an effect size which peaks at about 0.02 Hz above the basal HR.

HR responses locked to the player's own moves showed a qualitatively different pattern. First there was a decrease in HR which started almost 3 s before the move. This trend did not reach significance. HR then ramped before the move reaching an almost threefold increase in modulation compared to opponent moves, peaking at 0.06 Hz modulation of baseline activity.

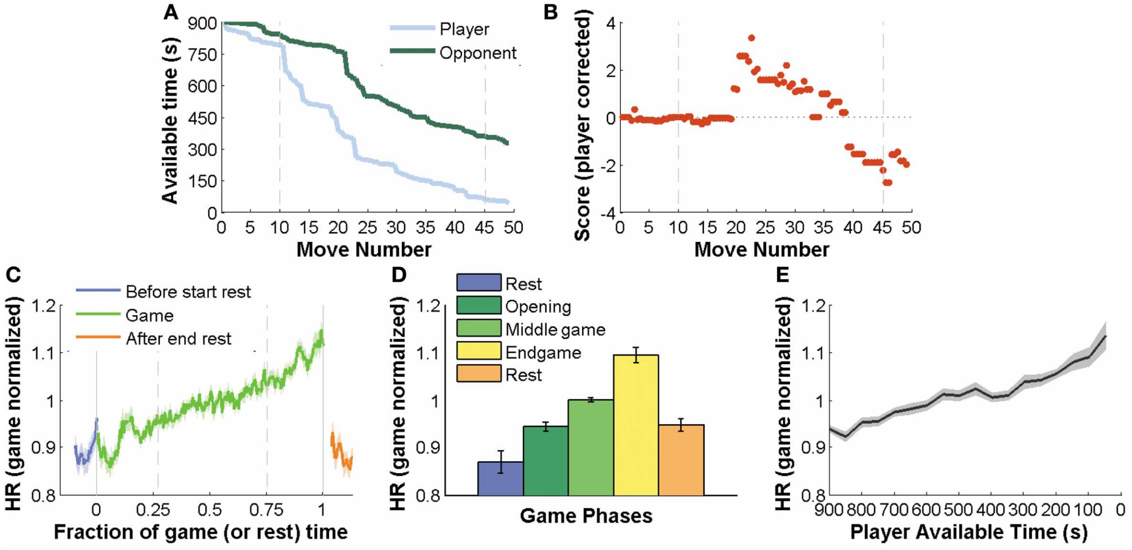


**FIGURE 1 | HR correlates with game state variables.** Main chess variables: player and opponent available time (AT), game phase and score. Game advance is represented by move number. For each move number there are two movements: one of the player and other of the opponent. **(A)** Player and opponent AT throughout a typical game. AT starts at 900 s (the budget for all the games in this experiment were 15 min per player) and decreases continuously. First it decreases slowly (opening moves are played fast) and as players devote more time to each move, AT decreases more rapidly. **(B)** Score progression throughout a typical game. Score is a measure of the value of each position. Positive or negative values indicate which player had advantage: positive values when the recorded player is better and negative for opponent advantage. **(C)** HR variations above the mean averaged for all games of the experiment. Before and after the game, we measured HR at rest. **(D)** Mean HR on both rest periods and game stages (opening, middle game, and endgame) for all games of the experiment. **(E)** HR as a function of AT averaged across all games. Vertical lines on panels **(A)** and **(B)** indicate stage transitions for this specific game. On panels **(C)**, vertical solid lines indicate rest to game transitions and vertical dashed lines represent the average time for each game stage transition.

Interestingly, the early deep prior to the move was the most sensitive to the contrast between blunders ($\Delta S \leq -1$) and correct moves $\Delta S \geq -0.3$) (**Figure 2C**). This modulation was virtually absent when the player blunders and was more pronounced when observing solely those trials in which there was not errors, but where errors were likely. We emphasize the difference between the light-blue-trace of **Figure 2B** (all own-moves, no blunders) with the light-blue-trace of **Figure 2C** (own moves, no blunders, but other variables matched to moves where blunders are made). The latter corresponded to a subset of the game, typically not including the opening, with less time available, where errors are more frequent, but selecting those cases in which errors were not made. In this specific filter which focused on difficult moments of the game, the early deep in HR was hence indicative, on average, of the quality of the move. In fact, this comparison (**Figure 2C**) revealed that only the $-2.5$ to $-1$ s interval showed a significant difference in HR for blunders versus non-blunders moves.

The comparison of blunders and non-blunders in opponent moves showed a very different pattern. Opponent blunders induced a higher HR increase than opponent non-blunder moves which was significant in the $-0.25$ to 2 s interval, almost entirely after the opponent move (**Figure 2D**).

### *HR dynamics in retrospectively reported cognitive events*

After the game, players reported in which moves they were engaged in the elaboration of a strategic plan or in calculation of variations. Players also reported their perceived errors and the moves that they ranked as especially good moves but we did not have sufficient records of these events to perform significant statistical analysis. For planning and calculation moves (**Figures 3A,B**) we found an increase in HR anticipating the move, compared to their respective matches. For planning moves, significant differences were found from $-1.5$ to 0 s (**Figure 3C**). For calculation moves, differences were found from $-2.5$ to $-0.5$ s (**Figure 3D**). Thus, both planning and calculation induced higher HR levels before the move.

## CLASSIFICATION ANALYSIS

Previous results pointed differences in mean HR for specific objective or subjectively labeled moves. In the vast majority of studies, this has been used as a statistical indicator that a variable (HR) indexes or informs about a state (blunder, calculation). Here we went beyond average difference estimators, observing the capacity of HR data to predict in a trial by trial basis, move observables such as its quality and the thought processes involved. It is clear that this analysis pushes the data to its limit since the HR is an intrinsically noisy signal. Specifically, we quantified the degree of separability of these temporal series training a linear decoder, using the SVM algorithm (Cristianini and Shawe-Taylor, 2000). We ran the classifier on data which were clearly before (3–1 s before the move) and after (1–3 s after the move) in the four comparisons described above. Classification was significant for all comparisons (**Table 1**), yielding classification values which were around 60% and revealing a trend which was consistent

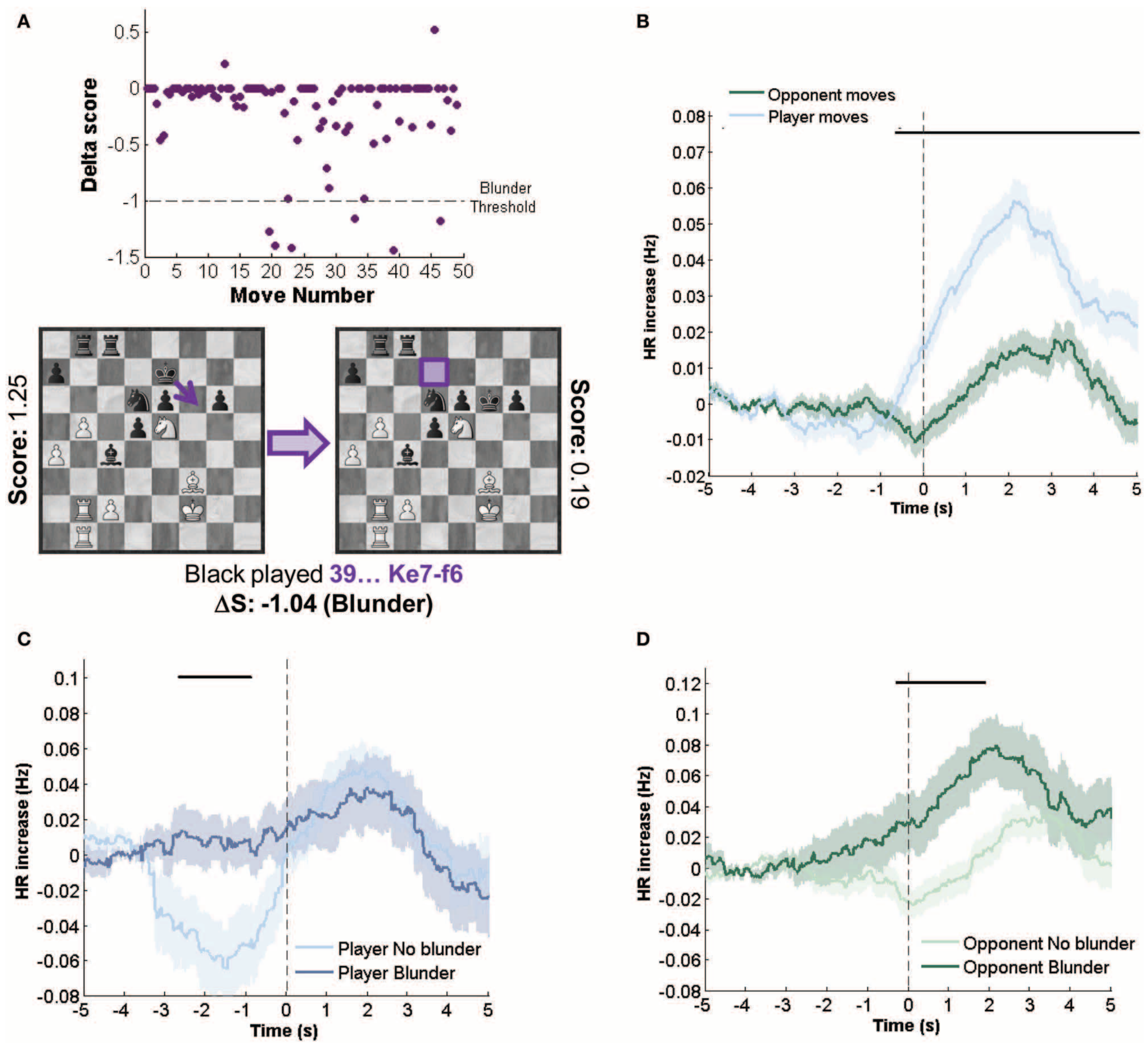


**FIGURE 2 | HR indexes objective errors and this depends on which player did the move. (A)** Upper panel: Delta Score ($\Delta S$) variation with move number for a typical game ($\Delta S \leq -1$ is a blunder). Lower panel, an example of two chess positions before and after a blunder or error move. Note that black's move allows white to fork the king and rook with the knight. **(B)** Temporal course of HR dynamics on own and opponent not error moves locked to move onset. **(C,D)** Dynamics of a player **(C)** and opponent **(D)** move for blunders and correct moves. For blunder versus non-blunder moves, matching conditions minimized differences in player AT and S. Black bars indicate intervals where differences between temporal series were significant ($p < 0.05$). Mean ± SEM across all games.

with the average data. This means that when using a leave-out procedure, where a subset of the data is used for training and another subset for testing, the performance of the classifier is correct in 60% of the test trials. Since this is a binary classification, chance level is at 50%. Calculation and planning were decoded better using data before the move and opponent blunders with data after the opponent move. The weakest classification was achieved for the player owns blunders, were classification was almost at chance. This is an interesting observation showing that fluctuations due to own blunders elicit a broad variety of changes and hence are less amenable to be captured by a linear classifier.

## DISCUSSION

Our work shows that beyond known modulations of body signals in decision making (Bradley, 2009; Jennings et al., 2009; Taelman et al., 2011), HR can signal relevant cognitive episodes including objective events such as the correctness of choice and subjective events tagged by retrospective reports such as engaging in a plan or in calculation relevant for multi-step cognition

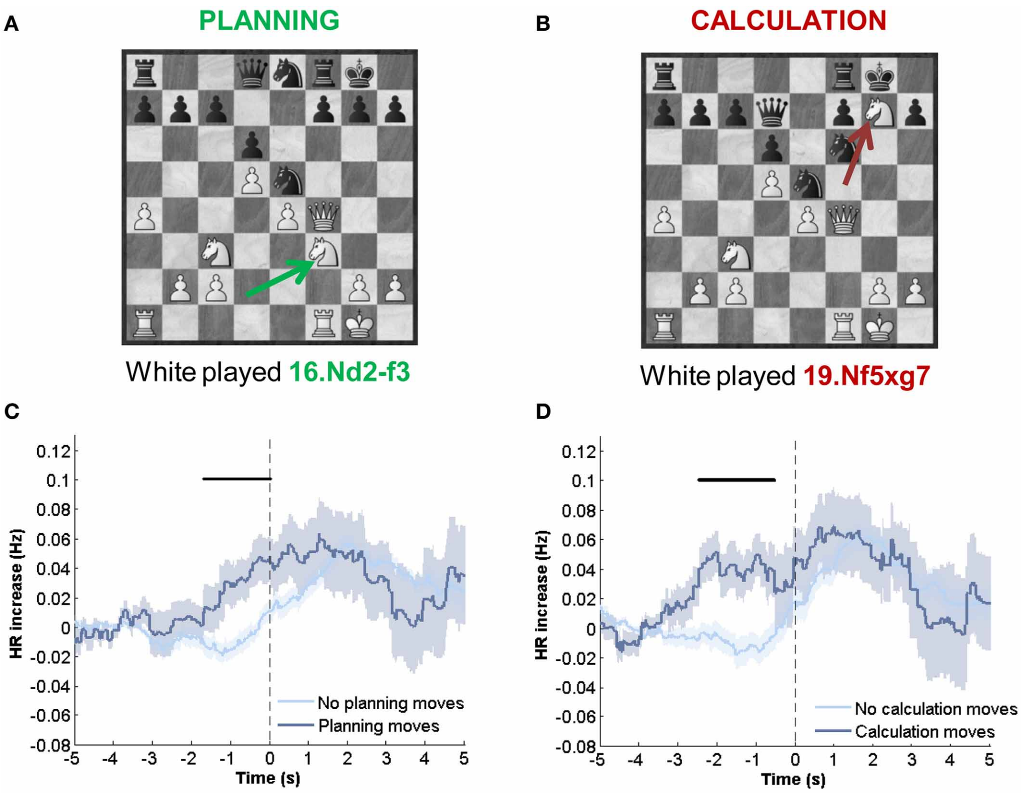


**FIGURE 3 | HR indexes subjective and retrospectively highlighted events. (A)** Example of a planning move. In the position, white player just made a planning move (retrospectively highlighted after the conclusion of the game). White player plan was to initiate an attack in the king's flank, starting with the transference of one of his/her knight to f5 square. Nd2-f3 (indicated with an arrow) was the first move of this plan. **(B)** Calculation move. Later in the same game, other move was highlighted as calculation (19.Nf5-g7). In this case, player was thinking about specific variations (main variation: 19.Nf5-g7 Kg8-g7, 20.Qf4-f6+ Kg7-g8, (a) 21.Rf1-f5 or (b) 21.Nc3-b5 or (c) 21.Ra1-a3). **(C)** HR dynamics comparison of planning moves and non-planning moves matched by other variables. **(D)** HR dynamics comparison of calculation moves and non-calculation moves matched by other variables. All moves were no blunders and matching conditions minimized differences in player AT and S. Those moves highlighted as both planning and calculation were not considered. Black bars indicate intervals where differences between temporal series were significant ($p < 0.05$). Mean ± SEM across all games.

**Table 1 | Linear classifier results.**

| Move type | Before (%) | After (%) | *N* (targets) |
|---|---|---|---|
| Player blunders | 54.94 ± 3.35 | 46.09 ± 1.45 | 24 |
| Opponent blunders | 56.50 ± 2.03 | 62.46 ± 1.55 | 34 |
| Planning | 56.87 ± 2.92 | 55.56 ± 3.23 | 15 |
| Calculation | 78.18 ± 2.96 | 58.14 ± 2.63 | 17 |

(Anderson and Lebiere, 1998). Thus, even if HR constitutes a relatively broad marker integrating a myriad of physiological variables, its dynamic was rich enough to reveal relevant episodes of inner thought.

The seminal work of Adriaan de Groot used chess as a vehicle to understand thought (De Groot, 1965). This work relied on introspection, using the methodology of thinking aloud as the main vehicle to identify episodes of thought. Here we showed that the HR signal carries information capable of indexing these episodes: increasing before player own blunders, planning and calculation moves, and reacting to opponent errors.

Previous studies have investigated how HR varies in longer chess games, consistently finding an increase in HR throughout the game (Pfleger et al., 1980; Hollinsky et al., 1997; Troubat et al., 2009). We replicated this effect suggesting a universality of this phenomenon observed in different time-scales, experimental setups and specific analytic measures of HR.

Our aim here was to understand how on top of this global trend, the HR signal is modulated by specific episodes which relate to strategies, calculations, and the outcomes of decisions made during the game.

First, we observed a very different dynamics when HR variations were locked to a player or to the opponent move. Changes in dynamics after the move could have two different origins which here we cannot disambiguate. First it could be simply the effect of the motor action. Second, and more interesting, it is possible that mechanisms of evaluation of one's own action engage a larger increase in HR than the evaluation of the opponent action.

For player moves, the significant differences between errors and good moves were found in the region corresponding to the recorded player's turn to play. We found a very robust marker in the HR signal which anticipated a correct decision, which involved a transient decrease in the HR prior to the move. This was apparent in all the data and much more strikingly when considering moves which were paired to errors (in the same parameters and state of the game) but in which the error was not made. This is consistent with the observation of HR decreases before making a good decision, only observed for risky choices and good performers on IGT (Crone et al., 2004). It is important to keep in mind that errors are typically made in tense situations of the game (time trouble, unbalanced score). In a simple and naive model in which HR indexes the load of rational thought (consistent with our observation of increased HR with calculation and planning) this deep could be understood as a comparable absence of rational thought. Following this logic and only as a driving hypothesis we suggest that in line with several behavioral observations, in such complicated situations, it might be better to follow hunches than rational and deliberate thought (Dijksterhuis et al., 2006). This hypothesis is consistent with the observations of a HR deceleration immediately before an action which has been related to inhibition of other actions and preparation for the imminent stimulus (Jennings and Van Der Molen, 2002).

Finally, HR was also altered by cognitive processes related with problem solving, as planning (setting a goal) and calculation (analysis of specific candidate moves and their variation). Planning and calculation moves (both are player moves) showed a similar pattern on HR compared to other matched moves: they had an HR increase before the signaled move. A particularly motivating challenge for future research is to understand the causal relation of this observation. As argued above, it may be that the load of rational thought induces transient increases of HR. Alternatively, pushing farther the SM hypothesis it is possible that SM do not only assist choice in overt actions but also signal internal episodes of a mental program (Duncan, 2010; Zylberberg et al., 2011). In other words, it is possible that the action by which a player makes a pause in the game, changes a plan, engages on deep calculation is flagged by internal somatic variables, like HR.

All the previous discussion was drawn analyzing how a factor affects the mean of a distribution. This is the most classic analysis by which inferences are drawn from significant global tendencies of the data. In HR data it seems difficult to go beyond these estimates because of the intrinsic high noise of the signal.

Here we made an effort in this direction, zooming in to single-trial analysis to inquire which factors produce reliable changes which serve to decode states from the data. We used a linear classifier procedure which essentially relies on a bisection of the data by a plane. This method effectively decodes when the factor produces a consistent (albeit noisy) perturbation in the data. If instead, a factor produces a myriad of different changes which when summed together produce a change in the mean, the decoder is not effective. Hence, one can see this analysis as a way to inquire the consistency of an effect. Our data showed a reliable classification for three of the four factors: planning, calculation, and opponent blunders. The most effective decoding was for calculation, when relying on data before the move, which reached levels above 75% which are considerably high for HR data which, as expected, has multiple sources of noise. Instead, the classification for own blunders was very modest, almost at chance levels. This is in fact a very robust result as even varying the parameters of the classifier; these numbers remain close to chance. We suggest that this data reflect that compared to planning, calculating or to the observation of an error of the opponent, one's own blunder may reflect many different internal processes which, in turn, affect the heart in different manner. Interestingly, decoding was effective in introspective variables which could not be measured without explicit reports.

## ACKNOWLEDGMENTS

Universidad de Buenos Aires, CONICET.

## REFERENCES

Anderson, J. R., and Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ariely, D. (2008). *Predictably irrational. The Hidden Forces That Shape Our Decisions*. NewYork, NY: Harper/HarperCollins Publishers.

Bechara, A., and Damasio, A. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336–372.

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–15.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.

Bradley, M. M. (2009). Natural selective attention: orienting and emotion. *Psychophysiology* 46, 1–11.

Cristianini, N., and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press.

Crone, E. A., Jennings, J. R., and Van Der Molen, M. W. (2003). Sensitivity to interference and response contingencies in attention-deficit/hyperactivity disorder. *J. Child Psychol. Psychiatry* 44, 214–226.

Crone, E. A., Somsen, R. J., Van Beek, B., and Van Der Molen, M. W. (2004). Heart rate and skin conductance analysis of antecendents and consequences of decision making. *Psychophysiology* 41, 531–540.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Grosset/Putnam.

De Groot, A. (1965). *Thought and Choice in Chess*. The Hague: Mounton.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., and Van Baaren, R. B. (2006). On making the right choice: the deliberation-without-attention effect. *Science* 311, 1005–1007.

Drucaroff, L. J., Kievit, R., Guinjoan, S. M., Gerschcovich, E. R., Cerquetti, D., Leiguarda, R., et al. (2011). Higher autonomic activation predicts better performance in iowa gambling task. *Cogn. Behav. Neurol.* 24, 93–98.

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* 14, 172–179.

Dunn, B. D., Dalgleish, T., and Lawrence, A. D. (2006). The somatic marker hypothesis: a critical evaluation. *Neurosci. Biobehav. Rev.* 30, 239–271.

Elo, A. (1978). *The Rating of Chess Players, Past and Present*. London: Batsford.

Hollinsky, C., Maresch, G., Hiller, M., Kohlberger, P., and Bieglmayer, C. (1997). Beeinflußt körperliche Fitneßdie Leistungsfähigkeit von Ranglistenschachspielern. *Ö J. Sportmed*. 27, 51–59.

Janssen, P., and Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* 8, 234–241.

Jennings, J. R., and Van Der Molen, M. W. (2002). Cardiac timing and the central regulation of action. *Psychol. Res.* 66, 337–349.

Jennings, J. R., Van Der Molen, M. W., and Tanase, C. (2009). Preparing hearts and minds: cardiac slowing and a cortical inhibitory network. *Psychophysiology* 46, 1170–1178.

Kotov, A. (1971). *Think like a GrandMaster*. London: B.T. Batsford Ltd.

Maia, T. V., and McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16075–16080.

Pfleger, H., Stocker, K., Pabst, H., and Haralambie, G. (1980). [Sports medical examination of top class chess players (author's transl)]. *MMW Munch. Med. Wochenschr.* 122, 1041–1044.

Sigman, M., Etchemendy, P., Slezak, D. F., and Cecchi, G. A. (2010). Response time distributions in rapid chess: a large-scale decision making experiment. *Front. Neurosci.* 4:60. doi: 10.3389/fnins.2010.00060

Slezak, D. F., and Sigman, M. (2011). Do not fear your opponent: suboptimal changes of a prevention strategy when facing stronger opponents. *J. Exp. Psychol.* 141, 527–538.

Taelman, J., Vandeput, S., Vlemincx, E., Spaepen, A., and Van Huffel, S. (2011). Instantaneous changes in heart rate regulation due to mental load in simulated office work. *Eur. J. Appl. Physiol.* 111, 1497–1505.

Tranel, D., Bechara, A., and Damasio, A. R. (1999). "Decision making and the somatic marker hypothesis," in *The New Cognitive Neurosciences, 2nd Edn*, ed M. S. Gazzaniga (London: The MIT Press), 1047.

Troubat, N., Fargeas-Gluck, M. A., Tulppo, M., and Dugue, B. (2009). The stress of chess players as a model to study the effects of psychological stimuli on physiological responses: an example of substrate oxidation and heart rate variability in man. *Eur. J. Appl. Physiol.* 105, 343–349.

Van Der Maas, H. L., and Wagenmakers, E. J. (2005). A psychometric analysis of chess expertise. *Am. J. Psychol.* 118, 29–60.

Zylberberg, A., Dehaene, S., Roelfsema, P. R., and Sigman, M. (2011). The human Turing machine: a neural framework for mental programs. *Trends Cogn. Sci.* 15, 293–300.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 09 August 2012; accepted: 17 September 2012; published online: 08 October 2012.*

*Citation: Leone MJ, Petroni A, Fernandez Slezak DF and Sigman M (2012) The tell-tale heart: heart rate fluctuations index objective and subjective events during a game of chess. Front. Hum. Neurosci.* ***6****:273. doi: 10.3389/fnhum.2012.00273*

## APPENDIX

### POST-GAME QUESTIONNAIRE (CUESTIONARIO POST-PARTIDA)

Player name (Nombre del jugador): ______________________ Game (Partida): ______________________
Round (Ronda): ______________________ Game Result (Resultado): ______________________
Opening theorical knowledge (YES/NO) (Conocimiento teórico de la apertura (SI/NO)): ______________________
Last move known (Hasta qué jugada): ______________________
Planning moments (Momentos de generación de plan): ______________________ (please, write the move corresponding to the plan start) (escribir la jugada propia correspondiente a su inicio).
Game style (Estilo de juego de la partida): ______________________
Does the style agree with your own style? (YES/NO) (¿Acorde al propio? (SI/NO)): ______________________
Tactical calculation moments (Momentos de cálculo táctico): ______________________

**Highlighted moves (Jugadas destacadas):**
- *Sacrifices* (Sacrificios): ______________________

**Errors (Errores):**

(1) **Own errors** (Propios)

a. Unconscious and opponent took advantage of it (concientes y aprovechados por el rival): ______________________
b. Conscious (after its execution) and opponent took advantage of it (concientes (luego de su realización) y aprovechados por el rival): ______________________
c. Conscious (after its execution) and opponent did NOT take advantage of it (concientes (luego de su realización) y NO aprovechados por el rival): ______________________

(2) **Opponent errors (Del adversario)**

You took advantage of it (Aprovechados): ______________________
You did NOT take advantage of it (NO aprovechados): ______________________

**Good moves (Buenas jugadas):**
Own moves (Propias): ______________________
Opponent moves (Del Adversario): ______________________

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 18 March 2013
doi: 10.3389/fnhum.2013.00006

# Susceptibility to emotional contagion for negative emotions improves detection of smile authenticity

***Valeria Manera*[1,2]*, *Elisa Grandi*[1] *and Livia Colle*[1]**

[1] *Center for Cognitive Science, Department of Psychology, University of Turin, Torino, Italy*
[2] *Stanford Psychophysiology Laboratory, Department of Psychology, Stanford University, Stanford, CA, USA*

***Edited by:***
*Maria Ruz, Universidad de Granada, Spain*

***Reviewed by:***
*Agneta H. Fischer, University of Amsterdam, Netherlands*
*Eva G. Krumhuber, Jacobs University Bremen, Germany*

****Correspondence:***
*Valeria Manera, Center for Cognitive Science, Department of Psychology, University of Turin, Via Po 14, 10123 Torino, Italy.*
*e-mail: valeria.manera@unito.it*

A smile is a context-dependent emotional expression. A smiling face can signal the experience of enjoyable emotions, but people can also smile to convince another person that enjoyment is occurring when it is not. For this reason, the ability to discriminate between felt and faked enjoyment expressions is a crucial social skill. Despite its importance, adults show remarkable individual variation in this ability. Revealing the factors responsible for these huge individual differences is a key challenge in this domain. Here we investigated, on a large sample of participants, whether individual differences in smile authenticity recognition are accounted for by differences in the predisposition to experience other people's emotions, i.e., by susceptibility to emotional contagion. Results showed that susceptibility to emotional contagion for negative emotions increased smile authenticity detection, while susceptibility to emotional contagion for positive emotions worsened detection performance, because it leaded to categorize most of the faked smiles as sincere. These findings suggest that susceptibility to emotional contagion plays a key role in complex emotion recognition, and point out the importance of analyzing the tendency to experience other people's positive and negative emotions as separate abilities.

**Keywords: smile authenticity, emotional contagion, simulation models, individual differences, positive and negative emotions**

## INTRODUCTION

A smile is a context-dependent emotional expression. A smiling face does not always signal the experience of enjoyable emotions: sometimes people smile to convince another person that enjoyment is occurring when it is not. This can be done for many different reasons, for example to hide, moderate, or justify something negative (e.g., a feeling of superiority or contempt, a manipulation, social embarrassment, or an inappropriate affect), or simply to coordinate conversation. The ability to discriminate between felt and faked enjoyment expressions is critical to effective social interaction, and to cope with the complexity of the human social world: recognizing a faked smile can prevent people from being deceived, from being inappropriate in a formal social situation, or from starting a potentially frustrating collaboration with a person with no cooperative intents (see Miles, 2009; Johnston et al., 2010). Despite recognizing the authenticity of smiles plays such a crucial role in dealing with everyday social interactions, there are striking individual differences in this ability: whereas some people are very good at distinguishing felt and simulated enjoyment expressions, others appear to lack this ability more or less completely (Frank et al., 1993; Gosselin et al., 2002; Del Giudice and Colle, 2007; McLellan et al., 2009; see Ekman, 2003 for a review). Which are the factors responsible for this remarkable individual variation? Despite unrevealing these factors is a key challenge in this domain, very few studies so far have directly addressed this question (see Manera et al., 2011).

A good way to derive workable hypotheses on the factors responsible for individual differences is starting from the mechanisms involved in smile authenticity detection. How can we distinguish felt from faked smiles by simply looking at other people's face? To date, research on smile recognition has almost exclusively focused on perceptual factors, and has shown that there are substantial perceptual differences between different types of smiles. Smiles judged as genuine involve the activation of specific muscle regions (e.g., the external strand of the *Obircularis Oculi* muscle, producing the narrowing of the eye aperture, and the appearance of *crow's feet* on the external side of the eye—also known as Duchenne marker, Frank et al., 1993), display smooth and more regular facial movements, and are longer in onset, apex and offset durations (Krumhuber and Manstead, 2009) compared to smiles that are judged as less genuine. Moreover, it has been shown that these dynamic and morphological cues are consistently used by observers in order to rate smile authenticity (Frank et al., 1993; Del Giudice and Colle, 2007; Miles and Johnston, 2007; Krumhuber and Manstead, 2009). However, we have recently shown that perceptual factors—such as the attention devoted to the eye region—do not seem to account for the striking variation in smile recognition accuracy across individuals (Manera et al., 2011). This suggests that other mechanisms are likely to be involved in smile authenticity detection.

Recently, it has been proposed that smile recognition relies on *embodied simulation* processes (Niedenthal et al., 2010). Embodied simulation models advance that observers automatically mimic other people's facial expressions, experience those emotions themselves, and consequently attribute them to the other person. Thus, in this account, smile recognition is based on first-person emotional experience (see Goldman and Sripada, 2005). This proposal allows to formulate new hypotheses concerning possible sources of individual variation in smile authenticity recognition. An intriguing hypothesis which—to our knowledge—has never been explored, is that the ability to distinguish between sincere and faked enjoyment expressions is affected by the predisposition to experience others' emotions, i.e., by susceptibility to emotional contagion. In the present paper we focused on this question, and we asked whether individual differences in smile authenticity recognition are accounted for by differences in the susceptibility to emotional contagion. In the following paragraphs we will first briefly review the embodied simulation account of smile recognition proposed by Niedenthal et al. (2010). We will then focus on emotional contagion, and we will finally introduce the present study.

### SIMULATION MODELS OF SMILE RECOGNITION

The Simulation of Smile Model (Niedenthal et al., 2010) advanced the hypothesis that we distinguish different categories of smile based on embodied simulation processes. Embodied simulation models posit that the perception of a facial expression automatically triggers the activation of the facial configuration associated with the observed emotion (*mimicry*), and this, in turn, induces in the observers the physiological activations and the subjective experience of the very same emotion (*emotional contagion)*. The experienced emotion is then attributed to the other person (Adolphs, 2006; see Goldman and Sripada, 2005 for a review of the different existing versions of face-based embodied simulation models). Even if not conclusive, there is consistent evidence in favor of simulation accounts of emotion recognition. We know that observers rapidly and automatically mimic other people's emotional expressions (Hess et al., 1998; Dimberg et al., 2000; Lishner et al., 2008), and that mimicry can aid emotion recognition. For instance, mimicking other's expressions results in faster emotion categorization (Stel and van Knippenberg, 2008), while blocking facial mimicry can impair emotion recognition (Oberman et al., 2007; Ponari et al., 2012; but also see Hess and Blairy, 2001). Furthermore, the observation of other people's emotional expressions is able to induce in the perceiver the experience of the corresponding emotions, i.e., emotional contagion (Hsee et al., 1992; Strayer, 1993; Schneider et al., 1994; Blairy et al., 1999; Wild et al., 2001; Lishner et al., 2008).

In a recent study, Maringer et al. (2011) provided preliminary evidence in favor of the SIMS model, demonstrating that mimicry can aid in smile authenticity recognition. Two groups of participants were presented with smiles with different dynamic qualities known to be associated with "true" or "false" smiles, and were asked to rate them on a scale of genuineness. During the task, half of the participants in each group were able to freely mimic the smiles, and the remaining half held a pencil in their mouths so as to block facial mimicry. In the mimicry condition, participants who saw true smiles rated them as more genuine compared to participants who saw false smiles. In the mimicry-blocked condition, no difference in the authenticity ratings was found between participants who saw true and false smiles. Although this study employed a between-subject design (not taking into account the striking individual differences in both smile genuineness recognition and tendency to mimic other people's emotions and behaviors), these results support the idea that simulation processes play a role in smile authenticity recognition. If this is the case, also emotional contagion should play a role in smile recognition.

### EMOTIONAL CONTAGION

Emotional contagion refers to the human tendency to automatically mimic and synchronize facial expressions, vocalizations, postures and movements with those of another person, and consequently converge emotionally with them (Hatfield et al., 1994). This basic form of empathy (Preston and de Waal, 2002; Singer, 2006) allows us to share the emotions of others directly, without any conscious effort and any form of cognitive mediation. Through emotional contagion we feel others' emotions directly in our body, as if they were our own emotions. Emotional contagion is not considered a single and undifferentiated mechanism; rather—since different emotions are characterized by distinct facial expressions, psycho-physiological patterns, and brain activations (LeDoux, 2000; Dalgleish, 2004)—it may be modulated by the specific emotional content of the observed stimuli (see Goldman and Sripada, 2005). In particular, emotional contagion for positive emotions and emotional contagion for negative emotions may be conceived as partially distinct mechanisms, as positive and negative emotions show differential facial and physiological activations (Schwartz et al., 1979; Davidson et al., 1990), and engage non-completely overlapping neural circuits (Adolphs et al., 1996). It has been proposed that the neural substrate responsible for emotional contagion is represented by the Mirror Neuron System for emotions, including the insula and the anterior mesial frontal cortex. This system is active both when we feel a specific emotion, and when we see another person's emotional expressions (see Keysers and Gazzola, 2009, and Rizzolatti and Sinigaglia, 2010 for reviews).

The SIMS model advanced the prediction that emotional contagion has an impact on smile authenticity recognition. Although it has never been empirically tested, this hypothesis is very plausible. There is evidence that felt (Duchenne) smiles are associated with the experience and physiological activations of positive emotions, while faked non-Duchenne smiles are associated with the experience and physiological activation of negative emotions (Davidson et al., 1990; Ekman et al., 1990; Soussignan, 2002). For this reason, it is possible to expect that felt and faked smiles recruit different components of emotional contagion.

In the present study we did not address directly this hypothesis on emotional contagion. As we were interested in individual differences, we focused instead on *susceptibility to emotional contagion*, i.e., a trait measure strongly associated to online measures of emotional contagion.

#### *Individual differences in the susceptibility to emotional contagion: the present study*

Susceptibility to emotional contagion, i.e., the predisposition to converge emotionally with other people, shows remarkable variations across individuals. While some people show a strong tendency to experience others' emotions, some others seem to be scarcely affected by the observation of others' emotional states, as testified by huge individual differences in self-reported emotional contagion (Doherty, 1997; Sonnby-Borgstrom, 2002). The individual predisposition to experience emotional contagion can be reliably measured through a number of self-report questionnaires, such as the Personal Distress subscale of the Interpersonal Reactivity Index (Davis, 1983), and the Emotional Contagion Scale (ECS) (Doherty, 1997). People scoring higher on these trait measures are more prone to experience other people's emotions, and mimic others' emotional expressions more consistently compared to people with a low susceptibility to emotional contagion (e.g., Hietanen et al., 1998; Blairy et al., 1999). Furthermore, when observing other people's emotional expressions, persons scoring higher in emotional contagion show stronger activation of brain areas in the mirror neuron system for emotions, such as the insula, inferior-parietal junction, and anterior cingulated cortex (Lawrence et al., 2006; Lamm et al., 2007; Pfeifer et al., 2007).

If simulation processes and emotional contagion are involved in smile authenticity recognition, as the SIMS model advanced, individual differences in the susceptibility to emotional contagion may account for individual differences in the ability to distinguish sincere and faked enjoyment expressions. Here we tested this hypothesis on a large sample of participants. Smile authenticity detection was assessed by means of the Smile Picture Set (SPS), a validated FACS-based task including sincere Duchenne smiles and simulated non-Duchenne smiles (Del Giudice and Colle, 2007). Susceptibility to emotional contagion was measured through the ECS (Doherty, 1997), a validated self-report questionnaire specifically designed to measure individual differences in susceptibility to emotional contagion. This self-report measure has been widely employed in a number of domains, and, to our knowledge, it is the only validated instrument tapping both positive and negative emotions. As discussed in the previous paragraphs, sincere Duchenne smiles are associated with the experience and physiological activations of positive emotions, while faked non-Duchenne smiles are associated with the experience and physiological activation of negative emotions (Davidson et al., 1990; Ekman et al., 1990; Soussignan, 2002). For this reason, it is plausible to expect that susceptibility to emotional contagion for positive and negative emotions play a different role in predicting individual differences in smile authenticity recognition.

## MATERIALS AND METHODS

### PARTICIPANTS

One hundred and eight undergraduate students (58 females, 50 males) from the University of Torino volunteered to participate in the study. The average age was 22 years (range = 18–34 years). All participants had normal or corrected-to-normal vision and were naive with respect to the purpose of the experiment. This research was approved by the local Ethical Committee in line with the Declaration of Helsinki.

### MATERIALS AND PROCEDURE

Participants were administered a FACS-based task evaluating the ability to detect smile authenticity from facial expressions (the *Smiles Picture Set*, Del Giudice and Colle, 2007), and a self-report questionnaire evaluating the tendency to experience emotional contagion. The tasks were administered individually in a randomized order, and took about 20 min to complete.

#### *Smile recognition*

Smile recognition was assessed through the SPS (Del Giudice and Colle, 2007), a validated FACS-based task consisting of 25 color pictures of an actor's face displaying smiles of variable intensity, either with closed lips (AU12) or bared teeth (AU12 + AU25). The set contains 11 Duchenne smiles with AU6 activation and 14 non-Duchenne smiles (seven with AU7 activation and seven with a neutral eye region). Complete FACS codings of the SPS pictures can be found in Del Giudice and Colle (2007). Previous research has demonstrated that the Duchenne smiles included in this set are rated as significantly more genuine compared to the non-Duchenne smiles (Del Giudice and Colle, 2007; Manera et al., 2011).

Stimuli were displayed on a 21″ monitor by means of Presentation 9.3 software (Neurobehavioral Systems), at a viewing distance of 70 cm. Pictures had a resolution of 1024 by 768 pixels, and subtended a vertical visual angle of 15° and a horizontal angle of 22°. The task started with a preliminary phase, in which four pictures of the same actor performing different expressions (anger, sadness, surprise, and disgust) were presented, each preceded by the picture of the actor's neutral face. This phase gave participants the opportunity to get accustomed to the actor's face. The 25 items were then shown in one of two randomized orders; each item had a duration of 3 s and started with a neutral face followed by the smiling face. After each stimulus presentation, participants were asked to decide whether the actor was really happy, or was pretending to be happy. Responses were given by pressing one of two keys on a keyboard. In order to reduce guessing, participants were allowed to abstain from responding if they were really unsure about their answer.

#### *Susceptibility to emotional contagion*

The individual tendency to experience emotional contagion was measured through the ECS (Doherty, 1997), a widely used self-report questionnaire. ECS is a 15-item scale that separately evaluates the susceptibility to emotional contagion for positive emotions (happiness and love, six items) and negative emotions (fear, anger and sadness, nine items). Examples are: "If someone I'm talking with begins to cry, I get teary-eyed" and "When someone smiles warmly at me, I smile back and feel warm inside." Participants were asked to rate their degree of agreement with each item on a 5-point scale, from "never" to "always."

### DATA ANALYSIS

Performance in the smile recognition task was assessed by means of the proportion of correct responses (accuracy) and the Signal Detection Theory parameter $d'$ (sensitivity). To calculate accuracy, each response was coded as correct (1 point) or incorrect (0 points). A correct response was scored for each "really happy" answer to Duchenne smile items, and for each "pretending to be happy" answer to non-Duchenne smile items.

"Don't know" answers were awarded 0 points. As in yes-no tasks the proportion of correct responses represents a biased measure (i.e., it does not consider systematic errors in performance), we also extracted Signal Detection Theory parameters. "Don't know" answers were coded as invalid trials, and were excluded from the analysis. The proportion of hits ("really happy" answer to valid Duchenne smile trials) and false alarms ("really happy" answer to valid non-Duchenne smile trials) were used to calculate the location of the criterion *c* (i.e., the general tendency to respond "really happy" or "pretending to be happy"; e.g., a value of zero indicates no bias) and the $d'$, an unbiased sensitivity index-independent of the criterion the participant is adopting (e.g., a value of zero indicates an inability to discriminate Duchenne smile trials from non-Duchenne smile trials, whereas larger values indicate a correspondingly greater ability to discriminate between them). Hits and false alarm proportions of zero were replaced with 0.5/N, and proportions of 1 were replaced with (N-0.5)/N (where N is the number of valid Duchenne smile and non-Duchenne smile trials for each participant).

To explore the effects of gender, accuracy, $d'$, *c* and emotional contagion scores were submitted to separate ANOVAs with participant's gender as within-subject factor. Pearson correlations between the two emotional contagion subscales were performed. To explore the impact of susceptibility to emotional contagion on smile recognition, accuracy, $d'$ and error type (false alarm rate and miss rate) were submitted to separate linear regressions, with the two emotional contagion scales (ECS_pos and ECS_neg) as regressors.

## RESULTS

Consistent with previous studies, major individual differences in smile authenticity detection were found. The proportion of correct responses in the SPS ranged from 0.28 to 1.00, with a mean of 0.68 ($SD = 0.13$). "Don't know" responses were less than 10%. $d'$ ranged from $-0.53$ to 3.53 ($M = 1.46$, $SD = 0.30$). Criterion *c* ranged from $-0.76$ to 1.82 ($M = 0.10$, $SD = 0.48$), and was significantly greater than 0 [$t_{(107)} = 2.13$, $p = 0.036$], thus indicating that participants, as a group, adopted a conservative criterion (i.e., had a slight tendency to respond "pretending to be happy"). No gender differences in accuracy, $d'$ and *c* were found [$F_{(1,\ 107)}$ ranging from 0.48 to 1.47, $p_s$ ranging from 0.23 to 0.49].

ECS_pos scores ranged from 13 to 30 ($M = 23.6$, $SD = 3.5$), and ECS_neg scores ranged from 12 to 37 ($M = 24.7$, $SD = 5.3$). Consistent with previous findings (e.g., Doherty, 1997), significant gender differences in susceptibility to emotional contagion were found, with females scoring higher than males [ECS_pos: $F_{(1,\ 107)} = 7.3$, $p = 0.01$; ECS_neg: $F_{(1,\ 107)} = 13.5$, $p < 0.001$]. A positive correlation between ECS_neg and ECS_pos was found [$r_{(106)} = 0.51$, $p < 0.001$].

Results of regression analysis for accuracy, $d'$, false positive and false negative error rate are reported in **Table 1**.

Emotional contagion was a significant predictor of performance for both accuracy [$F_{(2,\ 107)} = 3.77$, $p = 0.026$, $R^2 = 0.07$] and $d'$ [$F_{(2,\ 107)} = 6.02$, $p = 0.003$, $R^2 = 0.10$], suggesting that around 10% of the individual differences in smile authenticity recognition found in the present study were explained by differences in the susceptibility to emotional contagion. Interestingly, emotional contagion for negative emotions (ECS_neg) was found to be a positive predictor of participants' performance (accuracy: $p = 0.016$; $d'$, $p = 0.002$): the higher the scores on emotional contagion for negative emotions, the higher the performance in the smile detection task. On the contrary, emotional contagion for positive emotions (ECS_pos) was negatively correlated with accuracy and $d'$ (accuracy: $p = 0.023$; $d'$, $p = 0.006$), thus suggesting that the higher the score on emotional contagion for positive emotions, the lower the performance in the smile detection task. Regressions between error-type (miss and false alarm rate) and emotional contagion scores suggest a reason for this pattern of results: people scoring higher on ECS_pos performed worse in smile recognition because they made more false alarms ($p = 0.002$), i.e., they had a tendency to rate non-Duchenne smiles as sincere. The reverse pattern was found for ECS_neg, indicating that people scoring higher on ECS_pos made fewer false alarms ($p = 0.006$). No significant effect of emotional contagion on "miss" errors was found.

**Table 1 | Linear regression analyses of how emotional contagion scales predict smile authenticity recognition (accuracy and $d'$) and error types (false alarm rate and miss rate).**

| Emotional contagion | *B* | Beta | *T* | *P* |
|---|---|---|---|---|
| **DEPENDENT VARIABLE: *ACCURACY* - *F* TEST: *P* = 0.026; $R^2$ = 0.07** | | | | |
| ECS_pos | −0.24 | −0.25 | −2.31 | 0.023 |
| ECS_neg | 0.17 | 0.27 | 2.45 | 0.016 |
| **DEPENDENT VARIABLE: $d'$ - *F* TEST: *P* = 0.003; $R^2$ = 0.10** | | | | |
| ECS_pos | −0.07 | −0.30 | −2.78 | 0.006 |
| ECS_neg | 0.05 | 0.34 | 3.20 | 0.002 |
| **DEPENDENT VARIABLE: *FALSE ALARM RATE* - *F* TEST: *P* = 0.003; $R^2$ = 0.09** | | | | |
| ECS_pos | 0.02 | 0.35 | 3.22 | 0.002 |
| ECS_neg | −0.01 | −0.30 | −2.78 | 0.006 |
| **DEPENDENT VARIABLE: *MISS RATE* - *F* TEST: *P* = 0.348; $R^2$ = 0.001** | | | | |
| ECS_pos | 0.005 | 0.08 | 0.67 | 0.502 |
| ECS_neg | −0.007 | −0.16 | −1.46 | 0.148 |

## DISCUSSION AND FUTURE RESEARCH DIRECTIONS

Emotional contagion—the tendency to unconsciously mimic others' emotional expressions and, consequently, converge emotionally with them (Hatfield et al., 1994)—is a core aspect of human social functioning. People more susceptible to emotional contagion are more sensitive to others, have a higher self-esteem, and are more empathic compared to people who are less affected by others' emotions (Doherty, 1997). Hatfield et al. (1994) proposed that people who are more susceptible to emotional contagion are also better at reading others' emotional expressions. However, evidence in this respect is still sparse and contradictory (Riggio et al., 1989; Levenson and Ruef, 1992; Blairy et al., 1999). Here we showed for the first time that susceptibility to emotional contagion is related to the ability to discriminate sincere and faked enjoyment expressions: individual differences in emotional contagion accounted for around 10% of individual variation in smile authenticity recognition. Given that emotional contagion was assessed indirectly through self-reports, this percentage is notably high. Indeed, using the same smile recognition task, Manera et al. (2011) showed that perceptual factors (such as the attention devoted to the eye-region), which are well-known to play a role in smile recognition, explained less than 5% of inter-individual variation.

Interestingly, we demonstrated that susceptibility to emotional contagion for positive and negative emotions play a different, opposite role in smile authenticity detection. We found that participants with higher susceptibility to emotional contagion for negative emotions performed better in smile authenticity detection, and they made fewer "false positive" mistakes, that is, they rarely rated faked stimuli as sincere enjoyment expressions. On the contrary, participants with higher susceptibility to emotional contagion for positive emotions showed a reduced sensitivity in detecting emotion authenticity, specifically because they had a tendency to rate non-Duchenne smiles as sincere.

These findings support the hypothesis that susceptibility to emotional contagion is strongly influenced by its emotional content. Faked, non-Duchenne smiles are associated with the experience and physiological activation of negative emotions (Davidson et al., 1990; Ekman et al., 1990; Soussignan, 2002). People with high susceptibility to emotional contagion for negative emotions may be especially sensitive to cues of negative emotions (e.g., the absence of the Duchenne marker in the eye region of a smiling person), and specifically resonate with them, thus experiencing negative emotions. This focus on negative emotions may thus enhance their ability to detect faked smiles (which, in a task where they are asked to rate smile authenticity, translate into a better discrimination performance). Converging with this interpretation, it was found that socially rejected individuals, who have a tendency to focus on negative emotions such as sadness and anger, show an enhanced ability to determine whether an enjoyment facial expression is genuine or deceptive (Bernstein et al., 2008). On the contrary, people inclined to resonate with others' positive emotions may focus especially on the expressive cues signaling positive feelings (e.g., a big smile), experience positive feelings, and, as a consequence, they may overestimate others' happiness, and be easily deceived by subtly faked facial expressions. According to this interpretation, it has been found that older people—who show general biases in focusing on positive information (see Carstensen et al., 2003) have a greater bias toward reporting that any smiling individual is feeling happy, independently of whether he/she is displaying enjoyment or non-enjoyment smiles (Slessor et al., 2010). In order to test whether this interpretation is correct, it would be interesting to examine whether susceptibility to emotional contagion for positive and negative emotions predicts an increased attentional focus on positive and negative emotional cues, respectively, and/or whether this is related to online measures of mimicry and emotional contagion. Further studies would be needed in order to test these predictions. In particular, it would be important to directly investigate whether susceptibility to emotional contagion and online measures of emotional contagion (e.g., autonomic activity) provide converging results concerning the way they affect smile authenticity recognition.

Even if we focused on susceptibility to emotional contagion, which is a trait measures, we believe that our findings have important implications for research exploring the effect of contextual variables—such as participants' mood—on complex emotion recognition. For instance, people with a predisposition to experience others' positive emotions may be more often in a positive mood, as suggested by theories of mood-dependency (e.g., Bower, 1991; Gendolla, 2000). According to these theories, affective states serve as contextual cues that increase the availability of thoughts and memories of a similar hedonic tone; they create a preference for exposure to mood-congruent information, and a tendency to process mood-congruent information in greater detail. People in a positive mood have a tendency to recall mainly positive events and situations, to focus their attention on other's positive emotions, and to rate other's emotional expressions as more positive. The opposite is true for people in a negative mood (for reviews, see Matt et al., 1992; Clore et al., 1994; Forgas, 1995), as also confirmed by a wide range of data collected in clinical populations (see Teasdale et al., 1995; Gangemi et al., 2007). It would be interesting to investigate whether directly manipulating participants' mood affects the ability to detect faked enjoyment expressions. If our interpretation of the results is correct, then inducing a positive mood should worsen participant's ability to detect faked enjoyment expressions, while a negative mood should improve smile recognition performance.

Another interesting domain for future investigation is whether other forms of empathy are related to smile authenticity recognition. Empathy is an ability composed by a variety of neurocognitive processes of different complexity and mediated by distinct neural circuits (see Preston and de Waal, 2002; Decety and Lamm, 2006; Singer, 2006). Emotional contagion represents a very basic component of empathy, but at least two other components do exist, namely emotional and cognitive empathy. *Emotional empathy* refers to the cognitive and neural processes that produce a congruent emotion in the observer in response to others' directly perceived emotional displays, or to descriptions of others' emotion-laden experiences (Saxe, 2006). *Cognitive empathy* implies the conscious mental representation of the state of another individual, and it is often conflated with Theory of Mind, i.e., the ability to represent the mental states of others (Premack and Woodruff, 1978). It has been proposed that both emotional and cognitive empathy are related to emotion recognition ability, even if evidence in this respect is not straightforward (see Zaki et al., 2008). Investigating whether people with higher emotional and/or cognitive empathy are better in smile authenticity recognition would add significantly to the present findings, allowing a better characterization of how empathy affects smile authenticity recognition.

Our study has some limitations, which may have lead to underestimate the role of susceptibility to emotional contagion in smile recognition. First, our smile recognition task was based on the manipulation of a single perceptual cue—the Duchenne marker. Recent evidence indicates that symmetry and dynamic features can be even better predictors of participants' ability to assess smile sincerity, especially when dynamic facial displays are evaluated (e.g., Krumhuber and Manstead, 2009). As movement represents a cue aspect in emotion simulation

(Goldman and Sripada, 2005), it is plausible to expect that emotional contagion predicts an even greater amount of individual variation in smile recognition when using more naturalistic, dynamic facial expressions. Second, there is evidence that ratings of smile genuineness differ depending on the gender of the encoder (Krumhuber et al., 2007), with females' smiles rated as more faked compared to males' smiles. For this reason, it would be important to replicate these results with smile recognition tasks including male and female faces as stimuli.

## ACKNOWLEDGMENTS

This research was supported by a grant from the Regione Piemonte (Regione Piemonte, bando Scienze Umane e Sociali 2008, L.R. n. 4/2006). Valeria Manera was supported by the research fellowship "L'Oreal Italia per le Donne e la Scienza 2012" for the project "La predizione delle intenzioni nell'autismo: studi comportamentali e fMRI." The author would like to thank Marco Del Giudice for commenting on previous versions of the manuscript, and Sabrina Bhangal for proof-reading.

## REFERENCES

Adolphs, R. (2006). Perception and emotion: how we recognize facial expressions. *Curr. Dir. Psychol. Sci.* 15, 222–226.

Adolphs, R., Damasio, H., Tranel, D., and Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *J. Neurosci.* 16, 7678–7687.

Bernstein, M. J., Young, S. G., Brown, C. M., Sacco, D. F., and Claypool, H. M. (2008). Adaptive responses to social exclusion: social rejection improves detection of real and fake smiles. *Psychol. Sci.* 19, 981–983.

Blairy, S., Herrera, P., and Hess, U. (1999). Mimicry and the judgment of emotional facial expressions. *J. Nonverbal Behav.* 23, 5–41.

Bower, G. H. (1991). "Mood congruity of social judgment," in *Emotion and Social Judgment,* ed J. Forgas (Oxford, UK: Pergamon Press), 31–54.

Carstensen, L. L., Fung, H. H., and Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motiv. Emot.* 27, 103–123.

Clore, G. L., Schwarz, N., and Conway, M. (1994). "Affective causes and consequences of social information processing," in *Handbook of Social Cognition, Vol. 1: Basic Processes, 2nd Edn.*, eds R. S. J. Wyer and T. K. Srull (Hillsdale, NJ: Lawrence Erlbaum Associates), 323–417.

Dalgleish, T. (2004). The emotional brain. *Nat. Neurosci.* 5, 583–589.

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., and Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology I. *J. Pers. Soc. Psychol.* 58, 330–341.

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* 44, 113–136.

Decety, J., and Lamm, C. (2006). Human empathy through the lens of social neuroscience. *Sci. World J.* 6, 1146–1163.

Del Giudice, M., and Colle, L. (2007). Differences between children and adults in the recognition of enjoyment smiles. *Dev. Psychol.* 43, 796–803.

Dimberg, U., Thunberg, M., and Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychol. Sci.* 11, 86–89.

Doherty, R. W. (1997). The emotional contagion scale: a measure of individual differences. *J. Nonverbal Behav.* 21, 131–154.

Ekman, P. (2003). Darwin, deception and facial expression. *Ann. N.Y. Acad. Sci.* 1000, 205–221.

Ekman, P., Davidson, R. J., and Friesen, W. V. (1990). The Duchenne smile: emotional expression and brain physiology II. *J. Pers. Soc. Psychol.* 58, 342–353.

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychol. Bull.* 117, 39–66.

Frank, M. G., Ekman, P., and Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *J. Pers. Soc. Psychol.* 64, 83–93.

Gangemi, A., Mancini, F., and van den Hout, M. (2007). Feeling guilty as a source of information about threat and performance. *Behav. Res. Ther.* 45, 2387–2396.

Gendolla, G. H. E. (2000). On the impact of mood on behavior: an integrative theory and a review. *Rev. Gen. Psychol.* 4, 378–408.

Goldman, A. I., and Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition* 94, 193–213.

Gosselin, P., Perron, M., Legault, M., and Campanella, P. (2002). Children's and adults' knowledge of the distinction between enjoyment and non-enjoyment smiles. *J. Nonverbal Behav.* 26, 83–108.

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1994). *Emotional Contagion*. New York, NY: Cambridge University Press.

Hess, U., and Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *Int. J. Psychophysiol.* 40, 129–141.

Hess, U., Philippot, P., and Blairy, S. (1998). Facial reactions to emotional facial expressions: affect or cognition? *Cogn. Emot.* 12, 509–532.

Hietanen, J. K., Surakka, V., and Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology* 35, 530–536.

Hsee, C. K., Hatfield, E., and Chemtob, C. (1992). Assessments of the emotional states of others: conscious judgments versus emotional contagion. *J. Soc. Clin. Psychol.* 11, 119–128.

Johnston, L., Miles, L., and Macrae, C. N. (2010). Why are you smiling at me? Social functions of enjoyment and non-enjoyment smiles. *Br. J. Soc. Psychol.* 49, 107–127.

Keysers, C., and Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions and sensations. *Curr. Opin. Neurobiol.* 19, 1–6.

Krumhuber, E., and Manstead, A. S. R. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion* 9, 807–820.

Krumhuber, E., Manstead, A. S. R., and Kappas, A. (2007). Temporal aspects of facial displays in person and expression perception: the effects of Smile dynamics, head-tilt, and gender. *J. Nonverbal Behav.* 31, 39–56.

Lamm, C., Nusbaum, H. C., Meltzoff, A. N., and Decety, J. (2007). What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE* 2:e1292. doi: 10.1371/journal.pone.0001292

Lawrence, E. J., Shaw, P., Giampietro, V. P., Surguladze, S., Brammer, M. J., and David, A. S. (2006). The role of 'shared representations' in social perception and empathy: an fMRI study. *Neuroimage* 29, 1173–1184.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184.

Levenson, R. W., and Ruef, A. M. (1992). Empathy: a physiological substrate. *J. Pers. Soc. Psychol.* 63, 234–246.

Lishner, D. A., Cooter, A. B., and Zald, D. H. (2008). Rapid emotional contagion and expressive congruence under strong test conditions. *J. Nonverbal Behav.* 32, 225–239.

Manera, V., Del Giudice, M., Grandi, E., and Colle, L. (2011). Individual differences in the recognition of enjoyment smiles: no role for perceptual-attentional factors and autistic-like traits. *Front. Psychology* 2:143. doi: 10.3389/fpsyg.2011.00143

Maringer, M., Krumhuber, E., Fischer, A., and Niedenthal, P. M. (2011). Beyond smile dynamics: mimicry and beliefs in judgments of smiles. *Emotion* 11, 181–187.

Matt, G. E., Vázquez, C., and Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: a meta-analytic review. *Clin. Psychol. Rev.* 12, 227–255.

McLellan, T., Johnston, L., Dalrymple-Alford, J., and Porter, R. (2009). Sensitivity to genuine versus posed emotion specified in facial displays. *Cogn. Emot.* 24, 1277–1292.

Miles, L. (2009). Who is approachable? *J. Exp. Soc. Psychol.* 45, 262–266.

Miles, L., and Johnston, L. (2007). Detecting happiness: perceiver sensitivity to enjoyment and non-enjoyment smiles. *J. Nonverbal Behav.* 31, 259–275.

Niedenthal, P. M., Mermillod, M., Maringer, M., and Hess, U. (2010). The Simulation of Smiles (SIMS) Model: embodied simulation and the meaning of facial expression. *Behav. Brain Sci.* 33, 417–480.

Oberman, L. M., Winkielman, P., and Ramachandran, V. S. (2007). Face to face: blocking facial mimicry can selectively impair recognition of emotional expressions. *Soc. Neurosci.* 2, 167–178.

Pfeifer, J. H., Iacoboni, M., Mazziotta, M., and Dapretto, M. (2007). Mirroring others' emotions relates to empathy and interpersonal competence in children. *Neuroimage* 39, 2076–2085.

Ponari, M., Conson, M., D'Amico, N. P., Grossi, D., and Trojano, L. (2012). Mapping correspondence between facial mimicry and emotion recognition in healthy subjects. *Emotion* 12, 1398–1403.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a 'theory of mind'? *Behav. Brain Sci.* 4, 515–526.

Preston, S. D., and de Waal, F. B. M. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–72.

Riggio, R., Tucker, J., and Coffaro, D. (1989). Social skills and empathy. *Pers. Indiv. Differ.* 10, 93–99.

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–274.

Saxe, R. (2006). Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16, 235–239.

Schneider, F., Gur, R. C., Gur, R. E., and Muenz, L. (1994). Standardized mood induction with happy and sad facial expressions. *Psychiatry Res.* 51, 19–31.

Schwartz, G. E., Ahern, G. L., and Brown, S. L. (1979). Lateralized facial muscle response to positive and negative emotional stimuli. *Psychophysiology* 16, 561–571.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind-reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863.

Slessor, G., Miles, L., Bull, R., and Phillips, L. H. (2010). Age-related changes in detecting happiness: discriminating between enjoyment and nonenjoyment smiles. *Psychol. Aging* 1, 246–250.

Sonnby-Borgstrom, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scand. J. Psychol.* 43, 433–443.

Soussignan, R. (2002). Duchenne smile, emotional experience, and autonomic reactivity: a test of the facial feedback hypothesis. *Emotion* 2, 52–74.

Stel, M., and van Knippenberg, A. (2008). The role of facial mimicry in the recognition of affect. *Psychol. Sci.* 19, 984–985.

Strayer, J. (1993). Children's concordant emotions and contagions in response to observed emotions. *Child Dev.* 64, 188–201.

Teasdale, J. D., Taylor, M. J., Cooper, Z., Hayhurst, H., and Paykel, E. S. (1995). Depressive thinking: shifts in construct accessibility or in schematic mental models? *J. Abnorm. Psychol.* 104, 500–507.

Wild, B., Erb, M., and Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Res.* 102, 109–124.

Zaki, J., Bolger, N., and Ochsner, K. (2008). It takes two: the interpersonal nature of empathic accuracy. *Psychol. Sci.* 19, 399–404.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 12 October 2012; accepted: 07 January 2013; published online: 18 March 2013.*

*Citation: Manera V, Grandi E and Colle L (2013) Susceptibility to emotional contagion for negative emotions improves detection of smile authenticity. Front. Hum. Neurosci. 7:6. doi: 10.3389/fnhum.2013.00006*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 10 January 2013
doi: 10.3389/fnhum.2012.00355

# Description-based reappraisal regulate the emotion induced by erotic and neutral images in a Chinese population

***Jiaxin Peng*[1]*, Chen Qu*[2]*, *Ruolei Gu*[3] *and Yue-Jia Luo*[4]***

[1] *State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China*
[2] *Center for Studies of Psychological Application, South China Normal University, Guangdong, China*
[3] *Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China*
[4] *Sichuan Research Center of Applied Psychology, Chengdu Medical College, Chengdu, China*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*
*Jiongjiong Yang, Peking University, China*
*Jiang Qiu, Southwest University, China*

****Correspondence:***
*Yue-Jia Luo, Sichuan Research Center of Applied Psychology, Chengdu Medical College, Chengdu, CD 610500, China.*
*e-mail: luoyj@bnu.edu.cn*

*Chen Qu, Center for Studies of Psychological Application, South China Normal University, 55 West of Guangzhou Street, Guangdong, GD 510631, China.*
*e-mail: chenqu@scnu.edu.cn*

Previous emotion-regulation research has shown that the late positive potential (LPP) is sensitive to the down-regulation of emotion; however, whether LPP is also sensitive to the up-regulation of emotion remains unclear. The present study examined the description-based reappraisal effects on the up-regulation of positive emotions induced by erotic and neutral images in a Chinese population. Self-reported ratings and event-related potential (ERP) were recorded when subjects viewed pleasant and neutral images, which were shown after either a neutral or positive description. Self-reported results showed that images following positive descriptions were rated as more pleasant compared to images following neutral descriptions. ERP results revealed that the P2, P3, and slow wave (SW) components were larger for erotic pictures than for neutral pictures, while the positive description condition yielded attenuated erotic image-induced P2, P3 and SW and increased SW induced by neutral images. The results demonstrated that description-based reappraisal, as a method of reappraisal, significantly modulates the emotional experience and ERP responses to erotic and neutral images.

**Keywords: LPP, SW, up-regulation, emotion regulation, description-based reappraisal**

## INTRODUCTION

The inability to modulate emotions, i.e., dysregulation of emotion, may precipitate affective disorders. Since emotion regulation contributes significantly to mental health, an increasing number of studies have focused on this subject (Taylor and Liberzon, 2007; Amstadter, 2008; Mak et al., 2009). Emotion regulation refers to the process by which individuals regulate the category, timing, experience, and expression of one's emotions (Gross, 1998). In addition, emotion regulation may decrease, maintain, or increase levels of negative and/or positive emotions (Masters, 1991; Parrott and Schulkin, 1993; Langston, 1994). There are two regulation subtypes, i.e., down-regulation and up-regulation, which decrease or increase an emotional experience, respectively. Most previous research has focused on down-regulation, especially the down-regulation of negative emotions (Levesque et al., 2003; Ochsner et al., 2004). For example, using functional magnetic resonance imaging (fMRI), researchers have found that the prefrontal cortex and/or amygdala participate in the down-regulation of negative emotions (Ochsner et al., 2002; Levesque et al., 2003). In addition, event-related potential (ERP) studies have investigated the time course of the down-regulation of emotion processes and have shown that the late positive potential (LPP), an ERP component, is sensitive to the regulation of negative emotions, such that the magnitude of the LPP in response to emotional pictures is significantly attenuated when individuals are instructed to suppress their emotions (Moser et al., 2006; Foti and Hajcak, 2008). Foti and Hajcak (2008) found that both the LPP magnitude and arousal ratings were significantly reduced when unpleasant pictures were described in neutral rather than in negative terms.

To our knowledge, there are limited studies investigating up-regulation, especially the up-regulation of positive emotions. Understanding up-regulation of positive emotions is important for the pleasures of everyday interactions (Ochsner et al., 2004) because the pursuit of happiness may be the eternal theme throughout human life (Kringelbach and Berridge, 2009). For example, an up-regulation strategy could be used either to augment the joy at a wedding or to create a positive emotional experience in response to a previously appraised neutral stimulus (Langston, 1994). Researchers have tried to investigate the underlying mechanism of up-regulation of positive emotions due to the practical and theoretical implications of a positive emotional experience. These studies applied the reappraisal strategy, a form of cognitive strategy that modifies the affective feeling to emotion-eliciting situations by modulating the understanding of these situations. The reappraisal strategy is more effective than the suppression strategy in the down-regulation (Hajcak and Nieuwenhuis, 2006; Goldin et al., 2008) or the up-regulation of negative emotions (Ochsner et al., 2004). Unfortunately, these studies failed to detect significant differences between the up-regulation condition and the control condition regarding positive emotions (Moser et al., 2006; Krompinger et al., 2008).

This study applied the task paradigm of two previous studies (Moser et al., 2006; Krompinger et al., 2008) and modified the paradigm in two aspects. First, it is worth noting that in previous studies which used simple word instructions (e.g., "suppress" or "enhance") for emotion regulation, the ERP results revealed an attenuated LPP in the down-regulation condition but an unchanged LPP in the up-regulation condition (Moser et al., 2006; Krompinger et al., 2008). Accordingly, we suggest that the presentation of simple wording may not be appropriate for up-regulation. Instead, the current study used a description-based reappraisal strategy, which has been proven to be effective at down-regulating negative emotions and up-regulating neutral emotions (Foti and Hajcak, 2008; Macnamara et al., 2009, 2010).

Second, seeing that the current study was conducted in China, we used erotic pictures as positive stimuli. Compared to Western cultures, Chinese people adhere to more traditional moral values and are consequently less "open" to sexual freedom (Higgins et al., 2002). As a result, Chinese subjects reported less pleasure than Western subjects when viewing the same erotic pictures (Lang et al., 1999; Yen et al., 2010). In our opinion, finding a way to up-regulate Chinese people's positive emotions when facing erotic stimuli would have important implications in improving their quality of life and feeling of well-being.

In the current study to up-regulate emotion in Chinese participants, we selected pictures that depicted erotic couples as the positive stimuli and pictures of household objects as the neutral stimuli. A description-based reappraisal strategy was applied by using a brief neutral or positive description that was presented prior to the emotional pictures. In the regulation condition, the descriptions suggested that the persons in the picture were lovers conforming to traditional morality roles (positive description). Meanwhile, in the control condition, the descriptions indifferently mentioned the impersonal aspects of the pictures (neutral instruction). The present ERP research had three predictions: (1) the description-based reappraisal strategy would successfully raise the levels of pleasant feelings generated by erotic and neutral pictures among Chinese participants; (2) the LPP (P3 and slow wave) would be sensitive to the picture types, such that an increased LPP would be detected when viewing pleasant (erotic) pictures; and (3) the LPP (P3 and slow wave) would be sensitive to the instruction types, such that the LPP magnitude would be different between the up-regulation condition and the control condition. In addition to the LPP, an ERP component that is associated with several psychological processes and the most important ones are motivation and sustained attention (Ibanez et al., 2012). The LPP is enhanced for motivationally relevant stimuli (Schupp et al., 2000, 2004) and is sensitive to cognitive tasks (Polich, 2007; Frühholz et al., 2009; Hurtado et al., 2009; Ibáñez et al., 2011); we also analyzed the frontal P2 component, because previous studies suggested that the P2 was an index of emotional processing (Foti and Hajcak, 2008; Olofsson et al., 2008; Luo et al., 2010). Furthermore, because of well-known gender differences in the processing of erotic stimuli (Lykins et al., 2007), we also analyzed the gender difference in the up-regulation effects and the respective ERP correlates, even though gender differences were not the main emphasis of the present study.

## MATERIALS AND METHODS

### PARTICIPANTS

Forty subjects (20 females) were recruited from either the South China Normal University or Beijing Normal University of China. The average age of subjects was 21.43 years old ($SD = 2.24$). All participants were right-handed and had normal or corrected normal vision. In addition, all subjects were free from any neurological impairment. Subjects were paid for participation and provided written informed consent. The local ethics review board approved this study.

### STIMULI

The stimulus set was composed of 52 neutral (low-arousal) and 52 pleasant (high-arousal) color images. The neutral images (household objects) were selected from the Chinese Affective Picture System (Bai et al., 2005). The procedure of selection of the pleasant pictures was exactly the same as that used in the development of the Chinese affective picture system. All of the pleasant images depicted erotic couples that did not expose their sexual organs. A separate cohort of 35 Chinese participants evaluated the erotic pictures (arousal rating: $M = 6.02$, $SD = 0.56$, valence rating: $M = 5.88$, $SD = 0.30$, $N = 35$). Eight pleasant images(sport) and eight neutral images (antique) were presented as filler materials to keep the participants concentrating on the experiment; however, these filler materials were excluded from the subsequent analysis. Prior to each picture, a brief description of the upcoming picture was presented. There were two types of instructions (i.e., positive and neutral) for each picture. Positive instructions highlighted the positive aspects of the image, whereas neutral instructions described the images in neutral terms. A sample of participants ($N = 30$) evaluated all of the descriptions, and the positive descriptions were markedly more optimistic compared to the neutral descriptions.

A within-subjects 2 (picture type: pleasant and neutral) × 2 (instruction type: positive and neutral) × 2 (gender: males and females) mixed-factors ANOVA included a total of 240 trials. As in previous research (Moser et al., 2006; Krompinger et al., 2008), the 240 trials within the experimental conditions were randomly separated into different blocks, resulting in two pleasant picture blocks (the valence and arousal rating of the pleasant images in the positive and neutral instruction conditions were matched) and two neutral picture blocks (the stimuli in each condition were match as well). To ensure enough time for participants to rest, each block was divided into four sub-blocks, which resulted in a total of 16 sub-blocks. Each sub-block contained 15 pictures (8 positive sub-blocks contained 13 erotic couples pictures and 2 sporting pictures, 8 neutral sub-blocks contained 13 household object pictures and 2 antique pictures). The serial position of the four experimental condition blocks and of the stimuli in each block were counterbalanced across participants.

### EXPERIMENTAL TASKS AND PROCEDURES

Participants sat on a comfortable chair in front of a computer screen, which was located at eye level at a distance of 75 cm. After they finished the informed consent form, participants were attached to an electroencephalograph (EEG) sensor net and given the task instructions. Participants were told that prior to each picture, there would be a brief description about the upcoming picture. Then, participants were instructed to rate the valence of each proceeding picture on a 9-point scale: from 1 (very unpleasant) to 9 (very pleasant).

After a practice trial session, the formal experiment commenced, and all participants performed 16 sub-blocks with appropriate breaks. E-prime software was used to control the presentation and timing of all stimuli. As is shown in **Figure 1**, a fixation mark (+) was presented for 500 ms at the beginning of each trial to orient participants to the center of the screen. The brief description appeared 300–500 ms after the offset of the fixation cross and remained on the screen for 2000 ms. Thereafter, the target picture was presented for 1500 ms after a blank screen (randomized presentation between 1300 ms and 1500 ms). Next a prompt ("?") appeared in the center of the screen, which cued participants to rate the valence of the picture. After a 1500 ms delay, the next trial began. Color images were 8 × 11 cm, and each picture occupied approximately 6° of the horizontal visual angle and 8.4° of the vertical visual angle (or approximately 8.4° horizontally and 6° vertically).

### EEG RECORDING AND DATA ANALYSIS

The Brain Products system (Brain Products GmbH, Munich, Germany) continuously recorded EEG signals from 64 scalp electrodes based on the 10–20 system, with two electrodes placed on the left and right mastoids. Eye blinks and movements were recorded on an electrooculograph (EOG) from four facial electrodes: two approximately 1 cm above and below the participants' left eye, one approximately 1 cm to the left of the left eye, and one approximately 1 cm to the right of the right eye. The EEG was sampled at 500 Hz. All EEG/EOG electrode impedances were below 5 KΩ. Off-line analysis of data was performed using Brain Vision Analyzer software (Brain Products). All data were re-referenced to the average of the two mastoid recordings and band-pass filtered with cutoffs of 0.1 and 30 Hz. The EEG was segmented for each trial, beginning 200 ms before each picture onset (which served as the baseline) and continuing for 1000 ms.

Based upon the suggestions from Krompinger et al. (2008), we measured mean amplitudes in three windows following stimulus onset: 150–200 ms (P2), 330–400 ms (P3), and 400–800 ms (slow wave). For each window, data analysis involved repeated-measure analysis of variance (ANOVA) with the factors picture type (pleasant vs. neutral), instruction type (positive vs. neutral) and gender (male vs. female) as the main factors. Significance level was set at $p = 0.05$ for all analyses. *Post-hoc* testing of significant main effects was conducted using the Bonferroni method. Significant interactions were analyzed using simple-effects models.

## RESULTS

### SELF-REPORTED RATINGS RESULTS

**Table 1** shows the means and standard deviations for the four experimental conditions. As expected, experienced-pleasure ratings were higher for pleasant (erotic) images than neutral images, $F_{(1,\ 38)} = 41.98$, $p < 0.01$, indicating a sizable picture-type effect. The main effect of instruction-type was also significant, $F_{(1,\ 38)} = 18.35$, $p < 0.01$, indicating higher pleasure ratings for positive instructions compared to neutral instructions. The main effect of gender was not significant, $F_{(1,\ 38)} = 0.18$, $p = 0.67$. There were no other statistically significant effects among these three factors.

### ERP RESULTS

Consistent with previous research, P2 was largest at electrode Cz (4.80 μV) in all conditions (see **Figure 2**). Accordingly, the mean amplitudes of this electrode and eight adjacent electrodes (FCz, FC1, FC2, C1, C2, CPz, CP1, and CP2) were chosen for further analysis. P3 and the slow wave (SW) reached maximum (10.34 μV for P3, and 8.58 μV for slow wave) at electrode Pz, The

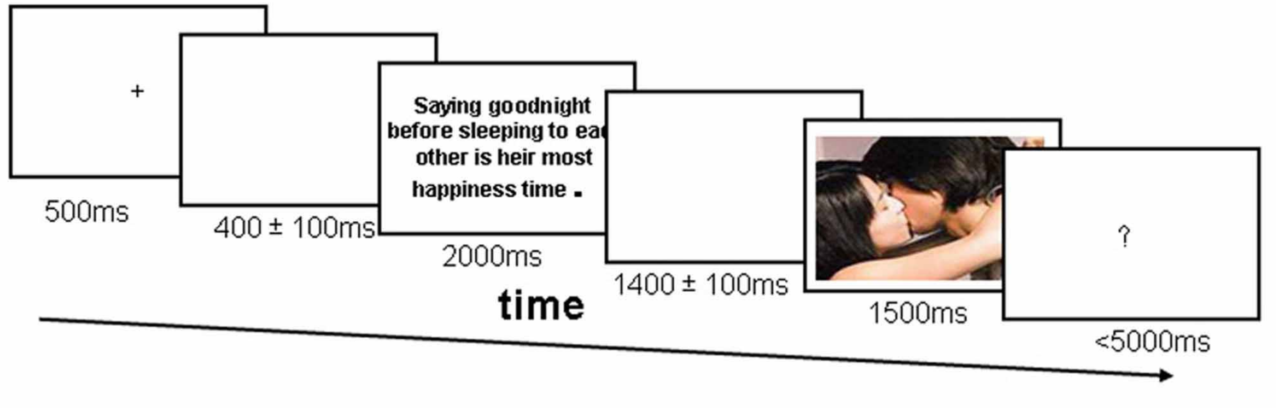


**FIGURE 1 | Experimental design for a single trial.** To give an example, the neutral description for an erotic picture might be "This is a scene of Gaby and Hellen in the room," the positive description for the same erotic picture might be "No matter where or when, the couple loves each other very much;" the neutral description for a neutral picture (e.g., a computer) might be "This desktop computer belongs to Dike and Amy," the positive description for the same neutral picture might be "It is the computer that gave them a chance to know each other and fall in love."

mean amplitudes of this electrode and 8 adjacent electrodes (CPz, CP1, CP2, POz, PO1, PO2, P1, and P2) were chosen for further analysis.

### *P2 (150–200 ms)*

**Figure 3** presents the mean amplitude for each component in each condition. The amplitude of P2 significantly differed between neutral pictures and pleasant (erotic) pictures, $F_{(1,\ 38)} = 7.91, p < 0.01$, indicating the amplitude of pleasant (erotic) pictures ($M = 4.97$) was significantly larger than neutral pictures ($M = 4.11$). The main effect of instruction-type was not significant, $F_{(1,\ 38)} = 2.62$, $p = 0.12$. However, an interaction effect between instruction type and picture type was found to be significant, $F_{(1,\ 38)} = 9.45$, $p < 0.01$. As **Figure 4** shows, the simple-effect analysis of this interaction revealed that for the pleasant (erotic) images, the amplitude of P2 evoked by those images following neutral descriptions ($M = 5.42$) was significantly larger than that of pleasant images following positive descriptions ($M = 4.52$), $p < 0.01$, but not significantly different among the neutral images, $p = 0.25$. The main effect of gender was not significant, $F_{(1,\ 38)} = 1.48$, $p = 0.23$, and gender did not have any significant interactions with the other factors.

**Table 1 | Results of pleasant ratings.**

| Instruction type | Picture type | M | | SD | |
|---|---|---|---|---|---|
| | | Male | Female | Male | Female |
| Positive | Pleasant | 7.13 | 6.92 | 1.02 | 0.95 |
| | Neutral | 5.74 | 6.26 | 0.77 | 1.08 |
| Neutral | Pleasant | 6.87 | 6.44 | 0.97 | 1.20 |
| | Neutral | 5.79 | 5.50 | 0.90 | 0.66 |

*Note: It is a 9-point scale: from 1 (very unpleasant) to 9 (very pleasant), which is wildly used by previous study (Lang et al., 1999; Bai et al., 2005).*

### *P3 (330–400 ms)*

As expected for the P3 component, there was a significant main effect of picture type, $F_{(1,\ 38)} = 141.70, p < 0.01$, indicating that the amplitude for pleasant (erotic) pictures ($M = 13.07$) was significantly larger than for neutral pictures ($M = 6.11$). There was also a significant main effect of instruction-type, $F_{(1,\ 38)} = 9.04, p < 0.01$. Interestingly, the P3 elicited by pictures

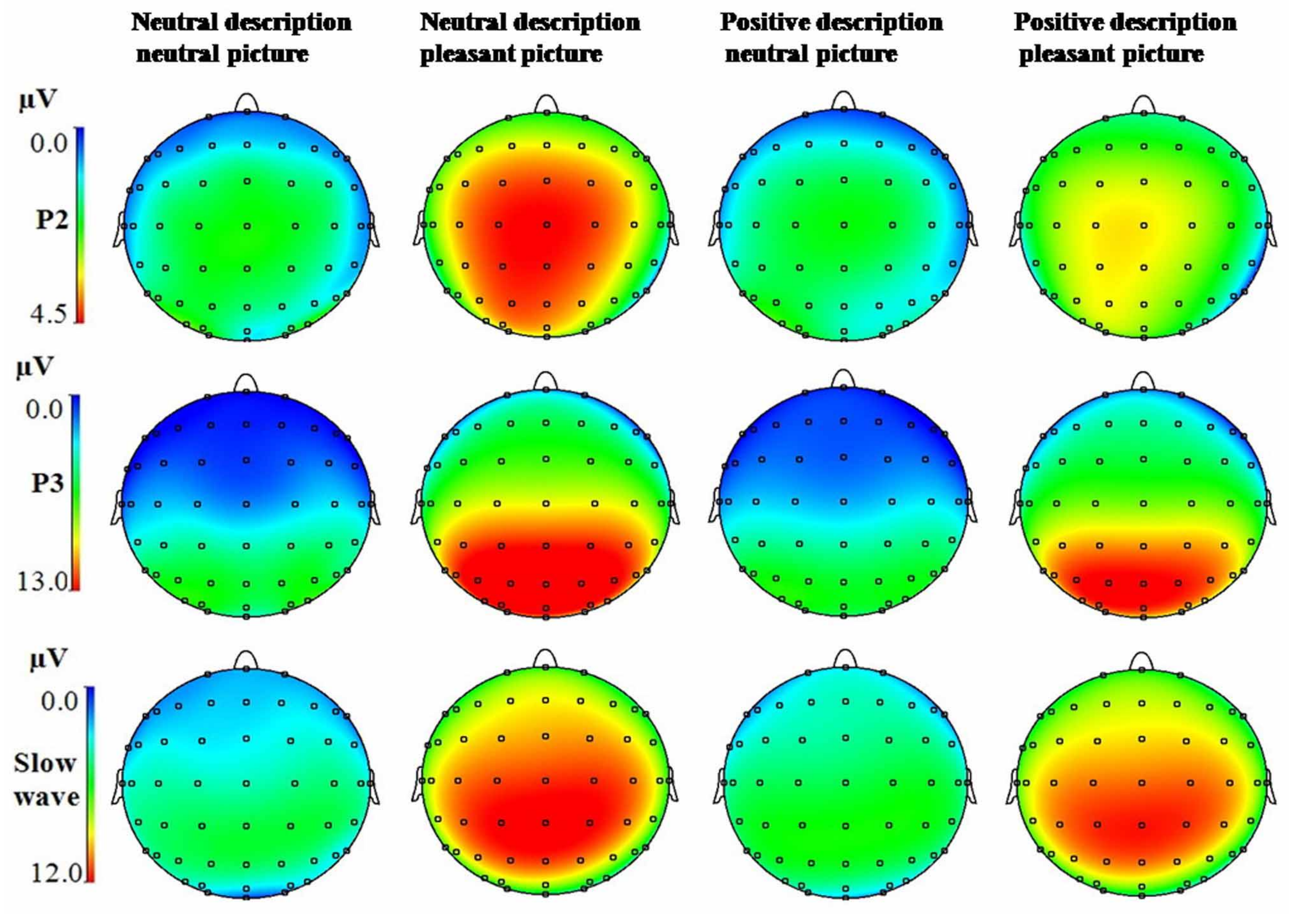


**FIGURE 2 | The topographic map of the P2 (150–200 ms), P3 (330–400 ms) and slow wave (400–800 ms) components for the four conditions.**

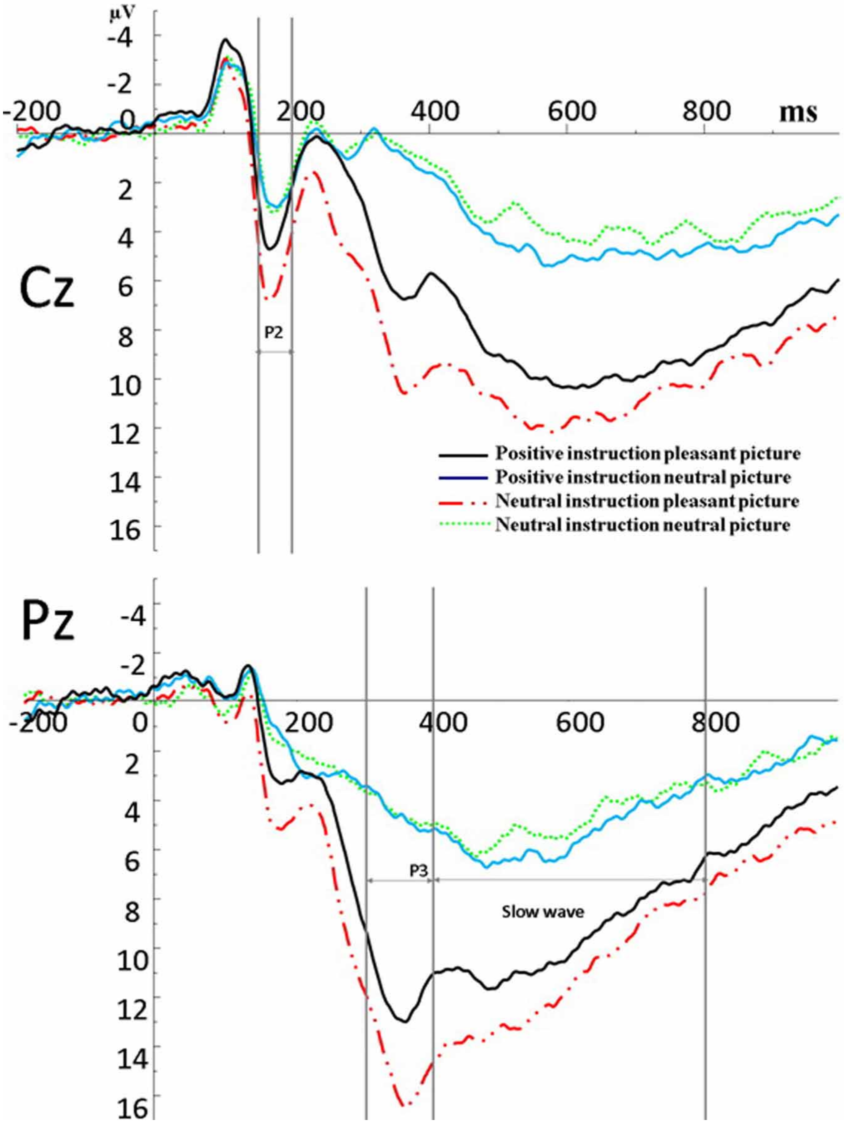


**FIGURE 3 | Stimulus-locked ERPs at Cz and Pz for the four conditions.** The vertical lines at Cz indicate the time windows (150–200 ms) submitted to statistical analysis of the P2 component. The vertical lines at Pz indicate the time windows (330–400 and 400–800 ms) submitted to statistical analysis of the P3 and Slow wave components.

in the neutral instruction condition ($M = 9.88$) was significantly larger than in the positive instruction condition ($M = 9.30$). The interaction between picture type and instruction type was also significant, $F_{(1,\ 38)} = 14.42$, $p < 0.01$. Simple-effect analysis of this interaction revealed that the influence of the instruction type on the P3 associated with pleasant (erotic) images was significant, $p < 0.01$, but not significant on the P3 associated with neutral images, $p = 0.16$ (see **Figures 3**, **4**). The main effect of gender was significant, $F_{(1,\ 38)} = 10.72$, $p < 0.01$, indicating that females had a greater response ($M = 4.97$) than males ($M = 4.11$) when viewing pictures. However, gender did **not** have any other statistically significant effects or interactions.

### *Slow wave (400–800 ms)*

For the SW component, a repeated ANOVA revealed a significant main effect of picture type, $F_{(1,\ 38)} = 118.84$, $p < 0.01$, such that the amplitude of the SW elicited by pleasant (erotic) pictures ($M = 10.10$) was significantly larger than that elicited by neutral pictures ($M = 5.46$). As a main effect, instruction type failed to reach significance, $F_{(1,\ 38)} = 0.16$, $p = 0.69$. Picture type and instruction type yielded a significant interaction, $F_{(1,\ 39)} = 14.87$, $p < 0.01$. Simple-effect analysis of this interaction revealed that for the pleasant (erotic) images, the amplitude of the SW in the positive instruction condition ($M = 9.80$) was significantly smaller than that in the neutral instruction condition ($M = 10.40$), $p < 0.05$. In contrast, the amplitude of the SW in

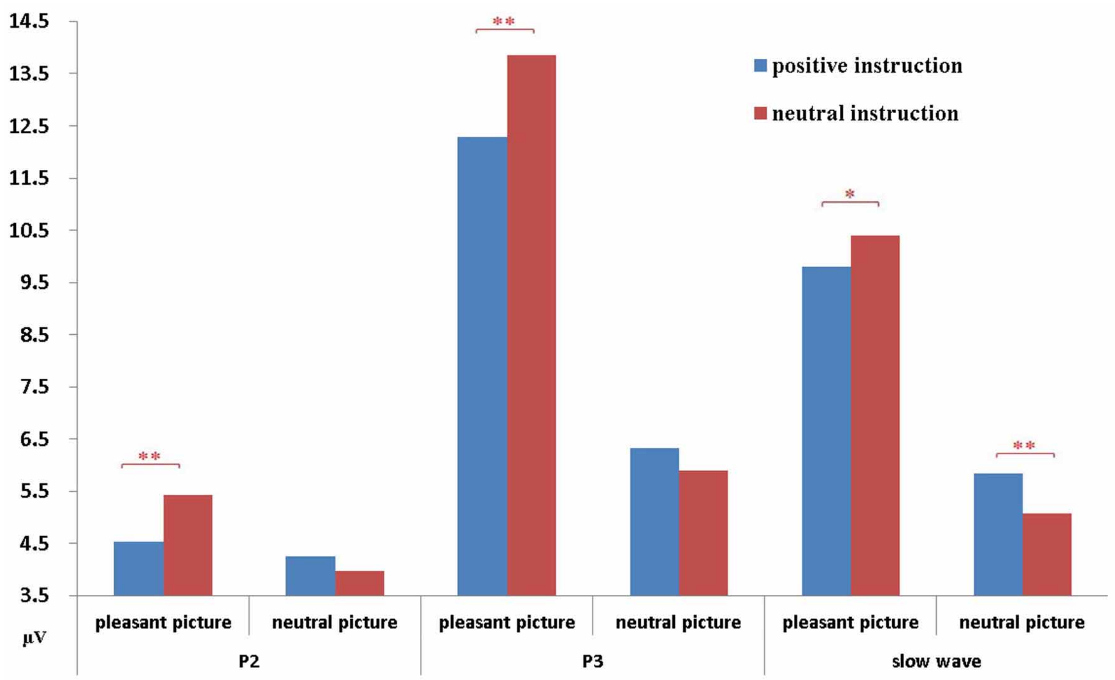


**FIGURE 4 | The effects of instruction-type on P2, P3, and the slow wave component.** Instruction-type had different effects on picture type for P2, P3, and the slow wave components; error bars represent standard errors. *$\alpha < 0.05$, **$\alpha < 0.01$.

the positive instruction condition ($M = 5.84$) was significantly larger than that in the neutral instruction condition for the pleasant (erotic) images, ($M = 5.08$), $p < 0.01$ (see **Figures 3**, **4**). Gender failed to establish a significant main effect, $F_{(1, 38)} = 2.67$, $p = 0.11$, and it did not have any significant interactions with the other factors.

## DISCUSSION

The present study investigated the effects of description-based reappraisal on the up-regulation of positive emotions induced by erotic or neutral images in a Chinese population. Our goal was to assess the electrophysiological index of positive-emotion enhancement with the application of a description-based reappraisal strategy. Self-reported measures revealed that erotic pictures were rated significantly more positive than neutral pictures, suggesting that erotic pictures induced positive emotions. In addition, pictures following positive instructions were rated more positively than those following neutral instructions, regardless of picture content (erotic or neutral). These results validate an enhancement of positive emotions, indicating that description-based reappraisal is sufficient for positive emotion up-regulation.

In regards to ERP, highly arousing erotic images evoked significantly larger mean amplitudes of P2 and LPP components (including P3 and the slow wave) compared to low-arousal neutral images, a finding that is consistent with prior ERP research (Hajcak and Nieuwenhuis, 2006; Moser et al., 2006; Foti and Hajcak, 2008). These results support the theory that the LPP amplitude is sensitive to the level of emotional arousal (Schupp et al., 2000, 2004; Keil et al., 2001).

Importantly, our data demonstrated that description-based reappraisal regulated the amplitudes of P2 and LPP to the erotic images and the SW response to the neutral images. More specifically, erotic images caused attenuated P2 and LPP responses in the up-regulation condition, while larger P2 and LPP signals were observed in the control condition; neutral images induced an increased SW in the un-regulation condition compared to the control condition. Description-based reappraisal influences participants' emotional experiences by providing a context in which the scene is re-evaluated, and thus affecting the allocation of attentional resources. In our opinion, the results of current study support the existence of the context-dependence effect which has been observed in many domains, such as visual perception, emotion perception and social cognition (Bar, 2004; Barrett et al., 2007; Fedota et al., 2012; Ibañez and Manes, 2012).

In the present study, we investigated the mechanisms of the up-regulation of pleasant emotion with a specific focus on the process of description-based reappraisal regulating pleasant emotions. In our opinion, description-based reappraisal helps an individual maintain a "moderate arousal state." It is widely recognized that a state of moderate arousal improves memory, while extreme or low arousal impairs it (Seybold, 2007; McMorris et al., 2011). Psychologists have suggested that individuals feel most contentment at the moderate level of arousal (Molinsky, 2007). It is also worth noting that highly pleasurable pictures are usually linked with moderate arousal levels (e.g., in the IAPS, highly pleasurable pictures, such as pictures of babies, had a valence of approximately 8 while their arousal levels were approximately 5, see Lang et al., 1999). Saarni suggests that the outcome of effective (positive) regulation is for an excited organism to return to a state of equilibrium (Saarni, 1997). In other words, positive regulation, which aims to regulate emotion optimistically, including up-regulation of high-arousal positive emotions, down-regulation of highly negative emotions, and up-regulation of low-arousal neutral images, will lead to a moderate arousal state. Numerous findings, including fMRI results, support the idea of the down-regulation of negative emotions to a moderate arousal state. For instance, Ochsner et al. (2004) used a reappraisal strategy to down-regulate highly negative emotions by asking participants to rethink the positive outcomes of negative scenes. The fMRI results revealed decreased activation of the amygdala (Ochsner et al., 2004), which is associated with the state of arousal (Gläscher and Adolphs, 2003). In terms of ERP studies, previous research found that when preceded by neutral descriptions, the amplitude of the LPP elicited by unpleasant images was decreased (Foti and Hajcak, 2008; Macnamara et al., 2009). Another study using moderately arousing pleasant pictures reported that the amplitude of the LPP in the up-regulation condition was smaller (but not significantly) than that observed in response to passive viewing of images (Krompinger et al., 2008). In the present study, highly arousing erotic pictures elicited extremely large LPP signals in the neutral-descriptions condition, demonstrating that the arousal level of the erotic pictures was extremely high, which is consistent with previous research carried out in a sample of Taiwanese (Yen et al., 2010). We suggest that when erotic pictures were reappraised to be more positive, participants would be less aroused to keep a moderate state of arousal rather than becoming more highly aroused.

Another notable finding in the present study is that neutral images induced an increased SW response in the up-regulation condition, which also supported the idea that positive regulation leads to a moderate state of arousal. The neutral images were non-emotional and non-arousing. Therefore, it is understandable that neutral images induced an extremely small SW in the control condition, which is in accordance with prior research (Hajcak and Nieuwenhuis, 2006; Moser et al., 2006; Foti and Hajcak, 2008; Krompinger et al., 2008). When neutral images were guided by positive descriptions in the up-regulation condition, positive emotions (e.g., love, contentment) emerged. Hence, neutral images were experienced as emotional stimuli and therefore induced an increased SW response (see also Macnamara et al., 2009). Nevertheless, it is a pity that we only asked participants to rate the valence of the pictures because we then lack information about the arousal levels of the pictures. Further studies should ask participants to rate the arousal levels of the pictures as well as the valence.

Regarding the influence of gender on emotion processing, self-reported measures failed to find any difference between males and females in the up-regulation of emotions. However, ERP results revealed that P3 was sensitive to gender such that females exhibited a larger P3 than males when processing either erotic or neutral pictures. Similar to previous findings (Whittle et al., 2011), it seems that females were more emotionally reactive than males. ERP results showed that the up-regulation effect was not different between males and females, suggesting that the description-based reappraisal strategy could effectively up-regulate emotion in both genders.

Finally, because the current study focused on whether the up-regulation strategy could produce more positive emotions, we asked the participants to rate the subjective pleasure level of each picture, although the arousal level remained untested. Consequently, it is unclear whether the factor of emotional arousal could account for our results. We admit that this is a limitation, and further studies interested in the potential influence of emotion-regulation instruction on the arousal level of stimuli should take this point into account. Another limitation is that we only used erotic images as pleasant stimuli. Consequently, the question of whether the significant results found in the present study could be applied to other types of pleasant images needs further study.

In summary, the self-reported ratings from our study demonstrated the description-based reappraisal strategy as an effective up-regulator of positive emotions in both genders. We found that the description-based reappraisal regulating the positive emotions induced by erotic images starts at the earlier phase of the ERP (P2, 150–200 ms) and results in decreased ERP responses to the erotic images. In contrast, the up-regulation of positive emotions induced by neutral images starts at the later phase of the ERP (the slow wave, 400–800 ms) and causes a larger SW.

## ACKNOWLEDGMENTS

The authors thank Agustin Ibanez, the editor of this paper, and the two reviewers for their constructive comments. This research was supported by the National Natural Science Foundation of China (30930031, 91132704), the Ministry of Science and Technology (973 Program, 2011CB711000), the National Key Technologies R&D Program (2009BAI77B01), the Global Research Initiative Program, United States National Institute of Health grants (1R01TW007897), the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (No. Y2CQ013005) and Humanities and Social Sciences project, Ministry of Education of China (09YJXLX008).

## REFERENCES

Amstadter, A. (2008). Emotion regulation and anxiety disorders. *J. Anxiety Disord.* 22, 211–221.

Bai, L., Ma, H., Huang, Y.-X., and Luo, Y.-J. (2005). The development of native Chinese affective picture system-A pretest in 46 college students. *Chin. Ment. Health J.* 19, 719–722.

Bar, M. (2004). Visual objects in context. *Nat. Rev. Neurosci.* 5, 617–629.

Barrett, L. F., Lindquist, K. A., and Gendron, M. (2007). Language as context for the perception of emotion. *Trends Cogn. Sci.* 11, 327–332.

Fedota, J. R., McDonald, C. G., Roberts, D. M., and Parasuraman, R. (2012). Contextual task difficulty modulates stimulus discrimination: electrophysiological evidence for interaction between sensory and executive processes. *Psychophysiology* 49, 1384–1393.

Foti, D., and Hajcak, G. (2008). Deconstructing reappraisal: descriptions preceding arousing pictures modulate the subsequent neural response. *J. Cogn. Neurosci.* 20, 977–988.

Frühholz, S., Fehr, T., and Herrmann, M. (2009). Early and late temporo-spatial effects of contextual interference during perception of facial affect. *Int. J. Psychophysiol.* 74, 1–13.

Gläscher, J., and Adolphs, R. (2003). Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *J. Neurosci.* 23, 10274–10282.

Goldin, P. R., McRae, K., Ramel, W., and Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* 63, 577–586.

Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Rev. Gen. Psychol.* 2:271. doi: 10.1037/1089-2680.2.3.271

Hajcak, G., and Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cogn. Affect. Behav. Neurosci.* 6, 291–297.

Higgins, L. T., Zheng, M., Liu, Y., and Sun, C. H. (2002). Attitudes to marriage and sexual behaviors: a survey of gender and culture differences in China and United Kingdom. *Sex Roles* 46, 75–89.

Hurtado, E., Haye, A., González, R., Manes, F., and Ibáñez, A. (2009). Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. *BMC Neurosci.* 10:69. doi: 10.1186/1471-2202-10-69

Ibanez, A. M., Melloni, M., Huepe, D., Helgiu, E., Rivera-Rei, A., Canales-Johnson, A., et al. (2012). What event-related potentials (ERPs) bring to social neuroscience? *Soc. Neurosci.* 7, 632–649.

Ibañez, A., and Manes, F. (2012). Contextual social cognition and the behavioral variant of frontotemporal dementia. *Neurology* 78, 1354–1362.

Ibáñez, A., Hurtado, E., Riveros, R., Urquina, H., Cardona, J. F., Petroni, A., et al. (2011). Facial and semantic emotional interference: a pilot study on the behavioral and cortical responses to the dual valence association task. *Behav. Brain Funct.* 7, 1354–1362.

Keil, A., Muller, M. M., Gruber, T., Wienbruch, C., Stolarova, M., and Elbert, T. (2001). Effects of emotional arousal in the cerebral hemispheres: a study of oscillatory brain activity and event-related potentials. *Clin. Neurophysiol.* 112, 2057–2068.

Kringelbach, M. L., and Berridge, K. C. (2009). Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn. Sci.* 13, 479–487.

Krompinger, J. W., Moser, J. S., and Simons, R. F. (2008). Modulations of the electrophysiological response to pleasant stimuli by cognitive reappraisal. *Emotion* 8, 132–137.

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1999). "International affective picture system (IAPS): instruction manual and affective ratings," in *The Center for Research in Psychophysiology,* (University of Florida).

Langston, C. A. (1994). Capitalizing on and coping with daily-life events: expressive responses to positive events. *J. Pers. Soc. Psychol.* 67, 11–12.

Levesque, J., Eug Ne, F., Joanette, Y., Paquette, V., Mensour, B., Beaudoin, G., et al. (2003). Neural circuitry underlying voluntary suppression of sadness. *Biol. Psychiatry* 53, 502–510.

Luo, W., Feng, W., He, W., Wang, N.-Y., and Luo, Y.-J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *Neuroimage* 49, 1857–1867.

Lykins, A. D., Meana, M., and Strauss, G. P. (2007). Sex differences in visual attention to erotic and non-erotic stimuli. *Arch. Sex. Behav.* 37, 219–228.

Macnamara, A., Foti, D., and Hajcak, G. (2009). Tell me about it: neural activity elicited by emotional pictures and preceding descriptions. *Emotion* 9, 531–543.

Macnamara, A., Ochsner, K. N., and Hajcak, G. (2010). Previously reappraised: the lasting effect of description type on picture-elicited electrocortical activity. *Soc. Cogn. Affect. Neurosci.* 6, 348–358.

Mak, A. K. Y., Hu, Z.-G., Zhang, J. X., Xiao, Z.-W., and Lee, T. M. C. (2009). Neural correlates of regulation of positive and negative emotions: an fMRI study. *Neurosci. Lett.* 457, 101–106.

Masters, J. C. (1991). "Strategies and mechanisms for the personal and social control of emotion," in *The Development of Emotion Regulation and Dysregulation*, eds J. Garber, and K. A. Dodge (New York, NY: Cambridge University Press), 182–207.

McMorris, T., Sproule, J., Turner, A., and Hale, B. J. (2011). Acute, intermediate intensity exercise, and speed and accuracy in working memory tasks: a meta-analytical comparison of effects. *Physiol. Behav.* 102, 421–428.

Molinsky, A. (2007). Cross-cultural code-switching: the psychological challenges of adapting behavior in foreign cultural interactions. *Acad. Manage. Rev.* 32, 622–640.

Moser, J. S., Hajcak, G., Bukay, E., and Simons, R. F. (2006). Intentional modulation of emotional responding to unpleasant pictures: an ERP study. *Psychophysiology* 43, 292–296.

Ochsner, K. N., Bunge, S. A., Gross, J. J., and Gabrieli, J. D. E. (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., et al. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage* 23, 483–499.

Olofsson, J. K., Nordin, S., Sequeira, H., and Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biol. Psychol.* 77, 247–265.

Parrott, W. G., and Schulkin, J. (1993). Neuropsychology and the cognitive nature of the emotions. *Cogn. Emot.* 7, 43–59.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148.

Saarni, C. (1997). Coping with aversive feelings. *Motiv. Emot.* 21, 45–63.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., and Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology* 37, 257–261.

Schupp, H. T., Junghöfer, M., Weike, A. I., and Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: an ERP analysis. *Psychophysiology* 41, 441–449.

Seybold, K. S. (2007). Physiological mechanisms involved in religiosity/spirituality and health. *J. Behav. Med.* 30, 303–309.

Taylor, S. F., and Liberzon, I. (2007). Neural correlates of emotion regulation in psychopathology. *Trends Cogn. Sci.* 11, 413–418.

Whittle, S., Yücel, M., Yap, M. B. H., and Allen, N. B. (2011). Sex differences in the neural correlates of emotion: evidence from neuroimaging. *Biol. Psychol.* 87, 319–333.

Yen, N.-S., Chen, K.-H., and Liu, E. H. (2010). Emotional modulation of the late positive potential (LPP) generalizes to Chinese individuals. *Int. J. Psychophysiol.* 75, 319–325.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 15 October 2012; accepted: 21 December 2012; published online: 10 January 2013.*

*Citation: Peng J, Qu C, Gu R and Luo Y-J (2013) Description-based reappraisal regulate the emotion induced by erotic and neutral images in a Chinese population. Front. Hum. Neurosci.* ***6****:355. doi: 10.3389/fnhum.2012.00355*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 18 January 2013
doi: 10.3389/fnhum.2012.00358

# Lateralized interactive social content and valence processing within the human amygdala

**Pascal Vrtička[1,2,3,4]*, David Sander[2,4] and Patrik Vuilleumier[2,3]**

[1] Center for Interdisciplinary Brain Sciences Research, Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University, Stanford, CA, USA
[2] Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland
[3] Laboratory for Neurology and Imaging of Cognition, Department of Neurology and Department of Neurosciences, University Hospital and Medical School, University of Geneva, Geneva, Switzerland
[4] Laboratory for the Study of Emotion Elicitation and Expression (E3 Lab), Department of Psychology, FPSE, University of Geneva, Geneva, Switzerland

**Edited by:**
Agustin Ibanez, Institute of Cognitive Neurology, Argentina

**Reviewed by:**
Reiko Graham, Texas State University, USA
Vladimir Lopez, Pontificia Universidad Catolica de Chile, Chile

***Correspondence:**
Pascal Vrtička, Center for Interdisciplinary Brain Sciences Research, Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University, 401 Quarry Rd., MC 5795, Stanford, CA 94305, USA.
e-mail: pvrticka@stanford.edu

In the past, the amygdala has generally been conceptualized as a fear-processing module. Recently, however, it has been proposed to respond to all stimuli that are relevant with respect to the current needs, goals, and values of an individual. This raises the question of whether the human amygdala may differentiate between separate kinds of relevance. A distinction between emotional (vs. neutral) and social (vs. non-social) relevance is supported by previous studies showing that the human amygdala preferentially responds to both emotionally and socially significant information, and these factors might even display interactive encoding properties. However, no investigation has yet probed a full 2 (positive vs. negative valence) × 2 (social vs. non-social content) processing pattern, with neutral images as an additional baseline. Applying such an extended orthogonal factorial design, our fMRI study demonstrates that the human amygdala is (1) more strongly activated for neutral social vs. non-social information, (2) activated at a similar level when viewing social positive or negative images, but (3) displays a valence effect (negative vs. positive) for non-social images. In addition, this encoding pattern is not influenced by cognitive or behavioral emotion regulation mechanisms, and displays a hemispheric lateralization with more pronounced effects on the right side. Finally, the same valence × social content interaction was found in three additional cortical regions, namely the right fusiform gyrus, right anterior superior temporal gyrus, and medial orbitofrontal cortex. Overall, these findings suggest that valence and social content processing represent distinct kinds of relevance that interact within the human amygdala as well as in a more extensive cortical network, likely subserving a key role in relevance detection.

**Keywords: social content, valence, human amygdala, trait anxiety, fMRI**

## INTRODUCTION

Thanks to the advancement of neuroimaging techniques and paradigms, our knowledge on human amygdala function has steadily increased during the last two decades. Yet, divergent views have emerged concerning stimulus properties that trigger amygdala responses. One classic notion is that the amygdala constitutes a *fear module* crucially involved in the automatic detection of threat-related information, fear reaction and fear learning (Ohman and Mineka, 2001). Other views suggest instead that the amygdala may preferentially process arousal (Anderson et al., 2003; Small et al., 2003) *or* valence (Murray, 2007) information. According to such notions, the amygdala would either represent a general *arousal indicator* or *valence processor*. Finally, besides these traditional accounts respectively derived from basic (Ekman, 1999) or bi-dimensional (Russell, 1980) theories of emotion, recent experimental evidence revealed amygdala activation to be determined by many additional factors (not confounded by arousal), such as eye gaze (N'Diaye et al., 2009; Sato et al., 2010), novelty (Blackford et al., 2010; Weierich et al., 2010; Balderston et al., 2011), social content (Norris et al., 2004; Britton et al., 2006; Goossens et al., 2009; Scharpf et al., 2010), context (Kim et al., 2003, 2004; Vrtická et al., 2008), personal impact (Ewbank et al., 2009), or individual differences in subjective evaluation (Schiller et al., 2009), motivational state (Canli et al., 2001; LaBar et al., 2001; Morris and Dolan, 2001) as well as various psychological traits (Canli et al., 2001; Bishop et al., 2004; Etkin et al., 2004; Sabatinelli et al., 2005; Dickie and Armony, 2008; Vrticka et al., 2008, 2012a; Vrticka and Vuilleumier, 2012). To integrate these different findings, a new account of human amygdala function has been put forward, primarily linking it with the appraisal of biological relevance (Sander et al., 2003; Sergerie et al., 2006; Adolphs, 2010; Pessoa and Adolphs, 2010). This concept of relevance detection has its origins in emotion psychology (Sander et al., 2003), in particular in appraisal theories of emotion (Sander et al., 2005), and refers to the preferential processing of events that are (biologically) relevant to major concerns/goals/needs and values of an individual at a specific moment in time (see Frijda, 2009; Reisenzein, 2009).

When considering human amygdala function in terms of biological relevance detection, the question arises what kind of information might be most relevant, and thus which stimulus properties are preferentially processed by the human amygdala. To address this issue, several studies have compared two different kinds of relevance, in particular emotional *vs.* social relevance (Norris et al., 2004; Britton et al., 2006; Harvey et al., 2007; Scharpf et al., 2010). On the one hand, *emotional relevance* was referred as to stimuli that are likely to be appraised so that they would elicit an emotional response (see Sander et al., 2005) and modulate cognitive processes such as attention (Vuilleumier, 2005; Brosch et al., 2008), independently of their social nature. On the other hand, *social relevance* was associated with stimuli conveying information about interpersonal interactions and conspecifics, regardless of their emotional value. This corroborates the view denoting a high relevance of social information for our species due to its direct link to guiding human behavior (Keltner and Kring, 1998; Hariri et al., 2002). It is also consistent with the so called social brain hypothesis (Dunbar, 1998), stating that the need for social skills led to functional specialization of new cognitive mechanisms and "*…fuelled the expansion of the human brain …*" (Adolphs, 2003, p. 166). Accordingly, the abovementioned investigations reported both emotional relevance effects (emotional vs. neutral) and social relevance effects (social vs. non-social) in the amygdala (Norris et al., 2004; Britton et al., 2006; Scharpf et al., 2010). The last study even suggests an interactive processing of emotional and social relevance (highest amygdala activation for emotional *and* social stimuli), although no formal test for such interactive processing was carried out (Scharpf et al., 2010). Such data corroborate the notion that socially relevant stimuli are most likely also emotionally relevant, in the sense that emotions are typically elicited in social situations or by taking into account the affective or motivational dimension of social contexts (Jakobs et al., 1997; Balderston et al., 2011). However, none of these studies differentiated the emotional relevance effect in terms of valence, distinguishing between positive *vs.* negative value (Morrison and Salzman, 2010), therefore leaving the question open of whether valence processing (e.g., threat) may (at least partly) determine the response to social relevance. According to the literature, negative stimuli should activate the amygdala stronger than positive ones due to their intrinsically higher biological relevance in terms of survival (Hariri et al., 2002). To clarify relevance detection in the human amygdala, it would therefore be useful to test for a full 2 (valence) × 2 (social content) interaction pattern, which was not possible in previous studies (Norris et al., 2004; Britton et al., 2006; Harvey et al., 2007; Scharpf et al., 2010). Consequently, the current functional magnetic resonance imaging (fMRI) study applied such experimental design, using both social and non-social *neutral* images serving as an additional baseline.

In addition, fMRI investigations in the field of emotion regulation have disclosed differential amygdala activation to positive and negative stimuli as a function of task instructions or viewing conditions, particularly through emotion regulation processes. The latter may involve cognitive (re-appraisal) or behavioral (expressive suppression) regulation strategies, aimed at either up- or down-regulating emotional states (Ochsner et al., 2002, 2004; Levesque et al., 2003; Kim and Hamann, 2007; Goldin et al., 2008). We therefore also included these two major regulation strategies (together with a natural viewing condition) in our experimental paradigm to further test for potential task effect on relevance detection mechanisms.

Furthermore, as already mentioned above, amygdala activity is consistently modulated as a function of motivational state (Canli et al., 2001; LaBar et al., 2001; Morris and Dolan, 2001) or personality traits (Canli et al., 2001; Bishop et al., 2004; Sabatinelli et al., 2005; Vrticka et al., 2008, 2012a,b; Vrticka and Vuilleumier, 2012). Therefore, we also included measures assessing individual differences in the current study, in order to probe for relations between relevance detection mechanisms and personal dispositions. First, we chose trait anxiety (STAI-T, see section "Methods"), because it has already previously been shown to modulate amygdala activity to threat- or fear-related stimuli (Bishop et al., 2004; Etkin et al., 2004; Dickie and Armony, 2008), and thus influence negative valence processing. Second, we assessed attachment style (secure vs. insecure—avoidant or anxious; Relationships Scales Questionnaire, see section "Methods"), because we have previously shown that this personality trait can influence social emotional processing within the human amygdala and increase blood-oxygen-level-dependent (BOLD) responses to social negative interaction scenarios (Vrticka et al., 2008, 2012a; Vrticka and Vuilleumier, 2012).

Finally, besides the question to what stimulus types or properties the amygdala would respond to preferentially, another issue on human amygdala activation concerns any *hemispheric lateralization* during social and/or emotional processing. Hemispheric lateralization of emotion in general has been proposed in different ways, either in terms of (1) fundamentally more right-lateralized emotion processing, regardless of valence; (2) a preferential representation of positive vs. negative emotions in the left vs. right hemispheres, respectively—and thus as a function of valence; or (3) a distinction between approach (left) vs. avoidance (right) behavioral tendencies, rather than valence (Sergerie et al., 2006). Regarding lateralization in the human amygdala more specifically, one prominent account has been related to language and proposed a differential representation of semantic (left) vs. non-semantic (right) information (Sergerie et al., 2006), whereas another hypothesis highlighted differential temporal dynamics and suggested faster emotional information processing with quicker habituation in the right as compared to the left amygdala (Sergerie et al., 2006). Amygdala lateralization has also been related to sex differences (Cahill, 2006), but here we chose to recruit only female participants to avoid confound related to this additional factor. However, assumptions on hemispheric lateralization have generally focused on the valence dimension of emotional information only, and it still remains to be seen how such lateralization accounts may be affected by social relevance.

According to the previous literature and theoretical considerations of the appraisal theory summarized above (Sander et al., 2003), we predicted the following findings. First, the amygdala should display a social *vs.* non-social activation difference (for both neutral and emotional stimuli), because social information is thought to be particularly relevant for our species. Second, this social relevance effect should interact with the representation of valence in terms of a negative *vs.* positive activation difference, as social threat is likely to be even more strongly relevant.

Third, personality (trait anxiety and/or attachment style) should modulate these valence and social effects in the human amygdala, because the appraisal of social emotional stimuli is strongly dependent on their significance for the observer in terms of personal values and needs. Thereby, social *and* emotional stimuli should be affected most, again due to their highest significance for the human species.

## METHODS

### SUBJECTS

We recruited 19 healthy paid volunteers (all right-handed women, mean age 24.82 ± 4.0), who all had a normal or corrected to normal vision, no history of neurological or psychiatric disease, and gave informed written consent according to the local ethical committee regulation. Only women were included in order to avoid any potential sex differences that could have modulated the effects of interest (see e.g., Cahill, 2006; Kim and Hamann, 2007; Vrticka et al., 2012a,b). fMRI data from the same study were previously reported in Vrticka et al. (2011) and Vrticka et al. (2012a), but the latter focused on whole-brain results comparing specific emotion regulation strategies and distinct attachment styles. In addition, behavioral data derived from stimulus validation (see section "Stimuli" below) was published in Vrticka et al. (2012b). Here, we specifically examine amygdala activity and brain areas that show main effects of social and emotional relevance, independently of any task effects, and we investigate the role of trait anxiety (rather than individual attachment style, see below).

### EXPERIMENTAL MATERIAL AND PROCEDURE

#### *Stimuli*

A total number of 360 emotional pictures were initially collected from the internet and from the International Affective Pictures System (IAPS). All were in colors, and adjusted to obtain similar size, contrast, and pixel resolution. Half of the pictures displayed scenes with a clear social content, such as two people fighting or a mother interacting with her baby. The other half represented animals, objects or landscapes that were not social, like a dead bird in industrial waste or a tropical island scene. All 360 pictures were rated in a separate behavioral study by 54 female students on three continuous rating scales (from 1 to 100), including PLEASANTNESS (from very negative to very positive), INTENSITY (from low to high arousal), and CONTROL (from absence to full presence of control over the emotional experience induced by viewing images (Vrticka et al., 2012b). According to the average rating results from this sample, 240 pictures were finally chosen for the fMRI study, and sorted by their SCENE CONTENT (either social or non-social) and VALENCE (either positive or negative). This gave rise to four stimulus categories (60 pictures each): Social Positive (SP) or Negative (SN), and Non-social Positive (NSP) or Negative (NSN). Negative images were rated as lower in pleasantness and control, but higher in intensity than positive images ($ps < 0.001$). However, there were no differences between social vs. non-social images ($ps > 0.25$), and no interactions ($ps > 0.11$). In addition, there were no differences in luminance overall ($ps > 0.098$), and social complexity for social images specifically (number of humans per image; $p = 0.5$). Please refer to Vrticka et al., (2011 and 2012a,b) for detailed values. Note that the differences in intensity between negative and positive stimuli could not be avoided in order to match pairs of social and non-social scenes in both valence conditions, because social material is otherwise typically judged as much more intense than non-social material (Ewbank et al., 2009). Finally, we also selected 40 neutral images from the IAPS database (20 including humans, 20 without humans) to be used in a baseline control condition (see below), with valence ratings situated between positive and negative images.

#### *Experimental conditions*

Before entering the fMRI scanner, all participants were told that the purpose of the experiment would be to investigate how the brain reacts to different types of images (e.g., real scenes *vs.* fiction scenes) and to which degree people can voluntarily influence the emotional impact of these images on them. Accordingly, the experimental layout comprised four different viewing conditions in which pictures were presented with different tasks to induce different emotion regulation strategies.

The first condition was used as a control baseline, and was introduced to the participants as "a photographic quality" judgment, where they had to indicate on each trial (by button press, using a 4-point scale—see below) whether the image was of good quality (e.g., well-focused or properly lighted). All images in this condition were neutral, but could display either scenes with humans (i.e., social content) or inanimate settings and landscapes (i.e., non-social content). This viewing condition was later used to provide a baseline for general differences in brain activation to social *vs.* non-social stimuli, irrespective of emotional processing demands and valence. It was presented as the first block of the first scanning run and the last block of the last run.

The three other viewing conditions included emotional images only, and comprised an emotion experience, a cognitive re-evaluation (re-appraisal), and a behavioral expressive suppression condition. Because there was no significant three-way interaction in any region of interest (see section "Results"), activations were collapsed during final data analysis. For more details regarding the experimental layout, please refer to the study where the emotion regulation effects were explored systematically as a function of the different viewing conditions (Vrticka et al., 2011).

The participant's task was to report their feeling state evoked by the preceding stimulus ("How did you feel while seeing the last image"?), using a 4-point scale (see below). All emotional images were counterbalanced across participants, so that the same images seen in one viewing condition by a given participant were seen in the other viewing conditions by different participants.

#### *Procedure*

The fMRI experiment was divided into three successive scanning runs. Each run included two of the three viewing conditions, presented in blocks of 40 emotional images (duration = 294 s per block), whereas the first and the last run also included an additional block of 20 neutral images (baseline condition, duration = 151 s). Within each block, images were pseudo-randomized and equally probable for the different stimulus categories (social *vs.* non-social content, positive *vs.* negative valence). The first and the third runs lasted approximately 13 min, and the middle run 10 min.

Each viewing condition block began with an instruction display (7 s), followed by images in pseudo-randomized order. Every individual trial started with a fixation cross at the screen center (average duration = 1125 ms jittered between 790 and 1485 ms), followed by an emotional or neutral image for 2 s, and then a response display probing for emotion ratings (4 s; see **Figure 1**). Ratings were made on a 4-button response box, according to a 4-point scale ranging from very and slightly negative (buttons 1 and 2, respectively) to slightly and very positive (buttons 3 and 4, respectively).

### *MRI acquisition*

MRI data were acquired on a 3 T whole-body scanner (Siemens TIM TRIO, Erlangen, Germany), using standard head-coil configuration. For each participant, a structural image was obtained with a MPRAGE T1-weighted sequence (TI/TR/TE/flip = 900/1900/2.32/9°, parallel acquisition (GRAPPA) with acceleration factor 2, FOV = 230 × 230 × 173 mm$^3$, Matrix = 256 × 246 × 192). Functional images (TR/TE/Flip = 2200 ms/30 ms/85°, parallel acquisition (GRAPPA) with acceleration factor 2, FOV = 235 mm × 235 mm, matrix = 128 × 84, resulting voxel size is 2.8 × 1.8 × 3.4 mm$^3$) covered the whole brain, composed of 36 contiguous 4 mm axial slices parallel to the inferior edge of the occipital and temporal lobes, and acquired continuously for a total of 975 images per participant (two sessions with 350 and one session with 275 images).

Image processing was performed with SPM2 (www.fil.ion.ucl.ac.uk) using standard procedures for realignment of the time-series, slice-timing correction, normalization to a standard brain template in MNI space, and smoothing with an 8 mm FWHM Gaussian kernel. Statistical analysis was performed using the general linear model implemented in SPM2, with a separate regressor for each event type convolved with a canonical hemodynamic response function. Twelve event types from the emotion regulation task (4 image categories: SP, NSP, SN, and NSN; for each of the three viewing conditions), plus two additional event types (social and non-social) from the baseline condition were modeled for each participant, using the three scanning runs in a fixed-effect analysis at the single-subject level. Movement parameters from realignment corrections were entered as additional covariates of no interest for each scanning run, in order to account for residual movement artifacts after realignment. Statistical parametric maps were then generated from linear contrasts between the different conditions in each participant.

A second-stage random-effect analysis was performed using one-sample *t*-tests on contrast images obtained in each subject for each comparison of interest. Because of *a-priori* predictions regarding amygdala, we opted to threshold amygdala activations at $p < 0.005$ (uncorrected, whole brain) and $k \geq 10$. All

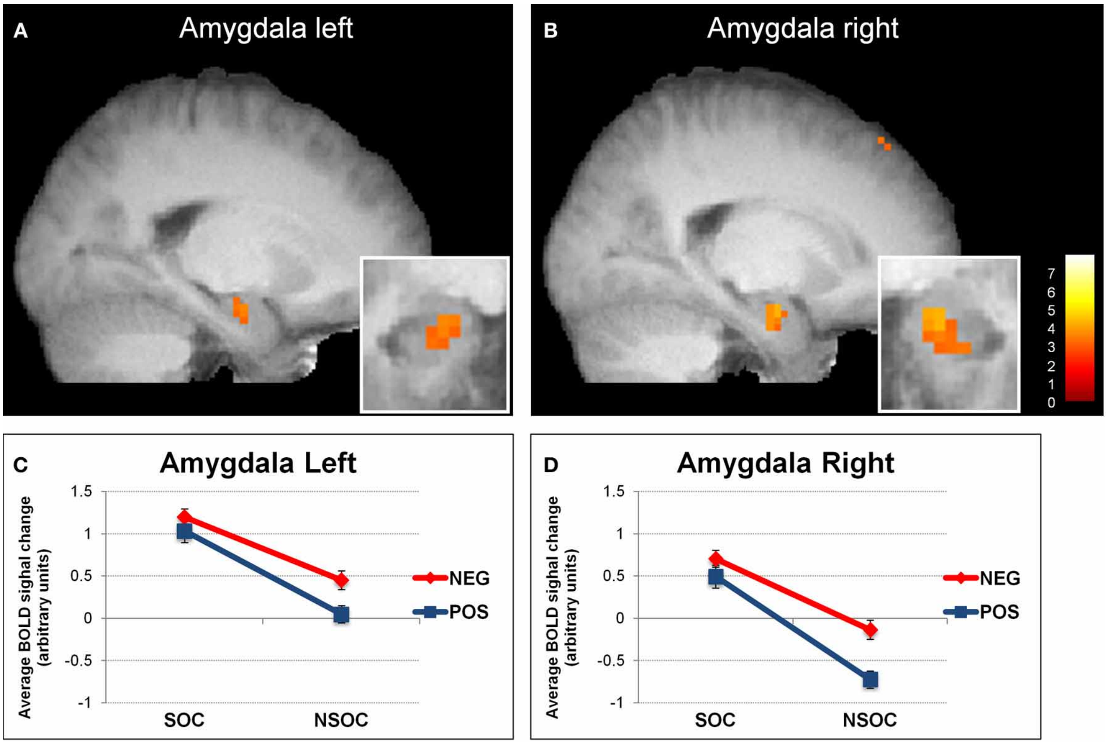


**FIGURE 1 | Interactive valence and social content processing in bilateral amygdala.** Statistical parametric maps (threshold $p = 0.005$) for the contrast social > non-social displaying increased activity in the left **(A)** and right **(B)** amygdala, (big windows: sagittal view; small windows: coronal view). Graphs depicting the extracted activation values (betas) averaged across voxels from the left **(C)** and right **(D)** amygdala, displaying a valence × social content interaction. NEG, negative; NSOC, non-social; POS, positive; SOC, social. Error bars reflect S.E.M.

other contrasts were thresholded at $p < 0.001$ (uncorrected, whole brain) and $k \geq 20$ (Lieberman and Cunningham, 2009). Average parameter estimates of activity (betas) for each condition were extracted from all voxels in regions of interest, defined by the full-extent clusters showing significant activation at a voxel level in the SPM group analysis (random-effect contrasts). These beta values were then used for repeated-measure ANOVAs, ANCOVAs and $t$-tests performed in SPSS (Version 18; http://www.spss.com/) with the factors of social content, valence, viewing condition, and personality, when appropriate.

### *Individual diffrence measures*

Trait anxiety was measured using a French version of the Spielberger State-Trait Anxiety Inventory (STAI-T) analyzed according to the author's manual (Spielberger, 1983). Attachment style was assessed by a validated French version (Guédeney et al., 2010) of the relation scales questionnaire (RSQ; Collins and Read, 1994), analyzed according to Kurdek (Kurdek, 2002) relying on an initial model by Simpson and colleagues (Simpson, 1990; Simpson et al., 2002)—see also (Vrticka et al., 2008, 2012a,b).

Influences of trait anxiety or attachment style on brain activity in regions of interest were assessed by a full 3 (viewing condition) × 2 (valence) × 2 (social content) × 1 (anxiety) or × 2 (attachment style) analysis of covariance (ANVCOVA), by including the respective personality measures for each subject in SPSS (www.ibm.com/software/analytics/spss/). To do so, the raw scores collected from STAI-T and attachment scales were centered to avoid multicollinearity problems in multiple regression (Aiken and West, 1991).

## RESULTS

### BEHAVIORAL RESULTS

A full 3 (viewing condition) × 2 (valence) × 2 (social content) ANOVA on the behavioral ratings revealed (1) a main effect of valence [$F_{(1,\ 18)} = 1149, p < 0.001$] because positive images were always rated as more pleasant than negative images, and (2) a viewing condition × valence interaction [$F_{(1,\ 18)} = 36.79$, $p < 0.001$], because both positive and negative images were rated as less pleasant and less unpleasant, respectively, during re-appraisal as compared to emotion experience and expressive suppression (Vrticka et al., 2011).

### MAIN EFFECTS AND INTERACTIONS OF THE fMRI ANALYSIS

In an initial step, we computed the four main effects contrast regarding valence (positive > negative and vice versa) and social content (social > non-social and vice versa). These analyses revealed significant BOLD signal change in the amygdala solely for the contrast social > non-social (see **Table 1** and **Figure 1**). Because the main aim of the present study was to investigate relevance detection within the amygdala (and only subsidiarily examine any additional brain areas displaying the same computational profile), all subsequent analysis steps are reported for this region first, and then for regions appearing in the same initial social > non-social comparison (functional region of interest determination; see **Table 1**).

**Table 1 | Brain areas activated in the main contrast SOCIAL > NON-SOCIAL, listed with peak coordinates and best estimates of anatomical location.**

| Region | BA | Voxel | *T*-Value | *x* | *y* | *z* |
|---|---|---|---|---|---|---|
| **SOCIAL > NON-SOCIAL** | | | | | | |
| Amygdala right* | | 26 | 4.21 | 21 | −6 | −18 |
| Amygdala left* | | 13 | 3.55 | −21 | −9 | −18 |
| Fusiform Gyrus right | 37 | 98 | 6.59 | 42 | −42 | −27 |
| Anterior STG right | 21 | 111 | 6.29 | 60 | −6 | −24 |
| mOFC | 11 | 40 | 5.26 | −3 | 54 | −18 |
| mPFC | 10 | 28 | 4.36 | 3 | 57 | 15 |
| Temporal Inferior left | 20/21 | 109 | 6.85 | −57 | −3 | −27 |
| PCC | 23 | 393 | 7.32 | 0 | −51 | 33 |
| pSTS left | 19 | 458 | 7.82 | −45 | −84 | 0 |
| pSTS right | 19 | 711 | 7.07 | 45 | −48 | 18 |
| Fusiform Gyrus left | 19/37 | 25 | 4.43 | −42 | −63 | −21 |
| Occipital cortex left | 17 | 331 | 7.74 | −6 | −102 | 9 |

*Statistical threshold was $p < 0.005$ (uncorrected, whole brain) and $k = 10$ for bilateral amygdala as indicated by $^*$ above, and $p < 0.001$ (uncorrected, whole brain) and $k \geq 20$ for remaining brain areas. STG, superior temporal gyrus; OFC, orbitofrontal cortex; PFC, prefrontal cortex; STS, superior temporal sulcus; PCC, posterior cingulate cortex; m, medial; dl, dorsolateral; dm, dorsomedial; p, posterior. The first five brain areas listed were the ones found to display a significant valence × social content interaction; BA, brodman area.*

### *Amygdala*

We first analyzed amygdala activation during the neutral control condition with a 2 (side) × 2 (social content) ANOVA. This confirmed a main effect of social content [social > non-social; $F_{(1,\ 18)} = 9.78$, $p = 0.006$], demonstrating that this effect was present even for non-emotional scenes (see **Figure 3A**).

We then analyzed amygdala activation to emotional images during the three different viewing conditions with a 2 (side) × 3 (viewing instruction) × 2 (valence) × 2 (social content) ANOVA. This revealed several main effects and interactions (see **Figure 1**). First, there was a main effect of side [$F_{(1,\ 18)} = 2699, p < 0.001$], because activity was overall higher in the left than in the right amygdala. Second, we found a main effect of valence [$F_{(1,\ 18)} = 7.37, p = 0.014$], as activity was overall higher for negative as compared to positive images. There was also a side × valence interaction [$F_{(1,\ 18)} = 7.15$, $p = 0.016$], because the activation difference between negative *vs.* positive images was greater in the right [$t_{(18)} = 2.73$, $p = 0.014$] as compared to the left [$t_{(18)} = 2.21$, $p = 0.04$] amygdala. Third, our data showed a main effect of social content [$F_{(1,\ 18)} = 38.39$, $p < 0.001$], because activity was overall higher for social than non-social scenes.

We also found a side × social content interaction [$F_{(1,\ 18)} = 6.25, p = 0.022$], which arose because of a more pronounced social > non-social activation difference in the right [$t_{(18)} = 5.83$, $p < 0.001$] as compared to the left [$t_{(18)} = 4.29$, $p < 0.001$] amygdala. These effects were accompanied with a valence × social content [$F_{(1,\ 18)} = 7.30$, $p = 0.015$] and a marginally significant side × valence × social content [$F_{(1,\ 18)} = 3.89$, $p = 0.064$] interaction. The valence × social content interaction showed that activity in both amygdalae was not significantly different between negative and positive social images [all $t_{(18)} < 1.51$, $p > 0.15$], but higher for negative than positive non-social images [all $t_{(18)} > 2.66$, $p < 0.016$], while the side × valence × social content interaction revealed that this effect was stronger in the right [$t_{(18)} = 10.24$, $p = 0.005$] as compared to the left [$t_{(18)} = 3.95$, $p = 0.062$] amygdala.

Finally, there were no interactions with viewing conditions {natural viewing, re-appraisal, and suppression; [all $F_{(1,\ 18)} < 1.12$, $p > 0.34$]}, suggesting that the abovementioned valence and social relevance detection pattern was independent of task instructions.

Overall, these data on amygdala activation revealed two main findings. Firstly, we observed a reliable preferential response to social > non-social information bilaterally, for both neutral and emotional images. Secondly, we also found an additional interactive processing of valence and social content, which was more right-lateralized.

### *Other areas in social brain networks*

The 3 × 2 × 2 ANOVA was similarly performed on extracted beta values from the other regions of interest that were identified by the same initial computation of the social > non-social main effects (see **Table 1**). Like in the amygdala, this revealed a significant valence × social content interaction in the right fusiform gyrus (FG), right anterior superior frontal gyrus (SFG), and medial orbito-frontal cortex (mOFC; see **Figure 2**). In all these three brain areas, there were also main effects of valence [negative > positive; all $F_{(19,\ 1)} > 5.67$, $p < 0.028$] and social content [social > non-social; all $F_{(19,\ 1)} > 65.63$, $p < 0.001$], but no interaction with the three different viewing conditions {natural viewing, re-appraisal, suppression; [all $F_{(19,\ 1)} < 2.39$, $p > 0.11$]}.

We also assessed activity in the FG, aSTG, and mOFC during the neutral control condition. This revealed a significant social > non-social activation difference in all three regions [all $F_{(1,\ 18)} > 16.03$, all $p = 0.001$] (see **Figures 3B,C,D**).

The findings in the FG, aSTG, and mOFC thus showed a very similar pattern to the amygdala, with preferential social vs. non-social processing, regardless of emotional content, but also interactive processing of valence and social content, independent of viewing conditions.

### CORRELATIONS WITH PERSONALITY MEASURES

We next assessed the possible influence of trait anxiety and adult attachment style on relevance detection using regression analyses with the corresponding questionnaire scores (see section “Methods” for details).

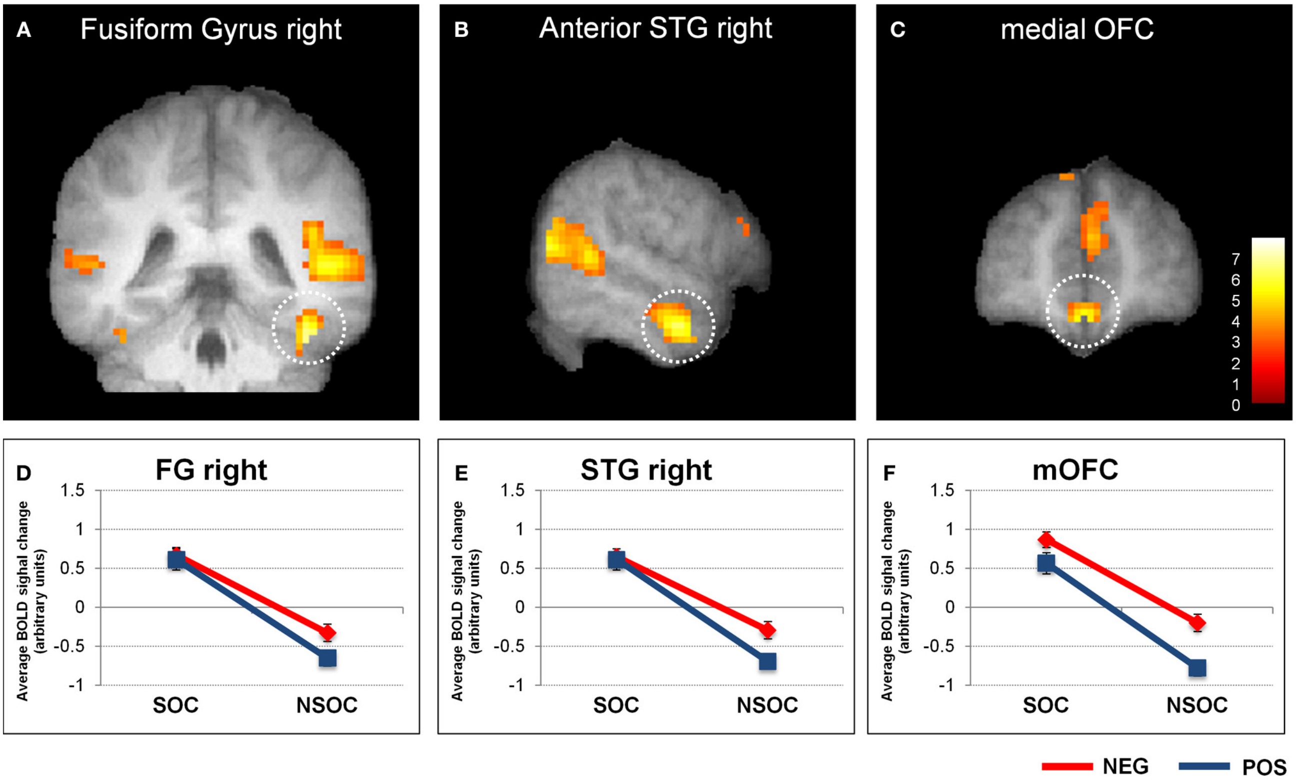


**FIGURE 2 | Interactive valence and social content processing in fusiform gyrus, right anterior superior temporal gyrus, and medial orbitofrontal cortex.** Statistical parametric maps (threshold $p = 0.001$) for the contrast social > non-social displaying significant activity in the right fusiform gyrus **(A)**, right anterior superior temporal gyrus **(B)**, and medial orbitofrontal cortex **(C)**. Graphs depicting averaged extracted raw activation values (betas) in right fusiform gyrus **(D)**, right anterior superior temporal gyrus **(E)**, and medial orbitofrontal cortex **(F)**, displaying a valence × social content interaction. NEG, negative; NSOC, non-social; POS, positive; SOC, social. Error bars reflect S.E.M.

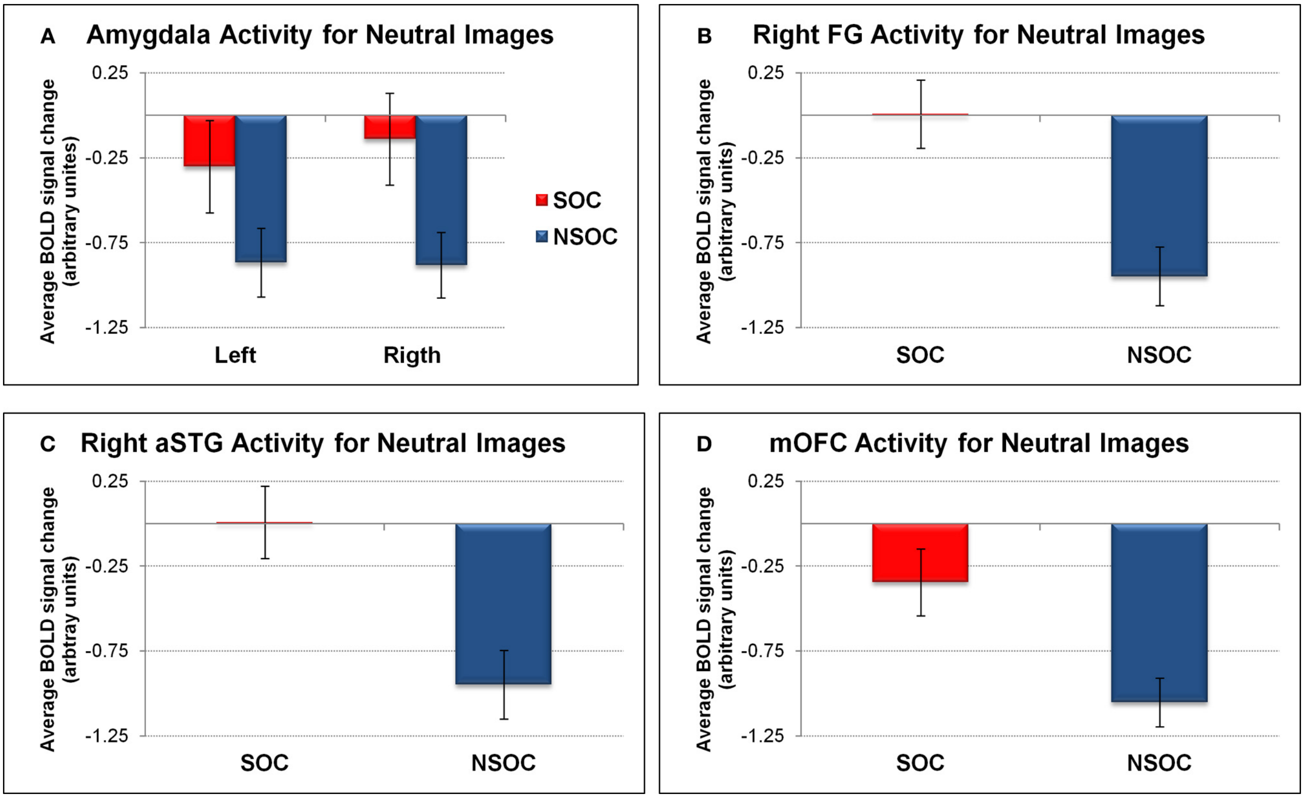


**FIGURE 3 | Social relevance effect for neutral images in regions of interest.** Social > non-social activation difference for neutral images in **(A)** bilateral amygdala, **(B)** right fusiform gyrus, **(C)** right anterior superior temporal gyrus, and **(D)** medial orbitofrontal cortex. FG, fusiform gyrus; NSOC, non-social; mOFC, medial orbitofrontal cortex; SOC, social; aSTG, anterior superior temporal gyrus. Error bars reflect S.E.M.

#### *Amygdala*

When considering individual differences using trait anxiety scores in an ANCOVA, we found a trend for a 2 (valence) 2 × (social content) × 1 (trait anxiety) triple interaction in the right amygdala $[F_{(1,\,18)} = 4.19,\ p = 0.059]$. There was no such effect in the left amygdala $[F_{(1,\,18)} = 1.63,\ p = 0.22]$, and there were no additional interactions with trait anxiety bilaterally. To visualize the direction of these effects, we performed a median split of activation parameters from the right amygdala according to individual trait anxiety scores (low: $n = 8$, mean AX score $= -0.86$; high: $n = 8$, mean AX score $= 0.93$). This visualization revealed that the right amygdala effect was explained by a significant valence × social content interaction solely for low $[F_{(1,\,7)} = 17.06,\ p = 0.004]$ but not high $[F_{(1,\,7)} = 0.68,\ p = 0.44]$ anxious subjects (see **Figures 4A,B**). To further elaborate on the finding that high anxious participants did not display any valence × social content interaction pattern in the right amygdala, we correlated activation values (betas) for each stimulus condition separately (SP, NSP, SN, NSN) with trait anxiety scores from all subjects. This revealed a selective negative association between trait anxiety and activation for social positive (SP) images ($r^2 = 0.24$, $p = 0.041$), in the sense that the higher the trait anxiety, the lower the right amygdala activation to SP images (see **Figure 4C**). No such relations were found for the three other image categories ($r^2 < 0.05$).

The valence × social content interaction for high anxious participants was unaffected by the different viewing conditions $[F_{(1,\,7)} = 0.02,\ p = 0.98]$. No influences of trait anxiety on valence and social content processing in the left amygdala were revealed.

No significant effects were found for attachment style measures.

#### *Other areas in social brain networks*

None of the additional regions of interest activated by social vs. non-social information showed significant effects of trait anxiety or adult attachment scores in the current analyses.

## DISCUSSION

The present fMRI study aimed at investigating relevance detection in the human amygdala by contrasting BOLD signal change in response to visual scenes as a function of both their valence (negative *vs.* positive) and their social (*vs.* non-social) content. Further, we examined the influence of different viewing conditions, implying different emotion regulation strategies, as well as individual differences. This design allowed us to extend previous fMRI studies with a similar purpose (Norris et al., 2004; Britton et al., 2006; Goossens et al., 2009; Scharpf et al., 2010) that did not differentiate between positive vs. negative valence of stimuli—thereby preventing the investigation of any valence × social

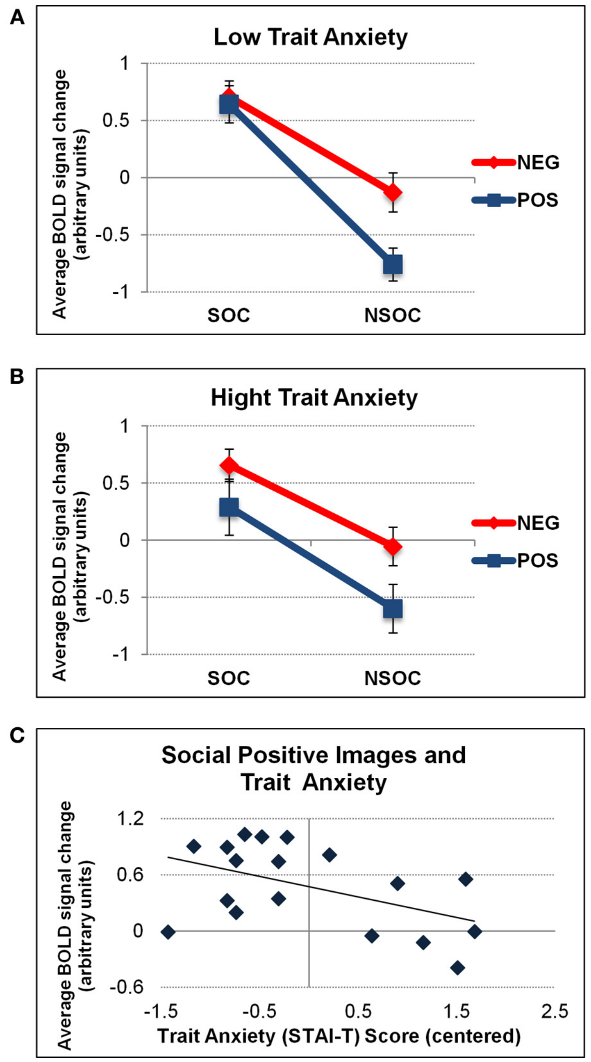


**FIGURE 4 | Interactive valence and social content processing in right amygdala as a function of trait anxiety.** Graphs depicting the average activation values (betas) extracted from the right amygdala as a function of low **(A)** and high **(B)** trait anxiety (STAI-T). A significant valence × social content interaction was only found for low anxious individuals. The latter was due to a selective effect of trait anxiety on activation to social positive (SP) images **(C)**. NEG, negative; NSOC, non-social; POS, positive; SOC, social. Error bars reflect S.E.M.

content interaction effects—and did not take into account task conditions and/or personality traits on relevance processing. Our analysis of amygdala activation revealed a significant interaction between the valence (positive vs. negative) and social content (social vs. non-social) of stimuli. Importantly, these effects were independent of the different viewing conditions (natural viewing, re-appraisal, suppression). Below we discuss these findings each in turn.

## SOCIAL CONTENT

In whole-brain analyses, bilateral amygdala showed a main effect for social vs. non-social images. This differential social effect was present for neutral as well as emotional stimuli (see **Figures 1** and **3A**). Such a preferential processing of images according to *social relevance* (Keltner and Kring, 1998; Hariri et al., 2002), even for scenes without any emotional value ("neutral"), converges with previous findings (Norris et al., 2004; Britton et al., 2006; Goossens et al., 2009; Scharpf et al., 2010) and thus extends the notion that the human amygdala is highly sensitive to the social significance of information, presumably due to the intrinsic relevance of social cues for our species.

Importantly, such preferential social vs. non-social processing in our study cannot be attributed to an arousal effect (Anderson et al., 2003; Small et al., 2003), since all social and non-social images were matched on this dimension. These results are consistent with those obtained by (Ewbank et al., 2009) showing that images with "personal impact" activate the amygdala when arousal is controlled for, and may be explained by the fact that social information has an intrinsic motivational value *per se* that may make social images particularly impactful (Morrison and Salzman, 2010). This value, however, is not synonymous with valence (negative or positive), because we found that social relevance and valence were processed interactively within the amygdala with no significant activation difference between positive and negative social images (see below). Consequently, it appears that social relevance is computed in the amygdala independently of the valence and arousal dimensions. This also accords with a key role of this region in processing faces (Todorov et al., 2011; Yang et al., 2012) or gaze (Kawashima et al., 1999; George and Conty, 2008; N'Diaye et al., 2009; Cristinzio et al., 2010), even when face expression is neutral.

It remains, however, to be better determined what are the social cues that preferentially drive amygdala responses, in particular whether they essentially reflect a differential tuning to faces, bodies, and other human features relative to other objects. Selective amygdala activation to faces or gaze has consistently been reported in human and other primate species (Gothard et al., 2007; Hoffman et al., 2007). Such an interpretation would accord with the fact that we observed the same activation pattern reflecting social relevance detection independent of valence in the fusiform gyrus, another brain region known to be preferentially involved in face processing (see also below). This suggests that one possible origin of the social relevance effect could be a more basic reactivity of the amygdala (and fusiform gyrus) toward facial cues or human body parts, (e.g., eyes, etc.) vs. objects or non-human scenes (Vuilleumier et al., 2004). Future studies should therefore address if social significance implied by other visual cues may also trigger a distinctive amygdala response to social relevance.

## VALENCE

Besides the social relevance effect, we also observed a valence effect in bilateral amygdala, with significantly higher activity during the processing of negative as compared to positive images overall (see **Figure 1**). This pattern accords with the view that negative information may have an intrinsically stronger relevance for the human organism, probably due to its more immediate implication for survival (Hariri et al., 2002). However, our study demonstrates that this valence effect was predominantly driven by a selective increase to negative information in *non-social* scenes, reflected by a significant valence × social content interaction. The difference between negative and positive social stimuli did not reach significance (see above). Consequently, the notion that negative or threat information has a generally higher value, and is thus linked with distinctive amygdala activation, seems partly inappropriate. According to our data, such "negativity bias" in amygdala responses may arise only (or predominantly) for non-social stimuli—whereas social significance may not need further negative affective significance to be behaviorally relevant. Furthermore, in our data, this valence effect was also likely to be independent of arousal (Morrison and Salzman, 2010), because the most arousing images (negative valence) were associated with heightened amygdala activation only when they were non-social, not otherwise.

### VALENCE × SOCIAL CONTENT INTERACTION

To our knowledge, by using a full factorial design, our study is the first to report direct evidence for this *interactive processing pattern of valence and social information* in the human amygdala. These results are consistent with the view that the amygdala is critical for relevance detection (Sander et al., 2003, 2005; Adolphs, 2010; Pessoa and Adolphs, 2010), whatever the nature of the relevance, be it social or affective. In future studies, it would therefore be interesting to examine the role of the human amygdala in processing stimuli with different kinds of relevance, including motivational significance (e.g., food or drugs; see Tang et al., 2012, for a recent review).

### HEMISPHERIC LATERALIZATION

The valence and social effects, as well as their interaction discussed above, all were found to display a significant hemispheric lateralization in the human amygdala. Whereas emotional images induced more activity in the left than right side in general, both the social *vs.* non-social and the negative *vs.* positive activation differences were stronger in the right than the left side, suggesting a more pronounced general valence × social content interaction in the right amygdala (see **Figure 1**). In the literature, different hemispheric lateralization accounts of human amygdala function have been proposed, related to language, temporal dynamics, gender (Sergerie et al., 2006), or even awareness (Morris et al., 1998). In our study, lateralization effects were independent of task conditions, suggesting little or no modulation by the degree of covert semantic processing (likely higher during cognitive reappraisal relative to natural viewing and expressive suppression). Hence, although indirect, this appears inconsistent with the first lateralization account.

On the other hand, the observed activation patterns could be compatible with differential temporal dynamics in amygdala, with possibly quicker habituation effects on the right side. This might result from larger habituation to emotional images in general, and to non-social positive scenes in particular, leading to both the main effect of valence and the valence × social content interaction for the right amygdala. Such habituation would also corroborate the relevance detector hypothesis (Sander et al., 2003, 2005; Adolphs, 2010; Pessoa and Adolphs, 2010) as it might be assumed that the more relevant information should be less prone to habituation, explaining why activity to non-social negative (*vs.* positive) images remained higher in the right amygdala. However, these interpretations remain speculative, as our study did not use direct tests for different lateralization accounts and did not examine differences in the time-course of activation during the experiment. Thus, our findings provide only indirect support to some lateralization in amygdala responses to social and emotional relevance, and more work is needed in the future to better clarify the sources of such asymmetries.

### PERSONALITY

Finally, our data also confirms an important role of individual personality differences in modulating amygdala reactivity. In the right amygdala, the valence × social content interaction was only present for low but not high anxious participants (see **Figure 4**). This was caused by a selective decrease of right amygdala response to social positive (SP) images as a function of trait anxiety. As noted above, this could potentially be attributed to some habituation of the right amygdala to these stimuli that might be further accelerated in high anxious participants because such images were perceived as less relevant to current concerns. This would accord with the negativity biases associated with anxiety and converge with previous findings showing increased processing of (socially) relevant information in the human amygdala as a function of trait anxiety (see Bishop et al., 2004; Sabatinelli et al., 2005; even though such effects were often bilateral in the latter studies). Although an association between STAI-T scores and the interactive response to valence × social content was selectively observed in the right amygdala, a formal ANOVA including SIDE as a factor was not significant. It therefore remains to be determined whether a true lateralization exists regarding personality effects on valence and social content processing in the human amygdala.

### RELEVANCE DETECTION IN CORTICAL BRAIN AREAS

Additional analysis of activation patterns for other areas within the social brain networks (i.e., displaying a main effect of social vs. non-social scenes) also revealed a significant valence × social relevance interaction in right fusiform gyrus (FG), right anterior superior temporal gyrus (aSTG), as well as medial orbitofrontal cortex (mOFC; see **Figure 2**). Thus, these regions showed a very similar profile of responses as the bilateral amygdala (see above), independent of viewing conditions. However, unlike for the amygdala, we did not find any influence of trait anxiety on right FG, right aSTG, or mOFC.

All of these three regions have previously been associated with privileged processing of social information, including preferential responses to (human) animate stimuli containing faces or bodies in the FG (Kanwisher et al., 1997; Peelen and Downing, 2005; Schwarzlose et al., 2005); representation of abstract social concepts/values and moral sentiments (Zahn et al., 2007), moral cognition (Moll et al., 2005), and social emotion processing (Wicker et al., 2003) in the aSTG; as well as social outcome monitoring (Amodio and Frith, 2006) and theory of mind (Gallagher and Frith, 2003) in the mOFC.

Negative valence effects in FG accord with previous evidence that this area is generally more activated by emotional (*vs.* neutral) and particularly negative (*vs.* positive) information (Vuilleumier et al., 2001, 2004; Hadjikhani and de Gelder, 2003;

Surguladze et al., 2003). Effects of valence in aSTG are less clear, because there is not much data regarding emotional and social processing in this area, even though a response to negative information—e.g., fear, anger, and particularly sadness—has often been reported (Levesque et al., 2003; Gillath et al., 2005; Moll et al., 2005). Finally, the current valence effect in mOFC is somewhat at odds with usual considerations that this brain area may primarily be involved in monitoring the positive, rewarding value of social interactions in terms of their probable outcomes (Rolls, 1996; Amodio and Frith, 2006). However, there is evidence that the mOFC is also involved in the attribution of negative value to initially neutral items (Goldstein et al., 2007), which implies that it may have a more general function in computing the subjective value of particularly social stimuli, even if negative, as in our task.

A valence × social content interaction has previously been reported in sensory areas such as the thalamus, superior temporal sulcus and middle occipito-temporal cortex, in addition to the anterior insula and lateral medial prefrontal cortex (Norris et al., 2004; Scharpf et al., 2010), but none of these brain regions were found in the present fMRI experiment. However, as already mentioned above, the emotional relevance effect examined in those studies comprised an emotional *vs.* neutral comparison, but not a 2 (valence) × 2 (social content) interaction, an important difference that is likely to account for the discrepancy between our new and previous findings. Because we looked for brain areas displaying a valence × social content interaction within regions of interest that were selected a priori based on the main effect social vs. non-social, our analysis was more selective. However, using a full 2 (valence) × 2 (social content) experimental design, we did find that valence and social content were processed interactively in several cortical areas including the FG, aSTG, and mOFC (in addition to bilateral amygdala), thus delineating a distributed subcortical-cortical network integrating emotional and social content processing. These regions may constitute a relevance detection network reciprocally interacting with the amygdala, as corroborated by previous findings showing (1) modulation of FG activity by inputs from the amygdala during emotional face perception in humans (Vuilleumier et al., 2001, 2004) and the existence of amygdala-visual cortex projections in macaques (Amaral et al., 2003), (2) simultaneous OFC and temporal cortex activation during social concept representation in humans (Zahn et al., 2007), (3) the presence of direct anatomical connections between the aSTG/middle and inferior temporal lobes and OFC in macaques (Kondo et al., 2003), and (4) functional correlations between amygdala and OFC activity during emotional conflict resolution (Etkin et al., 2006) and the implication of this circuitry in anxiety in humans as well as animals (Bishop, 2007). However, to draw any sound conclusions about causality and directionality of effects within this extended network, more specific investigations of relevance processing in humans within this subcortical-cortical network is necessary.

## CONCLUSION

The present fMRI study aimed at systematically investigating valence and social content processing within the human amygdala by dissociating between positive and negative, as well as between social and non-social stimuli, in addition to a "neutral" visual baseline. Moreover, we included three different viewing conditions representing "spontaneous" natural emotion processing (natural viewing), as well as cognitive (re-appraisal) and behavioral (expressive suppression) emotion regulation strategies, and also probed for individual differences.

Results revealed an interactive processing of valence and social content in the amygdala, more pronounced on the right than left side. This interaction was not modulated by the different task conditions, but depended on trait anxiety, being significant in low but not high anxious subjects. The latter was due to a selective decrease in amygdala activity to positive social images with higher anxiety scores. Overall, these data suggest that relevance detection in the amygdala operates at a task-independent processing level by integrating both valence and social content. The same valence × social content interaction was present in other cortical regions intimately connected to the amygdala, including the right fusiform gyrus, right anterior superior temporal gyrus, and medial orbito-frontal cortex, suggesting the existence of a distributed subcortical-cortical network for relevance detection in humans.

## ACKNOWLEDGMENTS

This research was supported by the National Center of Competence in Research (NCCR) Affective Sciences financed by the Swiss National Science Foundation (n° 51NF40-104897) and hosted by the University of Geneva, plus grants of the Swiss National Science Foundation to Pascal Vrtička, David Sander, and Patrik Vuilleumier.

## REFERENCES

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat. Rev. Neurosci.* 4, 165–178.

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Ann. N.Y. Acad. Sci.* 1191, 42–61.

Aiken, L. S., and West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage.

Amaral, D. G., Behniea, H., and Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience* 118, 1099–1120.

Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* 6, 196–202.

Balderston, N. L., Schultz, D. H., and Helmstetter, F. J. (2011). The human amygdala plays a stimulus specific role in the detection of novelty. *Neuroimage* 55, 1889–1898.

Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends Cogn. Sci.* 11, 307–316.

Bishop, S. J., Duncan, J., and Lawrence, A. D. (2004). State anxiety modulation of the amygdala response to unattended threat-related stimuli. *J. Neurosci.* 24, 10364–10368.

Blackford, J. U., Buckholtz, J. W., Avery, S. N., and Zald, D. H. (2010). A unique role for the human amygdala in novelty detection. *Neuroimage* 50, 1188–1193.

Britton, J. C., Phan, K. L., Taylor, S. F., Welsh, R. C., Berridge, K. C., and Liberzon, I. (2006). Neural correlates of social and nonsocial emotions: an fMRI study. *Neuroimage* 31, 397–409.

Brosch, T., Sander, D., Pourtois, G., and Scherer, K. R. (2008). Beyond fear – rapid spatial orienting toward positive emotional stimuli. *Psychol. Sci.* 19, 362–370.

Cahill, L. (2006). Why sex matters for neuroscience. *Nat. Rev. Neurosci.* 7, 477–484.

Canli, T., Zhao, Z., Desmond, J. E., Kang, E. J., Gross, J., and Gabrieli, J. D. E. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behav. Neurosci.* 115, 33–42.

Collins, N. L., and Read, S. (1994). "Cognitive representations of attachment: the structure and function of attachment models," in *Attachment Processes in Adulthood: Advances in Personal Relationships, Vol. 5*. eds K. Bartholomew and D. Perlman (London: Kingsley), 53–90.

Cristinzio, C., N'Diaye, K., Seeck, M., Vuilleumier, P., and Sander, D. (2010). Integration of gaze direction and facial expression in patients with unilateral amygdala damage. *Brain* 133, 248–261.

Dickie, E. W., and Armony, J. L. (2008). Amygdala responses to unattended fearful faces: interaction between sex and trait anxiety. *Psychiatry Res.* 162, 51–57.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthropol.* 6, 178–190.

Ekman, P. (1999). "Basic emotions," in *Handbook of Cognitive Emotion,* eds T. Dalgleish and M. Power (Sussex, UK: John Wiley and Sons, Ltd), 45–60.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., and Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51, 871–882.

Etkin, A., Klemenhagen, K. C., Dudman, J. T., Rogan, M. T., Hen, R., Kandel, E. R., et al. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron* 44, 1043–1055.

Ewbank, M. P., Barnard, P. J., Croucher, C. J., Ramponi, C., and Calder, A. J. (2009). The amygdala response to images with impact. *Soc. Cogn. Affect. Neurosci.* 4, 127–133.

Frijda, N. H. (2009). Emotions, individual differences and time course: reflections. *Cogn. Emot.* 23, 1444–1461.

Gallagher, H. L., and Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn. Sci.* 7, 77–83.

George, N., and Conty, L. (2008). Facing the gaze of others. *Neurophysiol. Clin.* 38, 197–207.

Gillath, O., Bunge, S. A., Shaver, P. R., Wendelken, C., and Mikulincer, M. (2005). Attachment-style differences in the ability to suppress negative thoughts: exploring the neural correlates. *Neuroimage* 28, 835–847.

Goldin, P. R., McRae, K., Ramel, W., and Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* 63, 577–586.

Goldstein, R. Z., Tomasi, D., Rajaram, S., Cottone, L. A., Zhang, L., Maloney, T., et al. (2007). Role of the anterior cingulate and medial orbitofrontal cortex in processing drug cues in cocaine addiction. *Neuroscience* 144, 1153–1159.

Goossens, L., Kukolja, J., Onur, O. A., Fink, G. R., Maier, W., Griez, E., et al. (2009). Selective processing of social stimuli in the superficial amygdala. *Hum. Brain Mapp.* 30, 3332–3338.

Gothard, K. M., Battaglia, F. P., Erickson, C. A., Spitler, K. M., and Amaral, D. G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *J. Neurophysiol.* 97, 1671–1683.

Guédeney, N., Fermanian, J., and Bifulco, A. (2010). Construct validation study of the Relationship Scales Questionnaire (RSQ) on an adult sample. *Encephale* 36, 69–76.

Hadjikhani, N., and de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr. Biol.* 13, 2201–2205.

Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., and Weinberger, D. R. (2002). The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage* 17, 317–323.

Harvey, P. O., Fossati, P., and Lepage, M. (2007). Modulation of memory formation by stimulus content: specific role of the medial prefrontal cortex in the successful encoding of social pictures. *J. Cogn. Neurosci.* 19, 351–362.

Hoffman, K. L., Gothard, K. M., Schmid, M. C., and Logothetis, N. K. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Curr. Biol.* 17, 766–772.

Jakobs, E., Fischer, A. H., and Manstead, A. S. R. (1997). Emotional experience as a function of social context: the role of the other. *J. Nonverb. Behav.* 21, 103–130.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., et al. (1999). The human amygdala plays an important role in gaze monitoring – A PET study. *Brain* 122, 779–783.

Keltner, D., and Kring, A. M. (1998). Emotion, social function, and psychopathology. *Rev. Gen. Psychol.* 2, 320–342.

Kim, H., Somerville, L. H., Johnstone, T., Alexander, A. L., and Whalen, P. J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *Neuroreport* 14, 2317–2322.

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., et al. (2004). Contextual modulation of amygdala responsivity to surprised faces. *J. Cogn. Neurosci.* 16, 1730–1745.

Kim, S. H., and Hamann, S. (2007). Neural correlates of positive and negative emotion regulation. *J. Cogn. Neurosci.* 19, 776–798.

Kondo, H., Saleem, K. S., and Price, J. L. (2003). Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. *J. Comp. Neurol.* 465, 499–523.

Kurdek, L. A. (2002). On being insecure about the assessment of attachment styles. *J. Soc. Pers. Relat.* 19, 811–834.

LaBar, K. S., Gitelman, D. R., Parrish, T. B., Kim, Y. H., Nobre, A. C., and Mesulam, M. M. (2001). Hunger selectively modulates corticolimbic activation to food stimuli in humans. *Behav. Neurosci.* 115, 493–500.

Levesque, J., Eugene, F., Joanette, Y., Paquette, V., Mensour, B., Beaudoin, G., et al. (2003). Neural circuitry underlying voluntary suppression of sadness. *Biol. Psychiatry* 53, 502–510.

Lieberman, M. D., and Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Soc. Cogn. Affect. Neurosci.* 4, 423–428.

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., and Grafman, J. (2005). The neural basis of human moral cognition. *Nat. Rev. Neurosci.* 6, 799–809.

Morris, J. S., and Dolan, R. J. (2001). Involvement of human amygdala and orbitofrontal cortex in hunger-enhanced memory for food stimuli. *J. Neurosci.* 21, 5304–5310.

Morris, J. S., Ohman, A., and Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature* 393, 467–470.

Morrison, S. E., and Salzman, C. D. (2010). Re-valuing the amygdala. *Curr. Opin. Neurobiol.* 20, 221–230.

Murray, E. A. (2007). The amygdala, reward and emotion. *Trends Cogn. Sci.* 11, 489–497.

N'Diaye, K., Sander, D., and Vuilleumier, P. (2009). Self-relevance processing in the human amygdala: gaze direction, facial expression, and emotion intensity. *Emotion* 9, 798–806.

Norris, C. J., Chen, E. E., Zhu, D. C., Small, S. L., and Cacioppo, J. T. (2004). The interaction of social and emotional processes in the brain. *J. Cogn. Neurosci.* 16, 1818–1829.

Ochsner, K. N., Bunge, S. A., Gross, J. J., and Gabrieli, J. D. E. (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., et al. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage* 23, 483–499.

Ohman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522.

Peelen, M. V., and Downing, P. E. (2005). Within-subject reproducibility of category-specific visual activation with functional MRI. *Hum. Brain Mapp.* 25, 402–408.

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* 11, 773–782.

Reisenzein, R. (2009). Emotions as metarepresentational states of mind: naturalizing the belief-desire theory of emotion. *Cogn. Syst. Res.* 10, 6–20.

Rolls, E. T. (1996). The orbitofrontal cortex. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 351, 1433–1443.

Russell, J. A. (1980). A circumplex model of affect. *J. Person. Soc. Psychol.* 39, 1161–1178.

Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R., and Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *Neuroimage* 24, 1265–1270.

Sander, D., Grafman, J., and Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Rev. Neurosci.* 14, 303–316.

Sander, D., Grandjean, D., and Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Netw.* 18, 317–352.

Sato, W., Kochiyama, T., Uono, S., and Yoshikawa, S. (2010). Amygdala integrates emotional expression and gaze direction in response to dynamic facial expressions. *Neuroimage* 50, 1658–1665.

Scharpf, K. R., Wendt, J., Lotze, M., and Hamm, A. O. (2010). The brain's relevance detection network operates independently of stimulus modality. *Behav. Brain Res.* 210, 16–23.

Schiller, D., Freeman, J. B., Mitchell, J. P., Uleman, J. S., and Phelps, E. A. (2009). A neural mechanism of first impressions. *Nat. Neurosci.* 12, 508–514.

Schwarzlose, R. F., Baker, C. I., and Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *J. Neurosci.* 25, 11055–11059.

Sergerie, K., Lepage, M., and Armony, J. L. (2006). A process-specific functional dissociation of the amygdala in emotional memory. *J. Cogn. Neurosci.* 18, 1359–1367.

Simpson, J. A. (1990). Influence of attachment styles on romantic relationships. *J. Pers. Soc. Psychol.* 59, 971–980.

Simpson, J. A., Rholes, W. S., Orina, M. M., and Grich, J. (2002). Working models of attachment, support giving, and support seeking in a stressful situation. *Pers. Soc. Psychol. Bull.* 28, 598–608.

Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., and Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39, 701–711.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory.* Palo-Alto, CA: Consulting Psychologists Press.

Surguladze, S. A., Brammer, M. J., Young, A. W., Andrew, C., Travis, M. J., Williams, S. C. R., et al. (2003). A preferential increase in the extrastriate response to signals of danger. *Neuroimage* 19, 1317–1328.

Tang, D. W., Fellows, L. K., Small, D. M., and Dagher, A. (2012). Food and drug cues activate similar brain regions: a meta-analysis of functional MRI studies. [Article]. *Physiol. Behav.* 106, 317–324.

Todorov, A., Said, C. P., Oosterhof, N. N., and Engell, A. D. (2011). Task-invariant brain responses to the social value of faces. *J. Cogn. Neurosci.* 23, 2766–2781.

Vrticka, P., and Vuilleumier, P. (2012). Neuroscience of human social interactions and adult attachment style. *Front. Hum. Neurosci.* 6. doi: 10.3389/fnhum.2012.00212

Vrticka, P., Andersson, F., Grandjean, D., Sander, D., and Vuilleumier, P. (2008). Individual attachment style modulates human amygdala and striatum activation during social appraisal. *PLoS ONE* 3:e2868. doi: 10.1371/journal.pone.0002868

Vrticka, P., Bondolfi, G., Sander, D., and Vuilleumier, P. (2012a). The neural substrates of social emotion perception and regulation are modulated by adult attachment style. *Soc. Neurosci.* 7, 473–493.

Vrticka, P., Sander, D., and Vuilleumier, P. (2012b). Influence of adult attachment style on the perception of social and non-social emotional scenes. *J. Soc. Pers. Relat.* 29, 530–544.

Vrticka, P., Sander, D., and Vuilleumier, P. (2011). Effects of emotion regulation strategy on brain responses to the valence and social content of visual scenes. *Neuropsychologia* 49, 1067–1082.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594.

Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., and Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat. Neurosci.* 7, 1271–1278.

Weierich, M. R., Wright, C. I., Negreira, A., Dickerson, B. C., and Barrett, L. F. (2010). Novelty as a dimension in the affective brain. *Neuroimage* 49, 2871–2878.

Wicker, B., Perrett, D. I., Baron-Cohen, S., and Decety, J. (2003). Being the target of another's emotion: a PET study. *Neuropsychologia* 41, 139–146.

Yang, J., Bellgowan, P. S. F., and Martin, A. (2012). Threat, domain-specificity and the human amygdala. *Neuropsychologia* 50, 2566–2572.

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., and Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6430–6435.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 19 October 2012; accepted: 26 December 2012; published online: 18 January 2013.*

*Citation: Vrtička P, Sander D and Vuilleumier P (2013) Lateralized interactive social content and valence processing within the human amygdala. Front. Hum. Neurosci.* ***6****:358. doi: 10.3389/fnhum.2012.00358*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 21 February 2013
doi: 10.3389/fnhum.2013.00043

# Emotional and non-emotional pathways to impulsive behavior and addiction

**Ana Torres**[1]**, Andrés Catena**[1]**, Alberto Megías**[1]**, Antonio Maldonado**[1]**, Antonio Cándido**[1]**,**
**Antonio Verdejo-García**[2,3] **and José C. Perales**[1]***

[1] *Learning, Emotion, and Decision Research Group, Mind, Brain and Behavior Research Center/Centro de Investigación Mente, Cerebro y Comportamiento (CIMCYC), Universidad de Granada, Granada, Spain*
[2] *Departamento de Personalidad, Evaluación y Tratamiento Psicológico, Universidad de Granada, Granada, Spain*
[3] *School of Psychology, Psychiatry and Psychological Medicine, Monash University, Melbourne, VIC, Australia*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*Estate M. Sokhadze, University of Louisville, USA*
*Maartje Luijten, Erasmus University Rotterdam, Netherlands*

****Correspondence:***
*José C. Perales, Learning, Emotion and Decision Research Group, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja, S/N. 18071 Granada, Spain.*
*e-mail: jcesar@ugr.es*

Impulsivity is tightly linked to addiction. However, there are several pathways by means of which impulsive individuals are more prone to become addicts, or to suffer an addiction more intensely and for a longer period. One of those pathways involves an inadequate appraisal or regulation of positive and negative emotions, leading to lack of control over hazardous behaviors, and inappropriate decisions. In the present work, we assessed cocaine-dependent individuals (CDI; $n = 20$), pathological gamblers (PG; $n = 21$), and healthy controls (HC; $n = 23$) in trait impulsivity measures (UPPS-P model's dimensions), and decision-making tasks (Go/No-go; delay-discounting task). During the Go/No-go task, electroencephalographic (EEG) activity was recorded, and Go/No-go stimuli-evoked potentials (ERP) were extracted. Theory-driven ERP analyses focused on the No-go > Go difference in the N2 ERP. Our results show that negative urgency is one of the several psychological features that distinguish addicts from HC. Nevertheless, among the dimensions of trait impulsivity, negative urgency is unique at independently covarying with gambling over-pathologization in the PG sample. Cocaine-dependent individuals performed more poorly than gamblers in the Go/No-go task, and showed abnormal Go/No-go stimuli-evoked potentials. The difference between the No-go stimulus-evoked N2, and the Go one was attenuated by severity and intensity of chronic cocaine use. Emotional dimensions of impulsivity, however, did not influence Go/No-go performance.

**Keywords: impulsivity, emotion, addiction, decision-making, delay discounting, Go/No-go, UPPS-P, N2 ERP**

## INTRODUCTION

Substance and behavioral addictions are characterized by loss of control over drives or habits, and by persistent preference for immediate rewards at the expense of relative net loss (Everitt and Robbins, 2000). Behavioral neuroscience has modeled these deficits by using motor inhibition (or impulsive action) and delay discounting (or impulsive choice) decision-making tasks (Winstanley et al., 2006; Verdejo-García et al., 2008). However, there are at least two mechanisms potentially accounting for impulsive actions and/or choices in addicts: one refers to pre-existing individual differences within impulsive personality traits, and the other to the impact of the degree of exposure to drug dosages or gambling episodes.

With regard to the first mechanism, there are multiple facets of trait impulsivity that may impact impulsive action and choice, and thus addiction vulnerability. The UPPS-P model of impulsivity has recently emerged as a successful factorial account of impulsive personality (Whiteside and Lynam, 2001; Cyders and Smith, 2008; Cándido et al., 2012). In this model, impulsivity is hypothesized to be composed of five separable dimensions. *Negative urgency* refers to the tendency to make poor decisions under conditions of negative affect; *positive urgency* refers to the tendency to make poor decisions under conditions of positive affect; *lack of premeditation* is the tendency not to think of the consequences of an action before engaging in it; *lack of perseverance* refers to the inability to stay focused on long, boring or difficult tasks; and *sensation seeking* represents the willingness to participate in exciting, new, and/or potentially dangerous activities. Most importantly, impulsivity dimensions in this model have been proved to correlate with future (prospective) aspects of addictive behavior, in a meaningful manner (see, for example, Cyders et al., 2007, 2009; Cyders and Smith, 2008).

With regard to the second mechanism, prolonged stimulants or gambling exposure can gradually boost impulsive traits and deter cognitive performance in response inhibition and delay discounting tests. In terms of behavioral performance, cocaine users seem disproportionately impaired in response inhibition tests (Verdejo-García et al., 2007b), whereas pathological gamblers (PG) display consistent preference for immediate rewards, in absence of significant influence of response inhibition deficits (Kertzman et al., 2011).

In other words, addiction, trait impulsivity, and performance in delay discounting and response inhibition tasks compound quite a complex, multifaceted triad, in which disentangling causes from effects is not always easy. The most straightforward way to do so is by means of prospective studies, in which personality and neuropsychological scores are used to predict the future emergence of addictive disorders. Cross-sectional studies, comparing addicts vs. non-addicts, are more limited in this regard. Nevertheless, in groups of addicts, it is possible to check for the existence of functional relations between the degree of exposure to the potentially neurologically damaging factor (e.g., the amount of drug consumed during the course of addiction) and the extent of personality and neuropsychological anomalies (including impulsive choice and impulsive action in lab tasks, and their neurophysiological correlates). Relationships of this type are strongly indicative of neurotoxic and/or neuroadaptive effects (see Albein-Urios et al., 2012a, for a more detailed presentation of this argument).

In the present work, we compared a group of PG against a group of cocaine dependent individuals (CDI), and a group of matched healthy controls (HC), in trait impulsivity dimensions (UPPS-P), impulsive choice (delay discounting), and impulsive action (motor inhibition Go/No-go task). Our first aim is to help to clarify the current state of affairs by carefully estimating the degree of exposure to cocaine/gambling activities, and checking for relationships, not only between the clinical category and personality/impulsive choice/impulsive action anomalies, but also between those anomalies and exposure estimates, within the clinical categories.

Complementarily, we have solid reasons to believe that some light can be shed on this complex pattern of interrelations by drawing two paths in it: one based on *emotion* appraisal and processing, and a second, *non-emotional* one, mostly related to other ("cool") components of executive functioning (Metcalfe and Mischel, 1999). On this regard, our first working assumption is the centrality of emotion in the characterization of different types of impulsivity. The segregation between emotional and non-emotional components of impulsivity has indeed a long history (see Barratt, 1993; Evenden, 1999; Perales et al., 2009), and is also implicit in the UPPS-P model (Whiteside and Lynam, 2001; Cyders et al., 2007; Cyders and Smith, 2008). Negative urgency, positive urgency, and sensation seeking, are by definition strongly loaded by emotional factors (Joseph et al., 2009). Among them, positive and negative urgency depend on inadequate appraisal of (and response to) emotions preceding decisions, whereas sensation seeking is more related to the anticipation of reward and the lack of apprehension against the risks involved in its attainment. The other two dimensions—lack of perseverance and lack of premeditation—are less dependent on emotion. In that sense, the UPPS-P model emerges as a tool to identify the importance of emotional vs. non-emotional ingredients of impulsivity in all kinds of hazardous, risky or pathological behavior.

More generally speaking, emotion is linked to impulsive behavior in at least two ways: via *behavior consequences*, and via *behavior antecedents*. In support of the first link, there is evidence of a psychological link between reward/punishment and emotion, so that the neural substrates for them are partially overlapping (Kringelbach, 2005; Murray, 2008; Quartz, 2009). Some individuals can be specially (in)sensitive to the affective value of present or delayed reward and punishment, so that, for example, punishment-sensitive people can be less prone to take risks, or reward-sensitive people can be more adventurous and sensation seeking (Dawe and Loxton, 2004; Franken and Muris, 2006; Reynolds, 2006; Torres et al., submitted). In the second case, emotions, acting as context for decisions, can turn such decisions into impulsive ones. Recent studies (see, for example, Albein-Urios et al., 2012b; Verdejo-García et al., 2013) report strong evidence on the involvement of inadequate appraisal of emotions (particularly negative ones) in hazardous behavior.

Our second working assumption is that the two neurobehavioral mechanisms theoretically linking brain function to rash action and addiction (delay discounting sensitivity and motor inhibition) are also differentially related to emotions. Separate pieces of previous evidence seem to support this idea. For example, Metcalfe and Mischel (1999) interpret findings from reward delay studies in terms of a hot/cool systems theory. The results reviewed support the idea that increased "cool" system activation increases the ability to delay gratification, whereas increased "hot" system activation decreases it. Importantly, emotions generated incidentally at the time of decision, but not directly related to the expected, to-be-delayed reward, as well as emotions consubstantial to the object of decision, can significantly influence choice (Schmeichel and Inzlicht, 2013).

On the other hand, core motor inhibition mechanisms seem to be mostly non-emotional, although this statement is more controversial. A broad exploration of the cognitive processes and brain areas involved in Go/No-go and other inhibition tasks can be found in Rubia et al. (2001). The areas identified comprise parts of the frontal, prefrontal, and parietal cortices, mostly—although not exclusively—coincident with the "cool" executive system (although see also Goldstein et al., 2007). According to Horn et al. (2003), differential Go/No-go stimulus-evoked early processing is mostly non-emotional, and reflects the core inhibitory component of the task. Conversely, more delayed processing recruits limbic and paralimbic structures (posterior orbitofrontal cortex, the temporal poles, and the posterior cingulate) which are likely to implement the emotional and motivational aspects of the task. Given this somewhat mixed pattern of results, providing correlations between different aspects of Go/No-go performance and emotional and non-emotional dimensions of impulsivity is theoretically valuable.

To sum up, with the present work we intend to add evidence on the two abovementioned assumptions and, complementarily, to make them converge to test the differential involvement of the emotional and non-emotional pathways in the two addictive disorders under scrutiny. With that intention in mind, we interviewed participants in their present and past drug-using and gambling habits, and we used the UPPS-P model to evaluate impulsive personality in the three groups, the Now-or-later test (Kirby et al., 1999) for delay discounting/impulsive choice; and the Go/No-go task (Verdejo-García et al., 2007b) for motor inhibition/impulsive action.

In addition, during the Go/No-go task, electroencephalographic (EEG) activity was recorded, and Go/No-go stimuli-evoked potentials (ERP) were extracted. Theory-driven ERP analyses focused on the N2 component, as observed in frontocentral electrodes (Jodo and Kayama, 1992; Mathalon et al., 2003; Miltner et al., 2003; Beste et al., 2011; Smith, 2011; Gajewski and Falkenstein, 2013) as the potentially best candidate to reflect inhibition-related cognitive activity during the Go/No-go task (Mathalon et al., 2003; see Folstein and Van Petten, 2007, for a review). Moreover, there are numerous reports of the relationship between N2 magnitude and addictive behavior, and, specifically, a reduction of the No-go > Go difference and the underlying source activations in samples of clinically addicted patients (Sokhadze et al., 2008; Dong et al., 2010; Luijten et al., 2011; Pandey et al., 2012). Complementarily, ERP measures are frequently more sensitive than behavioral ones, and can provide convergent evidence of the effect of key manipulations, especially when the behavioral effects of such manipulations are subtle (see, for example, Karayanidis et al., 2000; Hajcak et al., 2005).

For analysis and presentation purposes, our aim is divided into three specific research targets: (1) To quantify the influence of emotional and non-emotional dimensions of impulsivity, as measured by the UPPS-P model, on cocaine-use and gambling. Impulsivity dimensions are used, to *postdict* belongingness to clinical categories, but also to check for correlations with gambling episodes/cocaine dosage exposure estimates. (2) To check for the existence of differences between the clinical and non-clinical groups, as well as exposure-mediated effects on performance in Go/No-go and delay-discounting tasks (and on evoked EEG activity during the Go/No-go task). Finally (3), and in order to add evidence on the linkage between impulsive choice/action and emotional/non-emotional dimensions of impulsivity, we measured the differential impact of such dimensions on the two decision-making tasks, without taking addiction into account. The consequences of our results on the triadic mapping between impulsivity, decision-making, and addiction, and the role of emotions (particularly negative ones) in that mapping, will be discussed.

## MATERIALS AND METHODS

### PARTICIPANTS

Cocaine-dependent individuals (CDI; $n = 20$) were recruited from the Proyecto Hombre rehabilitation centers in Granada and Málaga (Spain) between January 2011 and July 2012. Pathological gamblers (PG; $n = 21$) were recruited from AGRAJER (Granadian Association of Gamblers in Rehabilitation, Granada, Spain) between October 2010 and July 2012. Most controls ($n = 23$) were recruited among non-drug using, non-gambling partners, and acquaintances (with no familiar linkage) of individuals in the clinical groups. The rest of them were recruited by incidental sampling, in such a way that their sociodemographic characteristics were not far from the clinical groups.

The inclusion criteria were (i) meeting DSM-IV criteria for cocaine dependence (CDI group) or pathological gambling (PG group)—as assessed by the Structured Clinical Interview for DSM-IV Disorders—Clinician Version (SCID; First et al., 1997); (ii) having a minimum abstinence interval of 15 days for all substances of abuse except nicotine, as determined by weekly urine toxicological tests (CDI) or cross validated therapist- and self-reports (PG). Exclusion criteria were: (i) the presence of any other Axis I or Axis II comorbid disorders with the exception of nicotine dependence; (ii) the presence of history of head injury or any diseases affecting the central nervous system.

The participants volunteering in this study were the same as those in Torres et al. (submitted). The specific socio-demographic characterization of the three groups is reported in Torres et al.'s abovementioned study. For the HC, PG, and CDI groups, respectively, sample sizes were 23, 21, and 20; proportions of females 0.09, 0.10, and 0.00; mean (SD) age 30.13 (8.64), 31.43 (5.92), and 34.75 (6.51); mean (SD) education years 14.55 (3.16), 13.90 (4.66), and 15.05 (4.21); and mean (SD) IQ 106.25 (10.22), 101.10 (9.07), and 105.35 (9.09). In all cases, differences between groups were not significant (min $p = 0.11$).

The study counted with explicit permission from the University of Granada's ethics committee. Prior to psychological and neuropsychological assessment, all participants were informed about the objectives and characteristics of the study, and signed an informed consent form. All of them were compensated with 36 € for their participation, independently of performance.

Upon consent, all participants were questioned about their age and number of education years, and were assessed using the Kaufman Brief Intelligence Test (K-BIT), the Interview for Research on Addictive Behavior (IRAB), the UPPS-P impulsivity scale, the Now-or-later task, and the Go/No-go task.

### PSYCHOMETRIC INSTRUMENTS

#### *UPPS-P impulsive behavior scale (Verdejo-García et al., 2010)*

In order to assess impulsivity, we used the Spanish version of the original 59-item questionnaire (Verdejo-García et al., 2010). This scale allows for a multidimensional assessment of impulsivity: positive, urgency, negative urgency, (lack of) premeditation, (lack of) perseverance, and sensation seeking.

#### *Interview for Research on Addictive Behaviors (IRAB, Verdejo-García et al., 2005)*

As noted in the introduction, a key factor in the present study is the degree of exposure to cocaine and gambling activities (in the CDI and PG groups, respectively). Most psychometric tools developed for clinical purposes do not measure exposure in an isolated manner (disregarding craving intensity, perception of lack of control over the addictive behavior, social and family problems, financial problems, and other symptoms and consequences of addiction).

All of those side factors are irrelevant to the current study. Actually, they would likely blur drug/gambling exposure effects. Hence, information about lifetime amount and duration of use of the different drugs was collected using the IRAB (Verdejo-García et al., 2005). The IRAB is inspired by applied and experimental behavior analysis, and was not developed to estimate the clinical significance of addiction, but to quantify the most important parameters of drug use behaviors (frequency, duration, amount), independently of the clinical status of the participant and the accompanying symptomatology. All the participants in the three groups went through the full IRAB interview. Here, however, we will consider only two composite measures yielded by the interview: (1) monthly amount of each drug consumed, in grams/month, and (2) severity, or estimated lifetime amount of drug consumed. In order to avoid extremely skewed distributions, monthly amount and severity were translated into within-design rank scores for all analyses. A more detailed display of IRAB results for the three groups can be found in Torres et al. (submitted, Appendix 1).

The IRAB has not been yet developed for gambling activities, so, in order to have equivalent measures for gambling and cocaine use, gamblers were questioned about the amount of money they used to gamble per month (in euros), and the lifetime duration of regular gambling (in months), across the whole course of their addiction. That is, the same questions used in the IRAB for registering drug use, were adapted to register the two key gambling parameters. Following the same procedure used with cocaine use parameters, severity of gambling was computed as the product of such factors, and then translated into within-design rank scores.

#### *The Kaufman Brief Intelligence Test (K-BIT, Kaufman and Kaufman, 1990)*

The K-BIT has been standardized and utilized widely, in both clinical and research settings, to assess cognitive abilities. It comprises measures of verbal and non-verbal intelligence and takes 10–30 min to administer. For control matching purposes, we will use only the compound IQ total score.

### DECISION-MAKING TASKS

#### *Go/No-go*

A computer-based implementation of the Go/No-go task was used (Verdejo-García et al., 2007b). The task consisted of 200 trials. In the first 100 trials (pre-switch), participants were asked to press any key as quickly as they could whenever the Go stimulus (a letter) was presented, and to withhold the response when the No-go stimulus (a different letter) was presented. The assignation of stimuli to the Go and No-go conditions was counterbalanced across subjects. In the second 100 trials of the task (post-switch), participants were asked to switch the assignation of the response from the Go to the No-go stimulus; in other words, they were asked to respond to the previously No-go stimulus and not to respond to the previously go stimulus. The proportion of Go vs. No-go trials on both phases (pre- and post-switch) was 80/20. The inter-trial interval (ISI) was set at 100 ms, and each stimulus was presented during 1000 ms. Auditory feedback (one of two distinctive sounds) was provided after each response to indicate whether that response had been right or wrong. If participants did not respond in the 1000 ms response window, the two same sounds were used as positive and negative feedback for not responding. That is, if no response was given, the same sounds indicated whether the absence of response had been right (No-go trials) or wrong (Go trials). Participants were instructed to respond as quickly as they could; however, in order to enforce time pressure, all responses given more than 400 ms after the stimulus were accompanied by the message "late," displayed on the center of the screen.

This version of the task was longer, but virtually identical to the one used by Verdejo-García et al. (2010). In previous implementations and pilot studies with the task, we have observed that false alarms (commission errors) and misses (omission errors) tend to become less frequent as the task progresses, which tends to generate performance floor and ceiling effects. The switch between phases was included to prevent such effects and boost the task's sensitivity to manipulations. Similarly, time pressure on responses (400 ms) prevents floor effects in false alarm rates' analyses, and tends to increase the magnitude of the No-go > Go N2 difference (Sokhadze et al., 2008).

#### *Now-or-later (Kirby et al., 1999)*

This paper-and-pencil task presents participants with 27 hypothetical two-option choices between an immediate small reward, and a delayed larger one (e.g., would you prefer 55 € now, or 110 € in 15 days?). The 27 items are pairs of imaginary monetary rewards varying in amount. Although responses allow for the calculation of the discount parameter ($k$), namely, the rate at which reward looses subjective utility across time, the calculation of such a parameter requires certain assumptions about the best-fitting time-utility function. Among the several theory-free possible dependent measures, we selected the simplest one, the total number of decisions favoring the immediate reward, as the main measure of reward-delay sensitivity.

### ERP EXTRACTION AND ANALYSIS

EEG activity was recorded exclusively during the Go/No-go task. EEGs were recorded from 62 scalp locations using tin electrodes arranged according to the extended 10–20 system mounted on an elastic cap (Brain Products, Inc.), and referenced online to FCz. Vertical and horizontal eye activity were recorded from one monopolar electrode placed below the left eye, and one monopolar electrode located in a straight line at the outer canthi of the right eye. Two scalp electrodes were attached to mastoids. All electrode impedances during recording were below 5 kΩ. EEG and EOG were sampled at 1000 Hz and amplified using a 0.016–1000 Hz band-pass filter. Subsequently, all EEG recordings were downsampled to 250 Hz, band-pass filtered using a 0.1–25 Hz 12 db/octave, re-referenced offline to average activity of the mastoids electrodes, and FCz activity was recovered. Offline signal preprocessing was done using EEGLAB software (Delorme and Makeig, 2004; freely available at http://sccn.ucsd.edu/eeglab).

EEG recordings were segmented from −200 to +300 ms, time-locked to the Go/No-go stimulus onset. Epochs were corrected for ocular artifacts by first computing the SOBI ICA decomposition (Belouchrani et al., 1993, 1997; Cardoso and Souloumiac, 1996, see also Tang et al., 2004) as identified by the ADJUST algorithm (Mognon et al., 2011). Other artifacts were subsequently removed using an automatic rejection procedure: segments were excluded for the remaining analyses when amplitudes were outside the ±100 μV range. Afterwards, segments were categorized as belonging to the Go or the No-go conditions. After the artifact correction procedure a minimum of 12 trials for the No-go and 57 for Go segments were retained for further processing.

The N2 amplitude was computed as the peak-to-peak difference between the most positive peak in the 160–220 ms time window and most negative peak in the 240–300 ms time window (see Nieuwenhuis et al., 2003; Smith, 2011, for a similar procedure). Given that the No-go > Go N2 difference is especially neat over frontocentral electrodes and independent of the reference (Jodo and Kayama, 1992; Mathalon et al., 2003; Miltner et al., 2003; Yeung and Nieuwenhuis, 2009; Smith, 2011), we used Fz, FCz, and Cz for testing between-group differences.

### SOURCE LOCATION

Standardized Low-Resolution Electromagnetic Tomography (sLORETA) was used for estimating the 3-D cortical distribution of current density underlying scalp activity. sLORETA, computations were done using the MNI152 template, with the 3D space solution restricted to cortical gray matter, according to the probabilistic Talairach atlas (Talairach and Tournoux, 1988). The cortical gray matter is partitioned in 6239 voxels at 5 mm spatial resolution. Brodmann anatomical labels are reported using MNI space. Standardized sLORETA current source densities with no regularization method were obtained from 61 channels (after recovering FCz) for each participant in each condition and for each time point in each feedback condition.

Brain localization analysis was carried out according to the following steps: first, a single measure of the activation of each voxel for the N2 interval (240–300 ms) was computed, by averaging voxel activations across the whole interval, and correcting by the average activity at the 160–220 interval. Second, we computed the correlation (across participants) between that averaged current density and the magnitude of the No-go > Go N2 difference, for each voxel and each feedback condition. And third, those areas in which at least 10 voxels were found to significantly correlate with the No-go > Go N2 difference score were identified. For a more detailed description of this source location rationale see Catena et al. (2012). For a virtually identical combination of peak-to-peak N2 amplitude computation, and time-window averaging for current source density estimation, see Nieuwenhuis et al. (2003).

## RESULTS

### RELATIONSHIPS BETWEEN IMPULSIVITY DIMENSIONS AND ADDICTIVE BEHAVIOR

#### *Impulsivity dimensions and belongingness to clinical groups*

Forward stepwise binomial regression analyses were carried out to test whether impulsivity dimensions were indicative of belongingness to the clinical groups. In the first one, the five UPPS-P dimensions entered the analysis as independent factors, and the clinical category [addicts (PG, CDI) vs. non-addicts (HC)] as dependent variable.

Stepwise regression is recommended over simultaneous regression for small samples. In the step 0 of this procedure the five impulsivity dimensions are included in a full regression model. In following steps, the variables are added one-by-one, in accordance with their predictive value. Only significant variables are introduced in the prediction equation. In our case, the process stopped at step 3, after including negative urgency, lack of premeditation and sensation seeking. The three-dimension binomial regression model correctly classified 82.8% of the cases. As displayed in **Table 1**, negative urgency, lack of premeditation, and sensation seeking were independently indicative of addiction. Both negative urgency and lack of premeditation were direct postdictors of addiction, whereas sensation seeking was a sign of non-addiction.

Secondly, a similar analysis was carried out to test whether impulsivity dimensions could distinguish between the two clinical categories (PG vs. CDI). Neither the full model nor any of the dimensions were indicative of the clinical category in this case (min $p = 0.16$).

**Table 1 | Forward binomial regression analysis (step 3) of clinical addiction [(CDI + PG) vs. HC] upon impulsivity dimensions.**

| | *B* | *SE* | *Wald* | *p* |
|---|---|---|---|---|
| Negative urgency | 1.084 | 0.334 | 10.517 | 0.001 |
| Lack of premeditation | 0.958 | 0.357 | 7.210 | 0.007 |
| Sensation seeking | –0.746 | 0.340 | 4.801 | 0.028 |

*B, standardized regression parameter; SE, standard error; Wald, contrast statistic; p, alfa error.*

#### *Impulsivity dimensions and addiction severity*

The model resulting from regressing gambling severity (in the PG group) over the five UPPS-P dimensions was significant ($R^2 = 0.25$), but only negative urgency was included in it, [$\beta = 0.50$, $t_{(20)} = 2.53$, $p = 0.02$]. The other four dimensions were far from significance (min $p = 0.22$). In the case of cocaine dependent individuals, neither the full regression model nor any of the UPPS-P dimensions was significantly predictive of cocaine use severity (computed for the CDI only; min $p = 0.27$).

### DIFFERENCES BETWEEN CLINICAL AND NON-CLINICAL GROUPS IN DECISION-MAKING TASKS

#### *Go/No-go Task*

***Behavioral results.*** A *hit rate* ($h$) was computed as the number of hits (responses to Go trials) divided by the sum of hits and misses (non-responses to Go trials), for each 25-trial block of the task. Similarly, a *false alarm rate* ($f$) was computed as the number of false alarms (responses to No-go trials) divided by the sum of false alarms and correct rejections (non-responses to No-go trials), for each 25-trial block of the task.

A group (HC, PG, CDI) × block (1–8) MANOVA on false alarm rates yielded a significant effect of group [Wilks' $\Lambda = 0.63$, $p = 0.05$]. Šidák-corrected *post-hoc* tests revealed a significant difference between PG and CDI in block 4 ($p = 0.02$), but not between PG and HC ($p = 0.83$), nor between HC and CDI ($p = 0.25$). Analogous group × block MANOVAs on hit rates and hit latencies did not yield any significant effect (min $p = 0.46$).

In order to test potential mediating effects of addiction severity we carried out MANCOVAs across blocks, with addiction severity as a continuous covariate, for mean false alarm rates, mean hit rates, and mean hit latencies, separately for PG and CDI. None of such analysis yielded any significant effect of severity or intensity (min $p = 0.12$).

***N2.*** **Figure 1** displays ERP waveforms, time-locked to the Go/No-go stimulus, separately for Go (left panel) and No-go stimuli (middle panel). The right panel represents the No-go > Go N2 difference for the three groups. No-go and Go N2 amplitudes were submitted to a 3 (group: HC, CDI, PG) × 2 (type of trial: No-go, go) × 3 (channel: FCz, Cz, Fz) repeated measures analysis of variance. Such an ANOVA yielded main effects of the type of trial, $F_{(1,\,61)} = 63.00$, $MSE = 5.01$, $p < 0.01$, and its interaction with group, $F_{(2,\,61)} = 6.65$, $p < 0.01$. No

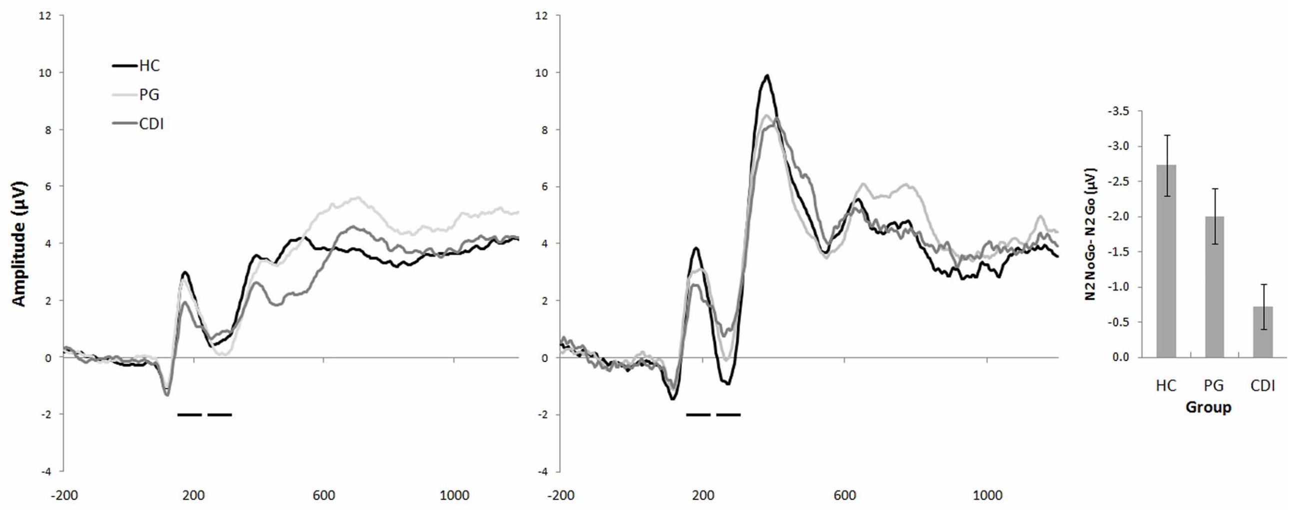


**FIGURE 1 | ERP waveform across groups (HC: Healthy Controls, PG: Pathological gamblers, CDI: Cocaine-dependent individuals) at FCz, time-locked to the Go stimulus (left panel), and the No-go stimulus (middle panel).** The right panel displays the No-go > Go N2 difference (the difference in N2 amplitudes between Go and No-go stimuli) for the three groups. The marks indicate the time frames used for the computation of the No-go > Go N2 difference.

other main or interaction effects were significant. Analyses within the interaction yielded a group effect for the No-go condition, $F_{(2,\,61)} = 3.83$, $p < 0.03$, but not for the Go one, $F_{(2,\,61)} = 0.05$. Bonferroni-corrected *post-hoc* comparisons indicated that negative amplitudes were larger for HC than for CDI ($p = 0.03$). No differences were observed between HC and PG ($p = 0.58$), or between PG and CDI ($p = 0.10$). No-go vs. Go differences were significant for all the groups, $t_{(23)} = 6.28$, $p < 0.01$, $t_{(19)} = 5.07$, $p < 0.01$, and $t_{(19)} = 2.24$, $p = 0.04$, respectively, for HC, PG, and CDI groups.

Correlations between cocaine use severity and monthly use and the No-go > Go N2 difference were both significant ($r = 0.40$, $p < 0.01$, and $r = 0.43$, $p < 0.01$, for severity and intensity, respectively). Restricting the analysis to the CDI group, the differential N2 effect was correlated only with use intensity ($r = 0.57$, $p < 0.01$), but not with use severity ($p = 0.45$). Given that the No-go > Go N2 effect has a negative sign, the direct correlations indicate that severity and intensity attenuated the difference between the No-go stimulus-evoked N2, and the Go N2.

Gambling intensity and severity did not correlate with the differential N2 effect (independently of whether the analysis was performed for the whole sample, or only for the PG group; all $p > 0.22$).

***Source location.*** Using the bootstrapping approach to control for multiple comparisons, included in the sLoreta package, we observed several clusters of voxels that significantly correlated with No-go > Go N2 difference scores in the whole sample of participants (**Table 2**; **Figure 2**). As expected, the size of the No-go > Go N2 difference was found to be associated to increased activation in a broad network of frontal and prefrontal cortices, predominantly in the left hemisphere.

Moreover, the correlation between the No-go > Go difference and peak voxel activation was significantly larger for the HC group than for the CDI group in the left BA10 area ($r = 0.57$ and –0.11, for HC and CDI, $p < 0.01$) and the left BA11 area ($r = 0.61$ and 0.11, for HC and CDI, $p = 0.03$). This suggests an abnormal functioning of these areas in the CDI group.

**Table 2 | Cortical areas identified to significantly correlate with No-go > Go N2 difference scores in the whole sample of participants.**

| Lobe | Structure | *BA* | *K* | *X* | *Y* | *Z* | *T* |
|---|---|---|---|---|---|---|---|
| L Limbic lobe | Anterior cingulate | 32 | 33 | −5 | 15 | 35 | −5.46 |
| L Limbic lobe | Cingulate gyrus | 24 | 23 | −5 | 10 | 30 | −6.17 |
| L Frontal lobe | Superior frontal gyrus | 11 | 20 | −20 | 65 | −10 | −5.48 |
| L Frontal lobe | Superior frontal gyrus | 10 | 60 | −20 | 65 | 0 | −6.04 |
| L Frontal lobe | Precentral gyrus | 9 | 43 | −35 | 5 | 40 | −5.09 |
| L Frontal lobe | Superior frontal gyrus | 8 | 56 | −20 | 25 | 50 | −5.72 |
| L Frontal lobe | Precentral gyrus | 6 | 108 | −35 | 0 | 30 | −5.35 |
| R Frontal lobe | Medial frontal gyrus | 8 | 14 | 0 | 20 | 50 | −4.70 |
| R Limbic lobe | Cingulate gyrus | 24 | 26 | 5 | 10 | 30 | −7.79 |
| R Limbic lobe | Cingulate gyrus | 32 | 24 | 0 | 15 | 35 | −5.69 |

*BA, Brodmann area; K, cluster size; X, Y, Z, Spatial coordinates; T, contrast statistic.*

*All p's < 0.001.*

### *Now or later task*

The number of trials in which the decision favored the immediate-reward option in the *now or later* task was taken as the main measure of sensitivity (intolerance) to reward delay. Mean (SD) scores were 10.52 (3.75), 17.00 (4.86), and 15.15 (4.59), for HC, PG, and CDI, respectively. The group effect was

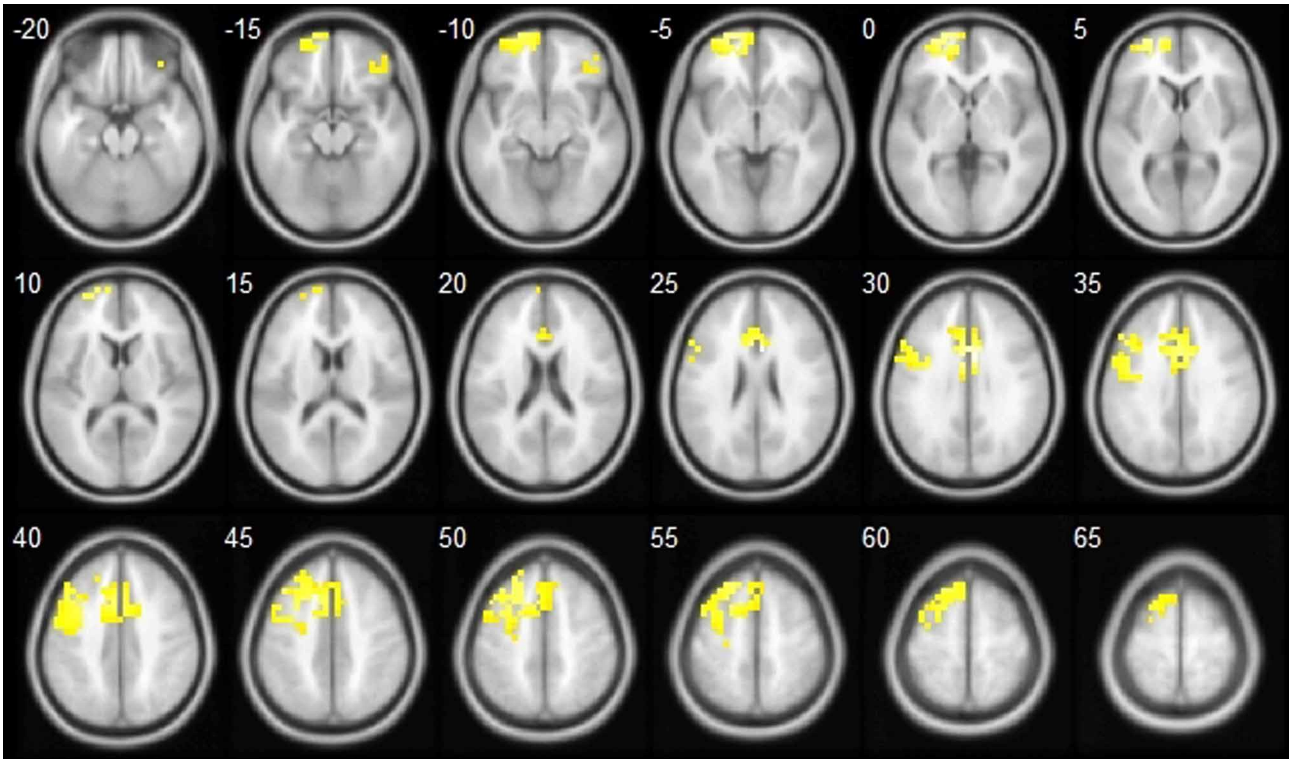


**FIGURE 2 | Topographical localization of the cortical areas identified to significantly correlate with No-go > Go N2 difference scores in the whole sample of participants.**

strongly significant [$F_{(2,\ 64)} = 12.72$, $MSE = 19.38$, $p < 0.01$]. Šidák-corrected *post-hoc* tests revealed significant differences between HC and PG ($p < 0.01$), and between HC and CDI ($p < 0.01$), but not between CDI and PG ($p = 0.46$).

Restricted linear regression analyses with severity as independent variable, and sensitivity to reward delay as dependent measure, were carried out for the PG and the CDI groups. In none of the two cases was severity significantly predictive of elevated sensitivity to reward delay [$t_{(20)} = 1.06$, $p = 0.30$; and $t_{(19)} = 1.89$, $p = 0.08$, for PG and CDI groups, respectively].

### RELATIONSHIPS BETWEEN IMPULSIVITY DIMENSIONS AND DECISION-MAKING TASKS

For the Go/No-go task, MANCOVAs across blocks, with the five UPPS-P dimensions as continuous covariates, were carried out on false alarm rates, hit rates, and hit latencies. No effects on false alarm and hit rates were close to significance. However, lack of premeditation and lack of perseverance exerted significant effects on latencies [Wilks' $\Lambda = 0.74$, $p = 0.04$, and Wilks' $\Lambda = 0.60$, $p < 0.01$, respectively]. The effect of lack of premeditation was restricted to block 4 [$F_{(1,\ 58)} = 4.78$, $MSE = 610.59$, $p = 0.03$], and the effect of lack of perseverance to block 2 [$F_{(1,\ 58)} = 9.81$, $MSE = 830.74$, $p < 0.01$]. The partial correlation between lack of premeditation and hit latency in block 4 (controlling for the other UPPS-P dimensions) was $r = 0.38$. Complementarily, the partial correlation between lack of perseverance and hit latency in block 2 was $r = 0.28$. In both cases, impulsivity significantly slowed decisions down.

Finally, the five UPPS-P dimensions were used as predictors of sensitivity to reward delay in a stepwise regression analysis. Negative urgency emerged as the only significantly predictive dimension [$\beta = 0.43$, $t_{(63)} = 3.74$, $p < 0.01$] included in the regression model.

## DISCUSSION

Our general research aim was to test whether emotional and non-emotional dimensions of impulsivity were differentially predictive of decision-making and addictive behavior in three samples of PGs, CDIs, and HCs.

The first specific research target was to use impulsivity dimensions to *postdict* belongingness to clinical categories, and to check for correlations between impulsivity dimensions and gambling episodes/cocaine dosage exposure. With that aim in mind, regression analyses were carried out to estimate the value of impulsivity dimensions as postdictors of the clinical category (addicts vs. non-addicts, and PG vs. CDI). Only negative urgency (but not positive urgency), lack of premeditation (in a positive direction), and sensation seeking (in a negative direction) were indicative of selective inclusion in the groups of addicted individuals.

Although none of the dimensions discriminated between the two clinical groups (CDI, PG), an important difference between them emerged when the same dimensions were used as postdictors of addiction severity. Negative urgency independently covaried with gambling severity in PG, but did not predict cocaine use severity in CDI. This is consistent with previous reports that negative urgency is a sign of *overpathologization* in addictive processes (e.g., Michalczuk et al., 2011), but not with those in which negative urgency has been linked to cocaine neurotoxic effects. Although some recent works hypothesize a relationship between cocaine dependence severity and negative urgency (Albein-Urios et al., 2012a; Cándido et al., 2012), direct evidence of such a relationship is sparse (Verdejo-García et al., 2007a). Thus, the question of whether negative urgency is involved in overpathologization of different types of addictive processes, or has a privileged role in gambling remains open. Still, the neat relationship between gambling severity and negative urgency discards the possibility that such a relationship is exclusively mediated by cocaine neurotoxicity.

Our second research target was to check for the existence of differences between clinical and non-clinical groups, as well as exposure-mediated effects on performance in Go/No-go and delay-discounting tasks, and on evoked EEG activity during the Go/No-go task. The results from previous studies suggest that cocaine use, but not pathological gambling, relates to performance deficits in response inhibition skills (see Fillmore and Rush, 2002; Verdejo-García et al., 2007b; Kertzman et al., 2011; Van Holst et al., 2012). Accordingly, our results show that CDIs, but not PGs, perform abnormally in the Go/No-go task. CDIs presented focal increases of false alarms (commission errors). In parallel, and most importantly, ERP results showed an abnormal pattern of Go/No-go stimuli-evoked EEG activity in the CDI group (the No-go > Go N2 difference in the CDI group was the smallest among the three groups). The deleterious effect on N2 was mediated by cocaine dosage exposure. This pattern of results is compatible with the well-known association between cocaine consumption and malfunctioning of motor inhibition mechanisms (Kaufman et al., 2003; Garavan et al., 2008), and also with the proposal that such malfunctioning is at least partially due to neurotoxic dosage exposure effects.

In contrast to this difference between clinical groups found in the Go/No-go task, both CDIs and PGs were more sensitive than HCs to reward delay, as measured by the now-or-later decision-making task. In this case, neither the two groups of addicts differed between them, nor the effect was mediated by cocaine use or gambling severity. Consequently, the finding by Kertzman et al. (2011) that gamblers discount reward more rapidly than other addicts has not been replicated.

Results on source localization complemented the ones on the N2 ERP. The areas involved in the generation of the No-go > Go N2 difference (in the whole sample) are mostly coincident with the ones described in previous works. Reporting the involvement of pre-SMA and anterior cingulate is common to virtually all relevant studies. A variety of other dorsal, lateral, and anterior areas of the prefrontal cortex have been reported to functionally interact with these (Rubia et al., 2001; Horn et al., 2003; Nieuwenhuis et al., 2003; Lavric et al., 2004; Ridderinkhof et al., 2004a,b; Tanji and Hoshi, 2008; Zheng et al., 2008; Smith et al., 2013) in the generation of the No-go > Go N2 difference.

Also with regard to source location, some mention needs to be made about group differences. As noted above, left BA10 and BA11 were unique at correlating with the N2 differential effect in the control group, but not in the CDI group (which is suggestive of abnormal functioning of these areas in cocaine addicts). Among the different areas involved in Go/No-go performance, Horn et al. (2003) specifically attributed a temporally early, inhibitory role to BA10/11. Zheng et al. (2008), on the other hand, attributed a similar inhibitory role to left BA10. Importantly, none of the areas mentioned by Horn et al. as potentially involved in the emotional aspects of Go/No-go performance (posterior orbitofrontal cortex, the temporal poles, and the posterior cingulate) are involved either in differential N2 generation, or in HC-CDI source location differences in the present study. This pattern of data thus suggests that the deleterious effect of cocaine addiction on the Go/No-go task mainly affects its inhibition component.

The last research target was aimed at completing the triadic mapping between addictions, impulsivity, and decision-making. Therefore, we measured the impact of impulsivity on decision making, without taking addiction into account. Lack of premeditation and lack of perseverance slowed down responses in the Go/No-go task. Although such a result seems counterintuitive, previous results show that, in speeded decision-making tasks, impulsivity interferes with response selection. For example, Expósito and Andrés-Pueyo (1997) found that "more impulsive Ss are more affected by stimulus-response incompatibility and therefore present higher latencies" (p. 696). Further implications of such a result go beyond the aims of the present study[1].

Still, the impulsivity dimensions involved in the latency effect (lack of premeditation and lack of perseverance), were only partially coincident with the ones related to addiction (lack of premeditation, sensation seeking, and negative urgency), and not coincident at all with the one predicting gambling severity (negative urgency). On the other hand, the impulsivity dimension predicting sensitivity to reward delay was the same one found to be predictive of gambling severity (negative urgency). Both the lack of involvement of emotional dimensions in Go/No-go performance, and the involvement of negative urgency in reward delay sensitivity are partially contradictory with Cyders and Coskunpinar's (2011a; although see Cyders and Coskunpinar, 2011b) findings on the relationship between trait and neuropsychological measures of impulsivity. Although our source location and behavioral results are fully congruent with each other, this apparent dissonance probably deserves further investigation.

In summary, among the emotional dimensions of impulsivity, negative urgency (but not positive urgency) has been observed to be selectively involved in addiction, independently of its type (pathological gambling, cocaine dependence). The tight link between negative urgency and emotionally-charged decision-making processes is reinforced by the fact that negative urgency

---

[1] In short, it seems to support the idea that impulsivity affects the decisional aspects of response selection in the Go/No-go task. This interpretation goes against the idea of motor impulsivity as mere disinhibition, namely, the malfunctioning of a central inhibitory mechanism (see also Donkers and Van Boxtel, 2004; Spinella, 2004; Perales et al., 2009, for discussions on the matter).

was the only dimension significantly predicting sensitivity to reward delay in the delay discounting task. Furthermore, negative urgency was specifically related to gambling overpathologization. This effect is compatible with the possibility that gambling (large amounts of money for long periods) is fueled by negative emotions and moods, and such emotions and moods might operate as gambling triggers (Oakes et al., 2011, 2012; Williams et al., 2012). And, the other way round, it is also compatible with the possibility that dosage-like exposure to gambling episodes generates sensitization and neuroadaptive effects (Robinson and Berridge, 2003; Mathewson, 2009).

Results on sensation seeking probably deserve a short digression. Rather counterintuitively, sensation seeking scores were lower in both samples of addicted individuals, and were independently and inversely indicative of belongingness to any of the clinical samples. As noted in the introduction, sensation seeking has been shown to be linked to reward delay sensitivity, and to be a significant predictor of recreational, non-clinical drug use (Cyders and Coskunpinar, 2011a,b; Torres et al., submitted). This can be taken as independent evidence of the transition from positive to negative emotion-driven impulsivity in drug-use pathologization.

Among the non-emotional dimensions of impulsivity, lack of premeditation was found to be simultaneously involved in the two types of addiction, and in delaying Go/No-go decisions. However, in agreement with previous evidence, the effect of addiction on the Go/No-go task was limited to CDI, and was of a different nature (namely, an increased focal false alarm rate, and abnormal Go/No-go stimulus evoked cortical activity). In other words, the effects of impulsivity on Go/No-go decision-making and addictive processes involve different neurocognitive mechanisms. Impulsivity is not by itself responsible for the decision-making anomalies in cocaine addicts revealed by the Go/No-go task.

This study holds relevant strengths and noteworthy limitations. Some limitations are: first, the relatively small sample size of the clinical groups, which may have impacted the statistical power of multivariate contrasts. Secondly, the number of valid trials in the Go/No-go task (particularly, in the case of No-go trials) was too small to analyze ERPs dynamically, that is, trying to capture the changes in cortical activity occurring during the task (in a block-by-block fashion, or across phases). Unfortunately, any attempts to analyze the task in parts rendered the signal-to-noise ratio too low to capture any significant effect. And finally, it is important to mention the limitation of the current design to capture causality in the associations between emotional and non-emotional pathways to impulsive personality and impulsive action and choice and drug use behavior. As acknowledged in the introduction, such causal relationship can only be tracked by longitudinal studies. On the other hand, its major strengths are the use of a multidimensional approach to impulsive traits and impulsive decisions, measuring both emotional and non-emotional pathways to impulsive action and impulsive choice, and combining behavioral and electrophysiological approaches. Another relevant strength is the comparison between two clinical groups of addicted individuals who were matched in terms of baseline cognition (IQ) and clinical features, but strikingly differed in their impulsive action patterns, which, as discussed, have relevant implications both from the basic science and the clinical perspectives.

## FINAL REMARKS

The results presented here make two relevant contributions to the current literature. First, they show that aspects of impulsivity with a core component of negative emotion processing and appraisal play a key role in decision making (reward delay sensitivity), and addiction (gambling severity and intensity). And, secondly, they show that some key aspects of decision-making and neurobehavioral anomalies in addicts are independent of such factors. CDIs, but not PGs, performed abnormally in the Go/No-go task, and showed an abnormal N2 signal. Source location analyses show the involvement of a broad network of frontal and prefrontal areas (but not of prelimbic and limbic areas) in the generation of N2, as well as left BA10 and BA11 abnormal involvement in N2 in the group of cocaine addicts. This pattern of neurophysiologic results is in agreement with the existence of an association between cocaine consumption and malfunctioning of inhibition mechanisms, at least partially due to neurotoxic, cocaine exposure effects.

## FUNDING

The research presented here has been funded by grants from the Ministerio de Ciencia e Innovación, MICINN (Spain), for Ana Torres/José C. Perales (ref. # PSI2009-13133), and Andrés Catena/Antonio Maldonado (ref. # PSI2009-12217); by a Junta de Andalucía (Spain) grant (ref. # PB09-SEJ4752) for Antonio Cándido: and by a RETICS (Redes Temáticas de Investigación Cooperativa en Salud) subprogramme grant (Ref. # RD12/0028/0017) from the Ministerio de Sanidad, Servicios Sociales e Igualdad (Spain), for José C. Perales/Antonio Verdejo-García.

## REFERENCES

Albein-Urios, N., Martinez-González, J. M., Lozano, O., Clark, L., and Verdejo-García, A. (2012a). Comparison of impulsivity and working memory in cocaine addiction and pathological gambling: implications for cocaine-induced neurotoxicity. *Drug Alcohol Depend.* 126, 1–6.

Albein-Urios, N., Verdejo-Román, J., Asensio, S., Soriano-Mas, C., Martínez-González, J. M., and Verdejo-García, A. (2012b). Reappraisal of negative emotions in cocaine dependence: dysfunctional corticolimbic activation and connectivity. *Addict. Biol.* 10, 1369–1600.

Barratt, E. S. (1993). "Impulsivity: integrating cognitive, behavioral, biological and environmental data," in *The Impulsive Client: Theory, Research, and Treatment,* eds W. McCowan, J. Johnson, and M. Shure (Washington, DC: American Psychological Association), 39–56.

Belouchrani, A., Abed Meraim, K., Cardoso, J.-F., and Moulines, E. (1993). "Second-order blind separation of temporally correlated sources," in *Proceeding of the International Conference on Digital Signal Processing,* (Cyprus), 346–351.

Belouchrani, A., Abed Meraim, K., Cardoso, J.-F., and Moulines, E. (1997). A blind source separation technique based on second order statistics. *IEES Trans. Signal Process.* 45, 434–444.

Beste, C., Ness, V., Falkenstein, M., and Saft, C. (2011). On the role of fronto-striatal neural synchronization processes for response inhibition – evidence from ERP phase-synchronization analyses in pre-manifest Huntington's disease gene mutation carriers. *Neuropsychologia* 49, 3484–3493.

Cándido, A., Orduña, E., Perales, J. C., Verdejo-García, A., and Billieux, J. (2012). Validation of a short spanish version of the Upps-p impulsive behavior scale. *Trastornos Adictivos* 14, 73–78.

Cardoso, J. F., and Souloumiac, A. (1996). Jacobi angles for simultaneous diagonalization. *SIAM J. Matrix Anal. Appl.* 17, 161–164.

Catena, A., Perales, J. C., Megías, A., Cándido, A., Jara, E., and Maldonado, A. (2012). The brain network of expectancy and uncertainty processing. *PLoS ONE* 7:e40252. doi: 10.1371/journal.pone.0040252

Cyders, M., and Coskunpinar, A. (2011a). Measurement of constructs using self-report and behavioral lab tasks: is there overlap in nomothetic span and construct representation for impulsivity? *Clin. Psychol. Rev.* 31, 965–982.

Cyders, M., and Coskunpinar, A. (2011b). The relationship between self-report and lab task conceptualizations of impulsivity. *J. Res. Pers.* 46, 121–124.

Cyders, M., and Smith, G. T. (2008). Emotion-based dispositions to rash action: positive and negative urgency. *Psychol. Bull.* 134, 807–828.

Cyders, M. A., Flory, K., Rainer, S., and Smith, G. T. (2009). The role of personality dispositions to risky behavior in predicting first-year college drinking. *Addiction* 104, 193–202.

Cyders, M. A., Smith, G. T., Spillane, N. S., Nichea, S., Fischer, S., Annus, A. M., et al. (2007). Integration of impulsivity and positive mood to predict risky behavior: development and validation of a measure of positive urgency. *Psychol. Assess.* 19, 107–118.

Dawe, S., and Loxton, N. J. (2004). The role of impulsivity in the development of substance use and eating disorders. *Neurosci. Biobehav. Rev.* 28, 343–351.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21.

Dong, G., Lu, Q., Zhou, H., and Zhao, X. (2010). Impulse inhibition in people with Internet addiction disorder: electrophysiological evidence from a Go/NoGo study. *Neurosci. Lett.* 485, 138–142.

Donkers, F. C. L., and Van Boxtel, G. J. M. (2004). The N2 in go/nogo tasks reflects conflict monitoring not response inhibition. *Brain Cogn.* 56, 165–176.

Evenden, J. L. (1999). Varieties of Impulsivity. *Pyschopharmacology* 146, 348–361.

Everitt, B. J., and Robbins, T. W. (2000). Second-order schedules of drug reinforcement in rats and monkeys: measurement of reinforcing efficacy and drug-seeking behaviour. *Psychopharmacology* 153, 17–30.

Expósito, J., and Andrés-Pueyo, A. (1997). The effects of impulsivity on the perceptual and decision stages in a choice reaction time task. *Pers. Indiv. Differ.* 22, 693–697.

Fillmore, M. T., and Rush, C. R. (2002). Impaired inhibitory control of behavior in chronic cocaine users. *Drug Alcohol Depend.* 66, 265–273.

First, M. B., Spitzer, R. L., Gibbon, M., and Williams, J. B. W. (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I)*. New York, NY: Biometric Research Department.

Folstein, J. R., and Van Petten, C. (2007). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170.

Franken, I. H. A., and Muris, P. (2006). Gray's impulsivity dimension: a distinction between reward sensitivity versus rash impulsiveness. *Pers. Indiv. Differ.* 40, 1337–1347.

Gajewski, P. D., and Falkenstein, M. (2013). Effects of task complexity on ERP components in Go/Nogo tasks. *Int. J. Psychophysiol.* Available online at: http://www.sciencedirect.com/science/article/pii/S016787601200579X

Garavan, H., Kaufman, J. N., and Hester, R. (2008). Acute effects of cocaine on the neurobiology of cognitive control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3267–3276.

Goldstein, M., Brendel, G., Tuescher, O., Pan, H., Epstein, J., Beutel, M., et al. (2007). Neural substrates of the interaction of emotional stimulus processing and motor inhibitory control: an emotional linguistic go/no-go fMRI study. *Neuroimage* 36, 1026–1040.

Hajcak, G., Moser, J. S., Yeung, N., and Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology* 42, 151–160.

Horn, N. R., Dolan, R., Elliott, R., Deakin, J. F. W., and Woodruff, P. W. R. (2003). Response inhibition and impulsivity: an fMRI study. *Neuropsychologia* 41, 1959–1966.

Jodo, E., and Kayama, Y. (1992). Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalogr. Clin. Neurophysiol.* 82, 477–482.

Joseph, J. E., Liu, X., Jiang, Y., Lynam, D., and Kelly, T. H. (2009). Neural correlates of emotional reactivity in sensation seeking. *Psychol. Sci.* 20, 215–223.

Karayanidis, F., Robaey, P., Bourassa, M., de Koning, D., Geoffroy, G., and Pelletier, G. (2000). ERP differences in visual attention processing between attention-deficit hyperactivity disorder and control boys in the absence of performance differences. *Psychophysiology* 37, 319–333.

Kaufman, A., and Kaufman, N. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, J. N., Ross, T. J., Stein, E. A., and Garavan, H. (2003). Cingulate hypoactivity in cocaine users during a GO-NOGO task as revealed by event-related functional magnetic resonance imaging. *J. Neurosci.* 23, 7839–7843.

Kertzman, S., Lidogoster, H., Aizer, A., Kotler, M., and Dannon, P. N. (2011). Risk-taking decisions in pathological gamblers is not a result of their impaired inhibition ability. *Psychiatry Res.* 188, 71–77.

Kirby, K. N., Petry, N. M., and Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J. Exp. Psychol.* 128, 78–87.

Kringelbach, M. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nat. Rev. Neurosci.* 6, 691–702.

Lavric, A., Pizzagalli, D. A., and Forstmeier, S. (2004). When "go" and "nogo" are equally frequent: ERP components and cortical tomography. *Eur. J. Neurosci.* 20, 2483–2488.

Luijten, M., Little, M., and Franken, I. H. A. (2011). Deficits in inhibitory control in smokers during a Go/NoGo task: an investigation using event-related brain potentials. *PLoS ONE* 6:e18898. doi: 10.1371/journal.pone.0018898

Mathalon, D. H., Whitfield, S. L., and Ford, J. M. (2003). Anatomy of an error: ERP and fMRI. *Biol. Psychol.* 64, 119–141.

Mathewson, S. (2009). *The Effect of Repeated Exposure to Unpredictable Reward on Dopamine Neuroplasticity*. Toronto, Canada: Graduate Department of Psychology, University of Toronto.

Metcalfe, J., and Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychol. Rev.* 106, 3–19.

Michalczuk, R., Bowden-Jones, H., Verdejo-García, A., and Clark, L. (2011). Impulsivity and cognitive distortions in pathological gamblers attending the UK National Problem Gambling Clinic: a preliminary report. *Psychol. Med.* 41, 2625–2635.

Miltner, W. H. R., Lemke, U., Weiss, T., Holroyd, C., Scheffers, M. K., and Coles, M. G. H. (2003). Implementation of error-processing in the human anterior cingulate cortex: a source analysis of the magnetic equivalent of the error-related negativity. *Biol. Psychol.* 64, 157–166.

Mognon, A., Bruzzone, L., Jovicich, J., and Buiatti, M. (2011). ADJUST: an Automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology* 48, 229–240.

Murray, E. (2008). The amygdala, reward and emotion. *Trends Cogn. Sci.* 11, 489–497.

Nieuwenhuis, S., Yeung, N., Van den Wildenberg, W., and Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cogn. Affect. Behav. Neurosci.* 3, 17–26.

Oakes, J., Pols, R., Battersby, M., Lawn, S., Pulvirenti, M., and Smith, D. (2011). A focus group study of predictors of relapse in electronic gaming machine problem gambling, part 1: factors that "push" towards relapse. *J. Gambl. Stud.* 28, 451–464.

Oakes, J., Pols, R., Battersby, M., Lawn, S., Pulvirenti, M., and Smith, D. (2012). A focus group study of predictors of relapse in electronic gaming machine problem gambling, part 2: factors that "pull" the gambler away from relapse. *J. Gambl. Stud.* 28, 465–479.

Pandey, A. K., Karamajan, C., Tang, Y., Chorlian, D. B., Roopesh, B. N., Manz, N., et al. (2012). Neurocognitive deficits in male alcoholics: an ERP/sLORETA analysis of the N2 component in an equal probability Go/NoGo task. *Biol. Psychol.* 89, 170–182.

Perales, J. C., Verdejo-García, A., Moya, M., Lozano, O., and Perez-García, M. (2009). Bright and dark sides of impulsivity: performance of women with high and low trait impulsivity on neuropsychological tasks. *J. Clin. Exp. Neuropsychol.* 31, 927–944.

Quartz, S. R. (2009). Reason, emotion and decision-making: risk and reward computation with feeling. *Trends Cogn. Sci.* 13, 209–201.

Reynolds, B. (2006). A review of delay-discounting research with humans: relations to drug use and gambling. *Behav. Pharmacol.* 17, 651–667.

Ridderinkhof, K. R., van den Wildenber, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004a). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004b). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Robinson, T., and Berridge, K. (2003). Addiction. *Annu. Rev. Psychol.* 54, 25–53.

Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., et al. (2001). Mapping motor inhibition: conjunctive brain activations across different versions of Go/No-Go and stop tasks. *Neuroimage* 13, 250–261.

Schmeichel, B. J., and Inzlicht, M. (2013). "Incidental and integral effects of emotions on self-control," in *Handbook of Cognition and Emotion*, eds M. D. Robinson, E. R. Watkins, and E. Harmon-Jones (New York, NY: Guilford Press).

Smith, J. L. (2011). To Go or not to Go, that is the question: do the N2 and P3 reflect stimulus- or response-related conflict. *Int. J. Psychophysiol.* 82, 143–152.

Smith, J. L., Jamadar, S., Provost, A. L., and Michie, P. T. (2013). Motor and non-motor inhibition in the Go/NoGo task: an ERP and fMRI study. *Int. J. Psychophysiol.* Available online at: http://www.sciencedirect.com/science/article/pii/S0167876012005508

Sokhadze, E., Stewart, C., Hollifield, M., and Tasman, A. (2008). Event-related potential study of executive dysfunctions in a speeded reaction task in cocaine addiction. *J. Neurother.* 12, 185–204.

Spinella, M. (2004). Neurobehavioral correlates of impulsivity: evidence of prefrontal involvement. *Int. J. Neurosci.* 114, 95–104.

Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain.* New York, NY: Thieme Medical.

Tang, A. C., Sutherland, M. T., and McKinney, C. J. (2004). Validation of SOBI components from high-density EEG. *Neuroimage* 25, 539–553.

Tanji, J., and Hoshi, E. (2008). Role of the lateral prefrontal cortex in executive behavioral control. *Physiol. Rev.* 88, 37–57.

Van Holst, R. J., Van Holstein, M., Van den Brink, W., Veltman, D. J., and Goudriaan, A. E. (2012). Response inhibition during cue reactivity in problem gamblers: an fMRI study. *PLoS ONE* 7:e30909. doi: 10.1371/journal.pone.0030909

Verdejo-García, A., Bechara, A., Recknor, E. C., and Pérez-García, M. (2007a). Negative emotion-driven impulsivity predicts substance dependence problems. *Drug Alcohol Depend.* 91, 213–219.

Verdejo-García, A., Perales, J. C., and Pérez-García, M. (2007b). Cognitive impulsivity in cocaine and heroin polysubstance abusers. *Addict. Behav.* 32, 950–966.

Verdejo-García, A., Contreras-Rodríguez, O., Fonseca, F., Cuenca, A., Soriano-Mas, C., Rodríguez, J., et al. (2013). Functional alteration in frontolimbic systems relevant to moral judgment in cocaine-dependent subjects. *Addict. Biol.* doi: 10.1111/j.1369-1600.2012.00472.x. [Epub ahead of print].

Verdejo-García, A., Lawrence, A. J., and Clark, L. (2008). Impulsivity as a vulnerability marker for substance-use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. *Neurosci. Biobehav. Rev.* 32, 777–810.

Verdejo-García, A., López-Torrecillas, F., Aguilar de Arcos, F. y Pérez-García, M. (2005). Differential effects of MDMA, cocaine, and cannabis use severity on distinctive components of the executive functions in polysubstance users, a multiple regression analysis. *Addict. Behav.* 30, 89–101.

Verdejo-García, A., Lozano, O., Moya, M., Alcázar, M. A., and Pérez-García, M. (2010). Psychometric properties of a spanish version of the UPPS–P impulsive behavior scale: reliability, validity and association with trait and cognitive impulsivity. *J. Pers. Assess.* 92, 70–77.

Whiteside, S. P., and Lynam, D. R. (2001). The five factor model and impulsivity: using a structural model of personality to understand impulsivity. *Pers. Indiv. Differ.* 30, 669–689.

Williams, A. D., Grisham, J. R., Erskine, A., and Cassedy, E. (2012). Deficits in emotion regulation associated with pathological gambling. *Br. J. Clin. Psychol.* 51, 223–238.

Winstanley, C. A., Eagle, D. M., and Robbins, T. W. (2006). Behavioral models of impulsivity in relation to ADHD: translation between clinical and preclinical studies. *Clin. Psychol. Rev.* 26, 379–395.

Yeung, N., and Nieuwenhuis, S. (2009). Dissociating response conflict and error likelihood in anterior cingulate cortex. *J. Neurosci.* 18, 14506–14510.

Zheng, D., Oka, T., Bokura, H., and Yamaguchi, S. (2008). The key locus of common response inhibition network for no-go and stop signals. *J. Cogn. Neurosci.* 20, 1434–1442.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 November 2012; accepted: 04 February 2013; published online: 21 February 2013.*

*Citation: Torres A, Catena A, Megías A, Maldonado A, Cándido A, Verdejo-García A and Perales JC (2013) Emotional and non-emotional pathways to impulsive behavior and addiction. Front. Hum. Neurosci. **7**:43. doi: 10.3389/fnhum.2013.00043*

frontiers in
HUMAN NEUROSCIENCE

MINI REVIEW ARTICLE
published: 27 March 2013
doi: 10.3389/fnhum.2013.00099

# Beyond human intentions and emotions

**Elsa Juan[1], Chris Frum[2], Francesco Bianchi-Demicheli[3], Yi-Wen Wang[4], James W. Lewis[5] and Stephanie Cacioppo[6]***

[1] Psychology Department, University of Geneva, Geneva, Switzerland
[2] Department of Physiology and Pharmacology, Center for Neuroscience, West Virginia University, Morgantown, WV, USA
[3] Geneva University Hospital, Geneva, Switzerland
[4] Academy of Psychology and Behavior, Tianjin Normal University, Tianjin, China
[5] Department of Neurobiology and Anatomy, Center for Neuroscience, West Virginia University, Morgantown, WV, USA
[6] High-Performance Electrical NeuroImaging Laboratory, Department of Psychology, Center for Cognitive and Social Neuroscience, The University of Chicago, Chicago, IL, USA

**Edited by:**
Agustin Ibanez, Institute of Cognitive Neurology, Argentina

**Reviewed by:**
Aldo Ragazzoni, University of Florence, Italy
Blas Couto, Institute of Cognitive Neurology, Argentina

***Correspondence:**
Stephanie Cacioppo, High-Performance Electrical NeuroImaging Laboratory, Department of Psychology, Center for Cognitive and Social Neuroscience, The University of Chicago, 5848 S. University Avenue, Chicago, IL 60637, USA.
e-mail: cacioppos@uchicago.edu

Although significant advances have been made in our understanding of the neural basis of action observation and intention understanding in the last few decades by studies demonstrating the involvement of a specific brain network (action observation network; AON), these have been largely based on experimental studies in which people have been considered as strictly isolated entities. However, we, as social species, spend much more of our time performing actions interacting with others. Research shows that a person's position along the continuum of perceived social isolation/bonding to others is associated with a variety of physical and mental health effects. Thus, there is a crucial need to better understand the neural basis of intention understanding performed in interpersonal and emotional contexts. To address this issue, we performed a meta-analysis using of functional magnetic resonance imaging (fMRI) studies over the past decade that examined brain and cortical network processing associated with understanding the intention of others actions vs. those associated with passionate love for others. Both overlapping and distinct cortical and subcortical regions were identified for intention and love, respectively. These findings provide scientists and clinicians with a set of brain regions that can be targeted for future neuroscientific studies on intention understanding, and help develop neurocognitive models of pair-bonding.

**Keywords: social neuroscience, embodied cognition, intention understanding, love, neuroimaging, fMRI meta-analysis**

## INTRODUCTION

Throughout the past three decades, a growing number of studies have shown that one understands actions and intentions of other people by shaping one's understanding and anticipation of the environment based on one's own motor system (Jeannerod, 2001; Van Overwalle and Baetens, 2009; Becchio et al., 2012). Theories on embodied cognition and simulation extend these findings by suggesting that reading the intention of others occurs through a direct and automatic matching process between observed and performed actions, and via the reactivation of the bodily states that were originally active during past self-related sensori-motor experiences—as if the observers were "reliving" the observed motor experiences (Rizzolatti et al., 2001; Rizzolatti and Craighero, 2004; Niedenthal et al., 2005; Lewis et al., 2006; Niedenthal, 2007; Rizzolatti and Sinigaglia, 2008; Grafton, 2009). The recent development in neuroimaging sheds lights on the neural activations mediating this mechanism (Grafton, 2009; Becchio et al., 2012). For instance, neuroimaging studies show that reading intentions of others recruit brain areas that are also activated when someone performs the same action (Grafton, 2009, for review; Ortigue et al., 2009; Rizzolatti and Sinigaglia, 2010). Interestingly, these brain areas can be divided into two functionally separable brain networks (Grafton, 2009; Ortigue et al., 2009). The first, referred to as the "action observation network" (AON), involves an inferior fronto-parietal network (FPN) and includes a subset of areas that are associated within the putative human Mirror Neuron System (hMNS; Grafton et al., 1996; Rizzolatti and Craighero, 2004; Desmurget et al., 2009; Grafton, 2009). This AON system is thought to be particularly important for integrating sensori-motor information during perceptual judgments about actions (Rizzolatti and Craighero, 2004), and also for understanding hand-object interactions and intentions on the basis of embodied cognitive mechanisms (Iacoboni et al., 2005; Rizzolatti and Sinigaglia, 2007; Grafton, 2009). The second brain network, referred to as the "social-network" (SN), recruits brain areas involved in social interaction (Grafton, 2009; Sugiura et al., 2009; Wakusawa et al., 2009; Canessa et al., 2012). This SN includes the medial prefrontal cortex, precuneate cortex, insula, and amygdala (Wheatley et al., 2007; Grafton, 2009). Both networks include the posterior part of the superior temporal sulcus (pSTS); the superior temporal gyrus (STG), the middle temporal gyrus (MTG), and the part of the angular gyrus that is near the ascending limb of STS (Allison et al., 2000; Pelphrey et al., 2004; Thompson et al., 2005, 2007; Pelphrey and Morris, 2006; Materna et al., 2008). Interestingly, the STS region, notably its posterior part (pSTS), is also recruited by relatively low level processes such as observation of visual biological motion (Jellema et al., 2000), auditory biological actions (Bidet-Caulet et al., 2005; Gazzola et al., 2006; Lewis et al., 2011), and other operations such as social inferential processing in tasks requiring mentalizing, and theory of mind (Grezes et al., 2004; Saxe et al., 2004; Schultz et al., 2004; Grossman et al., 2005; Frith and Frith, 2006; Brass et al., 2007; Van Overwalle and Baetens, 2009).

Although these findings provide valuable information about the brain mechanisms involved in the understanding of actions performed by strangers, they do not tell much about the brain mechanisms involved in the understanding of a significant other [a person with whom the participant intends to be with (i.e., a participant's partner in an intimate relationship or a best friend in an companionate relationship)]. To date, studies on intention understanding have been largely based on functional magnetic resonance imaging (fMRI) studies in which participants have been considered as strictly isolated entities i.e., focusing mostly on the action type rather than on the relationship with the agent and the observer. However, people typically spend most of their time in social settings interacting with significant others. Research shows that a person's position along the continuum of perceived social isolation/bonding to others is associated with a variety of physical and mental health effects (Cacioppo and Cacioppo, 2012). For instance, people who subjectively feel isolated live shorter lives than those who feel they have strong, dependable, meaningful social bonds (Cacioppo and Patrick, 2008; Cacioppo and Cacioppo, 2012 for review).

As a consequence, there is a health-related need to better understand the functional dynamic of our brain during actions performed in an interpersonal context. This is critical as we spend much of our lifetime interacting with significant others, acquaintances, as well as strangers.

A growing body of research in psychology highlights the importance of studying the processing of significant others by demonstrating the influence of implicit processing of significant others (compared to strangers) on the individual's perception and cognitive processes. For instance, evidence suggests that the emotional bond between an actor and a perceiver may facilitate mutual intention perception, with a stronger bond associated with faster intention understanding (Cutting and Kozlowski, 1977; Ortigue and Bianchi-Demicheli, 2008; Ortigue et al., 2010a).

Inspired by the theories on embodied cognition and simulation theories, one explanation for this facilitation effect is that intention understanding may be based, in part, upon mechanisms of self-expansion among significant others. Through self-expansion mechanisms, a collective unconscious mental representation may be formed among individuals who share self-characteristics, values, and actions in a common environment (Agnew and Etcheverry, 2006). In line with evolutionary theory's claim that intense emotional experiences during a lifetime (e.g., passionate love) may be a central human motivation to expand one's self (Aron and Aron, 1996; Barkow et al., 1992), the self-expansion theory of pair-bonding is a hallmark in dyadic relationships (Aron and Aron, 1996). As an illustration, couples in love often refer to one another as the "better half" or the "completion of oneself," and they refer to "We" rather than "I" (Hatfield and Sprecher, 1986; Hatfield and Rapson, 1993; Aron and Aron, 1996), therefore suggesting that there is a cognitive expansion of their self in a beloved, and vice versa an integration of the beloved's values and characteristics in their self (Hatfield and Walster, 1978; Aron and Aron, 1996; Ortigue and Bianchi-Demicheli, 2008). From a cognitive and social viewpoint, self-expansion means that each partner makes a decision (conscious or not) to include the significant other in their own mental self-representation (Aron and Aron, 1996), allowing the formation of a shared mental representation of the self and partner. Recent neuroimaging studies of love (Cacioppo et al., 2012, for review) provide further support in favor of this self-expansion model of love by demonstrating a recruitment of self-related brain network in people who are in love (see Cacioppo et al., 2012, for review). Based on these recent findings in social neuroscience and relationship science (Ortigue et al., 2009, 2010a,b; Cacioppo et al., 2012; Canessa et al., 2012) and based on the self-expansion theory of pair-bonding (Aron and Aron, 1996; Bianchi-Demicheli et al., 2006; Ortigue and Bianchi-Demicheli, 2008; Ortigue et al., 2010a,b), we hypothesized a common pattern (notably within SN) of activation between love and intention tasks (Ortigue and Bianchi-Demicheli, 2008). The rationale for identifying areas of overlap between tasks that involve love and intention is that love varies as a function of the extent to which an individual prefers or desires interaction with another person. This preference or desire, in turn, may activate networks associated more with behavioral intentions in everyday life.

To test this hypothesis, we statistically explored the neural similarities and differences of the neural bases between intention and passionate love for a partner by performing a meta-analysis of fMRI studies involving intention understanding and love, respectively.

## MATERIALS AND METHODS

### LITERATURE SEARCH

We performed a systematic review of functional neuroimaging studies of intention understanding and passionate love, respectively. All papers and books in the literature published up to May 2011 (inclusive) were considered for this review, subject to two general limitations: the publication had to be a manuscript, chapter or book, and the title and abstract had to be available in English. Materials were identified through computer-based search, as described below.

### SELECTION CRITERIA FOR INTENTION UNDERSTANDING LITERATURE

Our systematic computer-based search was based on the published literature of functional neuroimaging studies on intention understanding using MEDLINE library through PubMed database. Key words used for this search were "intention understanding," "action understanding," "fMRI," and "neuroimaging." Publications were selected on the basis of the following criteria: (1) fMRI neuroimaging studies; (2) with healthy adult participants, and (3) paradigms included stories (text or cartoons) or video-clips on intention understanding only. In all selected studies, participants' instruction was either to observe intentions, to infer the intentions of the actions performed by others, or to answer questions about the intention of the actions (i.e., "why"). In all the studies we classified as "intention understanding" agents were strangers (unfamiliar people) only. A list of the studies and contrast conditions for intention understanding are shown in **Table A1**.

## SELECTION CRITERIA FOR PASSIONATE LOVE LITERATURE

We similarly performed a computer-based search of functional neuroimaging studies on passionate love using MEDLINE library through PubMed database. Expanding on our earlier study (Ortigue et al., 2010b), we used the key words "love," "couple," "fMRI," and "neuroimaging." Publications were selected on the basis of the following criteria: (1) fMRI neuroimaging studies; (2) with healthy adult participants in love with a partner, (3) paradigms included viewing partner's face or partner's name, i.e., tasks related to their beloved partner. A list of the studies and contrast conditions for passionate love are shown in **Table A2**.

## META-ANALYSIS METHODS

To provide readers with a synthesized and statistical view of the common and different brain networks mediating intention understanding and passionate love, we analyzed the distribution of peak coordinates related to intention understanding (**Figure 1**, blue), passionate love (red), and regions common to both (purple). Using techniques reported previously by our group (Lewis, 2006; Ortigue et al., 2010b; Cacioppo et al., 2012), we adopted

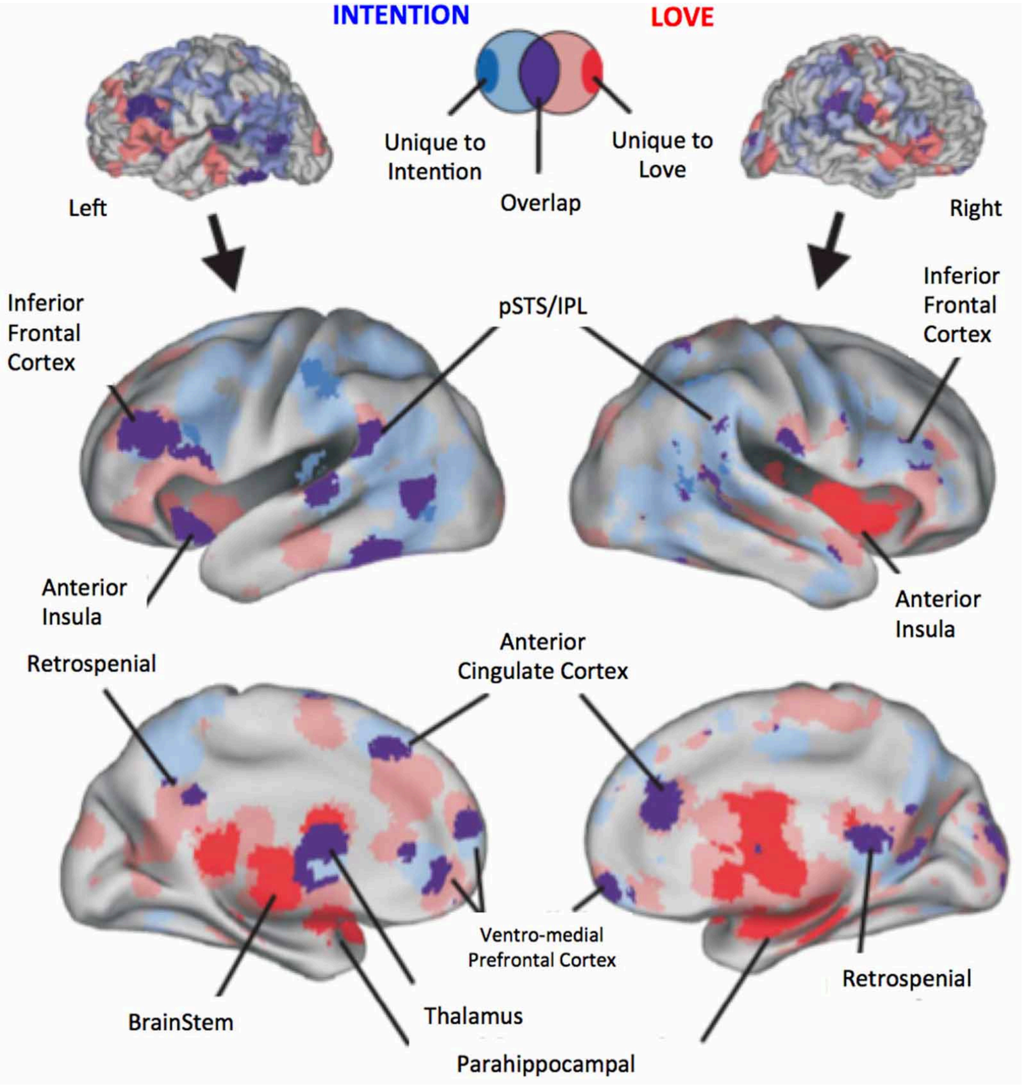


**FIGURE 1 | Meta-analysis results revealing brain regions and networks unique to either Intention Understanding (blue) and Passionate Love (red) and statistically significant overlap between the two (purple).** Upper panel illustrates a typical brain surface model. Lower panel illustrate a slightly inflated rendering of the PALS atlas cortical surface to facilitate visualization of the resulting activation foci. Brain activations significant at $p < 0.001$, FWER corrected. Variations from transparent to solid colors indicate the following terminology: Transparent red, all passionate love study only foci; Transparent blue, all intention study only foci; Purple, overlap of above conditions; Solid red, unique to passionate love only studies; Solid blue, unique to intention only studies.

a Multi-level Kernel Density Analysis (MKDA) approach (Wager et al., 2009). This approach quantitatively tested for consistency and specificity of regional activation across the two sets of studies: it minimizes biases such as having one study that reports many activation foci from dominating the meta-analysis, and it accounts for the smoothness of reported data, false-positive rates, and statistical power. Thus, the reported peak coordinates within a study's contrast maps are weighted in the meta-analysis by study quality and sample size. We calculated the number of statistical contrast maps that activated each voxel in the brain using 10 mm kernel (roughly matching the three dimensional spatial resolution of the reported data). Monte Carlo simulations (10,000 iterations) were used to obtain a threshold and establish statistical significance against a null hypothesis that activated regions in the resulting pair-wise contrast maps are not spatially consistent (i.e., that they are randomly distributed throughout the brain). The use of the distribution of maximum values provides a strong control of family wise error rate and is an established method for multiple comparisons correction (Nichols and Holmes, 2002). All voxels (which constitute the various brain region volumes) whose density exceeded the 99.9th percentile value under the null hypothesis were considered significant (i.e., Family Wise Error Rate corrected for spatial extent at $p < 0.001$). Brain coordinates obtained using the MKDA method above were entered into the SPM8 software program (Wellcome Trust Centre for Neuroimaging, London; http://www.fil.ion.ucl.ac.uk/spm/) using Anatomy Toolbox version 18 (Eickhoff et al., 2005) in order to facilitate identification and labeling of each activation peak. The MNI coordinates and volumes of significantly overlapping clusters (brain foci) were extracted using the AFNI software plug-in 3dcluster (Cox, 1996). Displayed localizations were further validated by visual inspection relative to the Duvernoy and Bourgoin brain atlas (Duvernoy and Bourgoin, 1999). To visualize the meta-analysis results (**Figure 1**), the significantly overlapping contrast indicator maps for the intention studies and the love studies, and the intention-overlap-love studies were projected (using MNI-Talairach coordinate space) onto the left and right hemisphere Population-Average, Landmark- and Surface-based atlas (PALS atlas), which is an atlas of cortical surfaces that represent the averaged cortical surfaces of 12 individuals (http://brainmap.wustl.edu; Van Essen, 2005). The left and right cortical surfaces were inflated to reveal major sulci of the brain to facilitate viewing of the data.

## RESULTS

Based on our search criteria, we found a total of 25 fMRI studies. This included 17 studies (21 experimental paradigms) for intention understanding (**Table A1**; Pelphrey et al., 2004; Walter et al., 2004; den Ouden et al., 2005; Iacoboni et al., 2005; Hamilton and Grafton, 2006, 2008; Wang et al., 2006; Brass et al., 2007; Buccino et al., 2007; Ciaramidaro et al., 2007; de Lange et al., 2008; Ortigue et al., 2009; Liew et al., 2010; Newman-Norlund et al., 2010; Ramsey and Hamilton, 2010; Carter et al., 2011; Jastorff et al., 2011) and eight studies (10 experimental paradigms) for passionate love (**Table A2**; Bartels and Zeki, 2000; Aron et al., 2005; Ortigue et al., 2007; Kim et al., 2009; Zeki and Romaya, 2010; Stoessel et al., 2011; Xu et al., 2011; Acevedo et al., 2012), involving a total of 457 participants. The number of participants included in each study ranged from 10 to 36 (for further details see **Tables A1**, **A2**).

Results confirm previous studies by demonstrating that understanding intentions of strangers involved the brain areas involved in both SN and AON, including areas sustaining embodied cognition, simulation, and self-other perception (such as vMPFC, BA6, MTG/STG, Angular gyrus, see **Table A3**; **Figure 1**, blue) as well as in the precuneus and the left anterior cingulate cortex, and the right middle and posterior cingulate cortex. In addition, activations in subcortical areas, such as putamen activity were also observed (see **Table A3** for further details). Similarly, results on passionate love activations reinforced previous studies (e.g., Acevedo et al., 2012 and Cacioppo et al., 2012) by demonstrating a distributed cortical and subcortical network of SN as well as brain areas known to be involved in partner preference, rewarding experiences, goal-directed actions, habit formation, and decision-making (for further details see **Table A4**).

### COMMON BRAIN NETWORK BETWEEN INTENTION UNDERSTANDING AND PASSIONATE LOVE

Our fMRI meta-analysis revealed a shared brain network between intention understanding and passionate love (**Figure 1**; purple; **Table A5**) that includes brain regions sustaining social cognition, embodied cognition, mentalizing about self-other, such as bilateral pSTS/IPL, bilateral inferior frontal cortex (IFC), ventromedial prefrontal cortex (vMPFC), anterior insula as well as brain regions involved in the mesolimbic and nigrostriatal dopaminergic pathways (caudate nucleus, thalamus, putamen, and parahippocampal area; **Table A5**).

## DISCUSSION

The present research highlights a shared network between love and intention, which includes (1) areas that overlap with areas related to dopamine circuits; and (2) several regions implicated in social cognition, embodied cognition, attachment, mental state representation, and self-representation. These results are consistent with previous studies indicating that both love and intention involve goal-directed and rewarding behaviors towards a specific partner. The recruitment of dorsal parts of the striatum, such as the caudate and putamen, which are innervated by dopamine coming from both the VTA and substantia nigra, is in line with recent work in animals showing that these brain areas are critical in the development of a pair bond and conditioned partner preference (Pfaus, 1999, 2009; Young et al., 2005, for review), as well as in the activation of stereotyped motor patterns (habits) related to conditioned incentive cues (Everitt and Robbins, 2005). Although dopamine transmission in the ventral striatum has been shown to be stimulated in the presence of both unconditioned and conditioned rewarding incentive cues stimuli (Pfaus et al., 2001; Aragona et al., 2003; Postuma and Dagher, 2006; Pfaus, 2009, for review), responses made to conditioned rewards appear to involve more dorsal striatal networks, among which are outputs of the dorsal striatum to cortical regions such as the insula (Postuma and Dagher, 2006)—a brain region which binds integrated visceral feedback with emotional and cognitive responses (Craig, 2002; Ibanez et al., 2010; Berntson et al., 2011; Cacioppo et al., 2012).

The overlap between passionate love and intention understanding in brain areas, such as the vMPFC, is consistent with a growing body of studies unraveling the recruitment of this brain area during tasks that require introspections about self and by tasks that require inferences about the minds of others perceived to be similar to self (Jenkins et al., 2008). The activation of the anterior insula is also consistent with earlier studies from our laboratory and others implicating a role in meta-representations of self (for further details see, Cacioppo et al., 2012). Other overlapping areas were also activated, such as left and right IFC and MTG/pSTS with a slight left hemisphere lateralization. These areas are in close proximity to classical Broca's and Wernicke's language regions, which entail frontal and temporo-parietal regions that are more highly interconnected via the arcuate fasciculus white matter fiber tracts in humans relative to chimpanzees and to monkeys (Rilling, 2008). One intriguing prospect is that the evolution of these interconnected inferior fronto-parietal regions may relate to pair-bonding and degree of social attachments, as well as other more uniquely human qualities such as language (Kim et al., 2010).

By identifying specific functional brain regions and networks in a large sample of healthy subjects, the present analysis reinforces the consistency and specificity of the brain regions that are being reported in the burgeoning body of studies on love and intention understanding. Interestingly, by revealing an additive brain network for both intention understanding and passionate love, the present findings offer a new way to look at the neurobiology of the loving mind during embodied cognition through the lens of a specific subset of AON and SN regions of interest. This provides scientists and clinicians with a unique and strong rationale to further investigate neurocognitive models that may explain the modulations of these common regions and brain networks in future studies on intention understanding in dyads.

One limitation of the present meta-analysis study was some of the fMRI studies of intention contrasted the condition of interest either to resting state(s) (e.g., Iacoboni et al., 2005; Wang et al., 2006; Newman-Norlund et al., 2010), while others used a variety of different control conditions. Consequently, the resulting activation patterns (**Figure 1**) must be interpreted with this caveat in mind. However, the goal of the present study was to provide a first pass glimpse at determining (and identifying) candidate brain regions or networks that may show preferential or selective activation to scenarios related to embodied cognition, pair-bonding, and self-expansion mechanisms. The present results do provide novel support for an overlapping brain network between these mechanisms. Future fMRI studies investigating the processing of significant others (compared to strangers) during intention understanding in the same paradigm is thus warranted/needed. The systematic neuroscientific study of the modulations of the neural network for understanding the intentions of significant others in healthy subjects, neurological patients after brain damage, and patients suffering from chronic social and affective disorders (such as autism), will provide critical insights on the spatio-temporal dynamics of self-expansion and embodied mechanisms that may mediate dyadic interactions, notably among couples.

## FINANCIAL DISCLOSURE/FUNDING

The authors disclosed receipt of the following financial support for the research and/or authorship of this article: Swiss National Science Foundation (Grant #PP00_1_128599/1 to Stephanie Cacioppo), the Mind Science Foundation (Grant #TSA2010-2 to Stephanie Cacioppo, Francesco Bianchi-Demicheli), NCRR NIH COBRE Grant E15524 (Grant #E15524 to the Sensory Neuroscience Research Center of West Virginia University to James W. Lewis), the MOE Project of Key Research Institute of Humanities and Social Sciences (Grant #12JJD190004 to Yi-Wen Wang), and the Program for New Century Excellent Talents in Universities (Grant #NCET-11-1065 to Yi-Wen Wang).

## REFERENCES

Acevedo, B. P., Aron, A., Fisher, H. E., and Brown, L. L. (2012). Neural correlates of long-term intense romantic love. *Soc. Cogn. Affect. Neurosci.* 7, 145–159.

Agnew, C. R., and Etcheverry, P. E. (2006). "Cognitive interdependence: considering self-in-relationship," in *Self and Relationships: Connecting Intrapersonal and Interpersonal Processes,* eds K. D. Vohs and E. J. Finkel (New York, NY: Guilford), 274–293.

Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278.

Aragona, B. J., Liu, Y., Curtis, J. T., Stephan, F. K., and Wang, Z. (2003). A critical role for nucleus accumbens dopamine in partner-preference formation in male prairie voles. *J. Neurosci.* 23, 3483–3490.

Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., and Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *J. Neurophysiol.* 94, 327–337.

Aron, E. N., and Aron, A. (1996). Love and expansion of the self: the state of the model. *Pers. Relat.* 3, 45–58.

Barkow, J. H., Cosmides, L., and Tooby, J. (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* New York, NY: Oxford University Press.

Bartels, A., and Zeki, S. (2000). The neural basis of romantic love. *Neuroreport* 11, 3829–3834.

Becchio, C., Manera, V., Sartori, L., Cavallo, A., and Castiello, U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Front. Hum. Neurosci.* 6:117. doi: 10.3389/fnhum.2012.00117

Berntson, G. G., Norman, G. J., Bechara, A., Bruss, J., Tranel, D., and Cacioppo, J. T. (2011). The insula and evaluative processes. *Psychol. Sci.* 22, 80–86.

Bianchi-Demicheli, F., Grafton, S. T., and Ortigue, S. (2006). The power of love on the human brain. *Soc. Neurosci.* 1, 90–103.

Bidet-Caulet, A., Voisin, J., Bertrand, O., and Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage* 28, 132–139.

Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Curr. Biol.* 17, 2117–2121.

Buccino, G., Baumgaertner, A., Colle, L., Buechel, C., Rizzolatti, G., and Binkofski, F. (2007). The neural basis for understanding non-intended actions. *Neuroimage* 36, 119–127.

Cacioppo, J. T., and Patrick, B. (2008). *Loneliness: Human Nature and the Need for Social Connection*. New York, NY: W. W. Norton and Company.

Cacioppo, S., Bianchi-Demicheli, F., Frum, C., Pfaus, J. G., and Lewis, J. W. (2012). The common neural bases between sexual desire and love: a multilevel kernel density fMRI analysis. *J. Sex. Med.* 9, 1048–1054.

Cacioppo, S., and Cacioppo, J. T. (2012). Decoding the invisible forces of social connections. *Front. Integr. Neurosci.* 6:51. doi: 10.3389/fnint.2012.00051

Canessa, N., Alemanno, F., Riva, F., Zani, A., Proverbio, A. M., Mannara, N., et al. (2012). The neural bases of social intention understanding: the role of interaction goals. *PLoS ONE* 7:e42347. doi: 10.1371/journal.pone.0042347

Carter, E. J., Hodgins, J. K., and Rakison, D. H. (2011). Exploring the neural correlates of goal-directed action and intention understanding. *Neuroimage* 54, 1634–1642.

Choi, J. C., Park, S. K., Kim, Y.-H., Shin, Y. W., Kwon, J. S., Kim, J. S., et al. (2006). Different brain activation patterns to pain and pain-related unpleasantness during the menstrual cycle. *Anesthesiology* 105, 120–127.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., et al. (2007). The intentional network: how the brain reads varieties of intentions. *Neuropsychologia* 45, 3105–3113.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.

Cutting, J. E., and Kozlowski, L. T. (1977). Recognizing friends by their walk: gait perception without familiarity cues. *Bull. Psychon. Soc.* 9, 353–356.

de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Curr. Biol.* 18, 454–457.

den Ouden, H. E., Frith, U., Frith, C., and Blakemore, S. J. (2005). Thinking about intentions. *Neuroimage* 28, 787–796.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811–813.

Duvernoy, H. M., and Bourgoin, P. (1999). *The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply*. Wien: Springer.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.

Eickhoff, S. B., Schleicher, A., Zilles, K., and Amunts, K. (2006). The human parietal operculum. I. Cytoarchitectonic mapping of subdivisions. *Cereb. Cortex* 16, 254–267.

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Frith, C. D., and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.* 16, 1824–1829.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Ann. N.Y. Acad. Sci.* 1156, 97–117.

Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. *Exp. Brain Res.* 112, 103–111.

Grezes, J., Frith, C., and Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *J. Neurosci.* 24, 5500–5505.

Grossman, E. D., Battelli, L., and Pascual-Leone, A. (2005). Repetitive TMS over posterior STS disrupts perception of biological motion. *Vision Res.* 45, 2847–2853.

Hamilton, A. C., and Grafton, S. (2006). Goal representation in human anterior intraparietal sulcus. *J. Neurosci.* 26, 1133–1137.

Hamilton, A. F., and Grafton, S. T. (2008). Action outcomes are represented in human inferior frontoparietal cortex. *Cereb. Cortex* 18, 1160–1168.

Hatfield, E., and Rapson, R. L. (1993). Historical and cross-cultural perspectives on passionate love and sexual desire. *Annu. Rev. Sex Res.* 4, 67–97.

Hatfield, E., and Sprecher, S. (1986). Measuring passionate love in intimate relationships. *J. Adolesc.* 9, 383–410.

Hatfield, E., and Walster, G. W. (1978). *A New Look at Love.* Lanham, MD: University Press of America.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3:e79. doi: 10.1371/journal.pbio.0030079

Ibanez, A., Gleichgerrcht, E., and Manes, F. (2010). Clinical effects of insular damage in humans. *Brain Struct. Funct.* 214, 397–410.

Jastorff, J., Clavagnier, S., Gergely, G., and Orban, G. A. (2011). Neural mechanisms of understanding rational actions: middle temporal gyrus activation by contextual violation. *Cereb. Cortex* 21, 318–329.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109.

Jellema, T., Baker, C. I., Wicker, B., and Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain Cogn.* 44, 280–302.

Jenkins, A. C., Macrae, C. N., and Mitchel, J. P. (2008). Repetition suppression of ventromedial prefrontal activity during judgements of self and others. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4507–4512.

Kim, P., Leckman, J. F., Mayes, L. C., Newman, M. A., Feldman, R., and Swain, J. E. (2010). Perceived quality of maternal care in childhood and structure and function of mothers' brain. *Dev. Sci.* 13, 662–673.

Kim, W., Kim, S., Jeong, J., Lee, K.-U., Ahn, K.-J., Chung, Y.-A., et al. (2009). Temporal changes in functional magnetic resonance imaging activation of heterosexual couples for visual stimuli of loved partners. *Psychiatry Investig.* 6, 19–25.

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., and Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct. Funct.* 214, 519–534.

Lewis, J. W. (2006). Cortical networks related to human use of tools. *Neuroscientist* 12, 211–231.

Lewis, J. W., Phinney, R. E., Brefczynski, J. A., and DeYoe, E. A. (2006). Lefties get it "right" when hearing tool sounds. *J. Cogn. Neurosci.* 18, 1314–1330.

Lewis, J. W., Talkington, W. J., Puce, A., Engel, L. R., and Frum, C. (2011). Cortical networks representing object categories and high-level attributes of familiar real-world action sounds. *J. Cogn. Neurosci.* 23, 2079–2101.

Liew, S. L., Han, S., and Aziz-Zadeh, L. (2010). Familiarity modulates mirror neuron and mentalizing regions during intention understanding. *Hum. Brain Mapp.* 32, 1986–1997.

Materna, S., Dicke, P. W., and Thier, P. (2008). The posterior superior temporal sulcus is involved in social communication not specific for the eyes. *Neuropsychologia* 46, 2759–2765.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., and Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage* 13, 684–701.

Newman-Norlund, R., van Schie, H. T., van Hoek, M. E. C., Cuijpers, R. H., and Bekkering, H. (2010). The role of inferior frontal and parietal areas in differentiating meaningful and meaningless object-directed actions. *Brain Res.* 1315, 63–74.

Nichols, T. E., and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25.

Niedenthal, P. M. (2007). Embodying emotion. *Science* 316, 1002–1005.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., and Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Pers. Soc. Psychol. Rev.* 9, 184–211.

Ortigue, S., and Bianchi-Demicheli, F. (2008). Why is your spouse so predictable? Connecting mirror neuron system and self-expansion model of love. *Med. Hypotheses* 71, 941–944.

Ortigue, S., Bianchi-Demicheli, F., Hamilton, A. F., and Grafton, S. T. (2007). The neural basis of love as a subliminal prime: an event-related functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 19, 1218–1230.

Ortigue, S., Patel, N., Bianchi-Demicheli, F., and Grafton, S. T. (2010a). Implicit priming of embodied cognition on human motor intention understanding in dyads in love. *J. Soc. Pers. Relat.* 27, 1001–1015.

Ortigue, S., Bianchi-Demicheli, F., Patel, N., Frum, C., and Lewis, J. W. (2010b). Neuroimaging of love: fMRI meta-analysis evidence toward new perspectives in sexual medicine. *J. Sex. Med.* 7, 3541–3552.

Ortigue, S., Thompson, J. C., Parasuraman, R., and Grafton, S. T. (2009). Spatio-temporal dynamics of human intention understanding in temporo-parietal cortex: a combined EEG/fMRI repetition suppression paradigm. *PLoS ONE* 4:e6962. doi: 10.1371/journal.pone.0006962

Pelphrey, K., Morris, J. P., and McCarthy, G. (2004). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *J. Cogn. Neurosci.* 16, 1706–1716.

Pelphrey, K. A., and Morris, J. P. (2006). Brain mechanisms for interpreting the actions of others from biological-motion cues. *Curr. Dir. Psychol. Sci.* 15, 136–140.

Pfaus, J. G. (1999). Neurobiology of sexual behavior. *Curr. Opin. Neurobiol.* 9, 751–758.

Pfaus, J. G. (2009). Pathways of sexual desire. *J. Sex. Med.* 6, 1506–1533.

Pfaus, J. G., Kippin, T. E., and Centeno, S. (2001). Conditioning and sexual behavior: a review. *Horm. Behav.* 40, 291–321.

Postuma, R. B., and Dagher, A. (2006). Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cereb. Cortex* 16, 1508–1521.

Ramsey, R., and Hamilton, A. F. (2010). Understanding actors and object-goals in the human brain. *Neuroimage* 50, 1142–1147.

Rilling, J. K. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* 11, 426–428.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670.

Rizzolatti, G., and Sinigaglia, C. (2007). Mirror neurons and motor intentionality. *Funct. Neurol.* 22, 205–210.

Rizzolatti, G., and Sinigaglia, C. (2008). Further reflections on how we interpret the actions of others. *Nature* 455, 589.

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–274.

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., and Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 42, 1435–1446.

Schultz, J., Imamizu, H., Kawato, M., and Frith, C. D. (2004). Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *J. Cogn. Neurosci.* 16, 1695–1705.

Stoessel, C., Stiller, J., Bleich, S., Boensch, D., Doerfler, A., Garcia, M., et al. (2011). Differences and similarities on neuronal activities of people being happily and unhappily in love: a functional magnetic resonance imaging study. *Neuropsychobiology* 64, 52–60.

Sugiura, M., Wakusawa, K., Sekiguchi, A., Sassa, Y., Jeong, H., Horie, K., et al. (2009). Extraction of situational meaning by integrating. multiple meanings in a complex environment: a functional MRI study. *Hum. Brain Mapp.* 30, 2676–2688.

Thompson, J. C., Clarke, M., Stewart, T., and Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *J. Neurosci.* 25, 9059–9066.

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., and Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage* 37, 966–973.

Van Essen, D. C. (2005). A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* 28, 635–662.

Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584.

Wager, T. D., Lindquist, M. A., Nichols, T. E., Kober, H., and Van Snellenberg, J. X. (2009). Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *Neuroimage* 45(1 Suppl.), S210–S221.

Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., et al. (2009). Neural correlates of processing situational relationships between a part and the whole: an fMRI study. *Neuroimage* 48, 486–496.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., and Bara, B. G. (2004). Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *J. Cogn. Neurosci.* 16, 1854–1863.

Wang, A. T., Lee, S. S., Sigman, M., and Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Soc. Cogn. Affect. Neurosci.* 1, 107–121.

Wheatley, T., Milleville, S. C., and Martin, A. (2007). Understanding animate agents: distinct roles for the social network and mirror system. *Psychol. Sci.* 18, 469–474.

Xu, X., Aron, A., Brown, L., Cao, G., Feng, T., and Weng, X. (2011). Reward and motivation systems: a brain mapping study of early-stage intense romantic love in Chinese participants. *Hum. Brain Mapp.* 32, 249–257.

Young, L. J., Murphy Young, A. Z., and Hammock, E. A. D. (2005). Anatomy and neurochemistry of the pair bond. *J. Comp. Neurol.* 493, 51–57.

Zeki, S., and Romaya, J. P. (2010). The brain reaction to viewing faces of opposite- and same-sex romantic partners. *PLoS ONE* 5:e15802. doi: 10.1371/journal.pone.0015802

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 15 December 2012; accepted: 08 March 2013; published online: 27 March 2013.*

*Citation: Juan E, Frum C, Bianchi-Demicheli F, Wang Y-W, Lewis JW and Cacioppo S (2013) Beyond human intentions and emotions. Front. Hum. Neurosci. 7:99. doi: 10.3389/fnhum.2013.00099*

## APPENDIX

**Table A1 | List of fMRI studies on intention understanding.**

| Study (first author listed) | Year | No of subjects | No of women | No of right-handed | Stimuli | Experimental comparisons |
|---|---|---|---|---|---|---|
| Brass | 2007 | 15 | 7 | 15 | Video-clips | Implausible > ordinary motion |
| Buccino | 2007 | 20 | 10 | 20 | Video-clips | Attend to intention > viewing |
| | | | | | Video-clips | Incorrect unintended motion > ordinary motion |
| Carter | 2011 | 17 | 8 | 15 | Video-clips | Human reaching movements: goal shift > goal miss |
| Ciaramidaro | 2007 | 12 | 6 | 12 | Comic strips | Intentional (communicative) > physical event |
| | | | | | Comic strips | Intentional (social) > physical event |
| | | | | | Comic strips | Intentional (private) > physical event |
| de Lange | 2008 | 19 | 10 | 19 | Pictures | Ordinary > extraordinary intentions |
| den Ouden | 2005 | 11 | 11 | Not specified | Scenarios | Intentional action > physical event |
| Hamilton | 2006 | 20 | 11 | 19 | Video-clips | Repetition suppression for action goal |
| Hamilton | 2008 | 20 | 13 | 20 | Video-clips | Repetition suppression for action outcome |
| Iacoboni | 2005 | 23 | 15 | 23 | Video-clips | Hand grasping > rest |
| Jastorff | 2011 | 15 | 9 | 15 | Video-clips | Non-rational > rational movements |
| Liew | 2010 | 18 | 8 | 18 | Video-clips | Hand gestures > still control |
| Newman-Norlund | 2010 | 18 | 10 | 18 | Video-clips | Object-directed actions > rest |
| Ortigue | 2009 | 24 | 0 | 24 | Video-clips | Repetition suppression for intention |
| Pelphrey | 2004 | 12 | 7 | 12 | Video-clips | Incorrect > correct motion |
| Ramsey | 2010 | 25 | 17 | 24 | Video-clips | Repetition suppression for object-goal |
| Walter (1) | 2004 | 13 | 7 | 13 | Comic strips | Intentional action > physical event |
| Walter (2) | 2004 | 12 | 6 | 12 | Comic strips | Intentional action > physical event |
| Wang | 2006 | 12 | 6 | 12 | Scenarios | Attend to face > rest |
| Total | | 306 | 161 | 291 | | |

**Table A2 | List of fMRI studies on passionate love.**

| Study (first author listed) | Year | No of subjects | No of women | No of right-handed | Stimuli | Experimental comparisons |
|---|---|---|---|---|---|---|
| Acevedo | 2012 | 17 | 10 | 17 | Faces | Partner > high familiar acquaintance |
| | | | | | Faces | Partner > close friend |
| Aron | 2005 | 17 | 10 | 17 | Faces | Beloved > familiar neutral acquaintance |
| Bartels | 2000 | 17 | 11 | 16 | Faces | Beloved > friend |
| Kim | 2009 | 10 | 5 | 10 | Faces | Beloved > friend (early) |
| | | | | | Faces | Beloved > friend (late) |
| Ortigue | 2007 | 36 | 36 | 36 | Names | Beloved > friend or stranger |
| Stoessel | 2011 | 12 | 6 | 12 | Pictures | Beloved > erotic (happy condition) |
| Xu | 2011 | 18 | 10 | 18 | Faces | Beloved > familiar neutral acquaintance |
| Zeki | 2010 | 24 | Not specified | Not specified | Faces | Beloved > neutral |
| Total | | 151 | 88 | 126 | | |

**Table A3 | Intention understanding activation peaks.**

| Label | Left hemisphere | | | | Right hemisphere | | | |
|---|---|---|---|---|---|---|---|---|
| | MNI coordinates | | | Number of voxels | MNI coordinates | | | Number of voxels |
| | *x* | *y* | *z* | | *x* | *y* | *z* | |
| vmPFC | | | | | 6 | 52 | −12 | 4000 |
| Inferior frontal gyrus (p. Opercularis) (BA 44: 20%) | | | | | 46 | 10 | 27 | 17904 |
| Superior frontal gyrus (BA 6: 50%) | −18 | 2 | 67 | 7048 | | | | |
| SMA (BA 6: 20%) | −6 | 20 | 51 | 6648 | | | | |
| BA 6 (20%) | | | | | 20 | −5 | 48 | 40 |
| Precentral gyrus (BA 6: 60%) | | | | | 36 | −12 | 58 | 4120 |
| Precentral gyrus (BA 44: 30%) | −46 | 7 | 33 | 29,424 | | | | |
| BA 45 (10%) | | | | | 37 | 33 | 0 | 3184 |
| Middle temporal gyrus/STG | −47 | −50 | 18 | 73,992 | | | | |
| Heschls gyrus (Insula/Ig1: 60%; TE1.1: 50%; OP2: 30%) | −33 | −28 | 13 | 464 | | | | |
| Gyrus ambiens | | | | | 46 | −6 | −23 | 7688 |
| Angular gyrus (IPC/PFm: 20%; PGa: 20%; hIP1: 10%) | | | | | 47 | −46 | 24 | 73,608 |
| Precuneus (SPL/7P: 20%; SPL/7A: 20%) | −2 | −55 | 47 | 17,744 | | | | |
| Anterior cingulate cortex | −3 | 49 | 11 | 10,424 | | | | |
| | | | | | 7 | 41 | 29 | 7928 |
| Middle cingulate cortex (BA 6: 10%) | | | | | 10 | −6 | 44 | 2864 |
| Posterior cingulate cortex | | | | | 12 | −41 | 10 | 7736 |
| Putamen | −26 | 10 | −6 | 4120 | | | | |
| Fornix | −4 | −2 | 10 | 4120 | | | | |
| Calcarine gyrus (BA 17: 50%; BA 18: 10%) | | | | | 16 | −82 | 14 | 4088 |
| Calcarine gyrus (BA 17: 70%; BA 18: 20%) | | | | | 23 | −99 | −2 | 3424 |
| Middle occipital gyrus (IPC/PGp: 30%) | −35 | −84 | 29 | 3024 | | | | |
| Optic radiations | −26 | −34 | 10 | 2336 | | | | |

**Table A4 | Passionate love activations peaks.**

| **Label** | **Left hemisphere** | | | | **Right hemisphere** | | | |
|---|---|---|---|---|---|---|---|---|
| | **MNI coordinates** | | | **Number of voxels** | **MNI coordinates** | | | **Number of voxels** |
| | ***x*** | ***y*** | ***z*** | | ***x*** | ***y*** | ***z*** | |
| vmPFC | | | | | 10 | 39 | −11 | 2744 |
| | −5 | 49 | −8 | 9384 | | | | |
| dlPFC | −7 | 51 | 22 | 5056 | | | | |
| SMA (BA 6: 60%; BA 4: 10%) | | | | | 10 | −18 | 61 | 24,760 |
| Postcentral gyrus (BA 3b: 60%; BA 1: 10%; BA 6: 10%) | | | | | 61 | −3 | 24 | 6784 |
| Superior temporal gyrus (TE3: 10%) | | | | | 63 | −14 | −6 | 3656 |
| Superior temporal gyrus (IPC/PFcm: 10%; OP1: 10%; TE1.1: 10%) | −56 | −28 | 8 | 4120 | | | | |
| Posterior middle temporal gyrus (V5: 10%) | −55 | −67 | 2 | 3312 | | | | |
| Inferior temporal gyrus | −40 | −8 | −28 | 4120 | | | | |
| Middle cingulate cortex | | | | | 10 | 5 | 32 | 16 |
| Supramarginal gyrus (IPC/PFm: 70%; IPC/PF: 30%; IPC/PGa: 20%) | | | | | 62 | −45 | 25 | 7040 |
| Supramarginal gyrus (IPC/PF: 60%; IPC/PFm: 40%; IPC/PFcm: 10%) | −63 | −46 | 25 | 4096 | | | | |
| Superior parietal lobule (SPL/7A: 40%; BA2: 30%; SPL/5L: 30%) | −23 | −50 | 53 | 3128 | | | | |
| PGa (10%) | | | | | 40 | −48 | 15 | 16 |
| Calcarine gyrus (BA 17: 100%; BA 18: 30%) | | | | | 9 | −90 | 1 | 22,440 |
| Cuneus | | | | | 18 | −70 | 24 | 6192 |
| Thalamus | −3 | −12 | 1 | | | | | 284,552 |
| Cerebellum | | | | | 42 | −57 | −28 | 7512 |
| | | | | | 34 | −45 | −48 | 6656 |
| | −36 | −78 | −40 | 3896 | | | | |
| **PASSIONATE LOVE ACTIVATIONS (USING SPHERICAL KERNEL OF 3 mm)** | | | | | | | | |
| Superior frontal gyrus | −24 | 12 | 66 | 648 | | | | |
| Posterior superior temporal gyrus (IPC: 50%) | −50 | −38 | 18 | 600 | | | | |
| Superior medial frontal gyrus | −9 | 27 | 59 | 1288 | | | | |
| | −10 | 52 | 16 | 632 | | | | |
| Middle orbital gyrus/orbitofrontal | | | | | 2 | 62 | −8 | 648 |
| Gyrus rectus (orbitofrontal) | −4 | 60 | −24 | 648 | | | | |
| Middle frontal gyrus | −30 | 54 | 2 | 648 | | | | |
| Superior anterior medial frontal gyrus | | | | | 8 | 54 | 22 | 1272 |
| | | | | | 11 | 65 | 5 | 1264 |

*(Continued)*

**Table A4 | Continued**

| **Label** | **Left hemisphere** | | | | **Right hemisphere** | | | |
|---|---|---|---|---|---|---|---|---|
| | **MNI coordinates** | | | **Number of voxels** | **MNI coordinates** | | | **Number of voxels** |
| | ***x*** | ***y*** | ***z*** | | ***x*** | ***y*** | ***z*** | |
| Inferior frontal gyrus (p. orbitalis) | | | | | 35 | 33 | −8 | 1200 |
| Inferior frontal gyrus (p. triangularis) | −37 | 28 | 0 | 16 | | | | |
| | −37 | 26 | −2 | 16 | | | | |
| | −33 | 35 | −2 | 824 | | | | |
| | | | | | 56 | 28 | 8 | 648 |
| Middle frontal gyrus | | | | | 30 | 47 | 4 | 1200 |
| | −30 | 44 | 11 | 344 | | | | |
| Anterior cingulate cortex | | | | | 8 | 40 | 9 | 1152 |
| | | | | | 8 | 28 | 16 | 648 |
| | −6 | 42 | −6 | 600 | | | | |
| Middle cingulate cortex | | | | | 4 | −22 | 42 | 648 |
| | | | | | 12 | −22 | 34 | 616 |
| | | | | | 14 | −20 | 43 | 40 |
| Posterior cingulate cortex | −1 | −32 | 26 | 2440 | | | | |
| Inferior temporal gyrus/fusiform area | | | | | 42 | −52 | −12 | 648 |
| Superior temporal gyrus | −50 | −28 | 2 | 648 | | | | |
| Temporal pole | | | | | 40 | 8 | −28 | 648 |
| | −34 | 18 | −22 | 648 | | | | |
| Caudate nucleus | | | | | 54 | −32 | 8 | 648 |
| | −8 | 20 | 0 | 648 | | | | |
| | | | | | 17 | 1 | 23 | 1024 |
| | | | | | 10 | 16 | 5 | 1136 |
| Insula | | | | | 39 | −6 | −7 | 2632 |
| | −42 | −9 | −6 | 2112 | | | | |
| | −36 | 18 | −4 | 648 | | | | |
| | −34 | −22 | 10 | 608 | | | | |
| SMA | −1 | 10 | 65 | 1296 | | | | |
| | | | | | 4 | 4 | 52 | 648 |
| Thalamus | −8 | −9 | −2 | 1288 | | | | |
| | | | | | 10 | −24 | 10 | 648 |
| Heschel gyrus | | | | | 45 | −17 | 8 | 1280 |

*(Continued)*

**Table A4 | Continued**

| Label | Left hemisphere | | | | Right hemisphere | | | |
|---|---|---|---|---|---|---|---|---|
| | MNI coordinates | | | Number of voxels | MNI coordinates | | | Number of voxels |
| | *x* | *y* | *z* | | *x* | *y* | *z* | |
| Cerebellum | | | | | 10 | −44 | −7 | 1168 |
| Cerebellar vermis | −2 | −58 | −10 | 648 | | | | |
| Putamen | −22 | 7 | −8 | 1024 | | | | |
| | −18 | 19 | −9 | 896 | | | | |
| Precuneus | | | | | 8 | −51 | 22 | 976 |
| Parahippocampal region (Amygdala: 10%) | | | | | 19 | 2 | −13 | 896 |
| Pallidum | | | | | 22 | −1 | 6 | 888 |
| Rolandic operculum | | | | | 60 | −14 | 12 | 648 |
| Postcentral gyrus | −62 | −4 | 22 | 600 | | | | |
| Lingual gyrus (Hippocampus: 40%) | −11 | −38 | −10 | 424 | | | | |

**Table A5 | Common activation peaks between intention understanding and passionate love studies.**

| Label | Left hemisphere | | | | Right hemisphere | | | |
|---|---|---|---|---|---|---|---|---|
| | MNI coordinates | | | Number of voxels | MNI coordinates | | | Number of voxels |
| | *x* | *y* | *z* | | *x* | *y* | *z* | |
| dlPFC | | | | | 2 | 41 | 36 | 32 |
| | −5 | 54 | 19 | 2000 | | | | |
| vmPFC | | | | | 9 | 45 | −11 | 280 |
| | | | | | 4 | 55 | −11 | 1904 |
| Superior frontal gyrus (BA 6: 40%) | | | | | 30 | −12 | 62 | 424 |
| Inferior frontal gyrus (p. Orbitalis) | | | | | 39 | 31 | −4 | 824 |
| Inferior frontal gyrus (p. Triangularis) (BA 45: 60%) | | | | | 50 | 30 | 16 | 632 |
| Inferior frontal gyrus (p. Triangularis) (BA 45: 60%; BA 44: 20%) | −46 | 24 | 21 | 5232 | | | | |
| Inferior frontal gyrus (p. Opercularis) | | | | | 39 | 18 | 13 | 232 |
| Inferior frontal gyrus (p. Opercularis) (BA 44: 40%) | | | | | 56 | 9 | 21 | 272 |
| Postcentral gyrus (IPC/PFop: 30%; OP4: 30%; OP3: 10%) | | | | | 62 | −16 | 23 | 1248 |
| Postcentral gyrus (BA 2: 80%; SPL/7PC: 20%; BA 3b: 10%) | | | | | 28 | −42 | 57 | 1104 |
| SMA (BA 6: 70%) | | | | | 4 | −5 | 48 | 48 |
| SMA | −6 | 21 | 47 | 2024 | | | | |
| SMA (BA 6: 60%) | −10 | −5 | 69 | 80 | | | | |
| Inferior temporal gyrus | −49 | −51 | −20 | 3032 | | | | |
| Middle temporal gyrus (V5: 10%) | −53 | −68 | 2 | 2312 | | | | |
| Superior temporal gyrus (OP1: 10%) | −56 | −30 | 6 | 2424 | | | | |
| Superior temporal gyrus (IPC/PF: 60%; IPC/PFm: 30%; PGa: 20%) | −60 | −46 | 23 | 1648 | | | | |
| Supra marginal gyrus (IPC/PFm: 70%; IPC/PF: 30%; IPC/PGa: 30%) | | | | | 61 | −45 | 24 | 5800 |
| Anterior cingulate cortex | −7 | 40 | −2 | 744 | | | | |
| | −7 | 30 | 6 | 224 | | | | |
| | | | | | 6 | 34 | 22 | 3696 |
| Posterior cingulate cortex | | | | | 7 | −40 | 14 | 3024 |
| | −2 | −53 | 29 | 1088 | | | | |
| Superior parietal lobule (SPL/7PC: 30%; SPL/7A: 20%; hIP3: 10%) | −26 | −53 | 53 | 1096 | | | | |
| Lingual gyrus (BA 17: 50%; BA 18: 10%) | | | | | 29 | −59 | 3 | 904 |
| Inferior temporal gyrus | | | | | 46 | −56 | −23 | 776 |
| Cuneus (BA 18: 10%; BA 17: 10%) | | | | | 16 | −77 | 18 | 672 |

*(Continued)*

**Table A5 | Continued**

| Label | Left hemisphere | | | | Right hemisphere | | | |
|---|---|---|---|---|---|---|---|---|
| | MNI coordinates | | | Number of voxels | MNI coordinates | | | Number of voxels |
| | *x* | *y* | *z* | | *x* | *y* | *z* | |
| Precuneus | | | | | 25 | −45 | 8 | 568 |
| Precuneus (SPL/5L: 40%; SPL/7a: 20%; SPL/5M: 10%) | −14 | −50 | 57 | 16 | | | | |
| Anterior insula (Id1: 40%) | | | | | 43 | −4 | −14 | 272 |
| Anterior insula | | | | | 35 | 28 | 5 | 256 |
| Hippocampus | | | | | 37 | −14 | −19 | 1224 |
| Putamen | −27 | 10 | −6 | 3488 | | | | |
| Fornix | −3 | −3 | 11 | 3368 | | | | |
| Calcarine gyrus (BA 17: 50%; BA 18: 10%) | | | | | 14 | −87 | 14 | 1424 |
| Calcarine gyrus (BA 17: 80%; BA 18: 10%) | | | | | 21 | −95 | −2 | 896 |
| Optic radiations | −26 | −36 | 11 | 1608 | | | | |
| | −33 | −54 | −2 | 56 | | | | |

*Note: Activations of 10 voxels or above are reported here. All brain areas reported are significant at $p < 0.001$, corrected. Abbreviations: dlPFC, dorsolateral Prefrontal Cortex; vmPFC, ventro-medial Prefrontal Cortex; SMA, Supplementary Motor Area; STG, Superior Temporal Gyrus; IPC/PF, part of Inferior Parietal Cortex; IPC/PFm, part of Inferior Parietal Cortex; IPC/PGp, part of Inferior Parietal Cortex; PGa, part of Parietal area; hIP1, human Intraparietal area 1 (Choi et al., 2006); hIP3, human Intraparietal area 1 (Choi et al., 2006); Insula/lg1 (Insular Lobe granular area; Kurth et al., 2010), 60%; TE1.1, part of the Primary Auditory Cortex (Morosan et al., 2001); OP1, part of Parietal Operculum (Eickhoff et al., 2006); OP2, part of Parietal Operculum (Eickhoff et al., 2006); SPL/7PC, part of Superior Parietal Lobule; SPL/7A, part of Superior Parietal Lobule; SPL/5M, part of Superior Parietal Lobule (Duvernoy and Bourgoin, 1999).*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 23 July 2013
doi: 10.3389/fnhum.2013.00378

# Your emotion or mine: labeling feelings alters emotional face perception—an ERP study on automatic and intentional affect labeling

**Cornelia Herbert[1]*†, Anca Sfärlea[1] and Terry Blumenthal[2]**

[1] *Department of Psychology, University of Würzburg, Würzburg, Germany*
[2] *Department of Psychology, Wake Forest University, Winston-Salem, NC, USA*

**Edited by:**
*Maria Ruz, Universidad de Granada, Spain*

**Reviewed by:**
*Stephan Moratti, Universidad Complutense de Madrid, Spain*
*Mareike Bayer, Georg-August-Universität Göttingen, Germany*

***Correspondence:**
*Cornelia Herbert, Department of Psychology, University of Würzburg, 97070 Würzburg, Germany*
*e-mail: c.herbert@dshs-koeln.de*

**†Present address:**
*Cornelia Herbert, Institute of Psychology, German Sport University Cologne, 50933 Cologne, Germany*

Empirical evidence suggests that words are powerful regulators of emotion processing. Although a number of studies have used words as contextual cues for emotion processing, the role of *what* is being labeled by the words (i.e., one's own emotion as compared to the emotion expressed by the sender) is poorly understood. The present study reports results from two experiments which used ERP methodology to evaluate the impact of emotional faces and self- vs. sender-related emotional pronoun-noun pairs (e.g., my fear vs. his fear) as cues for emotional face processing. The influence of self- and sender-related cues on the processing of fearful, angry and happy faces was investigated in two contexts: an automatic (experiment 1) and intentional affect labeling task (experiment 2), along with control conditions of passive face processing. ERP patterns varied as a function of the label's reference (self vs. sender) and the intentionality of the labeling task (experiment 1 vs. experiment 2). In experiment 1, self-related labels increased the motivational relevance of the emotional faces in the time-window of the EPN component. Processing of sender-related labels improved emotion recognition specifically for fearful faces in the N170 time-window. Spontaneous processing of affective labels modulated later stages of face processing as well. Amplitudes of the late positive potential (LPP) were reduced for fearful, happy, and angry faces relative to the control condition of passive viewing. During intentional regulation (experiment 2) amplitudes of the LPP were enhanced for emotional faces when subjects used the self-related emotion labels to label their own emotion during face processing, and they rated the faces as higher in arousal than the emotional faces that had been presented in the "label sender's emotion" condition or the passive viewing condition. The present results argue in favor of a differentiated view of language-as-context for emotion processing.

**Keywords: emotion regulation, language-as-context, affect labeling, face processing, event-related brain potentials, social context, social cognition, perspective taking**

## INTRODUCTION

Emotion perception in oneself and others is an important aspect of successful social interaction. It is important for emotional self-regulation, and is often compromised in affective and mental disorders such as autism, sociopathy, schizophrenia, and depression, as well as in disorders associated with emotional blindness (alexithymia).

Narrative writing has been shown to have positive effects on emotional self-regulation (Hayes and Feldman, 2004) and individual well-being, possibly by increasing self-referential processing and reappraisal of emotionally challenging events from different perspectives (Seih et al., 2011; for an overview: Pennebaker and Chung, 2011). There is strong evidence from cognitive emotion regulation research supporting reappraisal as one of the most effective cognitive strategies for intentional down-regulation of negative feelings experienced in real life situations, or in the laboratory during viewing of emotion inducing stimuli including pictures, faces, or films (Gross, 2002; John and Gross, 2004; Blechert et al., 2012). Similarly, several studies investigating the neural correlates of self-referential processing suggest that appraising emotional stimuli in terms of their personal relevance can lead to adaptive emotion processing (e.g., Ochsner et al., 2004; Northoff et al., 2006; Moran et al., 2009).

An important key component of narrative writing is affect labeling. During narrative writing people learn to "put their feelings into words" and to express or reframe them verbally. The success of narrative writing suggests that words can be powerful regulators of emotions. Empirical support for this suggestion comes from neurophysiological research investigating brain responses, in addition to subjective indicators of emotion processing in participants exposed to affective labels or verbal descriptions while viewing emotional stimuli. For example, Foti and Hajcak (2008), and Macnamara et al. (2009) recorded event-related brain potentials (ERPs) from the electroencephalogram (EEG). Participants viewed unpleasant pictures which were preceded by neutral or negative verbal sentences. Processing of neutrally framed unpleasant pictures decreased ratings of picture emotionality, self-reported negative affect, and amplitudes of the late positive potential (LPP). The LPP often shows larger amplitudes during processing of emotional stimuli than during the processing of neutral stimuli (Olofsson et al., 2008). Attenuated LPP amplitudes to verbally framed unpleasant pictures therefore suggest a decrease in the depth of emotional stimulus encoding.

Verbal reframing effects are not restricted to neutral cues or sentences. Hariri et al. (2000) and Lieberman et al. (2007) scanned brain activity patterns by means of functional imaging while individuals viewed faces and affective labels. Affective labels consisted of simple words which were presented together with unpleasant faces or socially stereotyped faces, including black and white people. The task was to indicate which of the words fit best with the emotion depicted in the face. Control conditions included passive stimulus viewing (without labels), non-verbal affect matching (i.e., verbal labels were replaced by faces), and shape matching of simple geometric figures. Affect labeling with words as cues was the only condition that decreased amygdala activation significantly. It also enhanced activity in the right ventrolateral cortex (Lieberman et al., 2007), an important prefrontal control area involved in a variety of tasks requiring executive control of attention, response-inhibition, and intentional emotion regulation (Cohen et al., 2013). Consistent with these neurophysiological observations affect labeling with words as cues decreased peripheral-physiologic responses of emotional arousal and, in another study, reduced self-reported negative affect and distress in response to unpleasant emotional pictures to a similar extent as did intentional emotion regulation by means of reappraisal (Lieberman et al., 2011). Along these lines, developmental research has demonstrated a relationship between language impairments in childhood and diminished self-control (Izard, 2001), and poor emotional competence and emotion regulation abilities during adulthood (Fujiki et al., 2004). Furthermore, reducing the accessibility of emotion words experimentally (via a semantic satiation procedure) has been shown to decrease emotion recognition accuracy during face processing in healthy subjects (Lindquist et al., 2006; Gendron et al., 2012), a finding that matches with clinical observations of decreased face recognition abilities in patients with aphasia, who experience extreme difficulties in naming words (Katz, 1980).

Together, all these findings support the theoretical view that language provides a conceptual context for emotion processing (Barrett et al., 2007, 2011). Specifically, they suggest that words as affective labels can improve emotion recognition and at the same time regulate emotion processing much like intentional emotion regulation strategies, lending support for the idea of an incidental emotion regulation process underlying language processing in emotional contexts (Lieberman et al., 2011). Therefore, affect labeling has been suggested as an additional technique for emotion regulation in clinical and therapeutic settings (Tabibnia et al., 2008; Lieberman, 2011; Kircanski et al., 2012), especially in patients who, due to the severity of their symptoms, are less sensitive to more complex cognitive behavioral interventions that often require that people are able and willing to reflect in detail upon their feelings and the logic of their maladaptive appraisals.

Two questions have not yet been answered: (1) are words (affective labels) equally effective in modulating emotion recognition and emotion processing when those labels directly relate to the participants' own emotion as compared to the emotion expressed in the sender's face? (2) would these effects be the same during intentional regulation as compared to automatic or uninstructed regulation? In other words, does it matter "what" is going to be labeled by the words (own emotion vs. emotion conveyed by the sender) and "how" (automatically or intentionally) this is done? One way to answer these questions would be to expose individuals to faces expressing discrete emotions, such as fear, anger, or happiness, and to instruct them to find words which best describe either their own emotions, or the emotion expressed by the sender's face. This would involve participants correctly identifying their own emotions and finding words to express them, and these abilities vary across individuals (Barrett et al., 2011). The appraisal of cortical and peripheral physiological changes associated with emotion processing are often limited to more basal emotional dimensions of perceived pleasantness (good-bad, like-dislike) or arousal intensity (calming-arousing). However, ERP methodology may be useful in identifying cue-driven emotion processing and intentional emotion regulation, using as affective cues pronoun-noun pairs that are either self-or sender-related.

Research on the processing of pronouns has shown that readers adopt a first person perspective (1PP) during reading of self-related pronouns and a third-person perspective (3PP) during reading of other-related pronouns (Borghi et al., 2004; Ruby and Decety, 2004; Brunye et al., 2009). Moreover, EEG studies have reported emotional pronoun-noun pairs describing the reader's own emotion (e.g., my fear, my fun) to be processed more deeply than pronoun-noun pairs making a reference to the emotion of others (e.g., his fear, his fun) or emotion words that contain no reference at all (e.g., the fun, the fear) (Herbert et al., 2011a,b). Further, in an imaging study, reading of self-related emotional pronoun-noun pairs selectively enhanced activity in medial prefrontal brain structures involved in the processing of one's own feelings (Herbert et al., 2011c), providing neurophysiological evidence that people spontaneously discriminate between the self and the other during reading (see also Walla et al., 2008; Zhou et al., 2010; Shi et al., 2011).

Self-related emotion labels (e.g., my fear, my happiness) provide a window to one's own emotions, linking one's own sensations with the emotion expressed by the sender's face. Sender-related labels (e.g., his/her fear, happiness) make a direct reference to the emotion expressed by the sender. Therefore, processing of self- and sender-related labels should have differential effects on how emotional information conveyed by a stimulus such as a face is decoded, and also which emotional reactions and feelings are experienced in return. While labels depicting one's own emotions should make facial expression more relevant to the self, thereby increasing attention capture by emotional faces, sender-related labels might specifically improve decoding of structural information from the face. As explained in more detail below, this should be accompanied by different modulation patterns of early brain potentials in the EEG, such as the face specific N170 and the early posterior negativity (EPN). In addition, self- and sender-related labels both contribute information that goes beyond the affective information available from the face. Both labels contain information required for appraising the meaning of the emotion expressed in the face. Making this information available to the subject could reduce processing resources required for appraising the emotional meaning of the faces. In the EEG, this should be reflected by modulation of late ERP components such as the LPP.

The aim of the present EEG-ERP study was to shed light on these questions by investigating how processing of self- and sender-related affective labels modulates emotional face processing across different stages of processing, i.e., from initial processing of emotional stimulus features to the more fine grained, in depth analysis of the emotional content of the presented faces. Two separate experiments were conducted. The first experiment investigated emotional face processing during an "unintentional labeling" task. This was done to investigate spontaneous effects of self- and sender-related affective labels on emotional face processing. The second experiment used an active emotion regulation context in which self- and sender-related affective labels served as cues for emotion regulation and faces as targets of emotion regulation. The active emotion regulation context was chosen to separate unintentional from intentional affect labeling processes, and to explore whether emotion regulation with affective labels and "unintentional" processing of affective labels would differentially modulate event-related brain potentials (ERPs) to emotional faces. Moreover, the active emotion regulation context will allow us to investigate whether emotion regulation with self- and sender-related affective labels will facilitate self-referential processing and cognitive reappraisal of emotional faces from different perspectives (1PP vs. 3PP). Specifically the active process of labeling and the intention to use these labels for emotion regulation should allow a person to get in touch with his/her feelings when self-related affective cues are being presented and to distance him- herself from the own feelings when faces are cued with sender-related affective cues.

By comparing ERPs elicited by the faces during the affective label conditions and during a control condition of passive face viewing it is possible to precisely determine the particular stages at which processing of self- and sender-related emotion labels impact emotional face processing, in which direction these effects will occur (up- vs. down-regulation), and if effects vary across the two processing conditions of spontaneous, unintentional processing (experiment 1) and active, intentional regulation (experiment 2). ERPs of interest included early and late ERP components, the P1, the face specific N170, the early posterior negativity potential (EPN), and the late positive potential (LPP) or slow wave (SW). These cortical components are thought to indicate stimulus-driven as well as sustained processing of emotional stimuli. The P1 reflects very early stimulus feature processing while the N170 reflects increased structural encoding and the EPN facilitated capture of attentional resources by stimuli of emotional relevance (Bentin and Deouell, 2000; Junghöfer et al., 2001; Schupp et al., 2004; Blau et al., 2007). Amplitude modulations of the LPP are thought to index sustained processing and encoding of emotional stimuli in functionally coupled, fronto-parietal brain networks (Moratti et al., 2011).

Furthermore, participants' subjective appraisals of the presented stimuli, their mood state, and their emotion perception and empathic abilities were assessed via self-report. As additional exploratory outcome measures, subjective measures could provide information about potential variables that mediate affect labeling and face processing.

## METHODS: EXPERIMENT 1 AND EXPERIMENT 2

### SUBJECTS

Twenty-one right-handed adults (16 females, 5 males), all native speakers of German, with a mean age of 22 years (*SD* = 3.1 years) participated in experiment 1. Seventeen right-handed adults (12 females, 5 males), all native speakers of German, with a mean age of 22 years (*SD* = 2.2 years) participated in experiment 2. Participants were recruited via the posting board of the University of Würzburg and received course credit or financial reimbursement of 15€ for participation. Exclusion criteria for participation were current or previous psychiatric, neurological, or somatic diseases, as well as medication for any of these. Participants reported normal audition, and normal or corrected to normal vision. Both experiments were conducted in accordance with the Declaration of Helsinki and methods were approved by the ethical committee of the German Psychological Society (http://www.dgps.de/en/).

Participants of experiment 1 and of experiment 2 had comparable scores on the Beck Depression Inventory (Hautzinger et al., 1994) (experiment 1: $M = 5.3$, $SD = 3.1$; experiment 2: $M = 3.4$, $SD = 2.6$). Both groups scored normally on the trait (experiment 1: $M = 44.7$, $SD = 5.1$; experiment 2: $M = 45.5$, $SD = 4.3$) and state (experiment 1: $M = 39.6$, $SD = 4.5$; experiment 2: $M = 42.3$, $SD = 5.7$) scales of the Spielberger State-Trait Anxiety Inventory (STAI, Laux et al., 1981). They reported more positive affect (experiment 1: $M = 37.0$, $SD = 6.1$; experiment 2: $M = 38.8$, $SD = 5.9$) than negative affect (experiment 1: $M = 19.4$, $SD = 6.2$; experiment 2: $M = 17.8$, $SD = 4.7$) on the PANAS mood assessment scales (Watson et al., 1988) and they did not differ in empathy (experiment 1: $M = 14.19$, $SD = 2.2$; experiment 2: $M = 15.1$, $SD = 1.1$), perspective taking (experiment 1: $M = 14.3$, $SD = 2.3$; experiment 2: $M = 15$, $SD = 2.2$), emotional intelligence (experiment 1: $M = 117.9$, $SD = 7.1$; experiment 2: $M = 120.1$, $SD = 7.5$), or emotional blindness (experiment 1: $M = 50.6$, $SD = 6.7$; experiment 2: $M = 47.9$, $SD = 7.3$). They did also not differ in self-esteem (experiment 1: $M = 39.4$, $SD = 2.8$; experiment 2: $M = 39.2$, $SD = 3.1$). Empathy and perspective taking were measured with the Saarbrückener Personality Questionnaire (Paulus, 2009), the German Version of the Trait Emotional Intelligence Questionnaire (TEIQue) was used for emotional intelligence (Freudenthaler et al., 2008) and emotional blindness was assessed with the German Version of the Toronto Alexithymia Scale (TAS-20, Bagby et al., 1994). Self-esteem was measured via the Frankfurter Self Concept Scale (FSSW, Deusinger, 1986). Habitual emotion regulation strategies including reappraisal or suppression were also assessed with a German translation of the emotion regulation questionnaire (ERQ, Gross and John, 2003). Again both groups reported comparable scores on the ERQ for using either reappraisal (experiment 1: $M = 5.0$, $SD = 0.7$; experiment 2: $M = 4.7$, $SD = 1.2$) or suppression (experiment 1: $M = 3.6$, $SD = 1.0$; experiment 2: $M = 3.4$, $SD = 0.7$) as an emotion regulation strategy in daily life.

### STIMULI (EXPERIMENT 1 AND EXPERIMENT 2)

Faces (fearful, angry, happy, and neutral) were taken from the Karolinska Directed Emotional Face database (KDEF, Lundqvist et al., 1998). Affective labels were sixty pronoun-noun pairs. Twenty of these pairs were related to fear, 20 to anger, and 20 to happiness. Stimuli were presented in six randomized blocks, three self-related blocks and three sender-related blocks. Each block contained twenty faces (half male/half female characters) from one emotion category (fear, anger, or happiness) and twenty labels. Labels matched the emotion depicted in the face and were related to the participants' own emotions (e.g., my fear, my pleasure, my anger) in the self-related blocks and to the emotion of the sender's face (e.g., his/her fear, his/her pleasure, his/her anger) in the sender-related blocks. Thus, each block consisted of 20 emotion congruent trials. Labels and faces were presented for 1.5 s each, and separated by a fixation cross of 500 ms duration. Trials were separated by an inter-trial interval lasting about 1 s and blocks were separated by a fixation cross indicating the beginning of a new block.

Blocks of passive viewing, in which 20 faces (half male/half female characters) of each emotion category as well as 20 neutral faces were preceded by random letter strings instead of affective labels, served as control condition (see **Figure 1**). Akin to the experimental condition, each trial consisted of 1.5 s letter presentation, a 500 ms fixation-cross period, and a picture viewing period of 1.5 s. Trials were separated by inter-trial intervals of about 1 s and blocks by inter-block intervals. Block order was randomized. Emotional faces were randomly assigned to the blocks, such that none of the faces was repeated across blocks (self vs. sender) and conditions (control vs. affective label conditions). In line with previous studies, the passive viewing conditions were always presented first to guarantee a neutral baseline of emotional face processing. Also in line with previous studies, neutral faces were presented in the control condition, only (e.g., Hajcak and Nieuwenhuis, 2006; Hajcak et al., 2006; Blechert et al., 2012).

Affective labels and letter strings were presented in black letters (font "Times"; size = 40) centered on a white background of a 19 inch computer monitor. Faces were presented in color, centered on the computer screen. Stimuli were presented at a visual angle of 4°. Experimental runs were controlled by Presentation software (Neurobehavioral Systems Inc.). An overview of the experimental design is shown in **Figure 1**.

## PROCEDURE

After arrival at the laboratory, participants were informed in detail about the EEG procedure; they were questioned about their handedness and health, and electrodes for EEG recording were attached before they received the following instructions.

### EXPERIMENT 1

In the control condition, participants were told to view the faces attentively without paying specific attention to the preceding letter strings. In the following "affective label" condition, they were told that again a series of faces would be presented, each of which would be preceded by a verbal cue describing either their own emotion to the face or the emotion of the person presented in the face. Participants were asked to attend to the stimulus pairs (cue and face), but received no instruction to appraise the stimuli in a specific way nor to intentionally regulate their emotions during face processing. Prior to the start of the experimental recording,

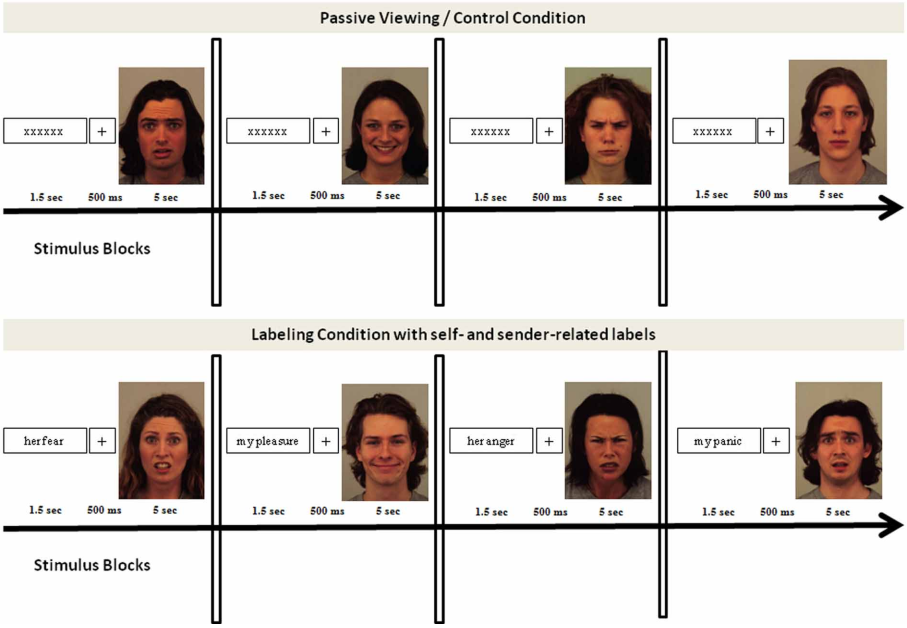


**FIGURE 1 | Experimental set-up of experiment 1 and experiment 2.** Both experiments used the same design, consisting of a control condition and an affective label condition. The control condition was always presented first and consisted of blocks of fearful, happy, angry, and neutral faces. The affective label condition consisted of six blocks, three self-related and three sender-related blocks. Block order was randomized and stimuli were randomly assigned to the conditions.

participants were given practice trials to familiarize them with the task. After the experimental recording, participants were asked to rate the stimuli for valence and arousal on a nine-point paper-pencil version of the Self Assessment Manikin (Bradley and Lang, 1994), they were questioned about their experience during picture viewing, and they filled in the additional questionnaires for perspective taking, emphatic concerns, emotion perception abilities, and habitual emotion regulation strategies. Finally, they were debriefed in detail about the purpose of the experiment.

### EXPERIMENT 2

Experiment 2 used the same stimuli and experimental set-up as experiment 1. However, in experiment 2, participants were asked to control their emotions during face processing by means of the cues presented prior to each face. When cues were self-related to their own emotion they should try to use the cues to get in touch with their feelings and label them when looking at the faces ("label own emotion" blocks). When cues were related to the emotion of the sender's face ("label sender's emotion" blocks) they should try to use the cues to distance themselves from their own emotions by labeling the emotion of the sender's face when looking at the faces. Participants received practice trials prior to the start of the experimental recording sessions to ensure that they understood the task and to familiarize them with cue-driven intentional regulation. In addition, they were asked to indicate their regulation success as well as their present feelings immediately after each regulation block on nine-point Likert scales. After the experiment, they were questioned in more detail about their regulation experiences. They rated the stimuli for valence and arousal on the Self-Assessment Manikin (Bradley and Lang, 1994). Afterwards, they filled in the questionnaires on perspective taking, emphatic concerns, emotion perception and habitual emotion regulation strategies, and were debriefed in detail about the purpose of the experiment.

## ELECTROPHYSIOLOGICAL RECORDINGS (EXPERIMENT 1 AND EXPERIMENT 2)

Electrophysiological data was recorded from 32 active electrodes using the actiCap system (Brain Products GmbH). For all electrodes impedance was kept below 10 kOhm. Raw EEG data were sampled at a rate of 500 Hz with FCz as reference. Off-line, EEG signals were digitally re-referenced to an average reference passing from 0.01 to 30 Hz, and corrected for eye-movement artifacts using the traditional algorithm by Gratton et al. (1983), implemented in the Analyzer2 software package (BrainProducts GMBH). Further artifacts due to head- or body-movements were rejected via a semi-automated artifact rejection algorithm. In total, this resulted in a loss of about 2–3 trails per block, leaving about 17 trials per block for averaging. Although signal to noise ratio increases with the number of averaged trials, recent research (Moran et al., 2013) has shown that differences in ERPs between experimental conditions can be reliably detected after a few averaged trials. This has been shown for late ERP components, which are more susceptible to background noise than early ERP components. Artifact-free EEG data were segmented from 500 ms before until 1500 ms after onset of the target faces using the 100 ms interval before target face onset for baseline correction. Baselines corrected epochs were then averaged for each experimental condition (affective label, control) and valence category (fear, anger, happy, neutral). Electrodes and time-windows for amplitude scoring of early (P1, N170, EPN) and late event-related brain potentials (LPP/slow wave) were determined in line with the previous literature on emotional face processing and by visual inspection of the grand mean average waveforms.

Amplitudes of early ERPs (P1, N170, and EPN) were analyzed at left and right posterior electrodes (O1, O2, PO10, PO9, P8, P7, TP10, TP9) from 80 to 120 ms (P1), from 140 to 180 ms (N170), and from 200 to 400 ms (EPN) post target face onset. Amplitudes of the late positive potential (LPP) were analyzed at parietal electrodes (P3, P4, Pz, P7, P8) in a time-window from 400 to 600 ms post target face onset. In experiment 2, intentional emotion regulation elicited a more pronounced cortical positivity (slow wave) over parietal electrodes compared to unintentional affect labeling. Akin to the LPP, amplitudes of this slow wave were analyzed at parietal electrodes (P3, P4, Pz, P7, P8), starting in a time-window from 400 to 800 ms from post-target face onset. ERP amplitudes were scored at each electrode as the averaged amplitude (in μV) in the respective time-window.

In addition, latencies of ERP amplitudes were defined via a semi-automatic peak detection algorithm of the Analyzer2 software package (BrainProducts GMBH). Latencies were analyzed in both experiments to determine if processing of affective labels had an influence on the speed of face processing.

Event-related potentials (P1, N1/EPN, and LPP) elicited during the processing of self- and sender-related affective labels were also analyzed from the epochs from 500 ms before until 1000 ms after word onset. The 100 ms interval before word onset was used for baseline correction. Amplitudes of early ERPs (P1, N1, and EPN) were analyzed at left and right posterior electrodes (O1, O2, PO10, PO9, P8, P7, TP10, TP9) from 80 to 120 ms (P1), 120 to 180 (N1), and from 200 to 400 ms (EPN) post word onset; amplitudes of the LPP were analyzed from 400 to 600 ms post word onset at the parietal electrodes P3, P4, and Pz.

## DATA REDUCTION AND STATISTICAL ANALYSIS

### ELECTROPHYSIOLOGICAL DATA—EXPERIMENT 1 AND EXPERIMENT 2

#### *Affective labels*

ERPs (P1, N1, EPN, LPP) elicited during the presentation of affective labels were analyzed with repeated measures analyses of variance (ANOVAs), which contained the factors *emotion (fearful, happy, angry)*, *label (self vs. sender)*, and *electrode location* as within-subject factors.

#### *Faces*

For faces the ANOVAs contained the factors *emotion (fearful, happy, angry)*, *condition (self vs. sender vs. passive viewing)*, and *electrode location* as within-subject factors. We also evaluated whether processing of emotional faces elicited larger ERP amplitudes than processing of neutral faces during passive viewing. The ANOVAs for the passive viewing comparisons included the factors *valence* (*fearful, happy, angry and neutral*) and *electrode location* as within-subject factors and were conducted for each ERP component of interest (P1, N170, EPN, and LPP/slow wave).

Where appropriate, p-values were adjusted according to Greenhouse and Geisser (1959). Significant main effects were decomposed by simple contrast test and results from these comparisons are reported uncorrected at $p < 0.05$. Interactions were followed up with planned comparisons within a row or column of the design matrix, to decrease the total number of comparisons by avoiding those that would not make theoretical sense. For example, an interaction between *condition* and *emotion* might involve contrasting $\text{fearful}_{\text{self}}$ vs. $\text{fearful}_{\text{other}}$, but would not involve a contrast between $\text{fearful}_{\text{self}}$ vs. $\text{angry}_{\text{other}}$, since this would involve a confound across levels of both variables. Again, results are reported $p < 0.05$, uncorrected.

## BEHAVIORAL DATA – EXPERIMENT 1 AND EXPERIMENT 2

### SUBJECTIVE RATINGS

Ratings were analyzed with repeated measures analyses of variance (ANOVAs). For faces, the ANOVAs contained the factors *emotion* (*fearful, happy, angry*) and *condition (passive viewing vs. self vs. sender)* as within-subject factors. Similar to the analysis of the ERPs, separate ANOVAs were calculated to consider differences in ratings between emotional and neutral faces. Ratings of the affective labels were analyzed with ANOVAs containing the factors *emotion* (*fearful, happy, and angry*) and *label (self vs. sender)* as within-subject factors. Self-report data including reports about changes in mood and regulation success given after each regulation block were also analyzed in separate ANOVAs, each containing the factors *emotion* (*fearful, happy, and angry*) and *label (self vs. sender)* as within-subject factors.

### CORRELATIONAL ANALYSIS: ERPs AND SELF-REPORT MEASURES

In both experiments, ERPs (P1, N170, EPN, and LPP/slow wave) were correlated with participants' self-report measures. Self-report measures of interest included positive and negative affect, depression, state and trait anxiety, and empathic concerns, perspective taking, self-esteem, and the ability or inability to describe and identify feelings as measured with the subscales of the Toronto Alexithymia Scale and the TEIQue emotional intelligence questionnaire. Although these analyses are exploratory in the present study, a mediating role of these variables could theoretically be expected based on clinical findings and the literature on individual differences (e.g., Herbert et al., 2011a,d; Moratti et al., 2011).

## RESULTS

### EXPERIMENT 1—LABELING WITHOUT INTENTIONAL REGULATION INSTRUCTION

#### *Electrophysiological data*

***Affective Labels.*** Processing of self- and sender-related labels did not differ in the P1, N1, and EPN time-windows. In the time-window of the LPP, amplitudes were more pronounced for self-related than for sender-related affective labels. The main factor *label* was significant, $F_{(1,\,20)} = 7.08$, $p = 0.02$.

***Faces.*** *Emotional faces: self vs. sender vs. control (passive viewing).* Emotional face processing differed significantly when preceded by affective labels as compared to during passive viewing. A first difference was observed in the P1 time-window and is indicated by the main factor *condition*, $F_{(2,\,40)} = 4.0$, $p = 0.027$. P1 amplitudes were significantly more positive for fearful, angry, and happy faces during the "affective label" conditions compared to during passive viewing. During passive viewing, P1 amplitudes did not differ significantly between fearful, angry, happy, and neutral faces.

N170 amplitudes showed a main effect of *emotion*, $F_{(2,\,40)} = 6.4$, $p = 0.003$: Fearful faces elicited significantly larger N170 amplitudes than angry faces, $F_{(1,\,20)} = 14.1$, $p = 0.001$. In addition, the interaction of the factors *emotion* and *condition* was significant, $F_{(4,\,80)} = 3.0$, $p = 0.05$: fearful faces elicited significantly larger N170 amplitudes when preceded by labels describing the sender's emotion as compared to when preceded by labels describing the viewer's own emotion, $F_{(1,\,20)} = 4.1$, $p = 0.05$, as well as compared to when presented without any labels (control condition), $F_{(1,\,20)} = 12.8$, $p = 0.002$. For emotional faces preceded by labels describing the viewer's own emotion, N170 amplitudes did not differ from passive viewing of emotional faces. During passive viewing, the factor *valence* was also significant, $F_{(3,\,60)} = 4.64$, $p = 0.02$: fearful faces elicited significantly larger N170 effects in comparison to neutral faces, $F_{(1,\,20)} = 7.03$, $p = 0.02$.

EPN amplitudes showed a significant main effect of *condition*, $F_{(2,\,40)} = 3.6$, $p = 0.05$. EPN amplitudes were significantly more pronounced for fearful, angry, and happy faces when they were preceded by self-related labels than when they were presented without any labels in the control condition, $F_{(1,\,20)} = 4.48$, $p = 0.046$. Cueing fearful, angry, and happy faces with emotion words describing the emotion of the sender's face did not change EPN amplitudes relative to when presented without any labels (control condition). EPN amplitudes did also not differ significantly between emotional and neutral faces during passive viewing. The factor *valence* was not significant, $F_{(3,\,60)} = 1.2$, $p = 0.32$.

The amplitude of the LPP was significantly modulated by the factor *condition*, $F_{(2,\,40)} = 3.8$, $p = 0.05$. LPP amplitudes were attenuated for emotional faces regardless of their valence (fearful, angry, and happy) when preceded by self-related and sender-related affective labels as compared to when preceded by letter strings (control condition). During the control condition, fearful, $F_{(1,\,20)} = 14.3$, $p = 0.001$, as well as happy faces, $F_{(1,\,20)} = 11.5$, $p = 0.002$, elicited significantly larger LPP amplitudes compared to neutral faces. The factor *valence* was significant, $F_{(3,\,60)} = 4.86$, $p = 0.004$.

In contrast to amplitude measures, cueing faces with affective labels had no significant effects on ERP latencies.

ERP results are summarized in **Figure 2**.

#### *Behavioral data*

***Ratings.*** Fearful and angry faces were rated as significantly more negative in valence compared to happy faces, regardless of whether faces were cued with affective labels or not, $F_{(2,\,40)} = 201.9$, $p < 0.01$. Ratings of emotional faces did not differ in terms of arousal, $F_{(2,\,40)} = 2.7$, $p = 0.08$, but emotional faces were rated as higher in arousal than neutral faces. This was true for faces shown in the control condition, $F_{(3,\,60)} = 26.7$, $p < 0.01$, and for faces shown in the "affective label" conditions,

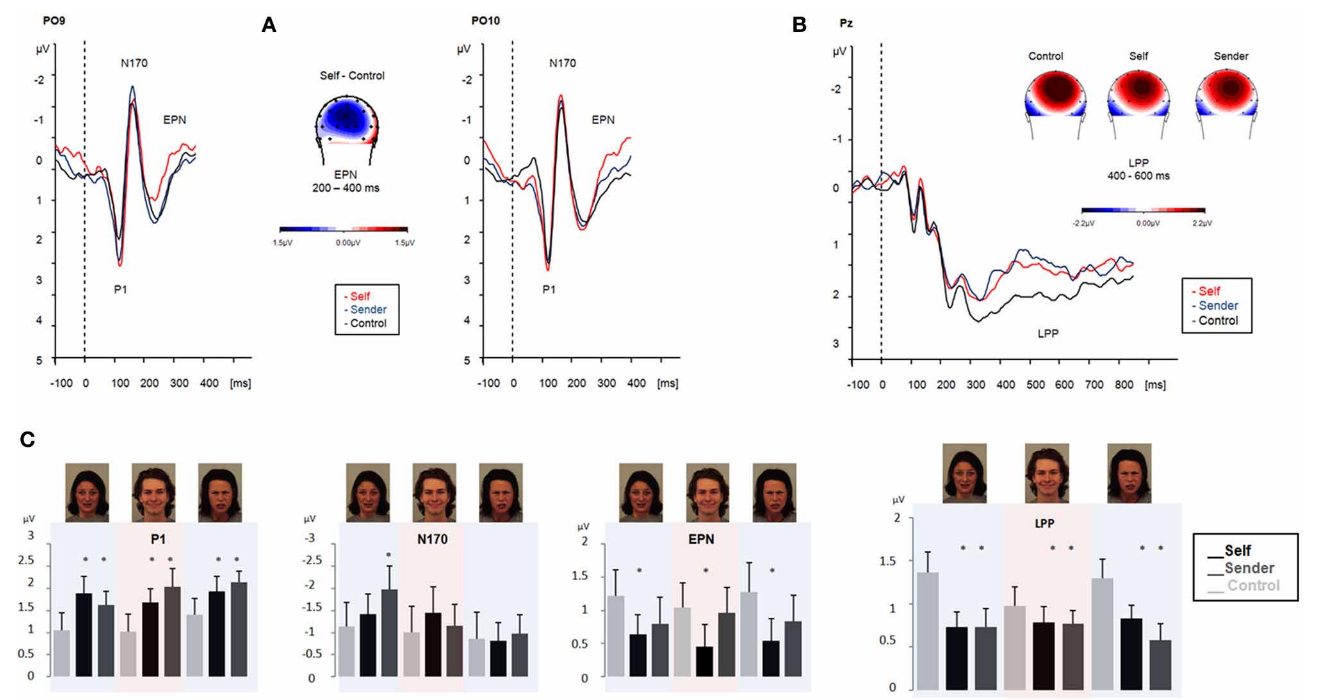


**FIGURE 2 | ERPs obtained during emotional face processing in experiment 1.** ERPs are collapsed across angry, fearful, and happy faces and contrasted with passive viewing of emotional faces (black line). Upper left panel **(A)** displays ERP modulation patterns in the early time-windows (P1, N170, and EPN, respectively). Upper right panel **(B)** displays amplitudes in the LPP time-window. Maps display the topographic distribution of the ERP patterns in μV. Lower panel **(C)**: overview of P1, N170, EPN and LPP modulation (means and SEMs) during emotional face processing in experiment 1. Color shadings highlight ERP modulation for each emotional category (fearful, angry, and happy).

*self*: $F_{(3,\,60)} = 26.2$, $p < 0.01$, and *sender*: $F_{(3,\,60)} = 29.9$, $p < 0.01$.

Self-related labels were rated as higher in arousal and as more relevant to the self compared to sender-related labels, *arousal:* $F_{(1,\,20)} = 11.1$, $p < 0.01$, *self-relevance*: $F_{(1,\,20)} = 37.0$, $p < 0.01$. This self-relevance effect was most pronounced for positive pronoun-noun pairs, $F_{(2,\,40)} = 10.2$, $p < 0.01$. Valence ratings confirmed a self positivity bias, $F_{(2,\,40)} = 5.1$, $p = 0.03$. Positive pronoun-noun pairs were rated more positive when related to the self.

Rating data are summarized in **Table 1** and **Table 2**.

***Manipulation check.*** None of the participants reported consistently using any strategy throughout the experiment. Some participants ($N = 14$) reported that they repeated the labels during face processing, but retrospectively none of them had the impression that this had reduced the emotionality of the faces or their own feeling state. Subjects reported no difficulties in understanding the intention of the labels, relating them spontaneously to either the self or the sender's face.

## EXPERIMENT 2—LABELING WITH INTENTIONAL REGULATION INSTRUCTIONS

### *Electrophysiological data*

***Affective Labels.*** P1, N1, and EPN amplitude modulations did not differ between self- and sender-related labels. However, akin to experiment 1, LPP amplitudes were more pronounced for self-related than for sender-related affective labels, *condition:* $F_{(1,\,16)} = 7.08$, $p = 0.02$.

**Table 1 | Experiment 1: rating data of emotional faces which were preceded by either self-related pronoun-noun pairs, sender-related pronoun-noun pairs, or no pronoun-noun pairs (control condition).**

| | Fear | Anger | Happiness | Neutral |
|---|---|---|---|---|
| **SELF-RELATED** | | | | |
| Valence | 3.40 (0.48) | 3.01 (0.52) | 6.69 (0.95) | |
| Arousal | 4.83 (1.59) | 4.99 (1.57) | 4.43 (1.37) | |
| **SENDER-RELATED** | | | | |
| Valence | 3.27 (0.65) | 3.04 (0.73) | 6.78 (0.82) | |
| Arousal | 4.89 (1.67) | 4.91 (1.53) | 4.29 (1.39) | |
| **CONTROLS** | | | | |
| Valence | 3.41 (0.59) | 2.99 (0.61) | 6.57 (0.77) | 4.67 (0.46) |
| Arousal | 4.68 (1.55) | 4.93 (1.76) | 4.21 (1.36) | 2.69 (1.41) |

*Scales ranged from 1 (extremely negative valence, extremely low arousal) to 9 (extremely positive valence, extremely high arousal). Standard deviations are in parentheses.*

***Faces.*** In the P1 time-window a main effect of *condition* was observed, $F_{(2,\,32)} = 3.7$, $p = 0.041$. Amplitudes of P1 were significantly enhanced for emotional faces during intentional regulation as compared to passive face processing, especially when

**Table 2 | Experiment1: mean valence, arousal, and self-relevance ratings of self-related and sender-related pronoun-noun pairs obtained after the experimental recordings.**

| | Fear | Anger | Happiness |
|---|---|---|---|
| **SELF-RELATED PRONOUN-NOUN PAIRS** | | | |
| Valence | 3.10 (0.95) | 3.21 (1.04) | 7.42 (0.99) |
| Arousal | 5.02 (1.80) | 4.77 (1.69) | 5.62 (1.36) |
| Self-Relevance | 5.54 (1.66) | 5.11 (1.35) | 7.03 (0.85) |
| **SENDER-RELATED PRONOUN-NOUN PAIRS** | | | |
| Valence | 3.40 (0.83) | 3.64 (0.81) | 6.67 (1.35) |
| Arousal | 4.16 (1.47) | 4.09 (1.70) | 4.31 (1.73) |
| Self-Relevance | 2.91 (1.51) | 3.11 (1.48) | 3.96 (1.82) |

*Scales ranged from 1 (extremely negative valence, extremely low arousal, extremely low self-relevance) to 9 (extremely positive valence, extremely high arousal, extremely high self-relevance). Standard deviations are in parentheses.*

using self-related labels for emotion regulation, *self*: $F_{(1,\ 16)} = 6.23$, $p = 0.023$, *sender*: $F_{(1,\ 16)} = 3.68$, $p = 0.073$. During passive viewing, P1 amplitudes did not differ between emotional and neutral faces, $F_{(3,\ 48)} = 1.8$, $p = 1.7$.

In the N170 time-window, the factor *condition* showed only a trend toward significance, $F_{(2,\ 32)} = 2.8$, $p = 0.079$, which indicated a slight reduction in N170 amplitudes to emotional faces during emotion regulation trials relative to passive viewing of emotional faces. During passive viewing, N170 amplitudes were more pronounced for emotional than for neutral faces, especially for fearful faces, $F_{(1,\ 16)} = 6.02$, $p = 0.026$. The factor *valence* was significant, $F_{(3,\ 48)} = 4.68$, $p = 0.046$.

In the EPN time-window a main effect of *emotion* was observed, $F_{(2,\ 32)} = 3.9$, $p = 0.037$: EPN amplitudes were more pronounced for fearful and angry faces than for happy faces, $F_{(1,\ 16)} = 4.94$, $p = 0.041$; $F_{(1,\ 16)} = 5.0$, $p = 0.04$. There was no significant interaction of the factors *emotion x condition*, $F_{(4,\ 64)} = 0.67$, $p = 0.61$. During passive viewing, EPN amplitudes did not differ significantly between emotional and neutral faces. The factor *valence* was not significant, $F_{(3,\ 48)} = 3.48$, $p = 0.18$.

In the LPP/slow wave time-window amplitudes were modulated by the main factor *condition*, $F_{(2,\ 32)} = 5.1$, $p = 0.021$. Amplitudes were significantly greater for emotional faces during the "label own emotion" condition compared to the passive viewing condition, $F_{(1,\ 16)} = 10.58$, $p = 0.005$, and also compared to the "label sender's emotion" condition, $F_{(1,\ 16)} = 7.0$, $p = 0.018$. During "label sender's emotion", amplitudes were attenuated compared to passive viewing of emotional faces. However, this attenuation for emotional faces during "label sender's emotion" was only significant when LPP amplitudes elicited during "label sender's emotion" were compared to passive viewing of neutral faces.

Akin to experiment 1, no significant differences were observed for ERP latencies.

ERP results of experiment 2 are displayed in **Figure 3**. Results obtained from both experiments are summarized in **Figure 4**.

### *Behavioral data*

***Ratings.*** Fearful and angry faces were rated as significantly more negative in valence compared to happy faces, $F_{(2,\ 32)} = 129.6$, $p < 0.01$. This was true regardless of whether faces were cued with affective labels or presented during the control condition. Valence ratings also differed between emotional faces and neutral faces (see **Table 3**). Again, this was true regardless of the condition in which the faces had been presented during the experiment. Arousal ratings were significantly modulated by the factor *condition*, $F_{(2,\ 32)} = 7.07$, $p < 0.01$. Emotional faces presented in the "label own emotion" condition were retrospectively rated higher in arousal than faces being presented in the control condition of passive viewing and the "label sender's emotion" condition.

Self-related pronouns were rated higher in arousal and in self-relevance than pronouns related to the sender, *arousal:* $F_{(1,\ 16)} = 14.2$, $p < 0.01$, *self-relevance*: $F_{(1,\ 16)} = 37.0$, $p < 0.01$. Positive pronoun-noun pairs were rated significantly higher in valence compared to pronoun-noun pairs describing fear and anger, as was indicated by a significant main effect of *emotion*, $F_{(2,\ 32)} = 129.1$, $p < 0.01$. A significant interaction of *emotion x condition*, $F_{(2,\ 32)} = 12.1$, $p < 0.01$ revealed a self-positivity bias: akin to experiment 1, valence ratings were higher for positive pronoun-noun pairs when related to the self than when related to the sender.

Rating data are summarized in **Table 3** and **Table 4**.

***Manipulation check.*** All participants were familiar with the task prior to the start of the experiment. All but 2 of them reported using a particular strategy to control their feelings when labeling their own or the sender's emotion. As did the participants in experiment 1, they reported rehearsing the labels during face processing. However, in contrast to the participants in experiment 1, they reported rehearsing the labels to experience the faces more intensively, for instance by linking faces with an autobiographical event in the "label own emotion" condition. During the "label sender's emotion" condition most subjects reported rehearsing the cues to increase the emotional distance between the self and the sender's face ($N = 10$ subjects). Some ($N = 6$) additionally tried to not show any feelings at all (emotion suppression) or to not empathize with the sender. Participants reported no difficulties in relating the cues to their own emotion or the emotion of the face. Furthermore, they reported that their feelings had increased during the "label own emotion" blocks and had the opposite impression in the "label sender's emotion" blocks. Ratings obtained after each block also indicated that subjects found it somewhat harder to regulate their emotions during the "label sender's emotion" blocks compared to during the "label own emotion" blocks. ANOVAs revealed a trend for the main factor *condition*, $F_{(1,\ 16)} = 3.5$, $p = 0.07$. In addition, a significant main effect of *emotion*, $F_{(2,\ 32)} = 3.6$, $p = 0.038$, indicated that amongst the to be regulated emotions (fear, anger, happiness), participants had the impression of putting more effort on the regulation of fear. Regarding the direction of their success, they reported that feelings became more positive, particularly when viewing happy faces and during the "label own emotion" blocks. This was indicated by a significant main effect of the factor

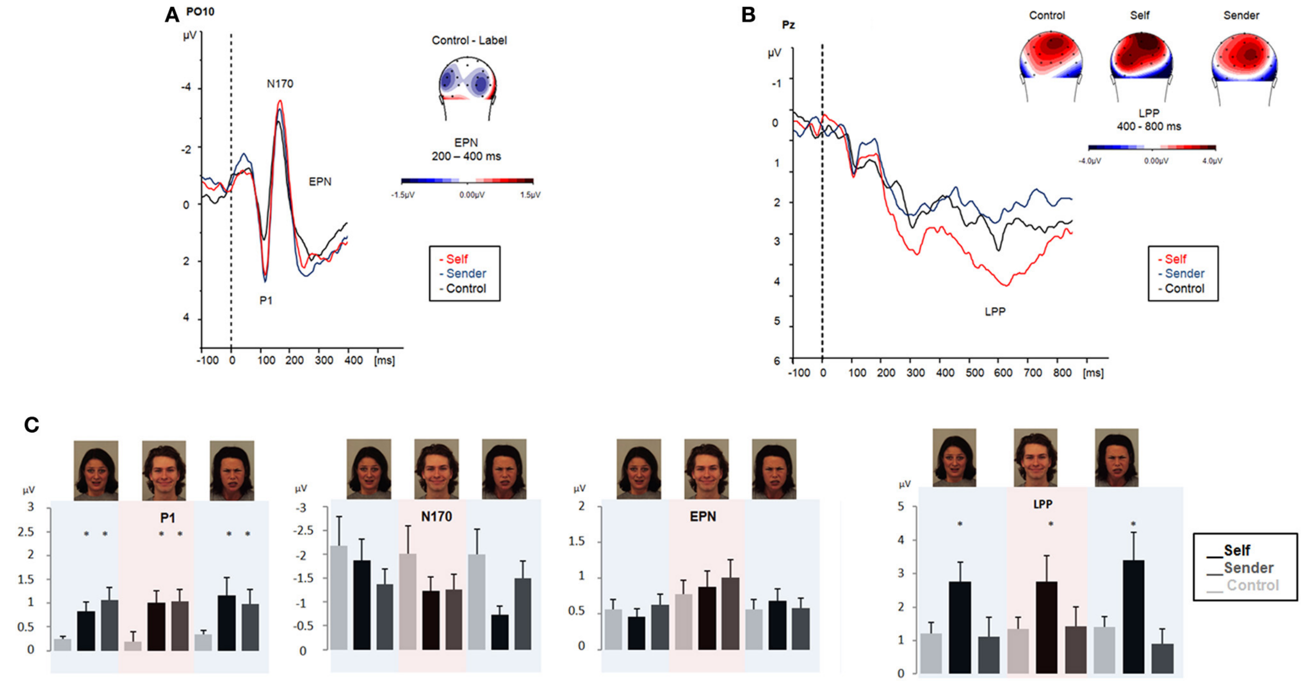
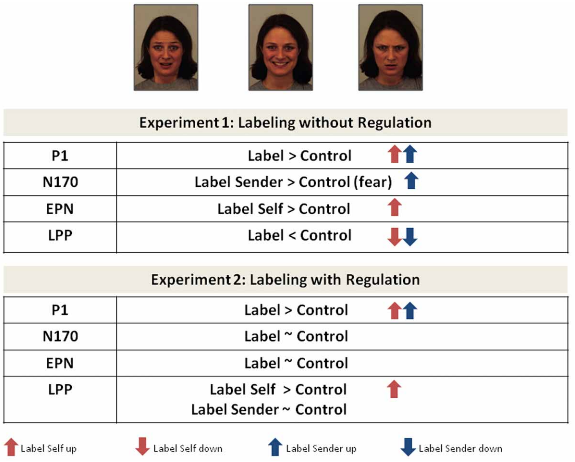


**FIGURE 3 | ERPs obtained during emotional face processing in experiment 2 (affect labeling with intentional regulation instruction).** Passive viewing of emotional faces (black line). Upper left panel **(A)** displays ERP modulation patterns in the early time-windows (P1, N170, and EPN, respectively). Upper right panel **(B)** displays amplitude modulation of the LPP. Maps display the topographic distribution of the ERPs in μV. Lower panel **(C)** overview of P1, N170, EPN and LPP modulation (means and SEMs) during emotional face processing in experiment 2. Color shadings highlight ERP modulation for each emotional category (fearful, angry, and happy).



| Experiment 1: Labeling without Regulation | |
|---|---|
| P1 | Label > Control ↑↑ |
| N170 | Label Sender > Control (fear) ↑ |
| EPN | Label Self > Control ↑ |
| LPP | Label < Control ↓↓ |

| Experiment 2: Labeling with Regulation | |
|---|---|
| P1 | Label > Control ↑↑ |
| N170 | Label ~ Control |
| EPN | Label ~ Control |
| LPP | Label Self > Control ↑<br>Label Sender ~ Control |

↑ Label Self up ↓ Label Self down ↑ Label Sender up ↓ Label Sender down

**FIGURE 4 | Modulation of event-related brain potentials obtained during experiment 1 and experiment 2.** Arrows indicate the direction of the modulation (up vs. down) compared to passive face processing. Only significant results are considered, trends are reported in the text.

**Table 3 | Experiment 2: rating data of emotional faces for the faces that were preceded by self-related pronoun-noun pairs, sender-related pronoun-noun pairs, or no pronoun-noun pairs (control condition).**

| | Fear | Anger | Happiness | Neutral |
|---|---|---|---|---|
| **SELF-RELATED** | | | | |
| Valence | 3.22 (0.91) | 3.02 (1.00) | 6.54 (1.15) | |
| Arousal | 6.02 (1.56) | 5.97 (1.79) | 5.00 (1.26) | |
| **SENDER-RELATED** | | | | |
| Valence | 3.37 (0.89) | 3.12 (1.11) | 6.72 (1.11) | |
| Arousal | 5.56 (1.48) | 5.74 (1.30) | 4.89 (1.49) | |
| **CONTROLS** | | | | |
| Valence | 3.32 (1.14) | 3.14 (1.23) | 6.44 (1.02) | 4.69 (0.94) |
| Arousal | 5.34 (1.40) | 5.83 (1.32) | 4.72 (1.48) | 3.41 (1.67) |

*Neutral faces were presented during the control condition only.*
*Scales ranged from 1 (extremely negative valence, extremely low arousal) to 9 (extremely positive valence, extremely high arousal). Standard deviations are in parentheses.*

*emotion*, $F_{(2,\ 32)} = 27.76$, $p < 0.01$, and a significant interaction of the factors *emotion* x *condition*, $F_{(2,\ 32)} = 8.56$, $p = 0.001$.

### CORRELATION ANALYSIS OF ERPs AND SELF-REPORT MEASURES

In experiment 1, correlation analyses revealed no significant results. Neither in the early nor in the late face processing time-windows about 1 s a significant relationship between ERPs and any of the selected self-report indices was found. In experiment 2, amplitudes of the slow wave showed a positive correlation with self-esteem (Pearson's $r = 0.4$, $p = 0.035$) during the "label own emotion" regulation blocks, and a significant negative correlation with empathy (Pearson's $r = -0.6$, $p = 0.004$) during

**Table 4 | Experiment1: mean valence, arousal, and self-relevance ratings of self-related and sender related pronoun-noun pairs obtained after the experimental recordings.**

| | Fear | Anger | Happiness |
|---|---|---|---|
| **SELF-RELATED PRONOUN-NOUN PAIRS** | | | |
| Valence | 2.82 (0.91) | 3.43 (0.84) | 7.56 (0.72) |
| Arousal | 5.09 (1.87) | 4.79 (1.74) | 6.16 (1.40) |
| Self-Relevance | 5.99 (1.56) | 6.08 (1.18) | 7.33 (0.88) |
| **SENDER-RELATED PRONOUN-NOUN PAIRS** | | | |
| Valence | 3.53 (1.02) | 3.79 (0.86) | 6.74 (0.95) |
| Arousal | 3.94 (1.87) | 3.95 (1.84) | 4.81 (1.89) |
| Self-Relevance | 3.21 (1.38) | 3.09 (1.43) | 4.02 (1.85) |

*Scales ranged from 1 (extremely negative valence, extremely low arousal, extremely low self-relevance) to 9 (extremely positive valence, extremely high arousal, extremely high self-relevance). Standard deviations are in parentheses.*

the "label sender's emotion" regulation blocks. Emotional intelligence was also negatively correlated with the amplitude of the slow wave during the "label sender's emotion" regulation (Pearson's $r = -0.43$, $p = 0.04$). However, when $p$-values were Bonferroni corrected for multiple comparisons ($p = 0.008$), only for empathy was the correlation still significant.

## DISCUSSION

Two separate experiments were conducted to investigate the impact of affective labels on face processing during spontaneous, automatic emotion processing or in an intentional emotion regulation context. EEG-ERP methodology was used to examine how decoding emotions from facial expressions changes when faces are preceded by verbal labels that vary in the extent to which they describe one's own emotion or the emotion expressed by the sender.

Processing of self- and sender-related affective labels increased perceptual processing of emotional faces as early as in the P1 time-window. The P1 component is assumed to reflect a global and coarse processing of facial stimulus features in the primary visual cortex. This stage temporally precedes a more fine grained, configural analysis of the structural features of the face, this later stage being reflected in amplitude modulations of the face specific N170 component (e.g., Bentin et al., 1996). Modulation of the P1 component by emotions and context has been reported in some but not all face processing studies (for an overview Righart and de Gelder, 2006). However, agreement exists that very early facial feature processing as reflected by the P1 can be modulated by task-related and context-dependent top-down processes (Heinze et al., 1994; Rauss et al., 2012). Processing of affective labels could influence early facial feature processing via top-down cognitive mechanisms of anticipation, or by activating conceptual processing of emotional faces in anticipation of their encounter. During passive viewing, faces were cued by letter strings instead of affective labels, but P1 amplitudes did not differ between emotional and neutral face conditions. Due to their semantic content, words as labels possess a greater anticipatory signal character than meaningless letter strings. P1 modulation by self- and sender-related affective labels occurred independently of the emotional valence of the faces (happy, fearful, or angry) and across experiments (automatic processing in experiment 1 vs. intentional emotion regulation in experiment 2), supporting the robustness of this observation.

The perceptual processing stages that followed the P1 and which were indexed by the N170 and the EPN component were influenced differently by self- and sender-related labels. As predicted, structural encoding of fearful faces (N170) increased significantly when verbal labels described the emotion of the sender's face, whereas cueing the faces with self-related labels describing the reader's own emotion facilitated attention capture by emotional faces in the time-window of the EPN component, after the N170. In contrast to the N170, the EPN is considered to reflect early conceptual and semantic analysis of a stimulus in the ventral visual processing stream (Schupp et al., 2006; Kissler et al., 2007). The N170 and the EPN effects were significant in experiment 1 and not observed in experiment 2, supporting their unintentional and implicit nature.

This corroborates the idea that "what" is to be labeled during affect labeling and "how" this is done (automatically or intentionally) influences both the direction and the intensity of our perceptual experiences. Decoding emotions from facial expressions is indeed not fully determined by the sensory information derived from the face, nor is it completely insensitive to contextual factors or independent from the experience on the perceiver's side. Also, the relationship between language and perception is stronger than traditionally assumed (Lindquist and Gendron, 2013). Not only does reading emotion words activate our sensory and motor systems, activation in these systems is also temporarily reduced when access to emotion concepts is blocked experimentally (Lindquist et al., 2006; Gendron et al., 2012) or when concept activation during face processing is changed by verbal negation (Herbert et al., 2012). The present observations, including P1, N170, and EPN modulations by affective labels, further emphasize an embodied view of language. They demonstrate that even minor linguistic variations that change the personal reference or ownership of a particular emotion concept can be powerful mediators of emotion perception. While spontaneous processing of affective labels describing one's own emotions increases the motivational relevance of emotional faces, be they happy, angry, or fearful (see EPN results), sender-related labels seem to improve emotion recognition more specifically by facilitating decoding of structural information from the face, especially from fearful faces (see N170 results).

Processing of affective labels modulated later stages of face processing as well. Again, results differed between the two experiments, supporting the notion of psychologically and physiologically different mechanisms underlying automatic and intentional affect labeling. Spontaneous processing of affective labels reduced the amplitude of the late positive potential (LPP) to fearful, happy, and angry faces relative to the control condition of passive viewing. Moreover, this was observed for emotional faces cued by self- and sender-related affective labels. The LPP and the slow wave are cortical correlates of sustained attention and depth of stimulus encoding (Kok, 1997; Schupp et al., 2000; Moratti et al., 2011).

Lower LPP amplitudes thus imply a reduction in depth of stimulus processing during automatic and hence unintentional affect labeling. This was not restricted to the processing of fearful or angry faces, applying to all faces expressing negative emotions. To the contrary, processing of self- and sender-related labels seemed to reduce processing costs for both, negative and positive emotional facial expressions.

A reduction in cortical processing depth, however, does not necessarily imply a dampening of affect on a subjective experiential level. When asked post-experimentally, none of the participants in experiment 1 had the impression that processing of affective labels had transiently dampened the own feeling state during the experiment. Ratings obtained for a subset of the faces after the experiment also indicated no changes in perceived valence or arousal. This underscores findings from previous studies suggesting that processing of affective labels has incidental effects on a bio-physiological level, but not necessarily on a subjective experiential level (Kircanski et al., 2012).

During active emotion regulation a different pattern emerged, distinguishing between automatic and intentional affect labeling processes. Unlike in experiment 1, amplitudes of the LPP/cortical slow wave in experiment 2 were enhanced for emotional faces when subjects used the self-related affective labels for emotion regulation during face processing. Likewise, they reported the feeling that the emotionality of the faces increased during the "label own emotion" regulation blocks and rated the faces afterwards as higher in arousal than the emotional faces that had been presented in the "label sender's emotion" or the passive viewing conditions. Increased processing in the "label own emotion" regulation condition contrasts with the view that making oneself aware of one's own emotions would transiently dampen the feeling state itself. It also runs counter to the assumption that affect labeling would always have a down-regulatory effect on emotional stimulus processing (Lieberman, 2007, 2011), independent of the personal reference properties of the label and the participants' intentions.

The results of experiment 2 suggest that intentional affect labeling can intensify encoding of emotional faces and make their content more intense and self-relevant. This is in line with observations from imaging studies on self-referential processing of emotional stimuli. Self referential processing of emotional stimuli activates medial prefrontal cortex regions, which are part of the salience network (Schmitz and Johnson, 2007). Parts of this network have been found to be active during reading of self-related emotional pronoun-noun pairs as well (Herbert et al., 2011c). In line with these observations, self-related emotional pronoun-noun pairs elicited larger LPP amplitudes relative to sender-related emotional-pronoun-noun pairs, which corroborates findings from recent studies showing similar effects (e.g., Walla et al., 2008; Zhou et al., 2010; Herbert et al., 2011a,b; Shi et al., 2011) and supports the notion that participants discriminated between the self and the other during reading of self- and sender-related labels.

Experiencing a strong sense of ownership can diminish self-other boundaries, such as when watching one's own face and the face of a sender being touched simultaneously (Maister et al., 2013). Labeling one's own emotion seems to provide another way to get in touch with one's own emotion while viewing someone else's face, particularly when there is the intention to do so, as seen in experiment 2. This could have effects similar to the resonance of being touched in real time as it might help synchronizing one's own feelings with the expressions of the sender's face.

Resonance between sender and observer might also play a role when attempting to regulate one's own emotion by labeling the emotion of the sender's face. In the present study, empathy and emotional intelligence were inversely related with amplitudes of the slow wave during the "label sender's emotion" conditions. In addition, self-esteem, which reflects the most fundamental appraisals about the self, was positively correlated with depth of face processing during the "label own emotion" blocks, corroborating theoretical conceptions that link self esteem with improved self consciousness or an improved self-awareness (Branden, 1969). Positive correlations between the personality measures found here might reverse when, instead of healthy subjects, clinically relevant samples with poor self-esteem, low empathy, and high emotional blindness are investigated. Gender likewise could play a role because more females than males took part in the present experiments. In any event, inter-individual differences should be taken into account in future research. In the present study, results on inter-individual differences can only be considered tentative due to their exploratory nature and the small sample sizes studied.

It has been debated to what extent interventions in which individuals are asked to focus on their own emotion and to become aware of them are helpful for emotion regulation. Looking at clinical disorders, internal self focus of attention can give rise to negative feelings of distress and heightened physiological arousal (Ingram, 1990). Similar observations of an increase in symptomatology (e.g., increase in distress, negative mood, and physical symptoms) have been reported immediately after expressive writing (Pennebaker and Beall, 1986; Pennebaker and Chung, 2007, 2011). In the present study, participants reported an increase in feelings during intentional affect labeling as well. However, this increase in feelings when labeling one's own emotion did not push subject's mood in a negative direction. A major difference between self monitoring in clinical disorders, expressive writing, and intentional affect labeling is that, during intentional affect labeling, labels provide a concrete context for appraising one's feelings and concomitant bodily changes, the latter often being verbally accessible only along simple physical dimensions of valence (good-bad) and arousal (calm-arousing). In this sense, using self- and sender-related labels actively and intentionally for emotion regulation seems to have comparable effects on emotion processing and cognitive reappraisal. A more speculative possibility is that using self- and sender-related labels actively and intentionally for emotion regulation might have facilitated verbal self-guidance and self-regulation by means of inner speech (Morin, 2005). Future studies could test this assumption.

## CONCLUSION

Language has long been considered as being somewhat independent from emotions, both with regard to its capacity to induce emotions and with regard to its potential to up- and down-regulate emotion processing in accord with situational demands.

The present study sheds light on the mechanisms underlying the emotional regulatory capacity of language. The results are the first to show that "what" is being labeled verbally (e.g., one's own emotion or the emotion of the sender's face) during face processing and how intentionally this is done matters. The present observations therefore pave the way for a differentiated view of language as a context for emotion processing. Some of our observations are specifically suggestive for future research, particularly the observation, that, with few exceptions, processing of affective labels with or without regulatory intent did not interact with the emotional valence of the faces, i.e., regardless of whether facial expressions were fearful, angry, or happy. That is, ERP modulations reflected the different experimental conditions of self vs. sender labeling, rather than the emotional content itself. Previous studies can provide limited information on this issue because stimulus material was limited to either fearful or angry or stereotypic material, whereas other researchers used emotional pictures instead of faces as stimuli. Similarly, much of the previous research outlined in this paper used functional imaging methods while the present studies delineate the influence of affective labels on face processing in real time with a resolution of milliseconds, by means of EEG methodology. Future studies using larger numbers of stimuli or different sets of stimuli, including emotion pictures or voices, are needed to test the valence specificity assumption. Inclusion of different stimulus materials besides faces will also show if effects occur across sensory modalities, which could improve our understanding of language-emotion-cognition interactions in real life social situations.

## ACKNOWLEDGMENTS

We thank Eunkyung Ho for help with data collection and analysis.

## FUNDING

Research was supported by the German Research Foundation (DFG, HE5880/3-1). This publication was funded by the German Research Foundation (DFG) and the University of Wuerzburg in the funding programme Open Access Publishing.

## REFERENCES

Bagby, R. M., Parker, J. D., and Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale–I. Item selection and cross-validation of the factor structure. *J. Psychosom. Res.* 38, 23–32. doi: 10.1016/0022-3999(94)90005-1

Barrett, L. F., Lindquist, K. A., and Gendron, M. (2007). Language as context for the perception of emotion. *Trends. Cogn. Sci.* 11, 327–332. doi: 10.1016/j.tics.2007.06.003

Barrett, L. F., Mesquita, B., and Gendron, M. (2011). Context in emotion perception. *Curr. Dir. Psychol. Sci.* 20, 286–290. doi: 10.1177/0963721411422522

Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565. doi: 10.1162/jocn.1996.8.6.551

Bentin, S., and Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cogn. Neuropsychol.* 17, 35–55. doi: 10.1080/026432900380472

Blau, V. C., Maurer, U., Tottenham, N., and McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behav. Brain Funct.* 3:7. doi: 10.1186/1744–9081-3-7

Blechert, J., Sheppes, G., Di Tella, C., Williams, H., and Gross, J. J. (2012). See what you think: reappraisal modulates behavioral and neural responses to social stimuli. *Psychol. Sci.* 23, 346–353. doi: 10.1177/0956797612438559

Borghi, A. M., Glenberg, A. M., and Kaschak, M. P. (2004). Putting words in perspective. *Mem. Cogn.* 32, 863–873. doi: 10.3758/BF03196865

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Branden, N. (1969). *The Psychology of Self-Esteem*. Palo Alto, CA: Nash Publishing Corp.

Brunye, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., and Taylor, H. A. (2009). When you and I share perspectives: pronouns modulate perspective taking during narrative comprehension. *Psychol. Sci.* 20, 27–32. doi: 10.1111/j.1467-9280.2008.02249.x

Cohen, J. R., Berkman, E. T., and Lieberman, M. D. (2013). "Intentional and incidental self-control in ventrolateral PFC," in *Principles of Frontal Lobe Function, 2nd Edn.*, eds D. T. Stuss and R. T. Knight (New York, NY: Oxford University Press), 417–440.

Deusinger, I. M. (1986). *Frankfurter Selbstkonzeptskalen (FSKN). Handanweisung.* Göttingen: Hogrefe.

Foti, D., and Hajcak, G. (2008). Deconstructing reappraisal: descriptions preceding arousing pictures modulate the subsequent neural response. *J. Cogn. Neurosci.* 20, 977–988. doi: 10.1162/jocn.2008.20066

Freudenthaler, H. H., Neubauer, A. C., Gabler, P.,Scherl, W. G., and Rindermann, H. (2008). Testing and validating the Trait Emotional Intelligence Questionnaire (TEIQue) in a German-speaking sample. *Pers. Individ. Diff.* 45, 673–678. doi: 10.1016/j.paid.2008.07.014

Fujiki, M., Spackman, M. P., Brinton, B., and Hall, A. (2004). The relationship of language and emotion regulation skills to reticence in children with specific language impairment. *J. Speech Lang. Hear. Res.* 47, 637–646. doi: 10.1044/1092-4388(2004/049)

Gendron, M., Lindquist, K. A., Barsalou, L., and Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion* 12, 314–325. doi: 10.1037/a0026007

Gratton, G., Coles, M. G. H., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* 55, 468–484. doi: 10.1016/0013-4694(83)90135-9

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823

Gross, J. J. (2002). Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology* 39, 281–291. doi: 10.1017/S0048577201393198

Gross, J. J., and John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *J. Pers. Soc. Psychol.* 85, 348–362. doi: 10.1037/0022-3514.85.2.348

Hajcak, G., Moser, J. S., and Simons, R. F. (2006). Attending to affect: appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion* 6, 517–522. doi: 10.1037/1528-3542.6.3.517

Hajcak, G., and Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cogn. Affect. Behav. Neurosci.* 6, 291–297. doi: 10.3758/CABN.6.4.291

Hariri, A. R., Bookheimer, S. Y., and Mazziotta, J. C. (2000). Modulating emotional response: effects of a neocortical network on the limbic system. *Neuroreport* 11, 43–48. doi: 10.1097/00001756-200001170-00009

Hautzinger, M., Bailer, M., Worall, H., and Keller, F. (1994). *Beck-Depressions-Inventar (BDI)*. Bern: Huber.

Hayes, A. M., and Feldman, G. (2004). Clarifying the construct of mindfulness in the context of emotion regulation and the process of change in therapy. *Clin. Psychol. Sci. Pr.* 11, 255–262.

Heinze, H. J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Munte, T. F., et al. (1994). temporal imaging and temporal imaging of brain activity during visual selective attention in humans. *Nature* 372, 543–546. doi: 10.1038/372543a0

Herbert, C., Deutsch, R., Platte, P., and Pauli, P. (2012). No fear, no panic: probing negation as a means for emotion regulation. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nss043. [Epub ahead of print].

Herbert, C., Pauli, P., and Herbert, B. M. (2011a). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Soc. Cogn. Affect. Neurosci.* 6, 653–661. doi: 10.1093/scan/nsq082

Herbert, C., Herbert, B. M., Ethofer, T., and Pauli, P. (2011b). His or mine? The time course of self-other discrimination in emotion processing. *Soc. Neurosci.* 6, 277–288. doi: 10.1080/17470919.2010.523543

Herbert, C., Herbert, B. M., and Pauli, P. (2011c). Emotional self-reference: brain structures involved in the processing of words describing one's own emotions. *Neuropsychologia* 49, 2947–2956. doi: 10.1016/j.neuropsychologia.2011.06.026

Herbert, B. M., Herbert, C., and Pollatos, O. (2011d). On the relationship between interoceptive awareness and alexithymia: Is interoceptive awareness related to emotional awareness? *J. Pers.* 79, 1149–1175. doi: 10.1111/j.1467-6494.2011.00717.x

Ingram, R. E. (1990). Self-focused attention in clinical disorders: review and a conceptual model. *Psychol. Bull.* 107, 156–176. doi: 10.1037/0033-2909.107.2.156

Izard, C. E. (2001). Emotional intelligence or adaptive emotions? *Emotion* 1, 249–257. doi: 10.1037/1528-3542.1.3.249

John, O. P., and Gross, J. J. (2004). Healthy and unhealthy emotion regulation: personality processes, individual differences, and life span development. *J. Pers.* 72, 1301–1333. doi: 10.1111/j.1467-6494.2004.00298.x

Junghöfer, M., Bradley, M. M., Elbert, T. R., and Lang, P. J. (2001). Fleeting images: a new look at early emotion discrimination. *Psychophysiology* 38, 175–178. doi: 10.1111/1469-8986.3820175

Katz, R. C. (1980). "Perception of Facial Affect in Aphasia," in *Clinical Aphasiology: Proceedings of the Conference 1980* (Bar Harbor, ME: BRK Publishers), 78–80.

Kircanski, K., Lieberman, M. D., and Craske, M. G. (2012). Feelings into words: contributions of language to exposure therapy. *Psychol. Sci.* 23, 1086–1091. doi: 10.1177/0956797612443830

Kissler, J., Herbert, C., Peyk, P., and Junghofer, M. (2007). Buzzwords: early cortical responses to emotional words during reading. *Psychol. Sci.* 18, 475–480. doi: 10.1111/j.1467-9280.2007.01924.x

Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biol. Psychol.* 45, 19–56. doi: 10.1016/S0301-0511(96)05221-0

Laux, L., Glanzmann, P., Schaffner, P., and Spielberger, C. D. (1981). *Das State-Trait-Angstinventar (STAI)*. Weinheim: Beltz.

Lieberman, M. D. (2007). "The X- and C-systems: the neural basis of automatic and controlled social cognition," in *Fundamentals of Social Neuroscience*, eds E. Harmon-Jones and P. Winkelman (New York, NY: Guilford), 290–315.

Lieberman, M. D. (2011). "Why symbolic processing of affect can disrupt negative affect: social cognitive and affective neuroscience investigations," in *Social Neuroscience: Toward Understanding the Underpinnings of The Social Mind*, eds A. Todorov, S. T. Fiske, and D. Prentice (New York, NY: Oxford University Press), 188–209.

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., and Way, B. M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. *Psychol. Sci.* 18, 421–428. doi: 10.1111/j.1467-9280.2007.01916.x

Lieberman, M. D., Inagaki, T. K., Tabibnia, G., and Crockett, M. J. (2011). Subjective responses to emotional stimuli during labeling, reappraisal, and distraction. *Emotion* 11, 468–480.

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., and Russell, J. A. (2006). Language and the perception of emotion. *Emotion* 6, 125–138. doi: 10.1037/1528-3542.6.1.125

Lindquist, K. A., and Gendron, M. (2013). What's in a word: language constructs emotion perception. *Emot. Rev.* 5, 66–71. doi: 10.1177/1754073912451351

Lundqvist, D., Flykt, A., and Öhman, A. (1998). *The Karolinska Directed Emotional Faces – KDEF.* Stockholm: CD-ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet. ISBN 91-630-7164-9.

Macnamara, A., Foti, D., and Hajcak, G. (2009). Tell me about it: neural activity elicited by emotional pictures and preceding descriptions. *Emotion* 9, 531–543. doi: 10.1037/a0016251

Maister, L., Tsiakkas, E., and Tsakiris, M. (2013). I feel your fear: shared touch between faces facilitates recognition of fearful facial expressions. *Emotion* 13, 7–13. doi: 10.1037/a0030884

Moran, J. M., Heatherton, T. F., and Kelley, W. M. (2009). Modulation of cortical midline structures by implicit and explicit self-relevance evaluation. *Soc. Neurosci.* 4, 197–211. doi: 10.1080/17470910802250519

Moran, T. P., Jendrusina, A. A., and Moser, J. S. (2013). The psychometric properties of the late positive potential during emotion processing and regulation. *Brain Res.* 1516, 66–67. doi: 10.1016/j.brainres.2013.04.018

Moratti, S., Saugar, C., and Strange, B. A. (2011). Prefrontal-occipitoparietal coupling underlies late latency human neuronal responses to emotion. *J. Neurosci.* 31, 17278–17286. doi: 10.1523/JNEUROSCI.2917-11.2011

Morin, A. (2005). Possible links between self-awareness and inner speech: Theoretical background, underlying mechanisms, and empirical evidence. *J. Conscious. Stud.* 12, 115–134.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain–a meta-analysis of imaging studies on the self. *Neuroimage* 31, 440–457. doi: 10.1016/j.neuroimage.2005.12.002

Ochsner, K. N., Knierim, K., Ludlow, D. H., Hanelin, J., Ramachandran, T., Glover, G., et al. (2004). Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J. Cogn. Neurosci.* 16, 1746–1772. doi: 10.1162/0898929042947829

Olofsson, J. K., Nordin, S., Sequeira, H., and Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biol. Psychol.* 77, 247–265. doi: 10.1016/j.biopsycho.2007.11.006

Paulus, C. (2009). *Der Saarbrücker Persönlichkeitsfragebogen SPF(IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index.* Available online at: http://psydok.sulb.uni-saarland.de/volltexte/2009/2363/

Pennebaker, J. W., and Beall, S. K. (1986). Confronting a traumatic event: toward an understanding of inhibition and disease. *J. Abnorm. Psychol.* 95, 274–281. doi: 10.1037/0021-843X.95.3.274

Pennebaker, J. W., and Chung, C. K. (2007). "Expressive writing, emotional upheavals, and health," in *Handbook of Health Psychology,* eds H. Friedman and R. Silver (New York, NY: Oxford University Press), 263–284.

Pennebaker, J. W., and Chung, C. K. (2011). "Expressive writing: connections to physical and mental health," in *The Oxford Handbook of Health Psychology*, ed H. Friedman (New York, NY: Oxford University Press), 417–437. doi: 10.1093/oxfordhb/9780195342819.013.0018

Rauss, K., Pourtois, G., Vuilleumier, P., and Schwartz, S. (2012). Effects of attentional load on early visual processing depend on stimulus timing. *Hum. Brain Mapp.* 33, 63–74. doi: 10.1002/hbm.21193

Righart, R., and de Gelder, B. (2006). Context influences early perceptual analysis of faces–an electrophysiological study. *Cereb. Cortex* 16, 1249–1257. doi: 10.1093/cercor/bhj066

Ruby, P., and Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *J. Cogn. Neurosci.* 16, 988–999. doi: 10.1162/0898929041502661

Schmitz, T. W., and Johnson, S. C. (2007). Relevance to self: a brief review and framework of neural systems underlying appraisal. *Neurosci. Biobehav. Rev.* 31, 585–596. doi: 10.1016/j.neubiorev.2006.12.003

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., and Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology* 37, 257–261. doi: 10.1111/1469-8986.3720257

Schupp, H. T., Flaisch, T., Stockburger, J., and Junghofer, M. (2006). Emotion and attention: event-related brain potential studies. *Prog. Brain Res.* 156, 31–51. doi: 10.1016/S0079-6123(06)56002-9

Schupp, H. T., Öhman, A., Junghofer, M., Weike, A. I., Stockburger, J., and Hamm, A. O. (2004). The facilitated processing of threatening faces: an ERP analysis. *Emotion* 4, 189–200. doi: 10.1037/1528-3542.4.2.189

Seih, Y.-T., Chung, C. K., and Pennebaker, J. W. (2011). Experimental manipulations of perspective taking and perspective switching in expressive writing. *Cogn. Emot.* 25, 926–938. doi: 10.1080/02699931.2010.512123

Shi, Z., Zhou, A., Liu, P., Zhang, P., and Han, W. (2011). An EEG study on the effect of self-relevant possessive pronoun: self-referential content and first-person perspective.

*Neurosci. Lett.* 494, 174–179. doi: 10.1016/j.neulet.2011.03.007

Tabibnia, G., Lieberman, M. D., and Craske, M. G. (2008). The lasting effect of words on feelings: words may facilitate exposure effects to threatening images. *Emotion* 8, 307–317. doi: 10.1037/1528-3542.8.3.307

Walla, P., Duregger, C., Greiner, K., Thurner, S., and Ehrenberger, K. (2008). Multiple aspects related to self-awareness and the awareness of others: an electroencephalography study. *J. Neural Transm.* 115, 983–992. doi: 10.1007/s00702-008-0035-6

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Zhou, A., Shi, Z., Zhang, P., Liu, P., Han, W., Wu, H., et al. (2010). An ERP study on the effect of self-relevant possessive pronoun. *Neurosci. Lett.* 480, 162–166. doi: 10.1016/j.neulet.2010.06.033

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 23 March 2013; accepted: 01 July 2013; published online: 23 July 2013.*
*Citation: Herbert C, Sfärlea A and Blumenthal T (2013) Your emotion or mine: labeling feelings alters emotional face perception—an ERP study on automatic and intentional affect labeling. Front. Hum. Neurosci. 7:378. doi: 10.3389/fnhum.2013.00378*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 17 June 2013
doi: 10.3389/fnhum.2013.00281

# Age and gender dependent development of Theory of Mind in 6- to 8-years old children

**Cecilia I. Calero[1,2]*†, Alejo Salles[1,2†], Mariano Semelman[1,2] and Mariano Sigman[1,2,3]**

[1] Laboratorio de Neurociencia Integrativa, Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
[2] IFIBA, CONICET, Buenos Aires, Argentina
[3] Universidad Torcuato Di Tella, Buenos Aires, Argentina

**Edited by:**
Agustin Ibanez, Institute of Cognitive Neurology, Argentina

**Reviewed by:**
David Huepe, Universidad Diego Portales, Chile
Marina L. Puglisi, University of Oxford, UK

***Correspondence:**
Cecilia I. Calero, Laboratorio de Neurociencia Integrativa, Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Intendente Güiraldes 2160, Pabellón I, Ciudad Universitaria, C1428EGA Buenos Aires, Argentina
e-mail: calero@gmail.com

†These authors have contributed equally to this work.

The ability to attribute different mental states to distinct individuals, or Theory of Mind (ToM), is widely believed to be developed mostly during preschool years. How different factors such as gender, number of siblings, or coarse personality traits affect this development is not entirely agreed upon. Here, we introduce a computerized version of the scaled ToM suite of tasks introduced by Wellman and Liu (2004), which allows us to meaningfully test ToM development on children 6 to 8-years old. We find that kids this age are still not entirely proficient in all ToM tasks, and continue to show a progression of performance with age. By testing this new age range, too, we are able to observe a significant advantage of girls over boys in ToM performance. Other factors such as number of siblings, birth order, and coarse personality traits show no significant relation with the ToM task results. Finally, we introduce a novel way to quantify the scaling property of the suite involving a sequence of set inclusions on one hand and a comparison between specially tailored sets of logistic models on the other. These measures confirm the validity of the scale in the 6- to 8-years old range.

**Keywords: Theory of Mind, scaling, mental states, development, gender differences**

## INTRODUCTION

Theory of Mind (ToM) is an important cognitive skill that refers broadly to our capacity to understand others' mental states including beliefs, desires, and knowledge, and the ability to comprehend that these may differ from our own (Premack and Woodruff, 1978). A paradigmatic example of a task requiring a well-developed ToM is that of false belief, which involves understanding that an agent might have a wrong representation of external reality, and act according to that representation (Baron-Cohen et al., 1985). ToM is deeply integrated with other cognitive domains and abilities. Among other factors, ToM development has been related to the success on executive function tests, and various assessments of language and social experience (Dunn et al., 1991; Dunn, 1995; Astington and Jenkins, 1999; Carlson and Moses, 2001; Apperly, 2012).

Over the past two decades, developmental changes in children's understanding of others' minds have been the focus of intense research (Wellman and Woolley, 1990; Dunn et al., 1991; Gopnik and Slaughter, 1991; Flavell, 1999; Astington and Jenkins, 1999; Meltzoff, 1999; Wellman et al., 2001). Most of the work performed to date was done in preschool children (3 to 5-years old) and coarsely agrees in that the basic aspects of ToM are mostly developed within this age range [Bartsch and Wellman, 1995; see Flavell (1999), for a review]. However, some studies argue that ToM continues to develop and change throughout life (Bosacki and Astington, 1999; Apperly, 2012; Devine and Hughes, 2012; Moran, 2013). Rai and Mitchell's (2004) study has shown that there is still considerable instability in understanding false beliefs in 5-years old, especially when the false belief scenario is framed in relation to a person's conscious choice or decision, rather than a physical object. Furthermore, Dumontheil et al.' (2010) results suggest that ToM improves between late adolescence and adulthood and even if ToM tasks are passed by age 4; their data indicate that the interaction between understanding others' mind and executive functions continues to develop in late adolescence (Dumontheil et al., 2010). Bosacki and Astington (1999) used ambiguous social vignettes followed by questions to assess the understanding of particular aspects of other's mental states in a study with preadolescent children. Their study was conceived from a ToM perspective in order to quantify preadolescents' mentalizing abilities and their results partially support that, individual differences in preadolescents' ability to understand the thoughts and emotions of others would be related to their social competence (Bosacki and Astington, 1999).

Most studies on ToM have not addressed issues of gender, family environment, and measures of temperament. A slight advantage of preschool girls on emotion understanding and false belief task performance has been observed before (Banerjee, 1997; Charman and Clements, 2002; Walker, 2005), nevertheless, most previous studies have found no significant gender differences on ToM development (Hughes and Dunn, 1998; Charman and Clements, 2002; Walker, 2005; Mathieson and Banerjee, 2011; Devine and Hughes, 2012). On the other hand, in studies carried out in preadolescence, girls performed significantly higher on the social understanding task (ToM) than boys independent of vocabulary ability (Bosacki and Astington, 1999). These results support Hatcher et al.' (1990) findings, in which girls scored higher than boys on social understanding tasks across grades 4 through 12. Further, the recent study in adolescents by Ibanez et al. (2013) presents a model that shows the direct effect of empathy, sex, and fluid intelligence on ToM. Only recent studies have found some relations between scores on false belief tasks and preschool children's family environments (Perner et al., 1994; Farhadian et al., 2010). However, there is no consensus on whether the amount of siblings or the birth order influences the development of ToM (Azmitia and Hesser, 1993; Lewis et al., 1996; Ruffman et al., 1998; Cutting and Dunn, 1999; Farhadian et al., 2010).

A complementary aim of most developmental studies of ToM consists in understanding the sequential unfolding of abilities underlying a full ToM. To date there is consensus in the notion that a child will correctly judge a person's desires before she can correctly judge her beliefs, and that she will be able to grasp that an agent might have a belief different from her own first if she doesn't know the true state of affairs, and only later if she does know what reality really is like (false belief). Cutting and Dunn (1999), Wellman and Liu (2004), Wellman et al. (2001) and, more recently, Wellman et al. (2011) have investigated this progression of abilities. In particular, Wellman and Liu (2004) proposed a suite of ToM tasks, based upon a meta-analysis of the literature of ToM developmental studies, and tested it in children from 3 to 5-years of age. Their results suggest that the abilities underlying ToM are attained progressively, and can thus be tested individually by an ordered suite such that a child capable of correctly performing a certain task in the suite should also be able to correctly perform all preceding tasks. Wellman and Liu's scaled suite of tasks has been subsequently employed to pinpoint cultural differences in the development of ToM (Wellman et al., 2006, 2011; Shahaeian et al., 2011).

In this work, we implement and test a computer version of Wellman and Liu's (2004) ToM suite in children in the 6 to 8-years old range. The aim of this work is threefold: first, we test whether it is possible to use the suite to test ToM development in older kids and check if the scaling seen in preschoolers is still valid in our version of the test. The age range chosen in this work is sometimes overlooked in the literature, even if it has been shown that children do not understand metaphor or irony before the age of six to seven (Ackerman, 1981)—-two behaviors that entail the capacity to go beyond the literal meaning of a statement—- and that they cannot reliably distinguish jokes from lies before age 6 to 7-years (Sullivan et al., 1995). In accordance, our first hypothesis is that ToM progression of Wellman and Liu's suite will still be present in children 6 to 8-years old.

Second, given that gender differences might be expected in ToM, and in line with the gender intensification hypothesis (Hill and Lynch, 1983), which establishes that gender differences increase in time because of growing pressure to conform to traditional gender-role stereotypes, we hypothesized that in a slightly older group of children—-in relation to the usually age range explore in the literature (3 to 5-years old)—- gender may have an appreciable effect on ToM performance.

The third and final aim of this work is to develop a novel analysis to quantify the validity of the scaling in the suite. The method we use has two parts: one uses set inclusions to quantify the extent to which the data differ from a perfect scaling, while the other involves the comparison between specially tailored sets of logistic models. The difference in prediction power among these sets of models gives another measure of the scaling quality.

## MATERIALS AND METHODS

### PARTICIPANTS

Seventy-six first (36) and second (40) graders [mean age: 7-years and 3 months (86.5 months); range from 6-years and 1 month to 8-years and 7 months] participated in the study. There were 42 boys and 34 girls in the sample, all of a high socioeconomic status and attending a well reputed private bilingual school in Buenos Aires. The school in which the study was performed approved the research and all children's parents or legal guardians gave signed voluntary consent. The consent form, presented to the caregivers supplemented with a note which explained the procedure, was previously authorized by the *Centro de Educación Médica e Investigaciones Clínicas "Norberto Quirno"* (CEMIC)'s Ethical Committee.

### ToM SUITE

Wellman and Liu's ToM suite of tasks is thoroughly described in the appendix of the original paper (Wellman and Liu, 2004). Briefly, the tasks involved in our version are: (1, DD) Diverse Desires: the child judges that two persons (her vs. someone else) have different desires about the same objects; (2, DB) Diverse Beliefs: the child judges that two persons (her vs. someone else) have different beliefs about the same object, while she does not know which belief is the right one; (3, KA) Knowledge Access: the child sees what is in a box and judges the knowledge of another person who does not see what is in it; (4, FB) Contents False Belief: the child judges another person's false belief about what is in a distinctive container while she (the child) knows what actually is inside the container; (5, EFB) Explicit False Belief: the child judges how someone will search, given that person's mistaken belief, and (6, BE) Belief vs. Emotion: the child judges how a person will feel given a belief that is mistaken. All tasks involve a control question which is used to make sure that the child understood the task, and a target question, which evaluates their performance. Although all six tasks were used for studying the influence of diverse factors in ToM development, we note that only the first four (DD, DB, KA, and FB) are involved in the progressive suite, and thus all scaling tests were performed only on these. Finally, we note that Wellman and Liu's original version of the suite included also a hidden emotion task. As this task involves two target questions instead of one target and one control, we chose to leave it out in order to facilitate direct comparison (the random choice performance baseline for this task is 33% instead of the rest of the tasks' two choice 50%). The implementation of the suite is depicted in **Figure 1**.

In Wellman and Liu's version of the suite, some tasks were presented using toy figurines while others involved drawings. This makes it difficult both to preserve the parallels among the different tasks and to compare the results with those involving other

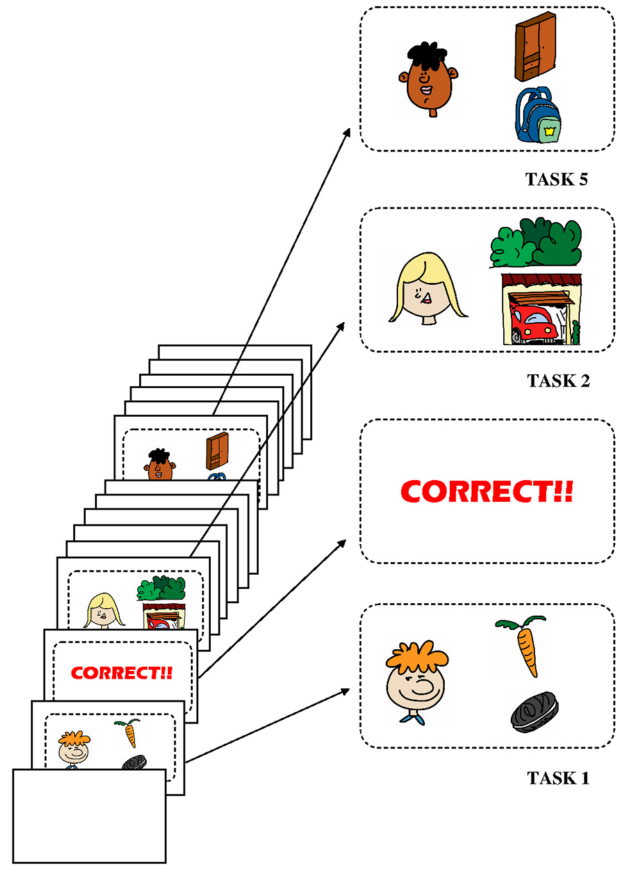


**FIGURE 1 | Description of the test.** The test was delivered to each child in a novel computer format. The experimenter first said to each child: "There are six levels; you will have to answer correctly in each one of them to win the game. Let's go!" Then, for each task, the experimenter explained the task to the child, who could choose any of the options given. There were no wrong answers, such that, independently of the answer, a "CORRECT" screen appeared next. Three of the tasks are shown in the figure.

methods found in the literature. Further, following our aim of carrying the scaled ToM suite to an older age range, we formulated it as an engaging computer game, which not only unifies the presentation format across all tasks, but also reduces the experimenter's involvement.

### PROCEDURE

Children were tested in a quiet room in the school by one of two adult experimenters. The six tasks were presented in the increasing difficulty order proposed by Wellman and Liu (2004). All the children that participated in the study correctly answered the control question in each task.

Teachers completed the short form of the Child Behavioral Questionnaire (CBQ) for all their students. This allowed us to measure child's temperament along three broad dimensions: (1) Extraversion/Surgency, (2) Negative Affectivity, and (3) Effortful Control (Rothbart et al., 2001; Putnam and Rothbart, 2006). They also filled a second form that included the family background data (birth order and number of siblings), age, and gender of each child.

### DATA ANALYSIS

In order to summarize the children's performance in the ToM suite we compute for each child her $z$-score, defined as the amount of correct target answers. Alternative measures were also tested, in which the contribution of each task to the total score was weighed either progressively with the task number or with a factor equal to one minus the observed mean performance for that task, and then all contributions summed together in a final score. The results obtained are insensitive to the scoring scheme chosen; we hence stick to the $z$-score above.

Throughout the analysis, we used non-parametric permutation tests in order to assess the significance of results. In each case, we randomly shuffled the assignments between predictor and dependent variables, to produce a surrogate version of the data. The relevant quantity (for instance correlation) was then evaluated for this surrogate data. By iterating this procedure many times (typically a thousand), we obtained the significance level of the result.

Apart from studying the correlations in the data, we built a logistic model to assess kids' responses for all tasks. With this model, we can study the effect of the different factors in the whole set of responses, without limiting the analysis to a particular definition of score. Nevertheless, a certain amount of independence among the factors entering the model is required in order to correctly interpret the results, and hence we restrict the model to age, gender, and number of siblings and birth order (considering these last two as a single factor in order to account for their interdependence). CBQ scores, on the other hand, were left out, since they correlate mildly with gender (ex. 0.4 linear correlation between surgency and gender).

Apart from gender, age, number of siblings, and birth order as predictors, the model includes dummy variables for discerning among the six tasks in the suite. Since each kid responds to all six tasks, we also need to index the subjects. As it turns out, age, sex, sibling amount, and sibling order taken together are almost enough to identify all subjects. There are however, five cases in which these repeat, so we resolve them by adding an extra indicator variable (i.e., order in which they took the experiment).

To measure the importance of each factor in predicting the results, we compare the full model with that with the factor in question removed. The difference between the log likelihoods of both models follows a $\chi^2$ distribution [apart from a factor of 2, see Stevenson (2004)], and we can hence evaluate the corresponding $p$-value. The $p$-values thus obtained are in agreement with those computed with the full model under the assumption of normally distributed errors.

## RESULTS

### ToM PERFORMANCE IN 6 TO 8-YEARS OLD CHILDREN

Despite the fact that some studies indicate that the development of a full ToM continues all throughout life (Devine and Hughes, 2012; Moran, 2013), almost all research has focused in 3 to 5-years old children. We took a step further to contribute to elucidate these notions and we examined age, gender, and family background influence on ToM development in the age range of 6 to 8 years old. **Figure 2** shows the average performance for the four tasks involved in the scaling for all children (thick black bars).

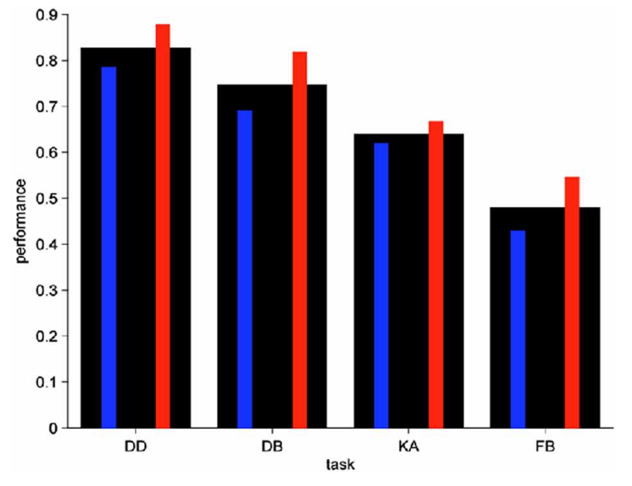


**FIGURE 2 | Kids' average performance.** Fraction correct, mean, thick black bars; boys, left thin blue bars; girls, right thin red bars.

The average performance for our first and second graders is in the same range as that of the preschoolers studied by Wellman and Liu, suggesting that the smaller intervention of the experimenter enabled by the computer platform allows for the testing of older kids without saturating the suite. Average performance for tasks EFB and BE was 0.64 and 0.6, respectively, also similar to that of Wellman and Liu's preschoolers. These tasks were not included in the graph to emphasize the progression effect in the first four.

In accordance with our first hypothesis that ToM progression of Wellman and Liu's suite will still be present in 6 to 8-years old, we observed an increase in ToM proficiency with age, with $z$-scores significantly correlated with age ($r = 0.334, p = 0.003$ permutation test), This result stresses even further the fact that the test is effective in the new age range, and shows that the general performance drop is not merely due to a statistical fluctuation.

Average performance divided by gender is also shown in **Figure 2** (boys in left blue thin bars and girls in right red thin bars). For task EFB we have average performances of 0.7 for girls and 0.6 for boys, while for the BE task we obtain 0.76 for girls and 0.48 for boys. We observe a clear effect of gender in ToM performance, with girls performing significantly better than boys in all tasks ($p = 0.037$ permutation test, grouped tasks). This gender effect is task-independent, as we show using the logistic model described in the following section.

We also studied the effect of family background and child temperament, factors whose potential influence in ToM performance has been previously discussed in the literature (Dunn et al., 1991; Perner et al., 1994; Dunn, 1995; Farhadian et al., 2010). Contrary to age and gender findings and in agreement with previous data (Lewis et al., 1996; Cutting and Dunn, 1999) we find no effect of sibling amount ($r = 0.034$, $p = 0.766$, permutation test) or birth order ($r = -0.043$, $p = 0.729$, permutation test) in ToM performance.

Similarly, coarse personality traits as evaluated by the CBQ do not correlate significantly with the $z$-score (Surgency: $r = -0.169$, $p = 0.149$; Negative Affect: $r = 0.197$, $p = 0.093$; Effortful Control: $r = 0.118$, $p = 0.315$; permutation tests).

The logistic model analysis fully supports the results discussed above. By taking one factor at a time out of the model, we can see how relevant each factor is in explaining the kids' responses (see Materials and Methods). The computed $p$-values are: age, $p = 0.002$; sex, $p = 0.015$; siblings (amount and order), $p = 0.978$. By testing an expanded model including a joint sex-task factor, we see that these two variables do not interact ($p = 0.47$). Hence, in accordance with our second hypothesis we observed a gender effect in ToM which was not observed for the same task in preschoolers. The other demographic or individual variables had no effect in performance.

## ToM SUITE PROGRESSION

We now turn to quantify the extent to which the ToM suite embodies a progressive test, that is, one in which in order to correctly resolve a given step, one must have the abilities required to solve all previous steps. As we mentioned before, we only consider tasks 1–4 (DD, DB, KA, and FB) for this part of the analysis, since it is only these that are involved in the hierarchy.

The progression in difficulty of ToM performance is evident from **Figure 2**, which shows a diminishing performance with increasing task number. An average decrease in performance, however, does not necessarily imply a sequential process. It is still possible that a group of children is capable of successfully performing Task 1 ($G_1$) and a smaller group is capable of successfully performing Task 2 ($G_2$) but that these groups have no intrinsic relation other than their difference in size, i.e., $|G_2| < |G_1|$, where $| \cdot |$ denotes set cardinality. The condition of strict sequential dependence in performance implies that all kids in $G_2$ are also in $G_1 (G_2 \subset G_1)$, in other words, that children succeeding in Task 2 also succeed in Task 1. This logic extends for all tasks from 1 to 4, such that $G_j \subset G_i$ for all $i < j$ in 1 to 4.

**Figure 3** depicts this sequence of inclusions for two extreme cases, **(A)** one in which there is no progression at all, and **(B)** one in which the hierarchy is perfect, along with our actual result **(C)**. In this figure, each task is represented by a square whose area encodes number of kids who successfully passed the task and area overlap indicates kids passing both corresponding tasks. For clarity, we only represent overlaps of successive tasks. Out of 76 kids, 62 passed DD, 56 passed DB, 48 passed KA, and 36 passed FB. Of these, 6 passed DB but not DD (11% of those that passed DB), 11 passed KA but not DB (23% of those that passed KA), and 10 passed FB but not KA (28% of those that passed KA). In a perfect hierarchy, there would be no such cases. These cases, however, constitute an expected fluctuation. In order to test this statistically, we take as a figure of merit the sum of cases in which a kid passed a task without passing the previous one, which is 27 in our case (lower values of this number represent better scalings). We then perform a bootstrap procedure in which we shuffle both kids and tasks, and obtain a surrogate value for this figure of merit. We find a better (lower) value in less than 1% of the cases, with an actual $p$-value of 0.059.

We refine this analysis by examining three sets of logistic models for the kids' responses, taken one task of the ToM suite at a time. The first is a set of Pointwise models, one for each task,

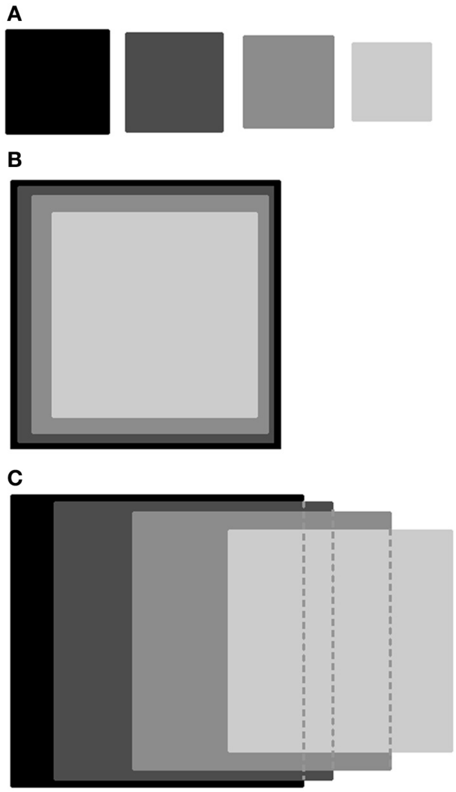


**FIGURE 3 | Task progression as set inclusions.** Tasks DD, DB, KA, and FB are represented as squares in decreasing gray level. Square area represents amount of kids successfully completing each task, and area overlap among successive tasks quantifies the amount of kids that responded correctly for both tasks. **(A)** Illustrative lack of scaling. **(B)** Illustrative perfect scaling. **(C)** Actual experimental results.

which include as predictors age, gender and sibling amount and order. The second is a set of Markov models, again one for each task, which comprise besides the factors in the Pointwise models the response for the previous task in the scale. For the first task, these two models are identical. Finally, the third set is one of Cumulative models, also one for each task, incorporating the same predictors as the Pointwise models plus the whole history of previous responses for each kid up to the current task. We point out that for the second task of the scale the Markov and Cumulative models coincide.

We compare pairwise the model sets described before, getting a *p*-value for the extra variables in a model to be explanatory. The summarized results are presented in **Table 1**. We find that, comparing the Pointwise and Markov models, the response to the first task is a good predictor of the response to the second. This is consistent with the scaling property: a child that did not pass the first task is very likely to also not pass the second, hence making the Markov model more powerful than the Pointwise. We note however, that even in a perfect scaling, we would have some amount of unpredictability, since we do not know at which point the kids will start failing the tasks, so that even having the previous task information will not help in predicting the outcome of the next. This is indeed what we see when comparing the Markov and Pointwise models for the subsequent tasks, where more kids begin to fail.

**Table 1 | Model sets comparison.**

| Task | *p* Markov vs. Cumulative | *p* Pointwise vs. Markov |
|---|---|---|
| 2 (DB) | – | 0.006 |
| 3 (KA) | 0.768 | 0.839 |
| 4 (FB) | 0.724 | 0.322 |

We can address this issue by further comparing the Markov and the Cumulative model sets (different only for task 3 and up). This comparison shows that the extent to which one can predict the outcome in a certain task by knowing the result for the previous task is not improved by knowing further previous results, as should be the case in a perfect hierarchy. In other words, knowing the result for the first task does not add information to knowing the result for the second task, if we are to predict the response to the third one, and similarly for task 4. This further stresses the scaling property of the suite.

## DISCUSSION

While many studies argue that by the age of four most normally developing children have already acquired an understanding of the mind; others instead have shown that ToM continues to mature at older ages Bosacki and Astington, 1999; Dumontheil et al., 2010; Devine and Hughes, 2012; Moran, 2013. In the present study we implement the ToM suite of tasks by means of a game in a computer platform, thus diminishing the experimenter's involvement. Corroborating Wellman and Liu (2004) previous results with preschool kids, we also found that the progressive and sequential effect of the suite could remain a major factor in older kids, revealing a hierarchy of nested processes of ToM in the 6 to 8-years range. However, because this ToM suite of tasks was never tested with the original age range and procedures in Argentina more studies need to be done.

We successfully apply the suite in older kids and to reach this conclusion we introduced a novel method to quantify the scaling property of the suite. This new proposal involves on one hand an intuitive quantification through set inclusions, and, on the other, a thorough comparison of a variety of logistic models including a varying amount of previous results as predictors for the outcome of a certain task. Both methods provide strong support for the scaling, and validate its use in the new age range tested.

Most studies on ToM have not addressed issues of gender. By testing older kids, we could examine the hypothesis that gender differences in ToM proficiency may develop late. This would be in accordance to the gender intensification hypothesis (Hill and Lynch, 1983), which predicts that gender differences increase in time because of increased pressure to conform to traditional gender-role stereotypes. Here, we conclusively showed a strong effect indicating that girls perform significantly better than boys for all ToM tasks in the age range tested.

Although some studies have proposed facilitative effects of (older) siblings that may operate via shared experiences of pretend play and deception, and talk about feelings and internal mental states (Perner et al., 1994; Ruffman et al., 1998;

Peterson and McAlister, 2006), other work has found no relation between ToM performance an number of siblings or birth order (Cutting and Dunn, 1999; Cole and Mitchell, 2000; Hughes and Ensor, 2005). In our study, the family background included in the analysis did not correlate with ToM performance, contributing to the idea that birth order and number of siblings would not be related to the development of ToM. However, we need to take into consideration that our results come from children in a new age range (6 to 8-years old), while previous evidence related to this particular topic was concentrated on preschoolers.

Similarly, there is some evidence supporting a relation between understanding of false belief and emotion and peer-related social competence (Dunn et al., 1991; Dunn, 1995), although the individual influence of child temperament on ToM has not yet been studied. The relation between ToM performance and emotion and temperament is thus surprisingly unclear, despite the importance of both domains to social interactions (Cutting and Dunn, 1999). Only recent studies have found some relations between a direct relation of gender and ToM mediated by empathy (Ibanez et al., 2013). Here, we evaluated the relation between three personality traits as quantified by the CBQ and ToM performance. None of these traits appears to have an impact in the understanding of other's minds. We should point out, however, that given the lack of memory or general intelligence measure, the use of non-standardized measures and the ethnical homogeneity of our sample, among other factors, interpretations of the present findings are to be made with caution.

Finally, we introduced a novel way in which to quantify the scaling property of the suite. This new proposal involves on one hand an intuitive quantification through set inclusions, and, on the other, a thorough comparison of a variety of logistic models including a varying amount of previous results as predictors for the outcome of a certain task. Both methods provide strong support for the scaling, and validate its use in the new age range tested.

## ACKNOWLEDGMENTS

This research was supported by a grant UBACyT 2010-2012 (N° 20020090100286), by a grant UBACyT joven (N° 20020110300047) and received support from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Mariano Sigman is sponsored by the James McDonnell Foundation. We gratefully acknowledge the helpful efforts of the school and, especially, the children who participated with the generous consent of their parents. Authors also thank Joaquín Ais and Ariel Zylberberg for fruitful discussions and useful comments on the manuscript.

## REFERENCES

Ackerman, B. (1981). Young children's understanding of a false utterance. *Dev. Psychol.* 31, 472–480. doi: 10.1037/0012-1649.17.4.472

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *Q. J. Exp. Psychol.* 65, 825–839. doi: 10.1080/17470218.2012.676055

Astington, J. W., and Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Dev. Psychol.* 35, 1311–1320. doi: 10.1037/0012-1649.35.5.1311

Azmitia, M., and Hesser, J. (1993). Why siblings are important agents of cognitive development: a comparison of siblings and peers. *Child Dev.* 64, 430–444. doi: 10.2307/1131260

Banerjee, M. (1997). Hidden emotions: pre-schoolers' knowledge of appearance-reality and emotion display rules. *Soc. Cogn.* 15, 107–132. doi: 10.1521/soco.1997.15.2.107

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46. doi: 10.1016/0010-0277(85)90022-8

Bartsch, K., and Wellman, H. M. (1995). *Children Talk about the Mind.* New York, NY: Oxford University Press.

Bosacki, S., and Astington, J. W. (1999). Theory of mind in preadolescence: relations between social understanding and social competence. *Soc. Dev.* 8, 237–255. doi: 10.1111/1467-9507.00093

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Charman, T., and Clements, W. (2002). Is there a gender difference in false belief development? *Soc. Dev.* 11, 1–10. doi: 10.1111/1467-9507.00183

Cole, K., and Mitchell, P. (2000). The role of social experience in the development of executive functions. *Br. J. Dev. Psychol.* 18, 279–295. doi: 10.1348/026151000165698

Cutting, A. L., and Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Dev.* 70, 853–865. doi: 10.1111/1467-8624.00061

Devine, R. T., and Hughes, C. (2012). Silent films and strange stories: theory of mind, gender, and social experiences in middle childhood. *Child Dev.* 84, 989–1003. doi: 10.1111/cdev.12017

Dumontheil, I., Apperly, I. A., and Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Dev. Sci.* 13, 331–338. doi: 10.1111/j.1467-7687.2009.00888.x

Dunn, J. (1995). Children as psychologists: the later correlates of individual differences in understanding of emotions and other minds. *Cogn. Emotion* 9, 187–201. doi: 10.1080/02699939508409008

Dunn, J., Bhown, J., Slomkowski, C., Tesla, C., and Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Dev.* 62, 1352–1366. doi: 10.2307/1130811

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M., Kumar, V., and Gazanizad, N. (2010). Theory of mind, birth order, and siblings among preschool children. *Am. J. Sci. Res.* 7, 25–35.

Flavell, J. H. (1999). Cognitive development: children's knowledge about the mind. *Annu. Rev. Psychol.* 50, 21–45. doi: 10.1146/annurev.psych.50.1.21

Gopnik, A., and Slaughter, V. (1991). Young children's understanding of charges in their mental states. *Child Dev.* 62, 98–110. doi: 10.2307/1130707

Hatcher, R., Hatcher, S., Berlin, M., Okla, K., and Richards, J. (1990). Psychological mindedness and abstract reasoning in late childhood and adolescence: an exploration using new instruments. *J. Youth Adolesc.* 19, 307–325. doi: 10.1007/BF01537075

Hill, J., and Lynch, M. (1983). "The intensification of gender-related role expectations during early adolescence," in *Girls at Puberty: Biological and Psychosocial Perspectives,* eds J. Brooks-Gunn and A. Peterson (New York, NY: Plenum), 201–228.

Hughes, C., and Dunn, J. (1998). Understanding mind and emotion: longitudinal associations with mental-state talk between young friends. *Dev. Psychol.* 34, 1026–1037. doi: 10.1037/0012-1649.34.5.1026

Hughes, C., and Ensor, R. (2005). Theory of mind and executive function: a family affair? *Dev. Neuropsychol.* 28, 645–668. doi: 10.1207/s15326942dn2802_5

Ibanez, A., Huele, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., and Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Pers. Individ. Dif.* 54, 616–621. doi: 10.1016/j.paid.2012.11.022

Lewis, C., Freeman, N. H., Kyriadkidou, C., Maridaki-Kassotaki, K., and Berridge, D. (1996). Social influences on false belief access: specific sibling influences or general apprenticeship? *Child Dev.* 67, 2930–2947. doi: 10.2307/1131760

Mathieson, K., and Banerjee, R. (2011). Peer play, emotion understanding, and socio moral explanation: the role of gender. *Br. J. Dev. Psychol.* 29, 188–196. doi: 10.1111/j.2044-835X.2010.02020.x

Meltzoff, A. (1999). Origins of the theory of mind, cognition and communication. *J. Commun. Disord.* 32, 251–269. doi: 10.1016/S0021-9924(99)00009-X

Moran, J. M. (2013). Lifespan development: the effects of typical aging on theory of mind. *Behav. Brain Res.* 237, 32–40. doi: 10.1016/j.bbr.2012.09.020

Perner, J., Ruffman, T., and Leekam, S. R. (1994). Theory of mind is contagious: you catch it from your siblings. *Child Dev.* 65, 1228–1238. doi: 10.2307/1131316

Peterson, C. C., and McAlister, A. (2006). Metal playmate: siblings, executive functioning and theory of mind. *J. Br. Psychol. Soc.* 24, 733–751.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Putnam, S. P., and Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *J. Pers. Assess.* 87, 103–113. doi: 10.1207/s15327752jpa8701_09

Rai, R., and Mitchell, P. (2004). Five-year-old children's difficulty with false belief when the sought entity is a person. *J. Exp. Child Psychol.* 89, 112–126. doi: 10.1016/j.jecp.2004.05.003

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., and Fisher, P. (2001). Investigations of temperament at three to seven years: the children's behavior questionnaire. *Child Dev.* 72, 1394–1408. doi: 10.1111/1467-8624.00355

Ruffman, T., Perner, J., Naito, M., Parkin, L., and Clements, W. (1998). Older (but not younger) siblings facilitate false belief understanding. *Dev. Psychol.* 34, 161–174. doi: 10.1037/0012-1649.34.1.161

Shahaeian, A., Peterson, C. C., Slaughter, V., and Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Dev. Psychol.* 47, 1239–1247. doi: 10.1037/a0023899

Stevenson, M. (2004). *An Introduction to Logistic Regression, Lecture Notes.* EpiCentre, New Zealand: IVABS Massey University.

Sullivan, K., Winner, E., and Hopfield, N. (1995). How children tell lie from joke: the role of second order mental state attribution. *Br. J. Dev. Psychol.* 13, 191–204. doi: 10.1111/j.2044-835X.1995.tb00673.x

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *J. Genet. Psychol.* 166, 297–312. doi: 10.3200/GNTP.166.3.297-312

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

Wellman, H. M., Fang, F., Liu, D., Zhu, L., and Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychol. Sci.* 17, 1075–1081. doi: 10.1111/j.1467-9280.2006.01830.x

Wellman, H. M., Fuxi, F., and Peterson, C. C. (2011). Sequential progressions in a theory of mind scale: longitudinal perspectives. *Child Dev.* 82, 780–792. doi: 10.1111/j.1467-8624.2011.01583.x

Wellman, H. M., and Liu, D. (2004). Scaling theory of mind. *Child Dev.* 75, 523–541. doi: 10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., and Woolley, J. D. (1990). From simple desires to ordinary beliefs: the early development of everyday psychology. *Cognition* 35, 245–275. doi: 10.1016/0010-0277(90)90024-E

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 March 2013; accepted: 29 May 2013; published online: 17 June 2013.*

*Citation: Calero CI, Salles A, Semelman M and Sigman M (2013) Age and gender dependent development of Theory of Mind in 6- to 8-years old children. Front. Hum. Neurosci. **7**:281. doi: 10.3389/fnhum.2013.00281*

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 26 June 2013
doi: 10.3389/fnhum.2013.00301

# Social modulation of decision-making: a cross-species review

***Ruud van den Bos*[1]*, Jolle W. Jolles[2] and Judith R. Homberg[3]***

[1] *Department of Organismal Animal Physiology, Faculty of Science, Radboud University Nijmegen, Nijmegen, Netherlands*
[2] *Department of Zoology, University of Cambridge, Cambridge, UK*
[3] *Department of Cognitive Neuroscience, Centre for Neuroscience, Donders Institute for Brain, Cognition, and Behaviour, UMC St. Radboud, Nijmegen, Netherlands*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*René San Martín, Duke University, USA*
*Koji Jimura, Tokyo Institute of Technology, Japan*

****Correspondence:***
*Ruud van den Bos, Department of Organismal Animal Physiology, Radboud University Nijmegen, Heyendaalseweg 135, NL-6524 AJ Nijmegen, Netherlands*
*e-mail: ruudvdbos@science.ru.nl*

Taking decisions plays a pivotal role in daily life and comprises a complex process of assessing and weighing short-term and long-term costs and benefits of competing actions. Decision-making has been shown to be affected by factors such as sex, age, genotype, and personality. Importantly, also the social environment affects decisions, both via social interactions (e.g., social learning, cooperation and competition) and social stress effects. Although everyone is aware of this social modulating role on daily life decisions, this has thus far only scarcely been investigated in human and animal studies. Furthermore, neuroscientific studies rarely discuss social influence on decision-making from a functional perspective such as done in behavioral ecology studies. Therefore, the first aim of this article is to review the available data of the influence of the social context on decision-making both from a causal and functional perspective, drawing on animal and human studies. Also, there is currently still a gap between decision-making in real life where influences of the social environment are extensive, and decision-making as measured in the laboratory, which is often done without any (deliberate) social influences. However, methods are being developed to bridge this gap. Therefore, the second aim of this review is to discuss these methods and ways in which this gap can be increasingly narrowed. We end this review by formulating future research questions.

**Keywords: decision-making, translational research, social environment, stress, psychological, humans, animals**

## INTRODUCTION

Decision-making plays a pivotal role in daily life and comprises a complex process of assessing and weighing short-term and long-term costs and benefits of competing actions. The output of the decision-making process, i.e., which action is to be taken, is determined by an interaction between impulsive or emotionally based systems, responding to immediate (potential) rewards and losses or threats, and reflective or cognitive control systems controlling long-term goals (Bechara, 2005; de Visser et al., 2011). Decision-making is influenced by many factors. However, whereas factors such as sex, age, genotype, and personality have been extensively investigated and discussed (reviews; Crone and van der Molen, 2004; Overman, 2004; Overman et al., 2004; de Visser et al., 2011; Homberg, 2012; van den Bos et al., 2013a), relatively little attention has been paid to the crucial moderating effect of social context on decision-making. This is all the more surprising given that decisions in real life are often strongly influenced by the social environment and involve direct and indirect social interactions.

The social environment may affect decision-making in different ways. For instance, decisions may directly involve social partners such as when deciding to share knowledge or goods with others or to provide support (review; Rilling and Sanfey, 2011). Furthermore, subjects may adjust their decisions depending on who is with them or who they consider as their reference-point at the time of the decision. For instance, in the case of so called "conformity behavior," subjects change their behavior to match that of the rest of the group (Morgan and Laland, 2012). Finally, the social environment may influence decisions globally by "setting the atmosphere." For instance, the social environment may breathe a tense or relaxed atmosphere, which influences the individual's emotional state and thereby its decisions (review; Starcke and Brand, 2012). While studies in the field of behavioral ecology have provided elaborate understanding of functional aspects of the social context of decision-making behavior, studies in the field of neuroscience have begun to provide information on the causal aspects and the neural substrate underlying decision-making behavior in a social context. Still, crosstalk between these fields rarely occurs. Researchers in both fields may benefit from insights from both domains that will enable progress toward a common understanding of the social modulating role on decision-making. Therefore, the first aim of our review is to discuss the influence of the social context on decision-making both from a causal and functional perspective, drawing on animal and human studies.

Currently, there is still a gap between decision-making in real life where influences of the social environment are extensive, and decision-making as measured in the laboratory, which is often done without any (deliberate) social influences. Subjects may for instance be less disturbed by stressful conditions when in company of friends or relatives with thereby little effect on their decisions in real life, while showing high levels of stress and concomitant effects on decision-making in the laboratory when tested singly. While these laboratory findings may be important for studying basic mechanisms of e.g., the effects of stress on decision-making (Preston et al., 2007; Lighthall et al., 2009; van den Bos et al., 2009), they hamper for instance assessing the value and general applicability of laboratory findings to the functioning of people, such as patients, in daily life. Furthermore, they miss out the important impact the social environment may normally play on individual and group decision-making. However, a major obstacle to assess the role of the social environment in decision-making under laboratory conditions in humans is that it is difficult to create ecologically valid conditions. Therefore, monitoring real life effects of the social environment on decision-making would be a significant step forward. In rodents, home-cage experimental set-ups, which allow for careful manipulation of brain-behavior relationships in social settings, have been developed as means of bridging precisely this gap. Therefore, a second aim of our review is to discuss these developments in methodology to address the question of the effect of the social environment on decision-making.

Given the foregoing, in the following sections we will discuss how the social environment may modulate decision-making and how this can be incorporated in experimental studies. In section Decision-making in a social context, we will discuss direct and indirect social influences on decision-making, while in section Social stress and decision-making the effects of social stress on decision-making are addressed. Where possible we link a causal and functional perspective and discuss underlying neural substrates. In section Laboratory studies and real-life studies we will (briefly) discuss ways to incorporate the social environment into studies of decision-making. We end this review (section Concluding remarks) with a brief summary of the main issues addressed and define (some) future questions.

## DECISION-MAKING IN A SOCIAL CONTEXT

Humans are an exceptionally successful species, both in the number of individuals and in our flexibility to expand to the range of environments and situations in which we live. A major factor underlying this success boils down to our complex social life as we have the ability to acquire valuable knowledge and skills from others through social learning and teaching and build upon this generation after generation (Boyd and Richerson, 1985; Laland et al., 2011). In our daily life we constantly make decisions based on our personal information and experience as well as that of others, i.e., social learning. Our behavior may be restricted through social conformity (Asch, 1956), or promoted or enhanced through facilitation (Zajonc, 1965). Furthermore, often the decisions of multiple individuals may result in collective behavior, such as the synchronization of applause (Néda et al., 2000), or have to be made jointly to reach a consensus (Conradt and Roper, 2005; Dyer et al., 2008). Living with others comes with the potential benefit of cooperation (Fehr and Fischbacher, 2004a) as well as costs of competition when resources are limited (Davies et al., 2012). Finally, an individual's decisions may be indirectly influenced by the social environment, by affecting an individual's emotional state. Importantly, the modulating role of the social environment is strongly affected by an individual's characteristics and personality as well as that of its group mates (Webster and Ward, 2011).

To fully understand the role of social modulation on decision-making, it is important to consider it from both a causal and functional perspective (Tinbergen, 1963; see e.g., Morgan and Laland, 2012). In neuropsychology, functional explanations are rarely taken into account while this behavioral ecological perspective may help to understand how the behavior of individuals is adapted to the social environment in which they live (Davies et al., 2012). A growing list of behaviors once described as uniquely human have now been described in a range of animals, such as teaching (Franks and Richardson, 2006; Thornton and McAuliffe, 2006), culture (see Laland, 2008; Laland et al., 2011), and conformity (Whiten et al., 2005; Galef and Whiskin, 2008; Jolles et al., 2011), which provide us with new insights into our own behavior. Therefore, the next few sections are focused on a behavioral ecological perspective with links to relevant human and animal laboratory studies. However, as the human literature on social decision-making has been reviewed elsewhere, we limit ourselves to the most relevant human experimental studies (see e.g., Fehr and Fischbacher, 2004a,b; Lieberman, 2007; Frith and Singer, 2008; Behrens et al., 2009; Rilling and Sanfey, 2011).

### SOCIAL LEARNING

For social species, like humans, the social environment plays a critical role in day-to-day decision-making, such as where to live, what to eat and with whom to mate, and may affect their emotional state (see section Observational fear learning). Decisions can be based on either personal experience and/or information gathered by others (Danchin et al., 2004) and through "social learning," individuals may for example learn how (observational learning) to deal with a resource or where it is located (local enhancement; Thorpe, 1956; Webster and Laland, 2012). Although social learning may involve several different learning mechanisms (Laland, 2008) only some rely on advanced cognitive abilities (Galef, 1988; Heyes, 1994) and most cases appear to result from very simple processes (Galef, 1988). Indeed, although social learning may seem particular to humans, animals from a broad range of species gather and exploit information generated by others (review; Galef, 1988; Heyes, 1994; Heyes and Galef, 1996).

A considerable part of the social learning literature has been performed with rats (review; Galef and Giraldeau, 2001; Galef, 2007) and has shown that rats use information from others to learn where, what, how and even when to eat (Galef and Giraldeau, 2001). Both the social information provided by visual and olfactory cues from conspecifics provide a strong basis for individual foraging decisions. Just by observing conspecifics, rats quickly locate food and join to feed with them (see Galef and Giraldeau, 2001). This is further intensified by deposited olfactory cues on both the food and the location of the food (Galef, 2007), which may for example enable young rats to learn what foods are best to eat as they may not be able to figure this out by themselves (see Galef, 2007). In particular the olfactory cues via the breath of conspecifics may result in these socially induced food preferences that may overrule personal preferences (Galef and Whiskin, 2008; Jolles et al., 2011) and even reverse learned aversions to foods (Galef, 1986).

To accurately make decisions, individuals need to constantly weigh the costs and benefits of private and social information and need to be selective when and whom to copy (Galef, 1995; Laland, 2004). Social learning may be beneficial as it allows individuals to acquire relevant information without having the risk or costs associated with individual learning. However, social information may be outdated, for example when the environment is highly variable, or less valuable, when the environment is very stable (Boyd and Richerson, 1985). Thus, relative reliance on social and individual learning can be viewed as involving a trade-off between accuracy and cost (Boyd and Richerson, 1985; Laland, 2004; Kendal et al., 2005). For example, Dally and colleagues (2008) showed that rooks selectively consumed the same food as a demonstrator when the foods were novel, but not when the foods were familiar. Likewise, Galef and Whiskin showed that the greater the discrepancy between private and social information, the less likely the subject is to behave in accord with the socially acquired information (Galef and Whiskin, 1998). Moreover, Brown and colleagues (2008) showed that personal and social information about spatial choices are combined in a rat's working memory and both the quality of the food available and the memory of a familiar conspecific's behavior affect an individual's tendency to visit spatial locations in a radial-arm maze.

The trade-off between accuracy and costs is nicely illustrated by the difference in public information use of two closely related species of sticklebacks. Coolen and colleagues (Coolen et al., 2003) showed that while nine-spined sticklebacks exploited public information and foraged at the areas they observed others to have better feeding rates, three-spined sticklebacks ignored this information and relied in their decisions on their own experience. This difference in social information use may be explained by the relative difference in costs of self-acquired information between the two stickleback species. The robust defenses that three-spined but not nine-spined sticklebacks have, such as large spines and armored body plates, allows them to sample alternative food patches directly in relatively better safety, as reflected by the increased time nine-spines spent hidden amongst vegetation (Laland, 2008).

When the presence of group mates affects the behavior of an individual, allowing or causing them to engage in certain behaviors at a different rate, or to perform behaviors that they would not perform at all if they were alone, this is called social facilitation (Zajonc, 1965). For example, in animals it has been shown that the presence of others may result in higher activity (Griffiths and Foster, 1998; Webster et al., 2007), increased foraging (Webster et al., 2007; Dally et al., 2008) and provide scrounging opportunities (review; Giraldeau and Caraco, 2000). For example, conform to human work, studies on rats have shown that the greater the number of models and the greater their uniformity in behavior, the more likely a naive subject will act in accord with the information that conspecifics provide (Galef and Whiskin, 1995). These changes in behavior can probably be ascribed to proximate mechanisms such as greater anti-predator benefits of larger groups (review; Krause and Ruxton, 2002), investment in vigilance and/or increased competition (review; Beauchamp, 2003). This is nicely illustrated by two studies in ravens (Stöwe et al., 2006a,b) which showed that when individuals were alone compared to in a group, they approached a novel object faster but spent less time close to and manipulating it. Although the social group enabled individuals to decrease time investment in vigilance, they may have a higher approach latency because individuals might wait for the other to take the risk and lead.

## CONFORMITY BEHAVIOUR

Social learning theory suggest that in most circumstances where natural selection favors reliance on social learning, conformity is favored and individuals, both humans and other animals, should adopt the behavior of the majority (Boyd and Richerson, 1985; Laland, 2004). This particular form of social modulation on decision-making is especially important as it has been argued to be a major mechanism in human cultural evolution (Boyd and Richerson, 1985; Efferson et al., 2008). One of the earliest described studies on human conformity was performed by Asch (1955, 1956). In a very influential paper, Asch (1955) described how adults would be willing to abandon their own perceptual judgment in a simple visual task and go with the overtly false alternative as a result of group normative behavior. Since then a huge number of studies has replicated these kinds of findings (see Bond and Smith, 1996; Morgan and Laland, 2012). Interestingly, the extent of conformity behavior seems to be strongly dependent on the situation. Namely, if a participant has to make a public response and is face-to-face with the majority, there is a strong normative influence of conformity, whereas it is weaker when participants make a private response and indirectly communicate with the majority (Bond, 2005). Furthermore, conformity behavior may be dependent on task difficulty and its importance (Baron, 1996), group size (Asch, 1955; Bond, 2005) and culture (Bond and Smith, 1996) among others.

Recently, several studies have addressed the neurobiological basis of conformity (see also Morgan and Laland, 2012). For instance, studies using mental rotation and auditory tasks (Berns et al., 2005, 2010) showed that social information may affect brain regions classically associated with perception as well as the processing areas associated with each task, suggestion that social information was affecting the subjects' perception as well as decision-making (see Morgan and Laland, 2012). Moreover, it has been shown that while cingulate areas are involved in monitoring the difference between private and public information (Klucharev et al., 2009), the ventral striatum is involved in the tendency to adjust one's behavior to the social information (Burke et al., 2010; Campbell-Meiklejohn et al., 2010), which may be related to rewarding aspects of being in line with the behavior of others (Klucharev et al., 2009; Burke et al., 2010; Campbell-Meiklejohn et al., 2010).

Conformity has been described in a wide range of animal species including fish (Laland and Williams, 1998; Day et al., 2001; Pike and Laland, 2010), rats (Galef and Whiskin, 2008; Jolles et al., 2011) and primates (Whiten et al., 2005; Dindo et al., 2009) (see Webster and Ward, 2011 for a review). For example, Laland and Williams (1997) showed that guppies preferentially chose a foraging route they had previously observed demonstrators use despite an equally valid available alternative. Individuals may base these kind of conformity decisions on heuristic rules of social attraction (Webster and Laland, 2012) such as to approach others (e.g., Laland and Williams, 1997), to approach larger over smaller groups (e.g., Lachlan et al., 1998; Day et al., 2001) and to approach groups that produce cues indicative of higher foraging success (e.g., Coolen et al., 2003, 2005). These tendencies are likely to benefit animals in most cases as it allows them to detect food without having to pay the costs of sampling the environment directly (see e.g., Pitcher et al., 1982; Day et al., 2001). However, sometimes this conforming to the behavior of others may come with "opportunity costs." For example, individual fish may discover a visually isolated food patch faster and exploit it for longer than when a group of conspecifics is present (Webster and Laland, 2012), and smaller groups may discover a hidden food patch more quickly than larger ones (Day et al., 2001). The reliance on social information may sometimes even result in individuals to base their decisions on maladaptive information, such as rats consuming less palatable and sodium-deficient diets based on the breath of conspecifics (Galef, 1986), and even after the source of information is removed, such as guppies that kept on using energetically costly routes to food patches despite shorter alternatives available (Laland and Williams, 1998).

Although conformity of the basic "follow the majority" kind has been demonstrated in a variety of species of which Pike and Laland's (2010) study on public information use in sticklebacks provides compelling evidence, only a few animal studies (Whiten et al., 2005; Galef and Whiskin, 2008; Jolles et al., 2011) have investigated the situation where conformity overrides the discovery of valid alternative means (cf. Asch, 1955, 1956). In a two-action diffusion study in chimpanzees, Whiten and colleagues (2005) showed that although some individuals discovered an alternative technique to free trapped food items to the one seeded in their group, they later re-converged on the norm of their group, demonstrating conformity in the face of discovering a functional alternative. Two recent studies also suggest the existence of this type of conformity in rats (Galef and Whiskin, 2008; Jolles et al., 2011). Rats were given the opportunity to learn that two diets differed in palatability. They were subsequently exposed to a demonstrator that had eaten the less palatable food and were thereafter exposed to the same diets again. By simply being exposed to the odors in the breath of a conspecific for 30 min, individuals considerably decreased their preference for the more palatable food. Interestingly, despite similar initial preferences and similar social information, some rats were more resistant to changing their preference in relation to private and social information than others (Jolles et al., 2011), suggesting a different sensitivity to conflicting information (cf. Klucharev et al., 2009).

## COLLECTIVE BEHAVIOUR AND GROUP DECISION-MAKING

Both humans and many group living animals exhibit complex, coordinated, group patterns, such as lanes of traffic flow in human crowds (Helbing and Molnar, 1995) and the three-dimensional movements of fish shoals (Couzin and Krause, 2003). Through collective action, individuals can enhance their capacity to detect and respond to salient features of the environment, resulting in more accurate decision-making (Couzin, 2009) without the need of explicit signals or complex communication (Couzin et al., 2005; Dyer et al., 2008). The common property of these phenomena is self-organization, suggesting that much of complex group behavior may be coordinated by relatively simple interactions among the members of the group (review; Couzin and Krause, 2003). Indeed, recently studies have begun to reveal that collective decision-making mechanisms across animal species, from insects to birds and even humans, seem to share similar functional characteristics (Couzin and Krause, 2003; Conradt and Roper, 2005; Sumpter, 2006). For example, Helbing and colleagues (Helbing and Molnar, 1995; Helbing et al., 2000) have shown that simple rules such as "try to minimize travel time," "avoid collisions" and "move in the direction of other people" may help explain pedestrian movements on busy streets and in life-threatening situations. Similar patterns have been described for non-human animals including the spectacular trails of ants on foraging trips (Couzin and Franks, 2003), the collective movements of starlings (Ballerini et al., 2008), and social interactions in shoaling fish (Herbert-Read et al., 2011).

In some cases group decisions are the result of a consensus reached by the individuals in the group (Conradt and Roper, 2005). Humans make these kinds of decisions all the time, from agreements in groups of a few people, to large-scale international conventions and political elections. However, also amongst non-human animals consensus decision-making is very common, such as travel routes in navigating birds use and the timing of activities (review; Conradt and Roper, 2003, 2005). In many situations conflicts may exist between the preferences of different individuals (Couzin et al., 2005). However, all individuals in the group have to decide on the same action because the group will fall apart unless a consensus is reached (Conradt and Roper, 2005), resulting in a loss of many of the advantages of group living (review; Krause and Ruxton, 2002). In line with theoretical predictions (Couzin et al., 2005), it has now been demonstrated that only a small proportion of knowledgeable individuals is needed to influence the direction of movement of the whole group, such as has been shown for nest site choice in social insect colonies (Franks et al., 2003; Seeley, 2003), the foraging movements in golden shiner fish (Reebs, 2000), and humans moving to a target without the use of verbal communications or obvious signaling (Dyer et al., 2008).

## COOPERATION AND COMPETITION

An important way to understand social decision-making in humans and other social animals is to look at it in terms of costs and benefits, not only to the actor as indicated above, but also to the recipient in the social context (Hamilton, 1964; West et al., 2007; Davies et al., 2012). For this it is important to keep in mind that via natural selection those genes are favored that increase an organism's ability to survive and reproduce (fitness). Therefore, individuals will often attempt to act in such a way as to receive immediate, selfish benefits, which may often result in competition or mutualistic cooperation. This is nicely illustrated by the Prisoner's Dilemma (PD; Axelrod and Hamilton, 1981) in which individuals can either cooperate or defect. Both individuals would benefit from mutual cooperation but both are also tempted to cheat, as it would be more rewarding to the individual. Therefore, irrespective of the other player's choice, it pays to defect. This raises the problem why cooperation is so common among human and animal societies (see West et al., 2007) and why individuals not act selfishly all the time and exploit the cooperative behavior of others (see Davies et al., 2012). In many cases, the cooperating individual simply acts selfishly and gains an immediate benefit, but thereby provides by-product benefits to its group mates, such as the benefits of an increased group size, i.e., reduced chance of predation, due to helping behavior in meerkats (Clutton-Brock, 2002). When on the other hand cooperation is altruistic—costly to the cooperator and beneficial to the recipient—cooperating individuals may still gain selfish benefits in the long term by using conditional strategies (Stevens and Hauser, 2004), such as cooperating only with relatives (kin selection; Hamilton, 1964), interacting only with those that have cooperated previously (reciprocity; Trivers, 1971; see Clutton-Brock, 2009), or under enforcement (Frank, 2003).

Individuals may help relatives as this may increase their genetic representation in future generations, and thus their fitness, as relatives share genes by common descent (see further Hamilton, 1964; West et al., 2007; Davies et al., 2012). If individuals preferentially help those that have helped them or those that help others, also known as reciprocity, the short-term cost of being cooperative is outweighed by the long-term benefit of receiving cooperation (Trivers, 1971). Although the PD has shown that when individuals meet only once it is better for individuals to defect than to cooperate, some form of cooperation may be stable if there is a chance both players will meet again because the long-term benefits of cooperation may outweigh the short-term benefit of defecting (Axelrod and Hamilton, 1981). Indeed, experimental work on both humans (Fehr and Fischbacher, 2003) and rodents (Rutte and Taborsky, 2007, 2008; Viana et al., 2010) has shown that individuals cooperate at higher levels in repeated interactions. For example, Rutte and Taborsky showed that rats that were trained to pull a stick in order to produce food for a partner pulled more often for an unknown partner after they were helped than if they had not received help before (generalized reciprocity; Rutte and Taborsky, 2007) and more often from a partner they received help from (direct reciprocity; Rutte and Taborsky, 2008). Furthermore, Schneeberger and colleagues (2012) showed that, similar to human PD studies, rats provided more food to cooperative partners than to defectors and that furthermore, this was dependent on costs: when rats experienced experimentally increased resistance to pull the stick of the apparatus and deliver food to the social partner, they reduced their help. It remains unclear, however, to what extent these behaviors may potentially be ascribed to simpler processes such as conditioned place preference. For example, rats have been shown to prefer a social partner over an empty space (Trezza et al., 2009) and to cooperate 80% of the time if they have the choice to act either alone or in cooperation with a social partner to obtain food pellets (Tsoory et al., 2012). Indeed, although reciprocity has attracted a huge amount of attention, it is thought to be generally unimportant outside humans (Hammerstein, 2003; Stevens and Hauser, 2004) as in most cases cooperation can be explained by more simple mechanisms such as by-product-benefits (Hammerstein, 2003; Clutton-Brock, 2009). Nevertheless, it shows that (lab) rodents may provide a good model system to investigate the mechanisms and development of cooperation (Łopuch and Popik, 2011).

Finally, enforcement or punishment may alter the benefit/cost ratio of helping and thereby favor cooperation (Frank, 2003). The consequences of punishment are nicely illustrated in cleaner fish. Cleaner fish remove parasites on the body of other species of fish that cannot remove the parasites themselves. Although, the cleaner fish prefer to eat parts of their clients' tissue they rarely perform this cheating behavior as their hosts may punish them by chasing them or by swimming away (Bshary and Grutter, 2002). What may be special about human cooperation is that we have the capacity to establish and enforce social norms (Fehr and Fischbacher, 2003, 2004a,b) because our societies are based on large-scale cooperation among genetically unrelated individuals (Henrich et al., 2003). For example, human research investigating the conditional cooperation on social norms has shown that subjects increase their contribution to the public good if the average contribution of the other group members increases (see Fehr and Fischbacher, 2004a). Moreover, third-party punishment experiments in which the PD is extended with a passive third party has shown that these individuals punish not-cooperating players despite a cost to themselves and that moreover, defection was punished much more severely if the other player cooperated than if they both defected (Fehr and Fischbacher, 2004b).

When individuals act selfishly under situations of limited resources, competition may occur between individuals. Competing individuals have to weigh the competitive efforts against expected benefits as well as the intensity of the conflict. Individuals may compete by exploitation and/or by resource defense (Davies et al., 2012). Importantly, the best way for an individual to behave often depends on what its competitors are doing (review; Davies et al., 2012), which will therefore result in a stable outcome of competition, also known as the evolutionary stable strategy (EES; Maynard Smith and Price, 1973). Under ideal free distributions in which individuals are free to go where they want and have complete information about the availability of resources (Fretwell, 1972), individuals will distribute themselves in such a way that all individuals have the same rate of resource acquisition. For example, people queuing at the check-out area of the supermarket will often decide to choose the shorter and faster queues, ultimately resulting in all queues being of more or less equal length. However, in most cases individuals may not be free to go where they want as better competitors will occupy the richer habitats. This situation is very common in the natural world (see Davies et al., 2012). For example, although ducks have been shown to occur in stable distributions of individuals among foraging sites (Harper, 1982), some ducks were better competitors than others and grabbed most of the food (Harper, 1982). Importantly, defense of a resource has costs as well as benefits and individuals should only behave territorial when the benefits are greater than the costs. This may also help explain why often variable competitive behavior can be found within a population, such as producers and scroungers in a foraging context (Giraldeau and Caraco, 2000), as the costs and benefits may be different between individuals.

Insight in the neural mechanisms underlying cooperation and competition is increasing (see Rilling and Sanfey, 2011; Huettel and Kranton, 2012). For example, a neuroimaging study of the Prisoner's Dilemma has shown that mutual cooperation led to increased activation in reward regions (Rilling et al., 2002), potentially explaining how cooperative social relationships may be sustained while inhibiting the impulse to act selfishly. Many social decision-making studies have used the Ultimatum Game in which two players split a sum of money, one player proposes a division, and the other can accept or reject this. For example, it has been shown that both unfair offers and their rejection elicited activity in brain areas related to emotion, such as the anterior insular cortex, suggesting an important role for emotions in social decision-making related to cooperation (Sanfey et al., 2003). Furthermore, alpha- and theta-oscillations in prefrontal areas have been found to be sensitive to social risk and to underlie fine-tuning regulation of social decisions (Billeke et al., 2012). A study investigating whether punishment of unfair offers might be affected by the relationship between the players has shown that when the proposer was a friend rather than an unknown person, unfair offers were much less frequently rejected. The anterior prefrontal cortex plays an important role in these kind of interpersonal economic interactions (Campanhã et al., 2011).

Rodent work has also provided interesting insights into the emotional and neurobiological bases of competition. For example, water-deprived rats in a pair competing for a single source of water quickly establish a firm relationship during which one rat drinks consistently more (the dominant rat) than the other (the submissive rat). However, interestingly, when the animals are exposed to severe stress, the dominants becomes less dominant, and when their submissive cagemates are administered anxiolytics, they increase their access to resources at the expense of that obtained by dominants (Joly and Sanger, 1991). One brain area in particular seems to play a central role in the cost-benefit decision making related to competition: the anterior cingulate cortex (ACC). For example, the ACC is implicated in action selection and action outcome and effort monitoring, as well as signaling the use of social information (Rudebeck et al., 2006). Hillman and Bilkey (2012) provided rats with a choice whether to physically compete with a peer for a large food reward or not to compete and to obtain a small reward. It was found that ACC neurons electrophysiologically responded to competitive effort costs, assisting the rats in goal-directed decision making under social competitive conditions (Hillman and Bilkey, 2012).

## OBSERVATIONAL FEAR LEARNING

Decision-making can be strongly influenced by the way the social environment affects an individual's emotional state. An important example of this is social learning of fear (reviewed by Olsson and Phelps, 2007). Learning about potentially harmful stimuli and events is important in shaping adaptive behavior, which may be less risky if learned socially through observation and social communication. Experimentally, social fear learning can be assessed by subjecting an observer to another individual who is undergoing cued threatening experiences, which may elicit physiological and behavioral responses in the observer as if undergoing the threat him/herself. Fear responses acquired through conditioning and observation of a distressed model were expressed to both seen and unseen (backwardly masked) conditioned stimuli, whereas, fear responses acquired through verbal communication were expressed only to seen conditioned stimuli (Olsson and Phelps, 2004). This indicates that the route of social information transmission affects how information is perceived. Also genotype affects social fear learning. Carriers of the low activity variant of the common serotonin transporter polymorphism displayed more cued fear responses compared to high activity variant carriers when subjected to an observational fear learning paradigm in which subjects had to view a movie in which models received shocks in the presence of a conditioned stimulus (Crişan et al., 2009). Furthermore, personality has been investigated as modulator of social fear learning using a paradigm in which participants watched mock panic attacks while emotional (e.g., fear and panic) and skin conductance levels were assessed. It was found that emotional avoidance and anxiety sensitivity were positively associated with more self-reported fear and more severe panic symptoms to the challenge procedure (Kelly and Forsyth, 2009). Similarly, Hooker et al. (2008) found that trait neuroticism enhanced social fear learning. Finally, there are sex differences in observational fear conditioning using modeled "mock" panic attacks as an unconditioned stimulus and an associated neutral cue as conditioned stimulus, as women reported more distress to the conditioned stimulus (Kelly and Forsyth, 2007). Mechanistically, social fear learning shares neural features with classical conditioned fear, including the involvement of the amygdala, but also requires higher-level reflective mental state attribution, like involvement by the anterior cingulate cortex and the anterior insula (see Olsson and Phelps, 2007; Olsson et al., 2007; Olsson and Ochsner, 2008).

Next to humans, observational fear learning has been shown in a large number of species (see Olsson and Phelps, 2007) but most animal studies have been performed with rodents, showing that both visual, auditory as well as olfactory stimuli play an important role in social transfer of fear. For example, Jeon et al. (2010) demonstrated that mice observing demonstrators undergoing foot-shock stress displayed increased contextual conditioned freezing when subsequently placed in the observing chamber. This process was reduced, but not occluded, when an opaque partition was placed between the observer and demonstrator. In rodents in particular, olfactory cues play an important role, especially related to alarm pheromones. These may change autonomic activity and increase defensive and risk assessment behaviors (Kiyokawa et al., 2004, 2006) and are excreted in the rat's perianal region, especially by allogrooming, as seen during the social interaction between the demonstrator and observer rats (Knapska et al., 2010). Also, distress vocalizations affect fear learning. For example, when a conditioned stimulus was coupled to aversive 22 KHz ultrasonic vocalizations (USVs), observers displayed conditioned freezing (Chen et al., 2009) and the number of 22 KHz-USVs emitted by a fearful demonstrator was positively associated with the conditioned freezing response displayed by the observer (Wöhr and Schwarting, 2008). In line with human studies, familiarity between the observer and demonstrator results in higher observational fear learning (Chen et al., 2009; Jeon et al., 2010). Interestingly, not only fear or distress itself can be socially transmitted amongst rats and mice, also the predictive value of the conditioned stimulus itself. Bruchey and colleagues (2010) demonstrated that observer rats acquire a freezing response by observing fear-conditioned demonstrators, i.e., being exposed to the conditioned stimulus in the absence of the foot-shock. Thus, the observers responded to the conditioned stimulus as if they had experienced foot-shocks themselves. Whereas fear can be socially transmitted by social interaction between a previously stressed demonstrator and a naive conspecific, it has also been demonstrated that observation of a non-fearful demonstrator mouse inhibited subsequent recall of a context-shock association in observers (Guzmán et al., 2009). Thus, it seems that previous experience with a fear-naive demonstrator 'buffered' fear conditioning in observers (Panksepp and Lahvis, 2011), providing strong evidence for socio-emotional influences on the behavioral response to threat.

### ROLE OF INDIVIDUAL CHARACTERISTICS AND PERSONALITY DIFFERENCES

Although the mere presence of others may affect the decisions an individual makes, such as via facilitation and conformity, this modulating effect is strongly influenced both by the characteristics of the individual as well as that of its group mates, for instance by social status (Nicol and Pope, 1999), familiarity (see above; Swaney et al., 2001; Jeon et al., 2010), sex (see Choleris and Kavaliers, 1999; Piyapong et al., 2010) and social relationships between individuals in the group (e.g., Beauchamp, 2000; Schwab et al., 2008; Jolles et al., 2013b). Furthermore, consistently expressed behavioral differences between individuals that are otherwise similar to one another in terms of age, size and sex—also known as personality types or coping styles (Réale et al., 2007; Koolhaas et al., 2010)—may play a particular large role on individual decision-making (Webster and Ward, 2011). For example, bold compared to shy individuals have been found to be less responsive to changes in their social environment (Magnhagen and Bunnefeld, 2009) and their partner's behavior (Harcourt et al., 2009; Schuett and Dall, 2009), have a lower tendency to join and follow conspecifics (Ward et al., 2004), base their decisions less on social information (Kurvers et al., 2010) and display greater initiative in leadership (Harcourt et al., 2009). It is especially the interplay between these personality traits, individual characteristics and the relationships between individuals that affects an individual's decisions (e.g., van Oers et al., 2005; Schuett and Dall, 2009; Jolles et al., 2013b). Importantly, in this way individual characteristics and heterogeneity within groups may ultimately impact the dynamics of group decisions and behavior and affect the way in which the group as a whole functions in relation to the environment (Webster and Ward, 2011). For example, individual differences in risk-taking strongly affect social feedback between individuals (Harcourt et al., 2009), individuals may not be uniformly distributed within groups (Jolles et al., 2013a), and certain individuals may take leadership positions and thereby determine group decisions (King et al., 2009; Nagy et al., 2010). Surprisingly, few studies have considered the impact of individual characteristics and personality traits on the social modulation of decision-making. For example, although sex differences have been described in a wide range of cognitive and behavioral processes, investigations of sex differences in social learning are still largely neglected (review: Choleris and Kavaliers, 1999). Furthermore, despite the surge of interest in personality traits in animals, only in recent times have studies started to consider personality in the context of the crucial moderating effect of the social environment (review: Webster and Ward, 2011). Finally, both human and non-human studies as well as models on group behavior still seldom consider the impact of such heterogeneity on the rules underlying their coordination (but see e.g., Jolles et al., 2013a).

## SOCIAL STRESS AND DECISION-MAKING

The social environment in which humans and animals live is not devoid of psychosocial stress. Stressors may entail among others potential or actual conflicts with conspecifics either in the context of dominance-submission or in competition over (valuable) resources, the sheer performance of a task in front of conspecifics, and experiencing or witnessing aggression and violence. To assess the effects of social stressors on decision-making in the laboratory, tests are needed which produce reliable and reproducible stress-related effects. One such psychosocial test is the Trier Social Stress Test (TSST; Kirschbaum et al., 1993) and its variants (e.g., group-wise TSST; Von Dawans et al., 2011).

The TSST has been shown to be very effective in inducing stress as measured by questionnaires regarding stress, mood and anxiety as well as parameters indicative of the activation of the two main stress axes, i.e., hypothalamus-pituitary-adrenocortical axis (HPA-axis; cortisol) and the sympatho-adrenomedullary axis (SAM-axis; (nor)adrenaline) (e.g., Kirschbaum et al., 1999; Kudielka and Kirschbaum, 2005; Nater et al., 2005, 2006; Starcke et al., 2008; Nater and Rohleder, 2009; van den Bos et al., 2009; Foley and Kirschbaum, 2010; Cornelisse et al., 2011; Starcke et al., 2011; Maruyama et al., 2012; Vinkers et al., 2013). This stress effect is related to the social-evaluative and uncontrollable elements of the task (Dickerson and Kemeny, 2004): subjects have to deliver a speech as well as do a difficult arithmetic in front of a panel that judges their performance without much *a priori* knowledge of the procedure. Even anticipating delivery of the speech is already stressful.

The activation of the SAM-axis is often measured by salivary alpha-amylase, while activation of the HPA-axis is often measured by salivary cortisol (Kirschbaum et al., 1999; Kudielka and Kirschbaum, 2005; Nater et al., 2005, 2006, 2007; van Stegeren et al., 2006, 2008; Nater and Rohleder, 2009; Foley and Kirschbaum, 2010; Thoma et al., 2012). While the SAM-axis is strongly activated during the TSST and returns to baseline immediately or quickly thereafter, HPA-axis activity peaks 10–20 min after the TSST and returns to baseline about 60 min thereafter (e.g., Nater et al., 2005, 2006; Cornelisse et al., 2011; Starcke et al., 2011; Maruyama et al., 2012; Thoma et al., 2012; Vinkers et al., 2013). Cortisol levels in men are generally higher than in women, while in women the menstrual cycle and contraceptives in addition have a modulatory effect (Kirschbaum et al., 1999; Kudielka and Kirschbaum, 2005; Foley and Kirschbaum, 2010; Nielsen et al., 2013; but see Kelly et al., 2008). Thus, the TSST seems to be a useful laboratory test to delineate the effects of psychosocial stress on decision-making, when decision-making tasks are delivered after the TSST. It should be noted that the Cold Pressor Test has been used as well to delineate the effects of stress on decision-making. As at first glance the results between this test and the TSST on decision-making were not different its effects on decision-making will be included in the following paragraphs.

## SOCIAL STRESS: EFFECTS ON DECISION-MAKING PARADIGMS

Following the TSST (as well as the Cold Pressor Test) several reward-based decision-making tasks have been shown to be affected (review; Starcke and Brand, 2012), i.e., the Iowa Gambling Task (IGT; Preston et al., 2007; van den Bos et al., 2009), the Balloon Analogue Risk Task (BART; Lighthall et al., 2009), the Game of Dice Task (Starcke et al., 2008), delay-discounting (Lempert et al., 2012) and a financial decision-making task (Porcelli and Delgado, 2009). Social stress paradigms have not been tested in animals with respect to reward-based decision-making. However, the data of other types of stress paradigms reveal similar effects as in humans: stress disrupts reward-based decision-making tasks in rats (Graham et al., 2010; Shafiei et al., 2012).

Thus far, only a few studies have been published on the effects of the TSST on social decision-making related paradigms. Social stress had no effects on moral decision-making, although in the stress group it was shown that the higher the salivary cortisol levels the more egoistic, and thus less altruistic, decisions were taken in highly emotional dilemmas (Starcke et al., 2011). Furthermore, social stress induced by the TSST increased pro-social behavior as measured by the Trust Game (reciprocal exchange) and the Dictator Game (altruism) (Takahashi et al., 2007; Von Dawans et al., 2012). Still in the latter game this effect seemed to be dependent of whether money was donated to a person or to an anonymous charity organization as Vinkers and colleagues (2013) observed that people donated less money to an organization following the TSST. Finally, altruistic punishment behavior in the Ultimatum Game was not affected immediately following the TSST (Von Dawans et al., 2012; Vinkers et al., 2013); however, it was affected when the task was administered 75 min after the TSST (Vinkers et al., 2013; see further below).

The overall impression from these studies is that differences are present in the consequences of social stress on paradigms that people play singly and those that involve interaction, even when virtually, with others. Von Dawans and colleagues (2012) suggest that this may be related to the workings of oxytocin, which would be released under stress and modulate the response in social decision-making in the direction of pro-social behavior and social support (see Taylor et al., 2000; Cousino Klein and Corwin, 2002; Heinrichs et al., 2003; Foley and Kirschbaum, 2010; Vinkers et al., 2013; see further below). The latter would lower the stress-response (Heinrichs et al., 2003; Foley and Kirschbaum, 2010). The data on the stress-related increase in pro-social behavior are in line with the observation that in primate species behaviors like reconciliation and consolation follow conflicts or social tension (e.g., Aureli et al., 1989; Koski et al., 2007; Fraser et al., 2008). These behaviors facilitate recovery from stress and counterbalance the negative consequences of social conflict on group-cohesion and may restore internal group-cohesion (Aureli et al., 1989; Fraser et al., 2008; but see Koski et al., 2007). For, maintaining internal cohesion is crucial as to maintain the benefits from group-living, which are related to increased possibilities to find and exploit food resources as well as lowering predation risk. Interestingly, oxycotin has been shown to promote in-group behavior and increase defensive aggression toward outsiders (De Dreu et al., 2010). To what extent this relates to the observation that altruism in the Dictator Game was enhanced following social stress depending upon whether it was in the context of persons or a charity organization (Takahashi et al., 2007; Von Dawans et al., 2012; Vinkers et al., 2013) remains to be studied. These data thus provide a link between causal mechanisms and functional mechanisms of pro-social behavior following social stress. The biological meaning of the data on reward-based decision-making is discussed in section Timing, coping styles and daily life (coping-styles).

## SOCIAL STRESS: SEX DIFFERENCES

Studies directed at dissecting sex differences showed that men displayed more risk-taking behavior following stress (IGT and BART), whilst women were more risk-aversive (BART) or became more task-focused (IGT). These studies also showed that sex differences were related to the levels of cortisol. The higher the levels of cortisol, the more risk-taking behavior was shown by men (IGT; van den Bos et al., 2009). Women, on the other hand displayed more risk-aversive or task-focused behavior with increasing levels of cortisol (BART; Lighthall et al., 2009; IGT; van den Bos et al., 2009). Data from the IGT also indicated that women became more risk-taking when levels were too high (van den Bos et al., 2009; see also Witbracht et al., 2012). Thus, overall these data indicate that stress has a different effect on reward-based decision-making in men and women with different underlying effects of cortisol. This was recently confirmed using the Cambridge Gambling Task and a job assessment procedure to induce stress: while salivary cortisol levels were positively correlated with risk-taking behavior in men, they were if anything weakly negatively correlated in women (van den Bos et al., 2013b). Interestingly, this study also revealed a different relationship between salivary alpha-amylase and risk-taking in men and women: while in women a positive relationship was found, a negative relationship existed in men (van den Bos et al., 2013b). These data underline that differences do exist between men and women regarding the relationship between stress, neuroendocrine changes and decision-making (see also de Visser et al., 2010; van den Bos et al., 2013a).

Studies on social decision-making have been mainly done in male only populations (Takahashi et al., 2007; Von Dawans et al., 2012; Vinkers et al., 2013) or do not mention potential sex-differences in the data set (Starcke et al., 2011), precluding therefore to discuss differences between men and women in this respect. Still, one study using the same social-decision-making tasks and stress protocol as applied in men, did not observe an effect of social stress on social-decision making in women (Koot, unpublished). None of the studies in men reported a relation with cortisol (Von Dawans et al., 2012; Vinkers et al., 2013). Furthermore, while one study reported a correlation between heart-rate and pro-social behavior (Von Dawans et al., 2012), other studies did not observe a correlation between salivary alpha-amylase and pro-social behavior (Takahashi et al., 2007; Vinkers et al., 2013).

The increase in risk-taking behavior in men in reward-related decision-making may be associated with a loss of top-down control of prefrontal over subcortical areas, such as mediated by the lateral orbitofrontal cortex and dorsolateral prefrontal cortex (Piazza and Le Moal, 1997; Arnsten, 1998, 2009; Erickson et al., 2003; Stark et al., 2006; Wang et al., 2007; Kern et al., 2008; Dias-Ferreira et al., 2009; Goldstein et al., 2010; Koot et al., 2011, 2013). Furthermore within the limbic system high levels of cortisol may shift the balance of the activity of the ventral striatum (reward-related behavior) and amygdala (punishment-related behavior) toward the ventral striatum (Piazza et al., 1993; Dellu et al., 1996; Piazza and Le Moal, 1997; Pruessner et al., 2004; Mather et al., 2010; Porcelli et al., 2012; see Wager et al., 2008). A recent study showed that increasing noradrenergic activity decreased amygdala activity and processing of fearful faces (Schwabe et al., 2013). Thus, it may be hypothesized that in men the prefrontal-subcortical balance is disrupted by acute stress. In line with this, it was recently observed that systemic injections of corticosterone in male rats in a rodent analogue of the Iowa Gambling Task (de Visser et al., 2011) disrupted decision-making performance, which was associated with changes in activity in prefrontal structures (Koot et al., 2011, 2013). Still it should be noted that such effects of corticosterone were not observed in other studies (Graham et al., 2010; Shafiei et al., 2012). However, as discussed by Koot et al. (2013) this may be due to the way these studies applied corticosterone and/or administered the task following corticosterone injections. In general, studies on stress in male rats have revealed that acute stress sensitizes the reward system (through corticosterone; e.g., Piazza and Le Moal, 1997). As mentioned above it has been suggested that stress induced release of oxytocin may have an effect on the way subjects engage into social decision-tasks. Currently no studies exist which have studied the interaction between stress, changes in neural structures and social decision-making tasks.

As to the underlying neural substrate in women it seems that top-down control may actually be increased under stress, related to levels of cortisol, with among others a lower striatal and a stronger amygdala activity (Stark et al., 2006; Wang et al., 2007; Mather et al., 2010; Porcelli et al., 2012). A recent study showed that increasing noradrenergic activity increased amygdala activity, decreased orbitofrontal activity (thereby decreasing top-down control) and increased processing of fearful faces (Schwabe et al., 2013). It has been suggested that the persistent activity in for instance the anterior cingulate cortex following a stressful experience in women may be associated with the development of depressive symptoms in women related to tendencies of ruminative thinking (Tamres et al., 2002; Wang et al., 2007). The menstrual cycle has a strong effect on the outcome of changes in neuronal activity (Goldstein et al., 2010; Ter Horst et al., 2013). Thus, at present changes in neural activity in women are less clear and straightforward than in men. However, by and large these changes in women seem compatible with a shift toward risk-aversive behavior. Like in men, currently studies in women are lacking which have looked at the interaction between stress, changes in brain structures and social decision-making tasks.

## TIMING, COPING STYLES AND DAILY LIFE

For a full understanding of the effects of social stress on decision-making paradigms three issues need further discussion: (1) short-term *versus* long-term effects of stress (timing), (2) relationship between stress, coping styles and task performance, and (3) consequences for daily life.

While most studies have applied decision-making tasks directly following the TSST, the data of several studies suggest that stress, notably cortisol, may have time-dependent effects on the balance between prefrontal and subcortical functioning. These timing effects may be related to non-genomic, rapid, and delayed, genomic, corticosteroid actions. For instance, when targeting these two time-domains specifically by administering cortisol in human subjects either shortly or several hours before behavioral testing, working memory was found to be improved by slow compared to rapid corticosteroid actions, and this improved performance was linked to enhanced activity in the dorsolateral prefrontal cortex (Henckens et al., 2011). These and other studies have led to the hypothesis that prefrontal cortical functioning is impaired by corticosteroids acting via rapid non-genomic pathways, but enhanced by slow corticosteroid actions (Joëls et al., 2012). Few studies have targeted these different time-windows thus far. In a recent study using the TSST and a social decision-making task it was shown that male subjects showed more acceptance of ambiguous offers when the task was administered 75 min after the TSST than when administered immediately thereafter, leading the authors to conclude that this may be due to enhanced cognitive control, although it should be mentioned that no direct relationship with cortisol levels was found (Vinkers et al., 2013). Accordingly, it may be suggested that the effect of psychosocial stress on decision-making may be different when tasks are administered immediately following a stressor or sometime thereafter. It is clear that this needs further study.

In men, it seems that high levels of cortisol following a stressor are related to risk-taking: high-cortisol responders show decreased IGT performance, while non/low-cortisol responders do not (van den Bos et al., 2009). However, these data seem to be in contrast with data on coping styles. Male subjects with a pro-active coping style are in general considered to be more risk-taking than male subjects with a reactive coping style (Koolhaas et al., 1999, 2010; Coppens et al., 2010). Subjects with a reactive coping style show a higher HPA-axis activity than subjects with a pro-active coping style, while subjects with a pro-active coping style show a higher SAM-axis activity than subjects with a reactive coping style (Koolhaas et al., 1999, 2010; Coppens et al., 2010). This apparent contradiction may be resolved when the stress response and the task are considered separately. Thus, the stress response in the TSST is directed to the challenge, i.e., the speech and arithmetical task in front of the panel. The physiological and neural changes would normally allow the organism to cope with this particular challenge. In case of the non/low-cortisol responders, which have a short-lasting SAM-axis activation during the TSST, this would be to take immediate action directed toward this particular stressor with an already learned response or routine (Koolhaas et al., 1999, 2010; Coppens et al., 2010). As the SAM-axis is strongly activated during the task with little effect thereafter, there is no after-effect when the IGT is administered. In case of the cortisol responders, which also have a short-lasting SAM-axis activation during the TSST and a long-lasting HPA-axis activation, the coping response is to "freeze" i.e., re-assess the situation by exploration and being sensitive to environmental cues (Koolhaas et al., 1999, 2010; Coppens et al., 2010). Given the observed effects of stress and cortisol on neural structures, their brain seems to be in a "exploratory mode," i.e., a decreased prefrontal activity and an increased ventral striatal (dopaminergic) activity in which risk-taking, as part of exploratory behavior, is included (Fiorillo et al., 2003). This "exploratory mode" seems to be set in motion during the TSST and remains for some time thereafter. This "exploratory mode" may be followed by an enhanced level of cognitive control (Vinkers et al., 2013), likely to be mediated by increased activity in the prefrontal cortex (Henckens et al., 2011), which may serve to store newly obtained information and/or regain homeostasis and cognitive control. Thus, when the IGT is administered shortly after the TSST, the ongoing coping response interferes with IGT task-performance. For, the IGT or other decision-making tasks require a delicate balance between cognitive control and reward/punishment-sensitivity or a prefrontal-subcortical balance (de Visser et al., 2011). As the brain of cortisol-responders is in a "exploratory/risk-taking mode" they are more sensitive to the immediately highly rewarding decks of cards, which is indeed the case as judged from their choice behavior (van den Bos et al., 2009). One may speculate that a different pattern emerges when the decision-making task in itself would be stressful. In this case, subjects with a pro-active coping style would probably take more risks than subjects with a reactive coping style as in this case the coping style is directly related to the task. It is clear that this needs further study. Data on rat behavior in resident-intruder paradigms in which pro-active copers show a fight-flight response and reactive copers a freezing response suggest such differences in risk-taking tendencies during interactions (Koolhaas et al., 1999, 2010).

Discussions on coping styles have nearly exclusively focused on differences in behavior in male subjects (Koolhaas et al., 1999, 2010; Coppens et al., 2010). While it has been suggested that male and female subjects in general differ in coping style, for instance fight-flight (pro-active) *versus* tend-and-befriend (reactive; Cousino Klein and Corwin, 2002) or problem-oriented (pro-active) *versus* emotion-oriented (reactive; see Tamres et al., 2002 for discussion), this is too limited a view as the differences in coping styles in male subjects already show. The current data rather suggest that a distinction in female subjects may also occur in coping styles, with in all likelihood "tend and befriend" as the most dominant one (Taylor et al., 2000; Cousino Klein and Corwin, 2002; Tamres et al., 2002) and likely to be a reactive coping style (Koolhaas et al., 2010). This coping style is related to the workings of oxytocin (Cousino Klein and Corwin, 2002), directed at caring for the young and/or seeking social support (Taylor et al., 2000; Tamres et al., 2002) and, possibly, leading to a reduction of HPA-axis activity (Cousino Klein and Corwin, 2002). Thus, the cortisol-mediated increase in risk-aversive behavior in reward-based tasks may be related to changes in neural structures toward safety and social support. Yet, it is clear that further studies are needed to substantiate this.

As risk-taking is in general considered to be a more disastrous mode of behavior than risk-aversiveness, the behavior of women following psychosocial stress may be considered as less detrimental than the behavior of men. However, this may not be entirely true. For instance, trying to rescue someone from a burning house with an objectively high risk of death or injury in the course of action is as disastrous as not-rescuing someone from a burning house with an objectively low risk of death or injury in the course of action. So it may be rather the context that gives one behavioral pattern an advantage over another or not than the attitude *per se*: today's hero may be tomorrow's fool and *vice versa*.

## LABORATORY STUDIES AND REAL-LIFE STUDIES

The laboratory environment offers the growing field of social neuroscience the opportunity to study general principles underlying brain-behavior relationships in a social context by using well-defined tasks tailored to be performed in and outside of scanners (review; Rilling and Sanfey, 2011). Similarly, the laboratory is well suited to study general effects of the social environment on individual behavior as for instance indicated in the previous section and as shown for instance by the effects of peers on risk-taking behavior (Gardner and Steinberg, 2005). However, at the same time it is often difficult, if not impossible, in laboratory settings to arrange the complete array of social contexts and/or social interactions. This becomes especially problematic when the behavior of the individual in relation to the social environment is the focus of attention, especially such as related to understanding the development of behavioral pathologies, intervention strategies and monitoring the success of therapies. For instance, in the case of the development of social conduct disorders probably not all social settings or social partners are equally likely to trigger a response, i.e., they may strongly differ between subjects and even across the life-span of subjects. As argued by others, *time*, i.e., when events relevant to the individual occur, and *context*, i.e., where and with whom events occur, are the limiting factors in laboratory studies (Johnson et al., 2009). Therefore, it would be ideal to study social interactions and their effect on the individual in real life as a complement to laboratory studies.

Among the most promising approaches to understanding time-limited behaviors in ecologically valid circumstances involves ambulatory monitoring through mobile technologies. This strategy is known alternately as Ecological Momentary Assessment (EMA) as well as the Experience Sampling Method (ESM), and it uses devices such as smart phones or other portable microcomputers to collect data at numerous intervals throughout the day. Like all methods, EMA/ESM also has its own limitations that include the necessity for all repeated assessments to remain brief as well as its reliance on subjective reports from the individual. However, extensive validation studies in diverse normal and psychiatric populations have demonstrated the feasibility and validity of this technique as a means of assessing psychological states and behavior in real time and in natural contexts (Granholm et al., 2008; Johnson et al., 2009; Husky et al., 2010). The major contribution of EMA/ESM is that it should provide a bridge between laboratory-based protocols with daily life behaviors that are otherwise inaccessible to the researcher. The value of EMA/ESM in investigating dynamic daily life risk factors has already been demonstrated relative to substance use behavior (Swendsen et al., 2000, 2011; Epstein et al., 2009), depressive cognitions (Swendsen, 1998), and many other "micro" processes of daily life.

Studying the dynamic character of social interactions including its long-term effects on the individual is standard in studies involving primates, either in the wild or in the laboratory. Still, it is clear that opportunities and possibilities for invasive neuroscience research are limited. In rodent studies, however, the opposite pattern seems to be present: while many experiments are directed at understanding the neural basis of for instance decision-making behavior in stand-alone tasks (de Visser et al., 2011), in general relatively little research is done in a long-term social context, an exception being for instance the work by Blanchard and Blanchard (1989), Blanchard et al. (1995, 2001, 2002). Recently home-cage tests have been developed allowing to address decision-making in a more naturalistic setting, i.e., to combine cognitive testing with a social/ecological-like environment (e.g., *Automated group-cage* (AGC), PhenoSys GmbH, Berlin, Germany; *Intellicage* (IC), Newbehavior AG, Zurich, Switzerland). In these home-cages, mice or rats are free to move and interact with each other but they can also voluntarily access operant modules situated inside or outside the home-cage. Using the *Home Cage Panels* (PRS Italia, Rome, Italy) Zoratto and colleagues (2013) developed a setting where animals were pair-housed, but could be tested singly. Furthermore, to assess social behavior in more detail programmes have been developed to analyse social interactions (De Chaumont et al., 2012). Thus, future studies in social neuroscience in rodents should develop protocols that combine new possibilities of studying the role of neural structures in behavior, such as by optogenetics, with well-defined and controlled home-cage social settings. This would allow inducing changes in behavior due to changes in neural structures within a social context in a controlled way.

## CONCLUDING REMARKS

A large extent of the exceptional success of the human species is related to our complex social life. It is therefore important to properly understand in what way the social environment may modulate decision-making. In the foregoing sections we have highlighted research on this topic, both from human and animal-based neuropsychological studies as well as insights from a behavioral ecology perspective. It is the combination of these top-down and bottom-up approaches that may enable us to fully understand decision-making and the social factors that affect it.

Although humans are outperforming animals regarding social skills, rodents and other animals bear some fundamental aspects of these skills as well, indicating the important role of social influences on decision-making in evolution. These comparisons may therefore not only enable us to better understand our own behavior, they may help us understand the ways in which our behavior may be fundamentally similar or different from that of other animals. However, with a few exceptions—like the observational fear learning task and Prisoner's Dilemma game—these kinds of translations between human and rodent based neuropsychological studies and behavioral ecology studies have rarely been done.

Future studies on social modulation of decision-making can therefore benefit by making the links between both fields of research and taking both a top-down and bottom-up approach. Furthermore, this may enable us to go beyond general social modulating effects and allow to understand how individual characteristics and heterogeneity within groups affect decisions that individuals make and the way this may ultimately affect group functioning, such as can be seen in human society. For instance, future experiments should focus on further dissecting the interaction between social stress, gender and decision-making taking timing, the kind of decision-making task and coping style into account. While in rats social stress has been mainly studied in the context of coping styles and neuro-endocrine changes (Koolhaas et al., 1999, 2010) as well as long-term consequences related to depression and buffering effects of the social environment hereon (e.g., Blanchard et al., 1995, 2001, 2002; Von Frijtag et al., 2000), no studies exist which have explicitly looked at decision-making either related to food, social partners or otherwise. Clearly such studies are needed in parallel to human studies to unravel short- and long-term effects of social stress in male and female subjects.

As argued, studies are needed in humans and animals under real-life conditions to assess the impact of (stressful) events on subsequent decisions. For instance, Newman and colleagues (2007) showed that daily hassles affected the decision to eat in female subjects in real life. Such studies clearly help in understanding both similarities and discrepancies between findings in real life and laboratory findings and thus enhance the application of laboratory-acquired knowledge in real-life conditions.

For a successful cross-species approach it is mandatory to show that the same principles underlie changes in decision-making behavior regarding social interactions and stress. This is all the more important when focusing e.g., on sex differences (van den Bos et al., 2013a). Thus, this requires understanding social behavior and its underlying principles in the context of behavioral ecology as outlined under section Decision-making in a social context for instance. Furthermore, it requires to use paradigms which in a species-specific way tap-off similar phenomena and/or are matched as closely as possible (see de Visser et al., 2011; van den Bos et al., 2012, 2013a; Jimura et al., 2013).

As a final remark, we do hope that this review may serve as a fruitful starting point for extending current studies and discussions of decision-making by incorporating the social environment.

## ACKNOWLEDGMENTS

The authors wish to thank Dr. Leonie de Visser for fruitful discussions. Furthermore, we wish to thank the reviewers whose comments helped to improve the focus of a number of sections of this review. The contribution of Jolle W. Jolles was supported by the Biotechnology and Biological Sciences Research Council (BBSRC).

## REFERENCES

Arnsten, A. F. T. (1998). The biology of being frazzled. *Science* 280, 1711–1712. doi: 10.1126/science.280.5370.1711

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat. Rev.* 10, 410–422.

Asch, S. E. (1955). Opinions and social pressure. *Sci. Am.* 193, 31–35.

Asch, S. E. (1956). Studies of independence and conformity: I a minority of one against a unanimous majority. *Psychol. Monograms* 70, 1–70.

Aureli, F., van Schaik, C. P., and van Hooff, J. A. R. A. M. (1989). Functional aspects of reconciliation among long-tailed macaques (*Macaca fascicularis*). *Am. J. Primatol.* 19, 39–51. doi: 10.1002/ajp.1350190105

Axelrod, R., and Hamilton, W. D. (1981). The evolution of cooperation. *Science* 211, 1390–1396. doi: 10.1126/science.7466396

Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., et al. (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1232–1237.

Baron, R. (1996). The Forgotten variable in conformity research: impact of task importance on social influence. *J. Pers. Soc. Psychol.* 71, 915–927.

Beauchamp, G. (2000). The effect of prior residence and pair bond on scrounging choices in flocks of zebra finches, *Taenopygia guttata*. *Behav. Processes* 52, 131–140.

Beauchamp, G. (2003). Group-size effects on vigilance: a search for mechanisms. *Behav. Processes* 63, 111–121.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat. Neurosci.* 8, 1458–1463.

Behrens, T. E., Hunt, L. T., and Rushworth, M. F. (2009). The computation of social behavior. *Science* 324, 1160–64. doi: 10.1126/science.1169694

Berns, G., Capra, C., Moore, S., and Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage* 49, 1–24. doi: 10.1016/j.neuroimage.2009.10.070

Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., and Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biol. Psychiatry* 58, 245–253.

Billeke, P., Zamorano, F., Cosmelli, D., and Aboitiz, F. (2012). Oscillatory brain activity correlates with risk perception and predicts social decisions. *Cereb. Cortex* doi: 10.1093/cercor/bhs269. [Epub ahead of print].

Blanchard, D. C., Spencer, R. L., Weiss, S. M., Blanchard, R. J., McEwen, B., and Sakai, R. R. (1995). Visible burrow system as a model of chronic social stress, behavioral and neuroendocrine correlates. *Psychoneuroendocrinoly* 20, 117–134. doi: 10.1016/0306-4530(94)E0045-B

Blanchard, R. J., and Blanchard, D. C. (1989). Antipredator defensive behaviors in a visible burrow system. *J. Comp. Psychol.* 103, 70–82.

Blanchard, R. J., McKittrick, C. R., and Blanchard, D. C. (2001). Animal models of social stress: effects on behaviour and brain neurochemical systems. *Physiol. Behav.* 73, 261–271.

Blanchard, R. J., McKittrick, C. R., Hardy, M. P., and Blanchard, D. C. (2002). "Social stress affects hormones, brain and behavior," in *Hormones and Behavior*, eds D. Pfaff, A. Arnold, A. Etgen, S. Fahrbach, R. Rubin (San Diego, CA: Academic Press), 735–772.

Bond, R. (2005). Group size and conformity. *Group Proces. Intergroup Relat.* 8, 331–354. doi: 10.1177/1368430205056464

Bond, R., and Smith, P. B. (1996). Culture and conformity: a meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychol. Bull.* 119, 111–137.

Boyd, R., and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: The University of Chicago Press.

Brown, M. F., Knight-Green, M. B., Lorek, E. J., Packard, C., Shallcross, W. L., Wifall, T., et al. (2008). Social working memory: memory for another rat's spatial choices can increase or decrease choice tendencies. *Learn. Behav.* 36, 327–340.

Bruchey, A. K., Jones, C. E., and Monfils, M. H. (2010). Fear conditioning by-proxy: social transmission of fear during memory retrieval. *Behav. Brain Res.* 214, 80–84.

Bshary, R., and Grutter, A. S. (2002). Asymmetric cheating opportunities and partner control in a cleaner fish mutualism. *Anim. Behav.* 63, 547–555. doi: 10.1006/anbe.2001.1937

Burke, C. J., Tobler, P. N., Schultz, W., and Baddeley, M. (2010). Striatal BOLD response reflects the impact of herd information on financial decisions. *Front. Hum. Neurosci.* 4:48. doi: 10.3389/fnhum.2010.00048

Campanhã, C., Minati, L., Fregni, F., and Boggio, P. S. (2011). Responding to unfair offers made by a friend: Neuroelectrical activity changes in the anterior medial prefrontal cortex. *J. Neurosci.* 31, 15569–15574. doi: 10.1523/JNEUROSCI.1253-11.2011

Campbell-Meiklejohn, D. K., Bach, D. R., Roepstorff, A., Dolan, R. J., and Frith, C. D. (2010). How the opinion of others affects our valuation of objects. *Curr. Biol.* 20, 1165–1170.

Chen, Q., Panksepp, J. B., and Lahvis, G. P. (2009). Empathy is moderated by genetic background in mice. *PLoS ONE* 4:e4387. doi: 10.1371/journal.pone.0004387

Choleris, E., and Kavaliers, M. (1999). Social learning in animals: sex differences and neurobiological analysis. *Pharmacol. Biochem. Behav.* 64, 767–776. doi: 10.1016/S0091-3057(99)00141-0

Clutton-Brock, T. H. (2002). Breeding together: kin selection, reciprocity and mutualism in cooperative animal societies. *Science* 296, 69–72. doi: 10.1126/science.296.5565.69

Clutton-Brock, T. H. (2009). Cooperation between non-kin in animal societies. *Nature* 462, 51–57. doi: 10.1038/nature08366

Conradt, L., and Roper, T. J. (2003). Group decision-making in animals. *Nature* 421, 155–158. doi: 10.1038/nature01294

Conradt, L., and Roper, T. J. (2005). Consensus decision making in animals. *Trends Ecol. Evol.* 20, 449–456. doi: 10.1016/j.tree.2005.05.008

Coolen, I., Van Bergen, Y., Day, R. L., and Laland, K. N. (2003). Species difference in adaptive use of public information in sticklebacks. *Proc. Biol. Sci.* 270, 2413–2419. doi: 10.1098/rspb.2003.2525

Coolen, I., Ward, A. J. W., Hart, P. J. B., and Laland, K. N. (2005). Foraging nine-spined sticklebacks prefer to rely on public information over simpler social cues. *Behav. Ecol.* 16, 865–870.

Coppens, C. M., de Boer, S. F., and Koolhaas, J. M. (2010). Coping styles and behavioural flexibility: towards underlying mechanisms. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 4021–4028. doi: 10.1098/rstb.2010.0217

Cornelisse, S., van Stegeren, A. H., and Joëls, M. (2011). Implications of psychosocial stress on memory formation in a typical male versus female student sample. *Psychoneuroendocrinology* 36, 569–578. doi: 10.1016/j.psyneuen.2010.09.002

Cousino Klein, L., and Corwin, E. J. (2002). Seeing the unexpected: how sex differences in stress responses may provide a new perspective on the manifestations of psychiatric disorders. *Curr. Psychiatry Rep.* 4, 441–448. doi: 10.1007/s11920-002-0072-z

Couzin, I. D. (2009). Collective cognition in animal groups. *Trends Cogn. Sci* 13, 36–43.

Couzin, I. D., and Franks, N. R. (2003). Self-organized lane formation and optimized traffic flow in army ants. *Proc. Biol. Sci.* 270, 139–146. doi: 10.1098/rspb.2002.2210

Couzin, I. D., and Krause, J. (2003). Self-organization and collective behavior in vertebrates. *Adv. Stud. Behav.* 32, 1–75.

Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature* 433, 513–516. doi: 10.1038/nature03236

Crişan, L. G., Pana, S., Vulturar, R., Heilman, R. M., Szekely, R., Druğa, B., et al. (2009). Genetic contributions of the serotonin transporter to social learning of fear and economic decision making. *Soc. Cogn. Affect. Neurosci.* 4, 399–408.

Crone, E. A., and van der Molen, M. W. (2004). Developmental changes in real life decision mak- ing: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Dev. Neuropsychol.* 25, 251–279.

Dally, J. M., Clayton, N. S., and Emery, N. J. (2008). Social influences on foraging by rooks (*Corvus frugilegus*). *Behaviour* 145, 1101–1124. doi: 10.1163/156853908784474470

Danchin, E., Giraldeau, L.-A., Valone, T. J., and Wagner, R. H. (2004). Public information: from nosy neighbors to cultural evolution. *Science* 305, 487–491. doi: 10.1126/science.1098254

Davies, N. B., Krebs, J. R., and West, S. A. (2012). *An Introduction to Behavioural Ecology*. Oxford: Wiley-Blackwell.

Day, R. L., Macdonald, T., Brown, C., Laland, K. N., and Reader, S. M. (2001). Interactions between shoal size and conformity in guppy social foraging. *Anim. Behav.* 62, 917–925. doi: 10.1006/anbe.2001.1820

De Chaumont, F., Coura, R. D.-S., Serreau, P., Cressant, A., Chabout, J., Granon, S., et al. (2012). Computerized video analysis of social interactions in mice. *Nat. Methods* 4, 410–417.

De Dreu, C. K. W., Greer, L. L., Handgraaf, M. J. J., Shalvi, S., van Kleef, G. A., Baas, M., et al. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science* 328, 1408–1411. doi: 10.1126/science.1189047

Dellu, F., Piazza, P. V., Mayo, W., Le Moal, M., and Simon, H. (1996). Novelty-seeking in rats - biobehavioral characteristics and possible relationship with the sensation-seeking trait in man. *Neuropsychobiology* 34, 136–145. doi: 10.1159/000119305

de Visser, L., Homberg, J. R., Mitsogiannis, M., Zeeb, F. D., Rivalan, M., Fitoussi, A., et al. (2011). Rodent versions of the Iowa gambling task: opportunities and challenges for the understanding of decision-making. *Front. Neurosci.* 5:109. doi: 10.3389/fnins.2011.00109

de Visser, L., van der Knaap, L. J., van der Loo, A. J. A. E., van de Weerd, C. M. M., Ohl, F., and van den Bos, R. (2010). Trait anxiety affects decision-making differently in healthy men and women: towards gender-specific endophenotypes of anxiety. *Neuropsychologia* 48, 1598–1606. doi: 10.1016/j.neuropsychologia.2010.01.027

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., et al. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* 325, 621–625. doi: 10.1126/science.1171203

Dickerson, S. S., and Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychol. Bull.* 130, 355–391.

Dindo, M., Whiten, A., and De Waal, F. B. M. (2009). In-group conformity sustains different foraging traditions in capuchin monkeys (*Cebus apella*). *PLoS ONE* 4:e7858. doi: 10.1371/journal.pone.0007858

Dyer, J. R. G., Ioannou, C. C., Morrell, L. J., Croft, D. P., Couzin, I. D., Waters, D. A., et al. (2008). Consensus decision making in human crowds. *Anim. Behav.* 75, 461–470. doi: 10.1016/j.anbehav.2007.05.010

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., and Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evol. Hum. Behav.* 29, 56–64. doi: 10.1016/j.evolhumbehav.2007.08.003

Epstein, D. H., Willner-Reid, J., Vahabzadeh, M., Mezghanni, M., Lin, J. L., and Preston, K. L. (2009). Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin and use. *Arch. Gen. Psychiatry* 66, 88–94. doi: 10.1001/archgenpsychiatry.2008.509

Erickson, K., Drevets, W., and Schulkin, J. (2003). Glucocorticoid regulation of diverse cognitive functions in normal and pathological emotional states. *Neurosci. Biobehav. Rev.* 27, 233–246. doi: 10.1016/S0149-7634(03)00033-2

Fehr, E., and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791. doi: 10.1038/nature02043

Fehr, E., and Fischbacher, U. (2004a). Social norms and human cooperation. *Trends Cogn. Sci.* 8, 185–190.

Fehr, E., and Fischbacher, U. (2004b). Third-party punishment and social norms. *Evol. Hum. Behav.* 25, 63–87. doi: 10.1016/S1090-5138(04)00005-4

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902. doi: 10.1126/science.1077349

Foley, P., and Kirschbaum, C. (2010). Human hypothalamus-pituitary- adrenal axis responses to acute psychosocial stress in laboratory settings. *Neurosci. Biobehav. Rev.* 35, 91–96. doi: 10.1016/j.neubiorev.2010.01.010

Frank, S. A. (2003). Perspective: repression of competition and the evolution of cooperation. *Evolution* 57, 693–705.

Franks, N. R., Dornhaus, A., Fitzsimmons, J. P., and Stevens, M. (2003). Speed versus accuracy in collective decision making. *Proc. Biol. Sci.* 270, 2457–2463. doi: 10.1098/rspb.2003.2527

Franks, N. R., and Richardson, T. (2006). Teaching in tandem-running ants. *Nature* 439, 153. doi: 10.1038/439153a

Fraser, D. N., Stahl, D., and Aureli, F. (2008). Stress reduction through consolation in chimpanzees. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8557–8562. doi: 10.1073/pnas.0804141105

Fretwell, S. D. (1972). *Populations in a Seasonal Environment.* Princeton, NJ: Princeton University Press

Frith, C. D., and Singer, T. (2008). The role of social cognition in decision making. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 3875–3886.

Galef, B. (1995). Why behaviour patterns that animals learn socially are locally adaptive. *Anim. Behav.* 49, 1325–1334. doi: 10.1006/anbe.1995.0164

Galef, B. G. (1986). Social interaction modifies learned aversions, sodium appetite, and both palatability and handling-time induced dietary preference in rats (Rattus norvegicus). *J. Comp. Psychol.* 100, 432–439. doi: 10.1037/0735-7036.100.4.432

Galef, B. G. (1988). "Imitation in animals: history, definition and interpretation of data form the psychological laboratory," in *Social Learning: Psychological and Biological Perspectives*, eds T. R. Zentall and B. G. Galef (Hillsdale: Erlbaum), 3–28.

Galef, B. G. (2007). "Social learning in rodents," in *Rodent Societies*. eds P. W. Sherman and J. Wolff (Chicago: University of Chicago Press), 207–215.

Galef, B. G., and Giraldeau, L.-A. (2001). Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. *Anim. Behav.* 61, 3–15. doi: 10.1006/anbe.2000.1557

Galef, B. G., and Whiskin, E. E. (1995). Learning socially to eat more of one food than of another. *J. Comp. Psychol.* 109, 99–101. doi: 10.1037/0735-7036.109.1.99

Galef, B. G., and Whiskin, E. E. (1998). Limits on social influence on food choices of Norway rats. *Anim. Behav.* 56, 1015–1020. doi: 10.1006/anbe.1998.0867

Galef, B. G., and Whiskin, E. E. (2008). "Conformity" in Norway rats. *Anim. Behav.* 75, 2035–2039. doi: 10.1016/j.anbehav.2007.11.012

Gardner, M., and Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood, an experimental study. *Dev. Psychol.* 41, 625–635.

Giraldeau, L.-A., and Caraco, T. (2000). *Social Foraging Theory*. Princeton, NJ: Princeton University Press.

Goldstein, J. M., Jerram, M., Abbs, B., Whitfield-Gabrieli, S., and Makris, N. (2010). Sex differences in stress response circuitry activation dependent on female hormonal cycle. *J. Neurosci.* 30, 431–438. doi: 10.1523/JNEUROSCI.3021-09.2010

Graham, L. K., Yoon, T., and Kim, J. J. (2010). Stress impairs optimal behavior in a water foraging choice task in rats. *Learn. Mem.* 17, 1–4.

Granholm, E., Loh, C., and Swendsen, J. (2008). Feasibility and validity of computerized ecological momentary assessment in schizophrenia. *Schizophr. Bull.* 34, 507–514.

Griffiths, R. A., and Foster, J. P. (1998). The effect of social interactions on tadpole activity and growth in the British anuran amphibians (*Bufo bufo, B. calamita, and Rana temporaria). J. Zool. Soc/ London* 245, 431–437.

Guzmán, Y. F., Tronson, N. C., Guedea, A., Huh, K. H., Gao, C., and Radulovic, J. (2009). Social modeling of conditioned fear in mice by non-fearful conspecifics. *Behav. Brain Res.* 201, 173–178.

Hamilton, W. D. (1964). The genetical evolution of social behaviour, I and II. *J. Theor. Biol.* 7, 1–52.

Hammerstein, P. (2003). *Genetic and Cultural Evolution of Cooperation*. Cambridge, UK: MIT Press.

Harcourt, J. L., Ang, T. Z., Sweetman, G., Johnstone, R. A., and Manica, A. (2009). Social feedback and the emergence of leaders and followers. *Curr. Biol.* 19, 248–252.

Harper, D. G. C. (1982). Competitive foraging in mallards: 'ideal free' ducks. *Anim. Behav.* 30, 575–558. doi: 10.1016/S0003-3472(82)80071-7

Heinrichs, M., Baumgartner, T., Kirschbaum, C., and Ehlert, U. (2003). Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biol. Psychiatry* 54, 1389–1398.

Helbing, D., Farkas, I., and Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature* 407, 487–490. doi: 10.1038/35035023

Helbing, D., and Molnar, P. (1995). Social force model for pedestrian dynamics. *Phys. Rev. E* 51, 4282–4286.

Henckens, M. J., van Wingen, G. A., Joëls, M., and Fernández, G. (2011). Time-dependent corticosteroid modulation of prefrontal working memory processing. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5801–5806.

Henrich, J., Bowles, S., Smith, E. A., Young, P., Boyd, R., Sigmund, K., et al. (2003). "The cultural and genetic evolution of human cooperation," in *Genetic and Cultural Evolution of Cooperation,* ed P. Hammerstein (Cambridge, MA: MIT Press), 445–468.

Herbert-Read, J., Perna, A., Mann, R. P., Schaerf, T. M., Sumpter, D. J. T., and Ward, A. J. W. (2011). Inferring the rules of interaction of shoaling fish. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18726–18731.

Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biol. Rev.* 69, 207–231.

Heyes, C. M., and Galef, B. G. (1996). *Social Learning in Animals: The Roots of Culture*. London: Academic Press.

Hillman, K. L., and Bilkey, D. K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. *Nat. Neurosci.* 15, 1290–1297.

Homberg, J. R. (2012). Serotonin and decision making processes. *Neurosci. Biobehav. Rev.* 36, 218–336.

Hooker, C. I., Verosky, S. C., Miyakawa, A., Knight, R. T., and D'Esposito, M. (2008). The influence of personality on neural mechanisms of observational fear and reward learning. *Neuropsychologia* 46, 2709–2724. doi: 10.1016/j.neuropsychologia.2008.05.005

Huettel, S. A., and Kranton, R. E. (2012). Identity economics and the brain: uncovering the mechanisms of social conflict. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 680–691.

Husky, M. M., Gindre, C., Mazure, C. M., Brebant, C., Nolen-Hoeksema, S., Sanacora, G., et al. (2010). Computerized ambulatory monitoring in mood disorders, Feasibility, compliance, and reactivity. *Psychiatry Res.* 178, 440–442. doi: 10.1016/j.psychres.2010.04.045

Jeon, D., Kim, S., Chetana, M., Jo, D., Ruley, H. E., Lin, S. Y., et al. (2010). Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. *Nat. Neurosci.* 13, 482–488.

Jimura, K., Chushak, M. S., and Braver, T. S. (2013). Impulsivity and self-control during intertemporal decision making linked to the neural dynamics of reward value representation. *J. Neurosci.* 33, 344–357. doi: 10.1523/JNEUROSCI.0919-12.2013

Joëls, M., Sarabdjitsingh, R. A., and Karst, H. (2012). Unraveling the time domains of corticosteroid hormone influences on brain activity: rapid, slow, and chronic modes. *Pharmacol. Rev.* 64, 901–938.

Johnson, E. I., Grondin, O., Barrault, M., Faytout, M., Helbig, S., Husky, M., et al. (2009). Computerized ambulatory monitoring in psychiatry, a multi-site collaborative study of acceptability, compliance, and reactivity. *Int. J. Methodol. Psychiatr. Res.* 18, 48–57.

Jolles, J. W., de Visser, L., and van den Bos, R. (2011). Male Wistar rats show individual differences in an animal model of conformity. *Anim. Cogn.* 14, 769–773.

Jolles, J. W., King, A., Manica, A., and Thornton, A. (2013a). Heterogeneous structure in mixed-species corvid flocks in flight. *Anim. Behav.* 85, 743–750.

Jolles, J. W., Ostojić, L., and Clayton, N. S. (2013b). Dominance, pair bonds and boldness determine social foraging tactics in rooks (*Corvus frugilegus*). *Anim. Behav.* 85, 1261–1269. doi: 10.1016/j.anbehav.2013.03.013

Joly, D., and Sanger, D. J. (1991). Social competition in rats: a test sensitive to acutely administered anxiolytics. *Behav. Pharmacol.* 2, 205–213.

Kelly, M. M., and Forsyth, J. P. (2007). Sex differences in response to an observational fear conditioning procedure. *Behav. Ther.* 38, 340–349.

Kelly, M. M., and Forsyth, J. P. (2009). Associations between emotional avoidance, anxiety sensitivity, and reactions to an observational fear challenge procedure. *Behav. Res. Ther.* 47, 331–338.

Kelly, M. M., Tyrkaa, A. R., Anderson, G. M., Pricea, L. H., and Carpentera, L. L. (2008). Sex differences in emotional and physiological responses to the Trier Social Stress Test. *J. Behav. Ther. Exp. Psychiatry* 39, 87–98.

Kendal, R. L., Coolen, I., Van Bergen, Y., and Laland, K. N. (2005). Trade-offs in the adaptive use of social and asocial Learning. *Adv. Study Behav.* 35, 333–379. doi: 10.1016/S0065-3454(05)35008-X

Kern, S., Oakes, T. R., Stone, C. K., McAuliff, E. M., Kirschbaum, C., and Davidson, R. J. (2008). Glucose metabolic changes in the prefrontal cortex are associated with HPA axis response to a psychosocial stressor. *Psychoneuroendocrinology* 33, 517–529. doi: 10.1016/j.psyneuen.2008.01.010

King, A. J., Johnson, D. D. P., and Van Vugt, M. (2009). The origins and evolution of leadership. *Curr. Biol.* 19, R911–R916. doi: 10.1016/j.cub.2009.07.027

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., and Hellhamer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosom. Med.* 61, 154–162.

Kirschbaum, C., Pirke, K. M., and Hellhamer, D. H. (1993). The 'Trier Social Stress Test' - a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* 28, 76–81. doi: 10.1159/000119004

Kiyokawa, Y., Kikusui, T., Takeuchi, Y., and Mori, Y. (2004). Partner's stress status influences social buffering effects in rats. *Behav. Neurosci.* 118, 798–804.

Kiyokawa, Y., Shimozuru, M., Kikusui, T., Takeuchi, Y., and Mori, Y. (2006). Alarm pheromone increases defensive and risk assessment behaviors in male rats. *Physiol. Behav.* 87, 383–387.

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., and Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron* 61, 140–151. doi: 10.1016/j.neuron.2008.11.027

Knapska, E., Mikosz, M., Werka, T., and Maren, S. (2010). Social modulation of learning in rats. *Learn. Mem.* 17, 35–42.

Koolhaas, J. M., De Boer, S. F., Coppens, C. M., and Buwalda, B. (2010). Neuroendocrinology of coping styles: towards understanding the biology of individual variation. *Front. Neuroendocrinol.* 31, 307–321. doi: 10.1016/j.yfrne.2010.04.001

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., et al. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neurosci. Biobehav. Rev.* 23, 925–935.

Koot, S., Baars, A., Hesseling, P., van den Bos, R., and Joëls, M. (2011). "Stress and decision-making in rats, effects of corticosterone," in *EBBS Conference Seville Spain September*, (Seville) 9-12 2011, Abstract D12–D49.

Koot, S., Baars, A., Hesseling, P., van den Bos, R., and Joëls, M. (2013). Time-dependent effects of corticosterone in a rodent model of the Iowa Gambling Task. *Neuropharmacology* 70, 306–315. doi: 10.1016/j.neuropharm.2013.02.008

Koski, S. E., Koops, K., and Sterck, E. H. M. (2007). Reconciliation, relationship quality, and postconflict anxiety: Testing the integrated hypothesis in captive chimpanzees. *Am. J. Primatol.* 69, 1–15.

Krause, J., and Ruxton, G. D. (2002). *Living in groups*. Oxford: Oxford University Press.

Kudielka, B. M., and Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biol. Psychol.* 69, 113–132.

Kurvers, R. H. J. M., Van Oers, K., Nolet, B., Jonker, R. M., Van Wieren, S. E., Prins, H. H. T., et al. (2010). Personality predicts the use of social information. *Ecol. Lett.* 13, 829–837.

Lachlan, R. F., Crooks, L., and Laland, K. N. (1998). Who follows whom. Shoaling preferences and social learning of foraging information in guppies. *Anim. Behav.* 56, 181–190. doi: 10.1006/anbe.1998.0760

Laland, K. N. (2004). Social learning strategies. *Learn. Behav.* 32, 4–14.

Laland, K. N. (2008). Animal cultures. *Curr. Biol.* 18, R366–R370. doi: 10.1016/j.cub.2008.02.049

Laland, K. N., and Williams, K. (1997). Shoaling generates social learning of foraging information in guppies. *Anim. Behav.* 53, 1161–1169. doi: 10.1006/anbe.1996.0318

Laland, K. N., and Williams, K. (1998). Social transmission of maladaptive information in the guppy. *Behav. Ecol.* 9, 493–499. doi: 10.1093/beheco/9.5.493

Laland, K. N., Atton, N., and Webster, M. M. (2011). From fish to fashion: experimental and theoretical insights into the evolution of culture. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 958–968. doi: 10.1098/rstb.2010.0328

Lempert, K. M., Porcelli, A. J., Delgado, M. R., and Tricomi, E. (2012). Individual differences in delay discounting under acute stress: the role of trait perceived stress. *Front. Psychol.* 3:251. doi: 10.3389/fpsyg.2012.00251

Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58, 259–289. doi: 10.1146/annurev.psych.58.110405.085654

Lighthall, N. R., Mather, M., and Gorlick, M. A. (2009). Acute stress increases sex differences in risk seeking in the balloon analogue risk task. *PLoS ONE* 4:e6002, 1–6. doi:10.1371/journal.pone.0006002

Łopuch, S., and Popik, P. (2011). Cooperative behavior of laboratory rats (Rattus norvegicus) in an instrumental task. *J. Comp. Psychol.* 125, 250–253. doi: 10.1037/a0021532

Magnhagen, C., and Bunnefeld, N. (2009). Express your personality or go along with the group: what determines the behaviour of shoaling perch. *Proc. Biol. Sci.* 276, 3369–3375. doi: 10.1098/rspb.2009.0851

Maruyama, Y., Kawano, A., Okamoto, S., Ando, T., Ishitobi, Y., Tanaka, Y., et al. (2012). Differences in salivary alpha-amylase and cortisol responsiveness following exposure to electrical stimulation versus the Trier Social Stress Tests. *PLoS ONE* 7:e39375. doi: 10.1371/journal.pone.0039375

Mather, M., Lighthall, N. R., Nga, L., and Gorlick, M. A. (2010). Sex differences in how stress affects brain activity during face viewing. *Neuroreport* 21, 933–937. doi: 10.1097/WNR.0b013e32833ddd92

Maynard Smith, J., and Price, G. R. (1973). The logic of animal conflict. *Nature* 246, 15–18. doi: 10.1038/246015a0

Morgan, T. J. H., and Laland, K. N. (2012). The biological bases of conformity. *Front. Neurosci* 6:87. doi: 10.3389/fnins.2012.00087

Nagy, M., Akos, Z., Biro, D., and Vicsek, T. (2010). Hierarchical group dynamics in pigeon flocks. *Nature* 464, 890–893. doi: 10.1038/nature08891

Nater, U. M., La Marca, R., Florin, L., Moses, A., Langhans, W., Koller, M. M., et al. (2006). Stress-induced changes in human salivary alpha-amylase activity - associations with adrenergic activity. *Psychoneuroendocrinology* 31, 49–58. doi: 10.1016/j.psyneuen.2005.05.010

Nater, U. M., and Rohleder, N. (2009). Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research. *Psychoneuroendocrinology* 34, 486–496. doi: 10.1016/j.psyneuen.2009.01.014

Nater, U. M., Rohleder, N., Gaab, J., Berger, S., Jud, A., Kirschbaum, C., et al. (2005). Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. *Int. J. Psychophysiol.* 55, 333–342. doi: 10.1016/j.ijpsycho.2004.09.009

Nater, U. M., Rohleder, N., Schlotze, W., Ehlert, U., and Kirschbaum, C. (2007). Determinants of the diurnal course of salivary alpha-amylase. *Psychoneuroendocrinology* 32, 392–401. doi: 10.1016/j.psyneuen.2007.02.007

Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., and Barabási, A. L. (2000). The sound of many hands clapping. *Nature* 403, 849–850. doi: 10.1038/35002660

Newman, E., O'Connor, D. B., and Conner, M. (2007). Daily hassles and eating behaviour: the role of cortisol reactivity status. *Psychoneuroendocrinology* 32, 125–132. doi: 10.1016/j.psyneuen.2006.11.006

Nicol, C. J., and Pope, S. J. (1999). The effects of demonstrator social status and prior foraging success on social learning in laying hens. *Anim. Behav.* 57, 163–171. doi: 10.1006/anbe.1998.0920

Nielsen, S. E., Segal, S. K., Worden, I. V., Yim, I. S., and Cahill, L. (2013). Hormonal contraception use alters stress responses and emotional memory. *Biol. Psychol.* 92, 257–266.

Olsson, A., and Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends Cogn. Sci.* 12, 65–71.

Olsson, A., Nearing, K. I., and Phelps, E. A. (2007). Learning fears by observing others: the neural systems of social fear transmission. *Soc. Cogn. Affect. Neurosci.* 2, 3–11. doi: 10.1093/scan/nsm005

Olsson, A., and Phelps, E. A. (2004). Learned fear of "unseen" faces after Pavlovian, observational, and instructed fear. *Psychol. Sci.* 15, 822–828. doi: 10.1111/j.0956-7976.2004.00762.x

Olsson, A., and Phelps, E. A. (2007). Social learning of fear. *Nat. Neurosci.* 10, 1095–1102.

Overman, W. H. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain Cogn.* 55, 134–147. doi: 10.1016/S0278-2626(03)00279-3

Overman, W. H., Frassrand, K., Ansel, S., Trawalter, S., Bies, B., and Redmond, A. (2004). Performance on the IOWA card task by adolescents and adults. *Neuropsychologia* 42, 1838–1851. doi: 10.1016/j.neuropsychologia.2004.03.014

Panksepp, J. B., and Lahvis, G. P. (2011). Rodent empathy and affective neuroscience. *Neurosci. Biobehav. Rev.* 35, 1864–1875. doi: 10.1016/j.neubiorev.2011.05.013

Piazza, P. V., Deroche, V., Deminiere, J.-M., Maccari, S., Le Moal, M., and Simon, H. (1993). Corticosterone in the range of stress-induced levels possesses reinforcing properties, Implications for sensation-seeking behaviors. *Proc. Natl. Acad. Sci. U.S.A.* 90, 11738–11742. doi: 10.1073/pnas.90.24.11738

Piazza, P. V., and Le Moal, M. (1997). Glucocorticoids as a biological substrate of reward: physiological and pathophysiological implications. *Brain Res. Rev.* 25, 359–372. doi: 10.1016/S0165-0173(97)00025-8

Pike, T. W., and Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biol. Lett.* 6, 466–468. doi: 10.1098/rsbl.2009.1014

Pitcher, A. T. J., Magurran, A. E., Winfield, I. J., and Pitcher, T. J. (1982). Fish in larger shoals find food faster. *Behav. Ecol. Sociobiol.* 10, 149–151. doi: 10.1007/BF00300175

Piyapong, C., Krause, J., Chapman, B. B., Ramnarine, I. W., Louca, V., and Croft, D. P. (2010). Sex matters: a social context to boldness in guppies (*Poecilia reticulata*). *Behav. Ecol.* 21, 3–8. doi: 10.1093/beheco/arp142

Porcelli, A. J., and Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychol. Sci.* 20, 278–283.

Porcelli, A. J., Lewis, A. H., and Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Front. Neurosci.* 6:157. doi: 10.3389/fnins.2012.00157

Preston, S. D., Buchanan, T. W., Stansfield, R. B., and Bechara, A. (2007). Effects of anticipatory stress on decision making in a gambling task. *Behav. Neurosci.* 121, 257–263.

Pruessner, J. C., Champagne, F., Meaney, M. J., and Dagher, A. (2004). Dopamine release in response to a psychological stress in humans and its relationship to early life maternal care: a Positron Emission Tomography study using [11C]-Raclopride. *J. Neurosci.* 24, 2825–2831.

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., and Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biol. Rev.* 82, 291–318.

Reebs, S. G. (2000). Can a minority of informed leaders determine the foraging movements of a fish shoal. *Anim. Behav.* 59, 403–409.

Rilling, J. K., and Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annu. Rev. Psychol.* 62, 23–48.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Rudebeck, P. H., Buckley, M. J., Walton, M. E., and Rushworth, M. F. (2006). A role for the macaque anterior cingulate gyrus in social valuation. *Science* 313, 1310–1312.

Rutte, C., and Taborsky, M. (2007). Generalized reciprocity in rats. *PLoS Biol.* 5:e196. doi:10.1371/journal.pbio.0050196

Rutte, C., and Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (Rattus norvegicus): direct vs generalised reciprocity. *Behav. Ecol. Sociobiol.* 62, 499–505.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Schneeberger, K., Dietz, M., and Taborsky, M. (2012). Reciprocal cooperation between unrelated rats depends on cost to donor and benefit to recipient. *BMC Evol. Biol.* 12:41. doi: 10.1186/1471-2148-12-41

Schuett, W., and Dall, S. R. X. (2009). Sex differences, social context and personality in zebra finches, *Taeniopygia guttata*. *Anim. Behav.* 77, 1041–1050.

Schwab, C., Bugnyar, T., Schloegl, C., and Kotrschal, K. (2008). Enhanced social learning between siblings in common ravens, *Corvus corax*. *Anim. Behav.* 75, 501–508.

Schwabe, L., Hoeffken, O., Tegenthoff, M., and Wolf, O. T. (2013). Opposite effect of noradrenergic arousal on amygdala processing of fearful faces in men and women. *Neuroimage* 73, 1–7.

Seeley, T. D. (2003). Consensus building during nest-site selection in honey bee swarms: the expiration of dissent. *Behav. Ecol. Sociobiol.* 53, 417–424.

Shafiei, N., Gray, M., Viau, V., and Floresco, S. B. (2012). Acute stress induces selective alterations in cost/benefit decision-making. *Neuropsychopharmacology* 37, 2194–2209.

Starcke, K., and Brand, M. (2012). Decision making under stress: a selective review. *Neurosci. Biobehav. Rev.* 36, 1228–1248.

Starcke, K., Polzer, C., Wolf, O. T., and Brand, M. (2011). Does stress alter everyday moral decision-making. *Psychoneuroendocrinology* 36, 210–219.

Starcke, K., Wolf, O. T., Markowitsch, H. J., and Brand, M. (2008). Anticipatory stress influences decision making under explicit risk conditions. *Behav. Neurosci.* 122, 1352–1360.

Stark, R., Wolf, O. T., Tabbert, K., Kagerer, S., Zimmerman, M., Kirsch, P., et al. (2006). Influence of the stress hormone cortisol on fear conditioning in humans: evidence for sex differences in the response of the prefrontal cortex. *Neuroimage* 32, 1290–1298.

Stevens, J. R., and Hauser, M. D. (2004). Why be nice. Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* 8, 60–65.

Stöwe, M., Bugnyar, T., Heinrich, B., and Kotrschal, K. (2006a). Effects of group size on approach to novel objects in ravens (*Corvus corax*). *Ethology* 112, 1079–1088.

Stöwe, M., Bugnyar, T., Loretto, M.-C., Schloegl, C., Range, F., and Kotrschal, K. (2006b). Novel object exploration in ravens (*Corvus corax*): effects of social relationships. *Behav. Processes* 73, 68–75.

Sumpter, D. J. T. (2006). The principles of collective animal behaviour. *Philos. Trans. R. Soc. B Biol. Sci.* 361, 5–22.

Swaney, W., Kendal, J., Capon, H., Brown, C., and Laland, K. N. (2001). Familiarity facilitates social learning of foraging behaviour in the guppy. *Anim. Behav.* 62, 591–598.

Swendsen, J. D. (1998). The helplessness-hopelessness theory and daily mood experience: an idiographic and cross-situational perspective. *J. Pers. Soc. Psychol.* 74, 1398–1408.

Swendsen, J. D., Tennen, H., Carney, M. A., Affleck, G., Willard, A., and Hromi, A. (2000). Mood and alcohol consumption: an experience sampling test of the self-medication hypothesis. *J. Abnorm. Psychol.* 109, 198–204.

Swendsen, J., Ben-Zeev, D., and Granholm, E. (2011). Real-time ambulatory monitoring of substance use and symptom expression in schizophrenia. *Am. J. Psychiatry* 168, 202–209.

Takahashi, T., Ikeda, K., and Hasegawa, T. (2007). Social evaluation-induced amylase elevation and economic decision-making in the dictator game in humans. *Neuroendocrinol. Lett.* 28, 662–665.

Tamres, L. K., Janicki, D., and Helgeson, V. S. (2002). Sex differences in coping behavior: a meta-analytic review and an examination of relative coping. *Personal. Soc. Psychol. Rev.* 6, 2–30.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., and Updegraff, J. A. (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychol. Rev.* 107, 411–429.

Ter Horst, J. P., Kentrop, J., de Kloet, E. R., and Oitzl, M. S. (2013). Stress and estrous cycle affect strategy but not performance of female C57BL/6J mice. *Behav. Brain Res.* 241 92–95.

Thoma, M. V., Kirschbaum, C., Wolf, J. M., and Rohleder, N. (2012). Acute stress responses in salivary alpha-amylase predict increase of plasma norepinephrine. *Biol. Psychol.* 91, 342–348.

Thornton, A., and McAuliffe, K. (2006). Teaching in wild meerkats. *Science* 313, 227–229.

Thorpe, W. I. (1956). *Learning and Instinct in Animals*. London: Methuen.

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie* 20, 410–433.

Trezza, V., Damsteegt, R., and Vanderschuren, L. J. (2009). Conditioned place preference induced by social play behavior: parametrics, extinction, reinstatement and disruption by methylphenidate. *Eur. Neuropsychopharmacol.* 19, 659–669.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.

Tsoory, M. M., Youdim, M. B., and Schuster, R. (2012). Social-cooperation differs from individual behavior in hypothalamic and striatal monoamine function: evidence from a laboratory rat model. *Behav. Brain Res.* 232, 252–263.

van den Bos, R., Harteveld, M., and Stoop, H. (2009). Stress and decision-making in humans, performance is related to cortisol-reactivity, albeit differently in men and women. *Psychoneuroendocrinology* 34, 1449–1458.

van den Bos, R., Homberg, J., and de Visser, L. (2013a). A critical review of sex differences in decision-making tasks, focus on the Iowa Gambling Task. *Behav. Brain Res.* 238, 95–108.

van den Bos, R., Taris, R., Scheppink, B., Haan, de L., and Verster, J. C. (2013b). Salivary cortisol and alpha-amylase levels during an assessment procedure correlate differently with risk-taking measures in male and female police recruits. Abstract ENP 2013, Lunteren, 29–31 May 2013, Lunteren, The Netherlands. Available online at: http://enpmeeting.org/2013/programme.php

van den Bos, R., Jolles, J. W., van der Knaap, L., Baars, A., and de Visser, L. (2012). Male and female Wistar rats differ in decision-making performance in a rodent version of the Iowa Gambling Task. *Behav. Brain Res.* 234, 375–379.

van Oers, K., Klunder, M., and Drent, P. J. (2005). Context dependence of personalities: risk-taking behavior in a social and a non-social situation. *Behav. Ecol.* 16, 716–723.

van Stegeren, A., Rohleder, N., Everaerd, W., and Wolf, O. T. (2006). Salivary alpha amylase as marker for adrenergic activity during stress: effect of betablockade. *Psychoneuroendocrinology* 31, 137–141.

van Stegeren, A. H., Wolf, O. T., and Kindt, M. (2008). Salivary alpha amylase and cortisol responses to different stress tasks: impact of sex. *Int. J. Psychophysiol.* 69, 33–40.

Viana, D. S., Gordo, I., Sucena, E., and Molta, M. A. (2010). Cognitive and motivational requirements for the emergence of cooperation in rat social game. *PLoS ONE* 5:e8483. doi: 10.1371/journal.pone.0008483

Vinkers, C. H., Zorn, J. V., Cornelisse, S., Koot, S., Houtepen, L. C., Olivier, B., et al. (2013). Time-dependent changes in altruistic punishment following stress. *Psychoneuroendocrinology* Avilable online at: http://dx.doi.org/10.1016/j.psyneuen.2012.12.012. doi: 10.1016/j.psyneuen.2012.12.012

Von Dawans, B., Fischbacher, U., Kirschbaum, C., Fehr, E., and Heinrichs, M. (2012). The social dimension of stress reactivity: acute stress increases prosocial behavior in humans. *Psychol. Sci.* 23, 651–660. doi: 10.1177/0956797611431576

Von Dawans, B., Kirschbaum, C., and Heinrichs, M. (2011). The Trier Social Stress Test for Groups (TSST-G): A new research tool for controlled simultaneous social stress exposure in a group format. *Psychoneuroendocrinology* 36, 514–522. doi: 10.1016/j.psyneuen.2010.08.004

Von Frijtag, J. C., Reijmers, L. G., Van der Harst, J. E., Leus, I. E., van den Bos, R., and Spruijt, B. M. (2000). Defeat followed by individual housing results in long-term impaired reward- and cognition-related behaviours in rats. *Behav. Brain Res.* 117, 137–146.

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., and Ochsner, K. N. (2008). Neural mechanisms of emotion regulation: evidence for two independent prefrontal-subcortical pathways. *Neuron* 59, 1037–1050. doi: 10.1016/j.neuron.2008.09.006

Wang, J., Korczykowski, M., Rao, H., Fan, Y., Pluta, J., Gur, R. C., et al. (2007). Gender difference in neural response to psychological stress. *Soc. Cogn. Affect. Neurosci.* 2, 227–239. doi: 10.1093/scan/nsm018

Ward, A. J. W., Hart, P. J. B., and Krause, J. (2004). The effects of habitat-and diet-based cues on association preferences in three-spined sticklebacks. *Behav. Ecol.* 15, 925–929. doi: 10.1093/beheco/arh097

Webster, M. M., and Laland, K. N. (2012). Social information, conformity and the opportunity costs paid by foraging fish. *Behavi. Ecol. Sociobiol.* 66, 797–809. doi: 10.1007/s00265-012-1328-1

Webster, M. M., and Ward, A. J. W. (2011). Personality and social context. *Biol. Rev.* 86, 759–773. doi: 10.1111/j.1469-185X.2010.00169.x

Webster, M. M., Ward, A. J. W., and Hart, P. J. B. (2007). Boldness is influenced by social context in threespine sticklebacks (*Gasterosteus aculeatus*). *Behaviour* 144, 351–371. doi: 10.1163/156853907780425721

West, S. A., Griffin, A. S., and Gardner, A. (2007). Evolutionary explanations for cooperation. *Curr. Biol.* 17, R661–R672. doi: 10.1016/j.cub.2007.06.004

Whiten, A., Horner, V., and De Waal, F. B. M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature* 437, 737–740. doi: 10.1038/nature04047

Witbracht, M. G., Laugero, K. D., Van Loan, M. D., Adams, S. H., and Keim, N. L. (2012). Performance on the Iowa Gambling Task is related to magnitude of weight loss and salivary cortisol in a diet-induced weight loss intervention in overweight women. *Physiol. Behav.* 106, 291–297.

Wöhr, M., and Schwarting, R. K. (2008). Maternal care, isolation-induced infant ultrasonic calling, and their relations to adult anxiety-related behavior in the rat. *Behav. Neurosci.* 122, 310–330.

Zajonc, R. B. (1965). Social facilitation. *Science* 149, 269–274. doi: 10.1126/science.149.3681.269

Zoratto, F., Laviola, G., and Adriani, W. (2013). Gambling proneness in rats during the transition from adolescence to young adulthood: a home-cage method. *Neuropharmacology* 67, 444–454. doi: 10.1016/j.neuropharm.2012.11.024

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 March 2013; accepted: 05 June 2013; published online: 26 June 2013.*

*Citation: van den Bos R, Jolles JW and Homberg JR (2013) Social modulation of decision-making: a cross-species review. Front. Hum. Neurosci. 7:301. doi: 10.3389/fnhum.2013.00301*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 04 July 2013
doi: 10.3389/fnhum.2013.00328

# Long-term changes in cognitive bias and coping response as a result of chronic unpredictable stress during adolescence

**Lauren E. Chaby**[1,2]*, **Sonia A. Cavigelli**[1,3], **Amanda White**[4], **Kayllie Wang**[4] and **Victoria A. Braithwaite**[1,2,4]

[1] Center for Brain, Behavior, and Cognition, Pennsylvania State University, University Park, PA, USA
[2] Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA, USA
[3] Department of Biobehavioral Health, Pennsylvania State University, University Park, PA, USA
[4] Department of Biology, Pennsylvania State University, University Park, PA, USA

**Edited by:**
Louise Barrett, University of Lethbridge, Canada

**Reviewed by:**
Carole Fureix, University of Guelph, Canada
Michael B. Hennessy, Wright State University, USA

***Correspondence:**
Lauren E. Chaby, Pennsylvania State University, 410 Forest Resources Building, University Park, PA 16802, USA
e-mail: chaby@psu.edu

Animals that experience adverse events in early life often have life-long changes to their physiology and behavior. Long-term effects of stress during early life have been studied extensively, but less attention has been given to the consequences of negative experiences solely during the adolescent phase. Adolescence is a particularly sensitive period of life when regulation of the glucocorticoid "stress" hormone response matures and specific regions in the brain undergo considerable change. Aversive experiences during this time might, therefore, be expected to generate long-term consequences for the adult phenotype. Here we investigated the long-term effects of exposure to chronic unpredictable stress during adolescence on adult decision-making, coping response, cognitive bias, and exploratory behavior in rats. Rats exposed to chronic unpredictable stress (e.g., isolation, crowding, cage tilt) were compared to control animals that were maintained in standard, predictable conditions throughout development. Unpredictable stress during adolescence resulted in a suite of long-term behavioral and cognitive changes including a negative cognitive bias [$F_{(1,\ 12)} = 5.000$, $P < 0.05$], altered coping response [$T_{(1,14)} = 2.216$, $P = 0.04$], and accelerated decision-making [$T_{(1,14)} = 3.245$, $P = 0.01$]. Exposure to chronic stress during adolescence also caused a short-term increase in boldness behaviors; in a novel object test 15 days after the last stressor, animals exposed to chronic unpredictable stress had decreased latencies to leave a familiar shelter and approach a novel object [$T_{(1,14)} = 2.240$, $P = 0.04$; $T_{(1,14)} = 2.419$, $P = 0.03$, respectively]. The results showed that stress during adolescence has long-term impacts on behavior and cognition that affect the interpretation of ambiguous stimuli, behavioral response to adverse events, and how animals make decisions.

**Keywords: adolescence, cognitive bias, coping, decision-making, chronic unpredictable stress, *Rattus norvegicus*, successive negative contrast**

## INTRODUCTION

Negative life experiences can have long-term effects on behavior and physiology (Sheriff et al., 2009; Archard et al., 2012). Stressful events (e.g., stressors) come in a variety of forms, but in vertebrates they are often considered to be unpredictable aversive stimuli that provoke a glucocorticoid hormone response mediated by the hypothalamic-pituitary-adrenal (HPA) axis (Spear, 2000; Koolhaas et al., 2011). Stages of development differ in sensitivity to stress, certain life stages have specific vulnerabilities that can lead to different, permanent changes in future responses to adverse events (McCormick and Mathews, 2008; Vidal et al., 2011). For example, in zebra finches (*Taeniopygia guttata)* exposure to excess heat during early-life enables the birds to modify their response to subsequent heat exposures in adulthood to minimize oxidative damage (Costantini et al., 2012). Similarly, rodent pups that experience isolation at different stages of development exhibit contrasting hormonal responses to stress in adulthood; rat pups separated from their mothers for 2 h a day at postnatal days 2–14 develop a hyper-responsive HPA axis, whereas pups isolated at postnatal days 15–16 develop a hypo-functioning HPA axis (Plotsky and Meaney, 1993; Sánchez et al., 1998; reviewed in Sánchez et al., 2001).

During the adolescent stage, glucocorticoid production in response to a stressor exceeds the adult hormone response in duration and intensity (McCormick et al., 2010). In comparison to adult rats, adolescent rats exposed to an acute stressor show a higher increase in both adrencorticotrophin hormone (ACTH) and glucocorticoids (reviewed in Romeo and McEwen, 2007; Foilb et al., 2011). Additionally, during adolescence various neural structures involved in stress and reward processing are still immature (Spear, 2000; McCormick and

Mathews, 2008). These characteristics of the adolescent life stage suggests that this period may be particularly vulnerable to effects from chronic glucocorticoid exposure (Romeo and McEwen, 2007; McCormick et al., 2010). Chronic exposure to elevated levels of glucocorticoid hormones has numerous effects on the brain including suppressed neurogenesis and enhanced dendritic pruning in the hippocampus, dendritic shortening in the medial prefrontal cortex, and enhanced dendritic growth in the amygdala, the fear center of the brain (reviewed in McEwen, 2005).

Adverse experiences during adolescence can impact the maturation of the central nervous system, shape future reward responses, and influence endocrine and behavioral function in adulthood (Romeo, 2003; McCormick et al., 2004; Andersen and Teicher, 2008; McCormick and Green, 2012). The changes that occur following stress exposure during adolescence are dynamic; some are immediate, some are short in duration, and some are long-term but only become apparent after an acute stressor is applied (McCormick and Green, 2012; McCormick et al., 2012; Saul et al., 2012). McCormick et al. (2012) found that exposure to unpredictable social instability and isolation in adolescent rats resulted in learning deficits in adulthood, but these effects were only apparent after an acute stressor was applied (McCormick et al., 2012). Others have reported that behavioral effects of stress during adolescence can be transient and fade over the lifetime of an animal (e.g., unpredictable isolation and novel social partner stressors during adolescence induce temporary changes in boldness, Mathews et al., 2008). Despite these important early studies, the long-term effects of stress during adolescence on emotion and cognition are not well-characterized. Yet, if we are to understand how animals cope with stress during development, and how early adverse experiences can prepare an animal to deal with subsequent stressors, we need to determine the long-term impacts of stress during the adolescent stage (Romeo, 2010).

A number of studies have demonstrated that stress during adolescence, including unpredictable chronic social and physical stress, can impact HPA axis function and glucocorticoid production in adulthood (McCormick and Mathews, 2008; Buwalda et al., 2011). The long-term consequences of stress during adolescence on cognition and behavioral coping response, however, remains unknown. A method to assess behavioral coping response and reward loss sensitivity, as mediated by glucocorticoid production, is the successive negative contrast (SNC) test (Mitchell and Flaherty, 1998; Gomez et al., 2009). SNC has been used for over 3 decades to evaluate an animal's response to the unexpected downshift of a familiar high-value reward to a novel low-value reward (Lombardi and Flaherty, 1978; Flaherty and Rowan, 1989). Recently SNC has been used as a measure of coping response to infer background emotional state in non-human animals (Burman et al., 2008; Gomez et al., 2009).

In humans, background emotional state can affect decision-making through a cognitive bias in stimulus interpretation that impacts stimulus perception, attention, and processing (Winkielman et al., 2007). Increasingly, measures of cognitive bias are being used as indicators of background emotional state in non-human animals (Burman et al., 2009; Brydges et al., 2011). Unlike most behavioral and physiological measures, cognitive bias tests can measure the valence of affect (positivity vs. negativity) rather than just arousal (Mendl and Paul, 2004). Prior studies have shown that adult rats can exhibit a negative cognitive bias, marked by an increased propensity to interpret ambiguous stimuli as threatening or aversive that can start during stress exposure and last up to several days after an aversive events (Harding et al., 2004; Burman et al., 2009). The potential longevity of a negative cognitive bias following exposure to stress, however, remains unclear (Mendl et al., 2009; but see Brydges et al., 2012). A previous focus on short-term changes in cognitive bias has meant that long-term changes have so far been underexplored (Brilot et al., 2010). In the current study, we addressed the long-term effects of chronic unpredictable stress during adolescence on behavior and cognition by evaluating changes in cognitive bias, decision-making, associative learning rate, coping response, and motivation to consume a reward in adulthood.

A range of behavioral tests were used to examine the consequences of stress during adolescence: (1) sucrose preference (Strekalova et al., 2004), (2) exploration of a novel object (Van Dijken et al., 1992; Cavigelli et al., 2009), (3) successive negative contrast (SNC), and (4) ambiguous judgment cognitive bias (Harding et al., 2004; Doyle et al., 2011). We measured exploratory behavior and motivation to consume a reward because alterations in these fundamental traits could potentially affect the interpretation of more complex reward or activity based tests including the cognitive bias and SNC tests. Stress can alter both the motivation to consume a reward and exploratory behavior; the magnitude of effects from stress are dependent upon the type and duration of the stressors and traits intrinsic to the animal (Zurita et al., 2000; Strekalova et al., 2004; Brilot et al., 2010). We hypothesized that stress during adolescence would induce a negative cognitive bias and stronger sensitivity to reward loss, both suggestive of a long-term negative background emotional state. Additionally, we predicted that stress during adolescence would result in altered decision-making, impaired associative learning, and decreased exploratory behavior in adulthood.

## METHODS

### SUBJECTS AND HOUSING

Sixteen male Long-Evans rats (Harlan Laboratory in Fredrick, Maryland, USA) were obtained at 21 days of age. Following transport, rats were given 7 days to settle before handling and behavioral testing commenced. A full timeline of all manipulations and behavioral tests is provided in **Figure 1**. Animals were pair-housed in plastic cages, 20 × 26 × 46 cm, with corn cob bedding and basic enrichment items: two 7.6 cm diameter PVC tubes hanging from the wire cage lid and two 2.5 × 2.5 × 8 cm pine wood blocks. Rats were kept on a 12:12 reversed light/dark cycle at 20–21°C and 41–42% relative humidity. Standard rat chow (LabDiet®) and tap water were available *ad-libitum* unless otherwise noted. To minimize disturbance, the experimenter was not in the room during data collection. Work was approved by the Pennsylvania State University IACUC committee, protocol #35761.

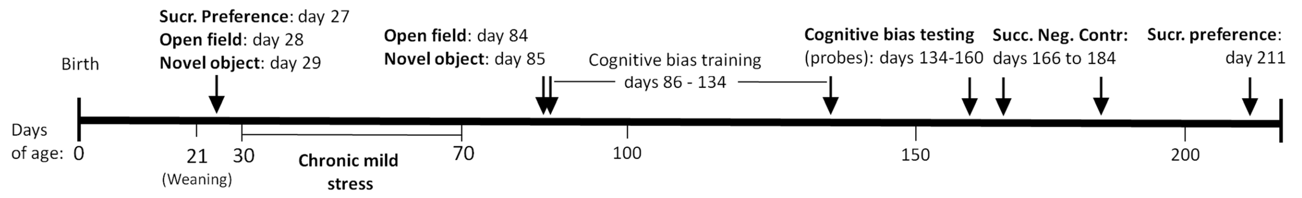


**FIGURE 1 | Timeline of procedures.** We evaluated the long-term effects of chronic unpredictable stress during the adolescent life stage on adult cognition and behavior using the following tests: cognitive bias, successive negative contrast, sucrose preference, novel object, and open field.

## ADOLESCENT CHRONIC UNPREDICTABLE STRESS

Four cages of pair housed rats ($n = 8$) were randomly assigned to the control condition and four cages ($n = 8$) to the stress treatment. For the latter group, stressors were presented daily from 30 to 70 days of age, with 8 days of rest occurring intermittently. Prior studies of adolescent-stress have varied in the duration of stress exposure, due in part to the large window of time during which adolescent ontogenetic changes occur. These changes are thought to conclude at approximately 55–60 days of age in male rodents (Spear, 2000). To cover the entirety of the ontogenetic window of adolescence, studies have included a postpubertal "sub-adult" period (Schmidt et al., 2007). Studies of adolescent-stress have used stress exposure periods spanning from 28 to 80 days of age (Spear, 2000; Sterlemann et al., 2010). As the current study evaluated behaviors mediated by the prefrontal cortex (i.e., decision-making, coping), and this region is still developing in early adulthood, the duration of stress exposure (30–70 days of age) included a postpubertal period in early adulthood (van Eden et al., 1990; Spear, 2000).

For the chronic unpredictable stress procedure both physical and social stressors were presented randomly across the light/dark cycle to maximize unpredictability. An average of three physical and three social stressors were presented between each rest day. Stressors noted to induce short-term changes in cognitive bias were used (e.g., cage tilt, damp bedding: Harding et al., 2004; e.g., crowding, confinement: Doyle et al., 2011; see **Table 1**). An additional stressor, isolation, was chosen because it has been associated with long-term changes in behavior following exposure during adolescence (McCormick et al., 2012).

To control for the influence of circulating corticosterone on tests mediated by glucocorticoid levels, such as the SNC, we controlled for daily rhythms in glucocorticoid production by avoiding testing during peak corticosterone production; all tests started a minimum of 2 h after the beginning of the dark cycle and were completed within 6 h of the start of the test (Mitchell and Flaherty, 1998). Weight and physical appearance were monitored; no changes in aggression or health related to either the unpredictable stress regimen or any of the behavioral tests were observed.

**Table 1 | Chronic unpredictable stressor descriptions.**

| | |
|---|---|
| **PHYSICAL** | |
| Smaller cage | Rat pairs were housed for 4 h in a cage with a 25% reduction in volume from the 20 × 26 × 46 cm standard home cage (Doyle et al., 2011). |
| Damp bedding | While rats were temporarily in an empty transfer cage, 200 ml of water was mixed into 2/3 of the bedding of the home cage. After 6 h in the damp bedding, pairs were transferred to a clean home cage (Zurita et al., 2000; Harding et al., 2004). |
| Cage tilt | Home cages were tilted at a 30° angle for 6 h (Zurita et al., 2000; Harding et al., 2004). |
| **SOCIAL** | |
| Isolation | Rats were housed individually for 1.5 h in a clean cage (20 × 26 × 46 cm) with a 7.6 cm diameter PVC tube and a 2.5 × 2.5 × 8 cm pine wood block (Zurita et al., 2000; McCormick et al., 2012). |
| Crowding | Sets of 2 rat pairs were combined into one clean cage (20 × 45 cm) for 4 h; iterations of pair combinations were balanced (Zurita et al., 2000; Harding et al., 2004; Doyle et al., 2011). |
| Foreign bedding | Experimental pairs were housed in the empty home cage of a pair of older conspecifics for 12 h. (Harding et al., 2004). |

## SUCCESSIVE NEGATIVE CONTRAST (SNC)

Coping response was evaluated from 166 to 184 days of age with an SNC test measuring response to an unexpected downshift in reward value (Burman et al., 2008; Gomez et al., 2009: see **Figure 2**). During the SNC test, individual animals were tested daily in an opaque, plastic container, 30.5 × 30.5 × 30.5 cm, for 5 min. A plastic bottle of sucrose solution was attached to the center of one wall. Motivation to consume sucrose solution was measured with a basic electronic device attached to a computer that registered each lick through the closing of a circuit, the computer then provided a record of licking rates. After an initial 12 days of trials with a 32% sucrose (w/v) reward, the solution concentration was decreased without warning to 4% (w/v). The lower concentration was administered for 7 days to monitor the recovery of lick rates. To ensure reward salience, 2 h of food deprivation preceded each trial. We defined animals as having learned the SNC task upon registering 10 licks in one session; pre-shift data were evaluated from the first day that more than 60% of the animals had learned the task (day 4) to the last day of 32% sucrose solution presentation (day 12).

## COGNITIVE BIAS, DECISION-MAKING, AND ASSOCIATIVE LEARNING

The ambiguous judgment task was used to assess the long-term impacts of stress during adolescence on cognitive bias, decision-making, and associative learning. Using a paradigm similar to

Brydges et al. (2011), animals were trained to associate a conditioned stimulus, a type of sandpaper (rough or smooth), with the location and color of a bowl containing an available food reward. To do this, individuals were placed in a 30 × 40 × 45 cm opaque plastic start box that was connected to a goal box by an 80 cm PVC pipe (see **Figure 3**). The goal box contained a white bowl and a black bowl separated by an opaque partition to ensure that the rats made a choice between the two bowls upon exiting the PVC tunnel. Of the two available bowls, one was associated with a high-value reward (3 Cheerios), the other with a low-value reward (1 Cheerio). To balance the scent cues, each bowl always contained three Cheerios, but the accessibility of the rewards varied depending on the trial condition. For a high-reward trial, 3 Cheerios were available in the high-reward bowl while the low-reward bowl contained 3 inaccessible Cheerios. For a low-reward trial, there was 1 accessible Cheerio (with two mesh covered inaccessible Cheerios) in the low-reward bowl, and all 3 Cheerios were inaccessible in the high-rewarded bowl.

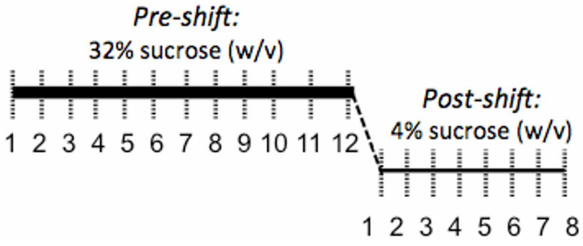


**FIGURE 2 | Illustration of the SNC procedure to assess coping response.** A high-value 32% sucrose reward was presented for 5 min per day for 12 days while licks were counted to evaluate motivation to consume a reward. Following this 12-day acclimation period, the reward was unexpectedly decreased to 4% sucrose. The resulting decrease in consumption of the 4% solution is interpreted as a negative reaction to the violation of positive expectations and a measure of coping response.

A tactile cue lining the PVC tunnel and goal box indicated which bowl had an available reward; one of two grades of silicon carbide waterproof sandpaper, coarse (P60) or fine (P1200), was paired with a specific reward type (e.g., coarse sandpaper signaled a low-reward in the black bowl on the left; Brydges et al., 2011). All sandpaper were of the same brand and were black in color. Pairings of sandpaper grade, bowl color, bowl side, and reward

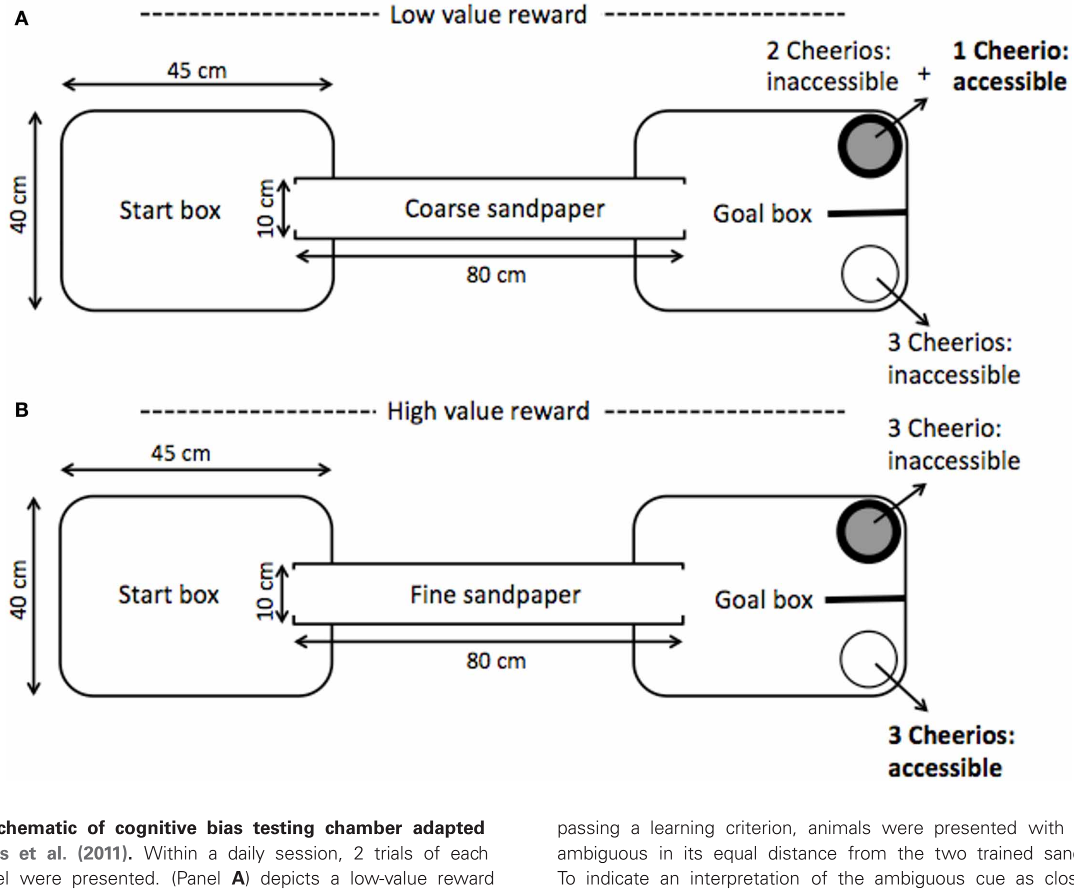


**FIGURE 3 | Schematic of cognitive bias testing chamber adapted from Brydges et al. (2011).** Within a daily session, 2 trials of each example panel were presented. (Panel **A**) depicts a low-value reward trial; the coarse sandpaper provided a cue that the low-value (1 Cheerio) reward was accessible whereas the fine sandpaper indicated a high-value reward (3 Cheerios) was present (Panel **B**). Sandpaper grade, bowl color, bowl side, and reward value pairings were counterbalanced. To balance scent cues from the rewarded and unrewarded bowls inaccessible Cheerios were present beneath a mesh barrier. After passing a learning criterion, animals were presented with a novel cue ambiguous in its equal distance from the two trained sandpaper cues. To indicate an interpretation of the ambiguous cue as closer to either a high or low-value reward the animal moved to either the high or low-reward location in the testing chamber. Interpretation of the ambiguous paper as closer to either the cue for the high or low-value reward conveyed a positive or negative value assignment from properties intrinsic to the animal, i.e., a positive or negative cognitive bias in the interpretation of ambiguity.

value were counterbalanced. All elements of the testing chamber were cleaned with 70% ethanol between each trial.

To learn the stimulus-reward associations rats were exposed to daily training sessions that consisted of 2 high-reward trials and 2 low-reward trials; the order of the 4 trials was randomized. Animals moved from the start box through the PVC tunnel into the goal box, and chose either the "correct" rewarded or "incorrect" unrewarded bowl. A choice was defined as a rat moving its nose or paw inside the bowl or touching the outside of the bowl with its nose or paw. If an animal chose the rewarded bowl first, the trial was counted as correct and the rat was allowed to consume the reward. If an animal chose incorrectly it was allowed to move to the correct bowl and consume the reward during the first 5 days of training. Decision-making was measured during the first 8 trials, after the rats had consumed at least one reward in the test chamber, by timing the latency between the incorrect selection of an inaccessible reward and the switch to choose the rewarded bowl. Starting the 6th day of training, the rat was removed immediately if it chose the incorrect side.

A learning criterion was set at 3 out of 4 trials with a correct first bowl choice for 4 out of 5 days. The number of days to reach the learning criterion was evaluated to determine if stress during adolescence impacts adult associative learning (Hammond et al., 2009). In both the stressed and control group 2 rats did not pass the learning criterion. After passing the criterion, probe trials were conducted where ambiguous/intermediate grade silicon carbide waterproof sandpaper (P220) was placed in the PVC tube connecting the start and goal boxes. On each day of probe testing a total of 5 trials were run; in addition to the 4 standard trials, one probe trial using ambiguous sandpaper was randomly inserted into the normal sequence, but the last trial was never a probe trial. A total of 5 probe trials were run over 5 days following the same design as Brydges et al. (2011). During probe testing all animals maintained the learning criterion.

After choosing a bowl during the ambiguous probe trial, the bowl choice was noted as either a high or a low-reward categorization of the ambiguous sandpaper cue and the animal was allowed to consume the reward. A number of studies using unrewarded probe trials found that animal stop responding during repeated probe trials, interpreted as a consequence of the animals learning that probes are not reinforced (Bateson and Matheson, 2007; Brilot et al., 2010; Doyle et al., 2010). To circumvent this, in the current study both high and low-rewards were present during probe trials to avoid cessation of response. A potential limitation of this design is that an initial probe interpretation may be reinforced if an animal encounters an expected reward, however, our data do not support this as all but 2 rats sampled both bowls during the 5 probe trials.

### SUCROSE PREFERENCE

To assess motivation to consume sucrose solution, a 24-h test for preference of 2% sucrose solution (weight/volume) relative to water was administered both prior to and after stress exposure (at 27 and 211 days of age) to determine whether preference changed over time (Strekalova et al., 2004). Animals had simultaneous access to 2 bottles placed side-by-side on the lid of their home cage: one with water and another with a 2% sucrose solution. To avoid effects from side preference, the location of the sucrose and water bottles was counterbalanced and halfway through the test, the sides of the two solutions were switched. Bottle weights were obtained before and after the 24-h period of unlimited access and relative preference was calculated: Sucrose preference $= 100 \times$ sucrose solution intake(g)/[sucrose solution intake(g) + water intake(g)]. To assess individual preferences, animals were singly-housed during the assay. To minimize social stress during testing, the individual cages of separated pairs were positioned next to each other and all rats had access to a 7.6 cm diameter PVC tube and a $2.5 \times 2.5 \times 8$ cm pine block for enrichment (Odberg, 1987; Gross et al., 2012). Following the sucrose preference test, pairs were recombined in a clean home cage.

### EXPLORATORY BEHAVIOR

Rats were given two tests to assess exploratory behavior, an open field and a novel object test. Both tests were administered at two time points, one before and one after stress exposure (Van Dijken et al., 1992; Cavigelli et al., 2009). To minimize the potential for the second set of behavioral tasks to be influenced by the first, the two test iterations were separated by 55 days and new stimulus objects were used during each novel object test. All tests were run in a $122 \times 122 \times 46$ cm opaque Plexiglas arena. Each task involved 5 min of free exploration during which latency to leave a 7.6 cm diameter PVC tube shelter was measured. All animals started both exploratory tests inside the PVC tube shelter; the tube was placed along the base of one arena wall in the same position and orientation for all tests.

#### *Exploratory task 1: open field*

To compare activity levels between adolescent-stress and control animals, activity in the arena was quantified with a video-recorded open field assay at two time points (28 and 84 days of age, pre and post chronic unpredictable stress). During video analysis an $8 \times 8$ grid was used to quantify activity by counting the number of squares crossed on the grid. Crossing of grid squares along the walls of the arena and the proportion of time spent in squares along the arena walls were quantified as indicators of thigmotaxis, a positive correlate of anxiety (Simon et al., 1994).

#### *Exploratory task 2: novel object*

Response to novelty was evaluated before and after chronic unpredicatable stress (at 29 and 85 days of age) with two behavioral measures: time to leave the PVC shelter (i.e., when all 4 feet were touching the arena floor) and latencies to physically contact the two novel objects in the arena with either a paw or nose. The novel objects varied in texture, color, and size. Several plastic objects were used including a translucent red triangle, an opaque matt yellow bowl, a shiny yellow cylinder, and a translucent shelter.

### DATA ANALYSIS

Sucrose preference and cognitive bias data conformed to the assumptions for parametric analyses. SNC data were square root transformed to achieve normality. Exploratory behavioral data from the novel object and open field assays were natural log transformed to achieve normality. To test whether chronic unpredictable stress during adolescence affected sucrose preference or exploratory behavior over time (pre and post chronic unpredictable stress), we used 2 factor (time and stress condition) repeated measures ANOVAs to compare across the 2 tests. For *post-hoc* analysis, independent samples two-tailed *t*-tests were used to compare the stress and control groups within the 2 test iterations. Only significant *post-hoc* findings are reported.

To evaluate ambiguity interpretations in the cognitive bias assay, we tested the first two ambiguous probe exposures separately using univariate general linear models because response to the ambiguous probe changes with repeated exposures; initial exposures are more reliable measures of affect (Bateson and Matheson, 2007; Brilot et al., 2010; Doyle et al., 2010). Following individual analysis of the first and second probes, all 5 probe exposures were evaluated with a repeated measures general linear model as in Brydges et al. (2011) and Burman et al. (2009). To determine whether activity or motivation to consume a reward impacted performance in the cognitive bias test, we included activity in the open field and sucrose preference as covariates in the ambiguous probe general linear models. Neither activity nor sucrose preference were significant factors in explaining variation in the data, so they were removed from the model [activity: $F_{(1,12)} = 0.929$, $P = 0.36$; sucrose preference: $F_{(1,12)} = 0.590$, $P = 0.46$]. The associative learning and decision-making data were analyzed with independent samples two-tailed *t*-tests. To determine whether animals that experienced adolescent stress had a stronger response to reward devaluation than control animals, repeated measures ANOVAs were used to assess behavior for pre-shift days 4–12 and post-shift days 13–18. To assess the impacts of activity and motivation to consume a reward on SNC scores locomotion in the open field and sucrose preference were included as covariates in the repeated measures ANOVAs; neither factor significantly explained variation in the data and were subsequently removed from the model [Pre-shift: activity: $F_{(1,12)} = 0.065$, $P = 0.81$; sucrose pref: $F_{(1,12)} = 0.092$, $P = 0.77$; Post-shift: activity: $F_{(1,12)} = 1.959$, $P = 0.20$; sucrose pref: $F_{(1,12)} = 1.984$, $P = 0.20$]. To evaluate response to the reward devaluation, lick numbers on the first day of post-shift were subtracted from the average of the last 3 days of pre-shift. These difference scores were tested with a two-tailed *t*-test. Analyses were run in SPSS; values are reported as means ± standard deviation.

## RESULTS

### SUCCESSIVE NEGATIVE CONTRAST

Response to the reward downshift was greater in the adolescent-stress group than in the control animals [$T_{(1,14)} = 2.216$, $P = 0.04$, $d = 1.02$, see **Figure 4**]. No differences were found in the pre-shift lick rates of the adolescent-stress and control animals [RM ANOVA, $F_{(1,\ 6)} = 0.092$, $P = 0.77$] or over time [$F_{(1,\ 6)} = 4.022$, $P = 0.37$], nor was there an interaction [$F_{(1,\ 6)} = 1.173$, $P = 0.61$]. Post-shift lick rates changed over time as rats returned to pre-shift licking rates [RM ANOVA, $F_{(1,\ 6)} = 9.911$, $P = 0.01$], but stress and control animals did not differ in their post-shift lick rates [RM ANOVA, $F_{(1,\ 6)} = 0.003$, $P = 0.95$], nor was there an interaction [$F_{(1,\ 6)} = 0.644$, $P = 0.70$].

### COGNITIVE BIAS ASSAY

All stressed animals interpreted the ambiguous cue as negative on the first day of probe testing, demonstrating a negative cognitive bias that differed from the control animals whose interpretations were half positive and half negative [$F_{(1,\ 12)} = 5.000$, $P < 0.05$, $R^2 = 0.33$, see **Figure 5**]. This difference in the interpretation of the ambiguous probe was not significant in subsequent trials [Day 2: $F_{(1,12)} = 0.000$, $P = 1.00$; GLM 5 days: $F_{(1,12)} = 0.471$, $P = 0.508$]. The total number of positive and negative probe interpretations from each group are depicted in **Figure 6**. Within the first two days of training, adolescent-stressed animals were faster to correct wrong decisions by abandoning the wrong bowl, reorienting, and choosing the correct bowl [*t*-test $T_{(1,14)} = 3.245$, $P = 0.01$, $d = 1.62$, see **Figure 7**]. However, animals stressed during

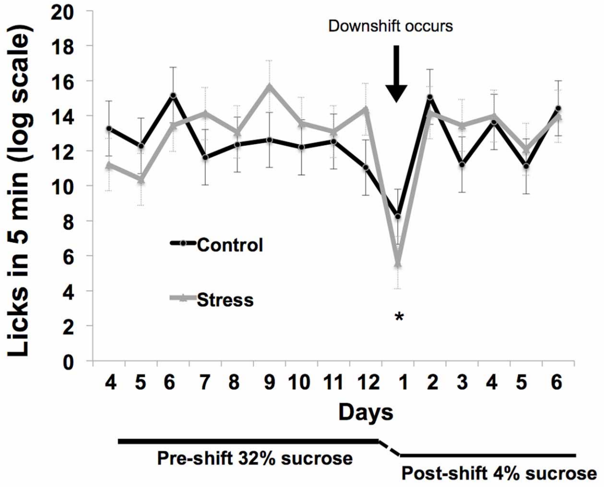


**FIGURE 4 | Coping response in a successive negative contrast test following stress during adolescence.** Animals exposed to chronic unpredictable stress during adolescence exhibited a long-term increase in reward loss sensitivity as measured by successive negative contrast (SNC). The asterisk indicates a significant difference in the downshift scores between the adolescent-stressed and control animals.

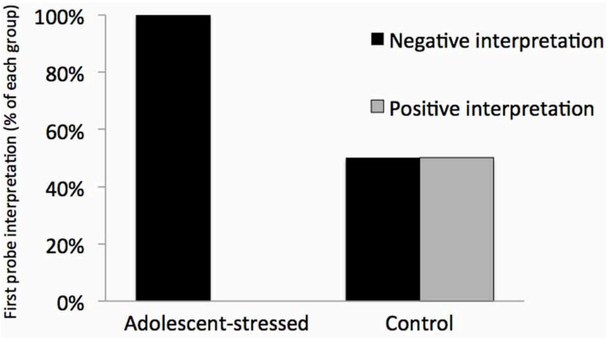


**FIGURE 5 | Interpretation of a novel ambiguous stimulus on the first day of an ambiguous judgment test for cognitive bias.** All animals that experienced stress during adolescence interpreted the ambiguous cue as negative, indicating a long-term negative cognitive bias, while control animals were equally likely to interpret the cue as positive or negative.

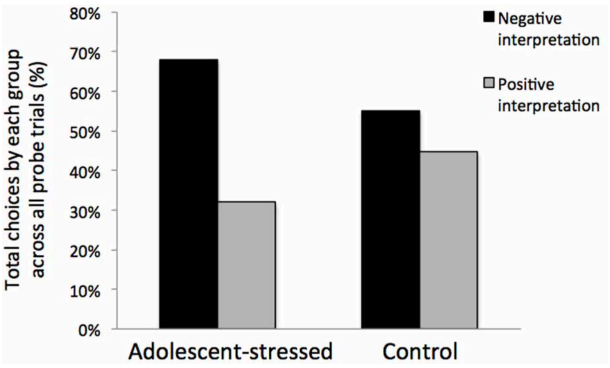


**FIGURE 6 | Interpretation of an ambiguous stimulus during all cognitive bias testing days.** Rats that experienced stress during adolescence trended toward a negative cognitive bias across the 5 day testing period.

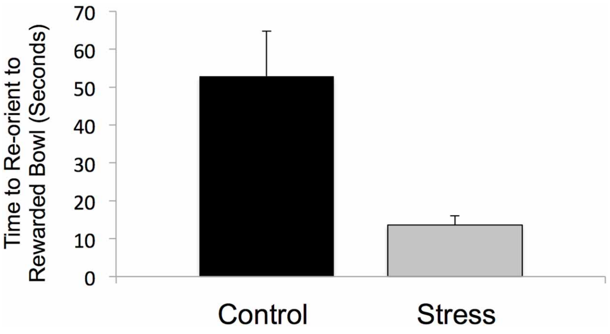


**FIGURE 7 | Latency to move to the correct, rewarded bowl after the first incorrect choice.** At the first encounter with the unrewarded bowl during cognitive bias training, prior to associative learning, animals that experienced stress during adolescence had a shorter latency to correct their choice and reorient to the rewarded bowl.

adolescence showed no difference in the number of days to learn the associative task compared with controls [*t*-test, stress: 26 ± 3 vs. control: 25 ± 5; $T(1, 10) = 0.419$, $P = 0.68$].

## SUCROSE PREFERENCE

Sucrose preference decreased over time [RM ANOVA effect of time: $F_{(1,\ 7)} = 5.680$, $P = 0.04$], but there was no effect of stress [RM ANOVA effect of stress: $F_{(1,\ 7)} = 0.417$, $P = 0.53$] and no interaction between stress condition and time [RM ANOVA stress × time interaction: $F_{(1,\ 7)} = 0.631$, $P = 0.42$; time 1: stress 83 ± 7% vs. control 82 ± 7%; time 2: stress 78 ± 7% vs. control 74 ± 8%].

## OPEN FIELD ACTIVITY SCORES

Activity increased over time, which is consistent with previous studies indicating that age influences exploratory behavior in an open field [RM ANOVA effect of time: $F_{(1,\ 7)} = 8.454$, $P = 0.01$; Bronstein, 1972]. There was no effect of stress [RM ANOVA effect of stress: $F_{(1,\ 7)} = 0.093$, $P = 0.77$], nor was there an interaction between stress condition and time [RM ANOVA stress × time interaction: $F_{(1,\ 7)} = 1.423$, $P = 0.25$; time 1, stress 244 ± 30 squares crossed vs. control 233 ± 86 squares crossed; time 2, stress 282 ± 29 squares crossed vs. control 303 ± 40 squares crossed].

Thigmotaxis decreased over time in all animals [RM ANOVA effect of time: $F_{(1,\ 7)} = 97.685$, $P < 0.00$]. There was no effect of stress condition [RM ANOVA effect of stress: $F_{(1,\ 7)} = 0.110$, $P = 0.75$], and no interaction between stress and time [RM ANOVA stress × time interaction: $F_{(1,\ 7)} = 0.359$, $P = 0.56$; time 1, stress 283 ± 12(s) vs. control 282 ± 10(s); time 2, stress 233 ± 15(s) vs. control 238 ± 21(s)].

## NOVEL OBJECT

The latency to approach a novel object decreased in animals exposed to stress during adolescence [RM ANOVA effect of stress: $F_{(1,\ 7)} = 4.682$, $P < 0.05$]. In the second test iteration rats exposed to adolescent-stress were faster to approach a novel object than control animals [latency to approach novel object at time 2, stress: 4.6 ± 2(s) vs. control: 16 ± 13(s); $T_{(1,14)} = 2.419$, $P = 0.03$, $d = 1.23$] with no baseline difference in the approach latency prior to stress exposure [latency to approach novel object at time 1, stress: 15.4 ± 19(s) vs. control: 16.7 ± 13(s); $T_{(1,14)} = 0.156$, $P = 0.88$]. While this difference appears to be a real biological effect, the variance between the groups was high which may explain the lack of interaction between time and treatment [RM ANOVA stress × time interaction: $F_{(1,\ 7)} = 1.544$, $P = 0.23$].

Latency to leave the PVC shelter decreased over time in all animals, which is congruous with previous findings that behavior in a novel object test changes as animals reach adulthood [RM ANOVA effect of time: $F_{(1,\ 7)} = 11.179$, $P = 0.01$; Saul et al., 2012]. During the novel object test, 15 days after the completion of the chronic unpredictable stress treatment, rats exposed to stress during adolescence left the PVC shelter faster than control animals [exit latency at time 2, stress: 2.3 ± 0.6(s) vs. control: 7.3 ± 1.8(s); $T_{(1,14)} = 2.240$, $P = 0.04$, $d = 3.73$]. There was no baseline difference in the latency to leave the PVC shelter prior to stress exposure [exit latency at time 1, stress: 2 ± 0.5(s) vs. control: 1.7 ± 1(s); $T_{(1,14)} = 0.344$, $P = 0.74$].

# DISCUSSION

Our results show that chronic unpredictable stress during adolescence has long-term effects on coping response, cognitive bias, and decision-making. Associative learning and sucrose preference, however, were not affected by stress exposure during adolescence. The novel object test showed increased boldness behaviors 15 days after completion of the chronic unpredictable stress paradigm. Activity and thigmotaxis in the open field were not affected by prior adverse experience. Stress-exposed rats were faster to leave a familiar shelter in an environment containing novelty and approached novel objects more quickly than control animals. The successive negative contrast test demonstrated that stress during adolescence induces a stronger response to the devaluation of an expected reward in adulthood. The sucrose preference test demonstrated that stress during adolescence does not alter motivation to consume a reward, confirming that the altered response to reward devaluation exhibited by animals exposed to stress during adolescence was not due to a difference in reward salience, but was a reaction to the downshift in reward value.

Exposure to stress during adolescence also decreased the latency to correct a choice and locate a food reward after an incorrect decision. In an early phase of training for the cognitive bias assay, adolescent-stressed animals were faster at abandoning, reorienting, and switching their choice of food bowl after encountering a bowl with an inaccessible reward than the control animals. The results from the sucrose preference test exclude the possibility that the shorter latency to find the reward after an incorrect decision is due to a difference in motivation to obtain the reward, as the preference test demonstrates that motivation to consume a reward is unchanged by stress during adolescence. Thus, the expediency of decision-making in stressed animals could be the result of decreased behavioral inhibition or increased impulsivity when compared to the control animals. Animals exposed to exogenous corticosterone during adolescence show a form of impulsivity marked by an increased preference for an immediate, small reward rather than a larger reward delivered after a variable delay (Torregrossa et al., 2012). It is possible that the decreased latency to abandon a first choice and transition to a second choice demonstrated by adolescent-stress animals also reflects increased impulsivity. Long-term changes in impulsivity behaviors may be underpinned by stress-induced changes in the brain. Stress may impair maturation processes that typically occur during adolescence, such as myelination in the prefrontal cortex, thereby prolonging an immature-like state of top-down connectivity into adulthood (McCormick, 2007). An immature-like state in prefrontal cortex could maintain increased impulsivity behaviors characteristic of the adolescent stage, and alter behavioral inhibition and decision-making in adulthood.

Our results showed that stress during adolescence induces a long-term negative cognitive bias. This finding, along with the SNC results demonstrating increased sensitivity to reward loss, indicate that stress during adolescence generates a long-term negative background emotional state (Burman et al., 2008; Mendl et al., 2009). A negative background emotional state can bias decision-making and expectations for the future; humans with a negative emotional state exhibit biases in attention (e.g., greater attention to threatening stimuli), memory (e.g., enhanced negative memory retrieval), and judgment (e.g., risk and ambiguity aversion, Paul et al., 2005). It is important, however, to keep in mind the potential ecological context of a negative cognitive bias induced by stress. For example, in sites of high predation, traits like threat bias and risk aversion may serve an adaptive function. If threat is prevalent in an environment, it may be advantageous to more readily treat ambiguity as negative or a potential threat (Mendl et al., 2009). Stress induced programming during adolescence for a long-term threat bias may serve to prepare an individual to cope with future exposure to a dangerous environment. Human studies suggest that the consequences of a negative cognitive bias are far reaching, but the full impacts of a negative cognitive bias in non-human animals are not yet clear (Winkielman et al., 2007).

The ambiguous judgment cognitive bias task used here captured differences in the interpretation of ambiguity as a result of stress during adolescence. The results of the 5 probe trials evaluated together, however, highlight a limitation of the ambiguous judgment test. During the first exposure to the ambiguous probe, animals interpret the novel ambiguous stimulus based only on their own biases and life history, whereas subsequent exposures to the ambiguous probe are influenced by previous interpretations of the probe. Thus, repeated probe tests can be subject to effects from learning (Doyle et al., 2010). In the current study the initial ambiguous probe trials were analyzed separately from subsequent probe trials similar to Brilot et al. (2010). Future studies that use the ambiguous judgment task should analyze initial probe exposures separately, as the use of repeated probe tests allows for learning and can yield misleading results (Brilot et al., 2010; Doyle et al., 2010).

The current study found that sensitivity to reward loss in adulthood is intensified by exposure to stress during adolescence, suggesting that animals exposed to adverse events during this period can undergo a long-term change in coping with challenge. This result could help explain an interesting phenomenon documented in previous studies: adolescent-stressed animals can appear to have unaltered behavior, temperament, and learning in adulthood, until they encounter a challenge, at which point behavioral differences become apparent (Watt et al., 2009; Vidal et al., 2011; McCormick et al., 2012). Our results suggest that the altered response to challenge demonstrated by adult animals exposed to stress during adolescence could arise from a long-term change in coping response that has behavioral and cognitive consequences that only become apparent upon subsequent exposure to stress.

Immediately following exposure to isolation and unpredictable housing during adolescence, exploratory behavior in an elevated plus maze is increased, however, a month following stress exposure exploratory behavior is decreased relative to controls (McCormick et al., 2008). Our results expand upon this finding to demonstrate that 2 weeks following physical and social stress during adolescence, male rats are faster to approach novelty, suggesting increased exploratory behavior. The contrast in effects of closely related stress paradigms emphasizes the need for longitudinal studies that evaluate the consequences of specific stress paradigms and span multiple life stages in order to more completely understand how resilience and vulnerability to stress change over the lifetime of an organism.

## ACKNOWLEDGMENTS

We thank the Huck Institute of the Life Sciences and the Eberly College of Science for support. This project was also funded, in part, under a grant with the Pennsylvania Department of Health using Tobacco CURE Funds. The Department specifically disclaims responsibility for any analyses, interpretations or conclusions. We would also like to thank the members of the Braithwaite group, with a special thanks to Steve Beri, Bryan Ferguson, and Cairsty Grassie for their patience and assistance. We are grateful to those who aided in data collection including Lindsay Bacik, Madelyn Rea, and Brianna Paterniani. The authors thank Rebecca Corwin and her group for their help and insight and Tyler Wagner for his guidance on data analysis. We are grateful to Nicholas Russell, Sound Technology, and Marc Dingman of the Vandenbergh group for their guidance and assistance constructing the lickometer.

## REFERENCES

Andersen, S. L., and Teicher, M. H. (2008). Stress, sensitive periods and maturational events in adolescent depression. *Trends Neurosci.* 31, 183–191. doi: 10.1016/j.tins.2008.01.004

Archard, G. A., Earley, R. L., Hanninen, A. F., and Braithwaite, V. A. (2012). Correlated behaviour and stress physiology in fish exposed to different levels of predation pressure. *Funct. Ecol.* 26, 637–645. doi: 10.1111/j.1365-2435.2012.01968.x

Bateson, M., and Matheson, S. M. (2007). Performance on a categorisation task suggests that removal of environmental enrichment induces "pessimism" in captive European starlings (*Sturnus vulgaris*). *Anim. Welfare* 16, 33–36.

Brilot, B. O., Asher, L., and Bateson, M. (2010). Stereotyping starlings are more "pessimistic". *Anim. Cogn*. 13, 721–731. doi: 10.1007/s10071-010-0323-z

Bronstein, P. M. (1972). Open-field behavior of the rat as a function of age: Cross-sectional and longitudinal investigations. *J. Comp. Physiol. Psychol.* 80, 335–341. doi: 10.1037/h0032986

Brydges, N. M., Hall, H., Nicolson, R., Holmes, M. C., and Hall, J. (2012). The effects of juvenile stress on anxiety, cognitive bias and decision making in adulthood: a rat model. *PLoS ONE* 7:e48143. doi: 10.1371/journal.pone.0048143

Brydges, N. M., Leach, M., Nicol, K., Wright, R., and Bateson, M. (2011). Environmental enrichment induces optimistic cognitive bias in rats. *Anim. Behav*. 81, 169–175. doi: 10.1016/j.anbehav.2010.09.030

Burman, O. H. P., Parker, R., Paul, E. S., and Mendl, M. (2008). Sensitivity to reward loss as an indicator of animal affect and welfare. *Biol. Lett.* 4, 330–333. doi: 10.1098/rsbl.2008.0113

Burman, O. H. P., Parker, R., Paul, E. S., and Mendl, M. (2009). Anxiety-induced cognitive bias in non-human animals. *Physiol. Behav.* 98, 345–350. doi: 10.1016/j.physbeh.2009.06.012

Buwalda, B., Geerdink, M., Vidal, J., and Koolhaas, J. M. (2011). Social behavior and social stress in adolescence: a focus on animal models. *Neurosci. Biobehav. Rev*. 35, 1713–1721. doi: 10.1016/j.neubiorev.2010.10.004

Cavigelli, S. A., Ragan, C. M., Michael, K. C., Kovacsics, C. E., and Bruscke, A. P. (2009). Stable behavioral inhibition and glucocorticoid production as predictors of longevity. *Physiol. Behav.* 98, 205–214. doi: 10.1016/j.physbeh.2009.05.012

Costantini, D., Monaghan, P., and Metcalfe, N. B. (2012). Early life experience primes resistance to oxidative stress. *J. Exp. Biol.* 215, 2820–2826. doi: 10.1242/jeb.072231

Doyle, R. E., Lee, C., Deiss, V., Fisher, A. D., Hinch, G. N., and Boissy, A. (2011). Measuring judgement bias and emotional reactivity in sheep following long-term exposure to unpredictable and aversive events. *Physiol. Behav.* 102, 503–510. doi: 10.1016/j.physbeh.2011.01.001

Doyle, R. E., Vidal, S., Hinch, G. N., Fisher, A. D., Boissy, A., and Lee, C. (2010). The effect of repeated testing on judgement biases in sheep. *Behav. Process.* 83, 349–352. doi: 10.1016/j.beproc.2010.01.019

Foilb, A. R., Lui, P., and Romeo, R. (2011). The transformation of hormonal stress responses throughout puberty and adolescence. *J. Endocrinol.* 210, 391–398. doi: 10.1530/JOE-11-0206

Flaherty, C. F., and Rowan, G. A. (1989). Rats (*Rattus norvegicus*) selectively bred to differ in avoidance-behavior also differ in response to novelty stress, in glycemic conditioning, and in reward contrast. *Behav. Neural Biol.* 51, 145–164. doi: 10.1016/S0163-1047(89)90782-6

Gomez, M. J., Escarabajal, M. D., de la Torre, L., Tobena, A., Fernandez-Teruel, A., and Torres, C. (2009). Consummatory successive negative and anticipatory contrast effects in inbred Roman rats. *Physiol. Behav.* 97, 374–380. doi: 10.1016/j.physbeh.2009.03.003

Gross, A. N., Richter, S. H., Engel, A. K. J., and Wurbel, H. (2012). Cage-induced stereotypies, perseveration and the effects of environmental enrichment in laboratory mice. *Behav. Brain Res.* 234, 61–68. doi: 10.1016/j.bbr.2012.06.007

Hammond, R., Mauk, R., Ninaci, D., Nelson, D., and Gibbs, R. B. (2009). Chronic treatment with estrogen receptor agonists restores acquisition of a spatial learning task in young ovariectomized rats. *Horm. Behav.* 56, 309–314. doi: 10.1016/j.yhbeh.2009.06.008

Harding, E. J., Paul, E. S., and Mendl, M. (2004). Cognitive bias and affective state. *Nature* 427, 312. doi: 10.1038/427312a

Koolhaas, J. M., Bartolomucci, A., Buwalda, B., de Boer, S. F., Fluegge, G., Korte, S. M., et al. (2011). Stress revisited: a critical evaluation of the stress concept. *Neurosci. Biobehav. Rev.* 35, 1291–1301. doi: 10.1016/j.neubiorev.2011.02.003

Lombardi, B. R., and Flaherty, C. F. (1978). Apparent disinhibition of successive but not of simultaneous negative contrast. *Anim. Learn. Behav.* 6, 30–42. doi: 10.3758/BF03211999

Mathews, I. Z., Wilton, A., Styles, A., and McCormick, C. M. (2008). Increased depressive behaviour in females and heightened corticosterone release in males to swim stress after adolescent social stress in rats. *Behav. Brain Res.* 190, 33–40. doi: 10.1016/j.bbr.2008.02.004

McCormick, C. M. (2007). An animal model of social instability stress in adolescence and risk for drugs of abuse. *Physiol. Behav.* 99, 194–203. doi: 10.1016/j.physbeh.2009.01.014

McCormick, C. M., and Green, M. R. (2012). From the stressed adolescent to the anxious and depressed adult: Investigations in rodent models. *Neuroscience*. doi: 10.1016/j.neuroscience.2012.08.063. [Epub ahead of print].

McCormick, C. M., and Mathews, I. Z. (2008). HPA function in adolescence: Role of sex hormones in its regulation and the enduring consequences of exposure to stressors. *Pharmacol. Biochem. Behav.* 86, 220–233. doi: 10.1016/j.pbb.2006.07.012

McCormick, C. M., Mathews, I. Z., Thomas, C., and Waters, P. (2010). Investigations of HPA function and the enduring consequences of stressors in adolescence in animal models. *Brain Cogn.* 72, 73–85. doi: 10.1016/j.bandc.2009.06.003

McCormick, C. M., Robarts, D., Gleason, E., and Kelsey, J. E. (2004). Stress during adolescence enhances locomotor sensitization to nicotine in adulthood in female, but not male, rats. *Horm. Behav.* 46, 458–466. doi: 10.1016/j.yhbeh.2004.05.004

McCormick, C. M., Thomas, C. M., Sheridan, C. S., Nixon, F., Flynn, J. A., and Mathews, I. Z. (2012). Social instability stress in adolescent male rats alters hippocampal neurogenesis and produces deficits in spatial location memory in adulthood. *Hippocampus* 22, 1300–1312. doi: 10.1002/hipo.20966

McCormick, C. M., Smith, C., and Mathews, I. Z. (2008). Effects of chronic social stress in adolescence on anxiety and neuroendocrine response to mild stress in male and female rats. *Behav. Brain Res.* 187, 228–238. doi: 10.1016/j.bbr.2007.09.005

McEwen, B. S. (2005). Glucocorticoids, depression, and mood disorders: structural remodeling in the brain. *Metabolism* 54, 20–23. doi: 10.1016/j.metabol.2005.01.008

Mendl, M., Burman, O. H. P., Parker, R. M. A., and Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: emerging evidence and underlying mechanisms. *Appl. Anim. Behav. Sci.* 118, 161–181. doi: 10.1016/j.applanim.2009.02.023

Mendl, M., and Paul, E. S. (2004). Consciousness, emotion and animal welfare: insights from cognitive science. *Anim. Welfare* 13, S17–S25.

Mitchell, C., and Flaherty, C. (1998). Temporal dynamics of corticosterone elevation in successive negative contrast. *Physiol. Behav.* 64, 287–292. doi: 10.1016/S0031-9384(98)00072-9

Odberg, F. O. (1987). The influence of cage size and environmental enrichment on the development of stereotypies in bank voles (*Clethrionomys glareolus*). *Behav. Process.* 14, 155–173. doi: 10.1016/0376-6357(87)90042-8

Paul, E. S., Harding, E. J., and Mendl, M. (2005). Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci. Biobehav. Rev*. 29, 469–491. doi: 10.1016/j.neubiorev.2005.01.002

Plotsky, P. M., and Meaney, M. J. (1993). Early, postnatal experience alters hypothalamic corticotropin-releasing factor (CRF) mRNA, median eminence CRF content and stress-induced release in adult rats. *Mol. Brain Res.* 18, 195–200. doi: 10.1016/0169-328X(93)90189-V

Romeo, R. D. (2003). Puberty: a period of both organizational and activational effects of steroid hormones on neurobehavioural development. *J. Neuroendocrinol.* 15, 1185–1192. doi: 10.1111/j.1365-2826.2003.01106.x

Romeo, R. D. (2010). Pubertal maturation and programming of hypothalamic–pituitary–adrenal reactivity. *Front. Neuroendocrinol.* 31, 232–240. doi: 10.1016/j.yfrne.2010.02.004

Romeo, R. D., and McEwen, B. S. (2007). Stress and the adolescent brain. *Ann. N.Y. Acad. Sci.* 1094, 202–214. doi: 10.1196/annals.1376.022

Sánchez, M. M., Aguado, F., Sánchez-Toscano, F., and Saphier, D. (1998). Neuroendocrine and immunocytochemical demonstrations of decreased hypothalamo-pituitary–adrenal axis responsiveness to restraint stress after long-term social isolation. *Endocrinology* 139, 579–587. doi: 10.1210/en.139.2.579

Sánchez, M. M., Ladd, C. O., and Plotsky, P. M. (2001). Early adverse experience as a developmental risk factor for later psychopathology: Evidence from rodent and primate models. *Dev. Psychopathol.* 13, 419–449. doi: 10.1017/S0954579401003029

Saul, M. L., Tylee, D., Becoats, K. T., Guerrero, B. G., Sweeney, P., Helmreich, D. L., et al. (2012). Long-term behavioral consequences of stress exposure in adolescent versus young adult rats. *Behav. Brain Res.* 229, 226–234. doi: 10.1016/j.bbr.2012.01.022

Schmidt, M. V., Sterlemann, V., Ganea, K., Liebl, C., Alam, S., Harbich, D., et al. (2007). Persistent neuroendocrine and behavioral effects of a novel, etiologically relevant mouse paradigm for chronic social stress during adolescence. *Psychoneuroendocrinology* 32, 417–429. doi: 10.1016/j.psyneuen.2007.02.011

Simon, P., Dupuis, R., and Costentin, J. (1994). Thigmotaxis as an index of anxiety in mice. Influence of dopaminergic transmissions. *Behav. Brain Res.* 1, 59–64. doi: 10.1016/0166-4328(94)90008-6

Sheriff, M. J., Krebs, C. J., and Boonstra, R. (2009). The sensitive hare: sublethal effects of predator stress on reproduction in snowshoe hares. *J. Anim. Ecol.* 78, 1249–1258. doi: 10.1111/j.1365-2656.2009.01552.x

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neurosci. Biobehav. Rev.* 24, 417–463. doi: 10.1016/S0149-7634(00)00014-2

Sterlemann, V., Rammes, G., Wolf, M., Liebl, C., Ganea, K., Müller, M. B., et al. (2010). Chronic social stress during adolescence induces cognitive impairment in aged mice. *Hippocampus* 20, 540–549.

Strekalova, T., Spanagel, R., Bartsch, D., Henn, F. A., and Gass, P. (2004). Stress-induced anhedonia in mice is associated with deficits in forced swimming and exploration. *Neuropsychopharmacology* 29, 2007–2017. doi: 10.1038/sj.npp.1300532

Torregrossa, M. M., Xie, M., and Taylor, J. R. (2012). Chronic corticosterone exposure during adolescence reduces impulsive action but increases impulsive choice and sensitivity to yohimbine in male Sprague-Dawley rats. *Neuropsychopharmacology* 37, 1656–1670. doi: 10.1038/npp.2012.11

Van Dijken, H. H., Van Der Heyden, J. A. M., Mos, J., and Tilders, F. J. H. (1992). Inescapable footshocks induce progressive and long-lasting behavioural changes in male rats. *Physiol. Behav.* 51, 787–794. doi: 10.1016/0031-9384(92)90117-K

van Eden, C. G., Kros, J. M., and Uylings, H. B. (1990). The development of the rat prefrontal cortex. Its size and development of connections with thalamus, spinal cord and other cortical areas. *Prog. Brain Res.* 85, 169–183. doi: 10.1016/S0079-6123(08)62680-1

Vidal, J., Buwalda, B., and Koolhaas, J. M. (2011). Differential long-term effects of social stress during adolescence on anxiety in Wistar and wild-type rats. *Behav. Process.* 87, 176–182. doi: 10.1016/j.beproc.2011.03.004

Watt, M. J., Burke, A. R., Renner, K. J., and Forster, G. L. (2009). Adolescent male rats exposed to social defeat exhibit altered anxiety behavior and limbic monoamines as adults. *Behav. Neurosci.* 123, 564–576. doi: 10.1037/a0015752

Winkielman, P., Knutson, B., Paulus, M., and Trujillo, J. L. (2007). Affective influence on judgments and decisions: moving towards core mechanisms. *Rev. Gen. Psychol.* 11, 179–192. doi: 10.1037/1089-2680.11.2.179

Zurita, A., Martijena, I., Cuadra, G., Brandão, M. L., and Molina, V. (2000). Early exposure to chronic variable stress facilitates the occurrence of anhedonia and enhanced emotional reactions to novel stressors: reversal by naltrexone pretreatment. *Behav. Brain Res.* 117, 163–171. doi: 10.1016/S0166-4328(00)00302-8

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 01 April 2013; accepted: 12 June 2013; published online: 04 July 2013.*

*Citation: Chaby LE, Cavigelli SA, White A, Wang K and Braithwaite VA (2013) Long-term changes in cognitive bias and coping response as a result of chronic unpredictable stress during adolescence. Front. Hum. Neurosci. **7**:328. doi: 10.3389/fnhum.2013.00328*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 August 2013
doi: 10.3389/fnhum.2013.00447

# Violence as a source of pleasure or displeasure is associated with specific functional connectivity with the nucleus accumbens

***Eric C. Porges* and Jean Decety***

*Department of Psychology, The University of Chicago, Chicago, IL, USA*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*Oliver C. Schultheiss, Friedrich Alexander University, Germany*
*Sandra Baez, Institute of Cognitive Neurology, Argentina*

****Correspondence:***
*Eric C. Porges, Department of Psychology, The University of Chicago, 5848 S. University Avenue, Chicago, IL 60637, USA*
*e-mail: eporges@uchicago.edu*

The appraisal of violent stimuli is dependent on the social context and the perceiver's individual characteristics. To identify the specific neural circuits involved in the perception of violent videos, forty-nine male participants were scanned with functional MRI while watching video-clips depicting Mixed Martial Arts (MMA) and Capoeira as a baseline. Prior to scanning, a self-report measure of pleasure or displeasure when watching MMA was collected. Watching MMA was associated with activation of the anterior insula (AI), brainstem, ventral tegmental area (VTA), striatum, medial, and lateral prefrontal cortex, orbitofrontal cortex, somatosensory cortex, and supramarginal gyrus. While this pattern of brain activation was not related to participants' reported experience of pleasure or displeasure, pleasurable ratings of MMA predicted increased functional connectivity (FC) seeded in the nucleus accumbens (NAcc) (a structure known to be responsive to anticipating both positive and negative outcomes) with the subgenual anterior cingulate cortex (ACC) and anterior insular cortex (AIC) (regions involved in positive feelings and visceral somatic representations). Displeasure ratings of MMA were related to increased FC with regions of the prefrontal cortex and superior parietal lobule, structures implicated in cognitive control and executive attention. These data suggest that functional connectivity is an effective approach to investigate the relationship between subjective feelings of pleasure and pain of neural structures known to respond to both the anticipation of positive and negative outcomes.

**Keywords: violence, emotion, pleasure, functional connectivity, nucleus accumbens, insular cortex, orbitalfrontal cortex, reward**

## INTRODUCTION

Across animal species, violence is a salient stimulus. It is an important environmental signal with survival consequences, conveying information about potential threats to one's health and physical integrity. Intraspecific violence is a near ubiquitous feature of the environment, and mammals, humans intraspecies, are no exception. Anthropologists agree that violence is pervasive, ancient, infinitely various and a central fact of human life (Whitehead, 2004). Like violence in general, violence in the context of entertainment (e.g., pankration, gladiators, boxing, etc.), is reported throughout the historical record of human behavior as one of the oldest forms of entertainment (Schechter, 2005). Examples of violent entertainment exist throughout contemporary society and are widely sought out; indicating that within the context of entertainment, a sizable portion of the population is attracted to viewing violence.

During the past decade, a number of functional neuroimaging studies have investigated the neural response while observing other individuals being hurt (Morrison et al., 2004; Jackson et al., 2005a; Cheng et al., 2007; Gu and Han, 2007; Lamm et al., 2007a,b; Moriguchi et al., 2007; Zaki et al., 2007; Michalska et al., 2013), and devoid of social context these studies have consistently revealed the activation of a specific neural network. This work implicates several brain regions that overlap with regions involved in processing physical pain, including the anterior cingulate cortex (ACC), anterior midcingulate cortex (aMCC), anterior insula (AI), and periaqueductal gray (PAG). Some of these studies have also reported involvement of the somatosensory cortex (Cheng et al., 2007, 2008; Moriguchi et al., 2007; Benuzzi et al., 2008; Lamm and Decety, 2008; Wood et al., 2010). Together, these regions are thought to be responsive to aversive stimuli, directing attention toward and supporting a representation of these stimuli, as well as supporting a response to a potential threat in the environment (Ogino et al., 2007; Benuzzi et al., 2008; Akitsuki and Decety, 2009; Decety and Michalska, 2010). Although this ensemble of regions recruited when one watches or experiences pain (sometimes referred to as the pain matrix) is well established (for review see Lamm et al., 2011), much less is known for social situations that are both violent and socially appropriate.

A limited number of studies have investigated neural responses to violence in the context of socially acceptable entertainment such as television shows (e.g., Murray et al., 2006a). In these studies, a similar pattern of neural recruitment was observed, as in previous research with non-entertainment stimuli depicting pain and injury, with the notable inclusion of the striatum. In rodents the nucleus accumbens (NAcc), a structure within the striatum, has populations of cells that are selectively responsive to the anticipation of aversive outcomes (Badrinarayan et al., 2012). Other investigations have identified distinct regions within the NAcc to be particularly responsive to uncertain, but salient events, with outcomes that could be either positive or negative (Anselme et al., 2013). In humans, while no functional segregation has been demonstrated, the NAcc has similarly been reported to be particularly responsive to the anticipation of salient events where the outcome is uncertain but could be either positive or negative (Jensen et al., 2003; Cooper and Knutson, 2008).

NAcc recruitment associated with both positive and negative outcomes is not unexpected, as it is still unclear whether separate neural systems are involved in aversive and appetitive processing in the brain. In fact, at the systems level, there is extensive overlap throughout the neural circuitry and chemistry involved in processing appetitive and aversive stimuli (Leknes and Tracey, 2008). Recently, progress has been made in disentangling the specific circuits that underlie this complex behavior in animal models (Smith et al., 2011).

Describing a psychological construct as synonymous with neuroanatomy, such as "pleasure" always being concurrent with NAcc recruitment, obscures the nuances of brain reactivity. For example, when rodents were placed in a stressful or threatening environment, the regions of the NAcc that were activated in response to aversive stimuli, roughly double in size compared to baseline (Reynolds and Berridge, 2008). However when rodents were housed in their preferred, safe environment, regions associated with appetitive behaviors expanded dramatically while those associated with responses to fear shrunk to one third of their baseline size.

Given the importance of the information conveyed in observed violence, neural structures sensitive to salience, such as portions of the ventral striatum including the NAcc, should be responsive to viewing violence independent of the subjective appraisal. Only after this appraisal, which occurs in the context of the viewer's relationship with the stimuli, can an evaluation and behavioral strategy be generated to observed violence. These behavioral responses, emerging from the connectivity of neural structures' responsive to salience, would be dependent on the individual's bias to perceive violence as pleasurable or not pleasurable.

Previous functional neuroimaging research has shown that a small shift in context or in the instructions given to a subject modulates neural responses to the observation of another in pain; these shifts have been observed both centrally and peripherally. For example, inducing an individual to adopt the perspective of another person experiencing pain increases the magnitude of facial muscle contractions associated with first person pain, and was associated with increases heart rate (Lamm et al., 2008). Presenting a participant with stimuli depicting a person being injured by another under two instructions sets—"imagine you are harming the other" and "imagine you are being harmed by the other"—produces a relatively larger response in pain matrix, as well as in regions associated with autonomic arousal (Decety and Porges, 2011). More importantly, in the previous study, the most reliable and significant changes in brain response were detected in functional connectivity (FC) analyses.

The work cited above has either implicitly or explicitly characterized observed violence as strictly aversive stimuli (Jackson et al., 2005a, 2006; Lamm et al., 2007a; Lamm and Decety, 2008) and demonstrated changes in functional recruitment based on shifting the context of the violence within subject (Lamm et al., 2008; Decety and Porges, 2011) or selecting subject populations that would contextualize the violent stimuli in different ways, such as physicians viewing a medical procedure as compared to lay subjects viewing the same procedure (Cheng et al., 2007). However, violence has a long history of being a source of entertainment and pleasure, and today there are still numerous socially appropriate ways for people to watch and consume violence, indicating that observing violence is not always aversive, but can also be desired and contain pleasurable aspects for the viewer. These appetitive aspects are apparent in nearly all forms of popular media: television, movies, video games, etc. where violent content is overtly marketed. We hypothesize that the neural mechanisms engaged when an individual views violence in the context of entertainment they derive pleasure from viewing will be functionally distinct from those activated when an individual observes they find unpleasurable to view.

The present study tests this hypothesis by focusing on the dynamic response of the NAcc and the functional relationship this region has with other neural targets, as a function of individual differences in the relationship to a given stimulus. To test this hypothesis, the study displayed violent stimuli to participants, who had previously reported the degree of pleasure or displeasure they derived from watching similar videos. Based on this self-reported measure, a FC analysis was run, seeded from the NAcc with pleasure ratings used as a predictor. Drawing on previous research, we anticipated that whole brain responses would include the activation of regions implicated in the perception of pain in others (Jackson et al., 2005b; Murray et al., 2006a; Decety and Porges, 2011). We further predicted that findings in the FC analysis seeded from the NAcc would increase with regions associated with positive somatic and visceral representations and pleasurable states, including the AI (Damasio and Damásio, 1994; Craig, 2009) and frontal cortex (e.g., Moll et al., 2007).

## METHODS

### PARTICIPANTS

Forty-nine male individuals were recruited from the Chicago metropolitan area (mean age = 25 years; range = 18–35 years). Flyers and online advertisements were posted in the region that targeted individuals who either did or did not enjoy watching Mixed Martial Arts (MMA). Participants had no neurological diagnosis and had normal or corrected to normal vision. Participants' written informed consent was obtained, and this study was approved by the Institutional Review Board at the University of Chicago.

### MATERIALS

Stimuli consisted of 10 s video clips of the target condition (MMA) or control (Capoeira), 20 clips per condition. MMA is a full contact combat sport similar to boxing and kick boxing, but also involves the use of other martial arts techniques such as wrestling. MMA contestants utilize these tools to knock unconscious, force their opponent to "give up," or be designated the winner by a panel of judges. MMA has grown increasingly popular since its start in 1993, with the parent company of UFC valued at approximately $1 billion in 2008 (Gregory and Osborne, 2009). MMA is a sport regularly shown on both general broadcast television and basic cable and all clips used in this study were in line with standards for these media. Non-MMA, control condition stimuli consisted of Capoeira, a form of dance, derived from Brazilian martial arts, which like MMA involves kicks and hand and arm strikes, such as punches, between two participants of the same gender, but no contact is made between participants and no injury or intent to injure takes place. This control condition was selected because it was non-violent, but in all other ways, is very similar in physical and visual characteristics to MMA.

## PROCEDURE

To ensure all participants were run under consistent conditions, the same 2 investigators, dressed in white lab coats, were responsible for all subject interactions. Prior to scanning, subjects were asked to report "How pleasurable do you find watching MMA?" on a 5-point Likert scale ranging from −2 "extremely unpleasurable" to 0 "Neither pleasurable or unpleasurable" to + 2 "Most pleasurable" ($m = 0.45$, std $= 1.04$). Stimuli were presented using E-Prime 2.0 Professional (Psychology Software Tools, Inc., Pittsburgh, PA) and displayed in-scanner via an Avotec rear projection system (Avotec Inc., Stuart, FL). Prior to functional runs, subjects were shown sample stimuli of both classes in the manner the actual task would employ. During functional runs, each participant viewed one of four possible stimuli permutations. Two of the permutations were randomly generated with manual oversight to prevent the same class of stimuli being displayed more than twice in a row; these permutations were inverted for a total of four permutations. Prior to the onset of all stimuli, a one-word text cue describing the stimuli that were to appear was displayed for 1 s. A 10 s inter-block interval preceded and followed all video clips. Due to the engaging nature of the task and the anticipation of motor, motor planning, and somatosensory recruitment, a decision was made to give participants no task during functional runs other than to attend to the stimuli. Compliance was confirmed via in scanner camera that allowed for observation of subjects eyes. Following collection of functional data, a structural scan was collected.

## MRI SCANNING PARAMETERS AND ANALYSIS

### *Functional acquisition*

Whole brain functional MRI data were collected with an 8-channel Phillips Sense head-coil in a Philips Achieva 3T scanner using Phillips provided T2* weighted EPI sequence (32 slices, TR 2 s, TE 25 ms, FOV 240 × 240 ×127.5 mm, 80 × 80 ×32 mm matrix, flip angle 80°, in plane resolution of 3 × 3 mm, slice thickness 3.5 mm, 0.5 mm skip). To recover signal in the orbital frontal region, a Philips provided Z-Shim sequence was utilized.

### *Structural MRI*

Anatomical scans were acquired in the axial plane using a Philips SENSE-Ref, T-1 weighted sequence (300 slices, 1.2 mm, −0.6 mm gap, FOV: 250 × 250 × 180 mm, 240 × 240 × 300 matrix).

### *Image processing and analysis*

Using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK) in Matlab (Mathworks Inc., Sherborn, MA, USA), functional volumes were motion realigned, with movement parameters saved for use at the first level as repressors of no interest. Functional images were co-registered to the segmented structural images (gray matter, white matter, and cerebral spinal fluid), and these structural images were normalized to the MNI template with these normalization parameters applied to the co-registered functional images. Normalized functional images were resliced as 2 mm isotropic voxels and smoothed with an 8 mm full-width at half-maximum Gaussian kernel. Hemodynamic responses were assessed by setting up fixed effects general linear models (GLM) for each subject. Regressors of interest modeling the experimental conditions and the cue displayed prior to all stimuli, epochs were defined and convolved with the canonical hemodynamic response function (HRF). All models included a high-pass filter with a cut-off at 128 s in order to remove scanner drifts. Following model estimation, contrasts were calculated for each subject to assess differences between conditions, and, in relation to the implicitly modeled fixation, baseline data were assessed. The resulting first-level contrast images were entered into second-level random effects (RFX) analyses to assess differences between conditions with population inference. Activation differences between conditions were assessed using a voxel-level threshold of $p = 0.01$ and a spatial extent threshold of $k = 10$, corrected for multiple comparisons across the whole volume using the false discovery rate (FDR, $p < 0.05$) approach (Genovese et al., 2002). Results were visualized in xjView toolbox (http://www.alivelearn.net/xjview) and MRIcron (http://www.mccauslandcenter.sc.edu/mricro/mricron/install.html).

ROI analysis of activation while watching MMA vs. an implicitly modeled baseline condition (fixation cross) was extracted. This was conducted using RFXplot (Gläscher, 2009), with a bilateral, 10 mm sphere mask created in WFU Pickatlas (Maldjian et al., 2003) of the NAcc. Results were extracted in percent signal change. This sphere was bilateral as laterality of activation has not been widely reported in previous investigations of NAcc activation to either reward or salience (Cooper and Knutson, 2008). The mask was centered at widely reported MNI coordinates for the NAcc ±10, 12, −2 (Luijten et al., 2012).

FC analysis was performed using the Conn FC toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012), using images preprocessed as previously described in SPM8. The Conn toolbox employs "aCompCor," anatomically informed component-based noise correction, to correct for physiological and other sources of noise by regressing signal from the white matter and cerebral spinal fluid as well as movement parameters (Behzadi et al., 2007). FC was assessed using how pleasurable subjects found watching MMA as a covariate for a "seed to voxel" methodology; this produced Fisher transformed correlation coefficients for all voxels in the whole brain relative to same bilateral 10 mm sphere seed region used for NAcc region in the ventral striatum in the ROI analysis. A whole brain Family Wise Error (FWE) corrected threshold of $p < 0.05$ was used to identify significant clusters.

## RESULTS

### WHOLE BRAIN ANALYSIS

#### *MMA> control*

The whole-brain analysis, FDR corrected at $p < 0.05$, contrasting the MMA condition vs. the control condition (Capoeira) revealed greater activation in a large number of clusters extending from the supplementary motor area (SMA) to the medial aspect of the superior frontal gyrus (SFG), to right dorsolateral prefrontal cortex (dlPFC). A large activation was detected in the right middle temporal gyrus extending to the superior temporal sulcus. Bilateral activation of the AI, precentral, post central gyrus, and superior parietal and temporal parietal junction (TPJ) were observed as well as bilateral regions in the visual cortex and cerebellum. Posterior and middle cingulate cortex were robustly recruited as well as extensive thalamic and brainstem recruitment such as the PAG, ventral tegmental area (VTA), and regions in the striatum including the head of caudate (see **Table 1** and **Figure 1**).

#### *Control> MMA*

A whole brain analysis of the control condition (Capoeira) greater than MMA, FDR corrected at $p < 0.05$, revealed bilateral posterior insula cortex, precentral gyrus, and parahippocampal gyrus, were more active. In addition, increased activation was seen in regions of the visual cortex and a large cluster focused in the posterior portions of the SMA that extended to the bilateral postcentral gyrus, extending on the right side to the superior temporal gyrus.

### SELF REPORTED PLEASURE FROM MMA

Subjects' reports of the pleasure they derived from watching MMA was used as a covariate of interest in our in the whole

**Table 1 | Brain regions that show a significant hemodynamic increase when participants watched MMA videos as compared with control videos (Capoeira).**

| Region | Abbreviation | Side | *x* | *y* | *z* | *t*-value |
|---|---|---|---|---|---|---|
| Anterior insula cortex | AIC | l | −36 | 24 | −8 | 4.82 |
| Anterior insula cortex | AIC | r | 46 | 22 | 6 | 5.41 |
| Anterior thalamus | | r | 10 | −4 | −2 | 5.09 |
| Brainstem | | l/r | −6 | −20 | −20 | 4.76 |
| Brainstem | | r | 8 | −14 | −12 | 3.65 |
| Brainstem | | l/r | −2 | −22 | −20 | 4.74 |
| Caudate head | Ventral striatum | r | 8 | 10 | 4 | 3.84 |
| Caudate head | Ventral striatum | r | 10 | 8 | 2 | 3.83 |
| Caudate body | | l | −8 | 2 | 12 | 2.81 |
| Dorsal lateral prefrontal cortex | DLPFC | r | 28 | 50 | 18 | 4.56 |
| midcingulate cortex | MCC | l/r | −4 | −22 | 36 | 4.04 |
| Inferior frontal gyrus | IFG | r | 38 | 24 | −20 | 4.49 |
| Inferior parietal lobule | | l | −38 | −40 | 46 | 3.71 |
| Inferior parietal lobule | | r | 30 | −44 | 56 | 5.25 |
| Medial frontal gyrus | MFG | r | 54 | 16 | 30 | 3.15 |
| Medial frontal gyrus | MFG | r | 38 | 20 | 32 | 4.63 |
| Midtemporal gyrus | STS | r | 52 | −32 | −10 | 5.11 |
| Periaqueductal gray | | l/r | 2 | −28 | −8 | 3.41 |
| Posterior cingulate cortex | PCC | l/r | 4 | −50 | 30 | 3.56 |
| Postcentral gyrus | | r | 52 | −22 | 34 | 3.1 |
| Postcentral gyrus | | l | −56 | −24 | 32 | 7.31 |
| Precentral gyrus | | l | −42 | −2 | 46 | 2.89 |
| Precentral gyrus | | r | 32 | −8 | 58 | 3.51 |
| Superior frontal gyrus | SFG | l/r | 4 | 38 | 50 | 4.54 |
| Superior medial frontal cortex | SupMPFC | l/r | 2 | 46 | 40 | 5.11 |
| Superior medial frontal cortex | SupMPFC | l/r | 4 | 52 | 24 | 4.88 |
| Supplementary motor area | SMA | l/r | 8 | 24 | 56 | 4.39 |
| Superior parietal lobule | SPL | l | −28 | −58 | 62 | 8.3 |
| Superior parietal lobule | SPL | r | 26 | −60 | 60 | 7.99 |
| Temporal parietal junction | Supramarginal gyrus | l | −62 | −48 | 36 | 3.89 |
| Temporal parietal junction | Supramarginal gyrus | r | 58 | −44 | 34 | 2.86 |
| Temporal parietal junction | Supramarginal gyrus | r | 54 | −56 | 34 | 3.46 |
| Temporal pole | | r | 44 | 6 | −42 | 3.24 |
| Ventral tegmental area | VTA | l/r | −2 | −22 | −20 | 4.74 |

*L, left hemisphere; R, right hemisphere. MNI (x, y, z) coordinates.*

brain analysis, FDR corrected at $p < 0.05$, $k = 5$. No regions were significant.

### ROI ANALYSIS

In preparation for the FC analysis, a region of interest analysis was extracted for bilateral NAcc against baseline. Activation in that region was not correlated with self-reported pleasure from watching MMA ($r = -0.031$) (**Figure 2**).

### FUNCTIONAL CONNECTIVITY

A whole brain, FWE corrected threshold of $p < 0.05$ was used to identify significant clusters. A large cluster that extended from the subgenual cingulate cortex to the left anterior insular cortex (AIC) increased in connectivity with the NAcc as self-reported pleasure from watching MMA increased (**Table 1**). Conversely, as self-reported pleasure from watching MMA decreased, multiple clusters emerged identifying a cluster extending over the left dlPFC, another in the right superior parietal lobule, and a third in the left hemisphere of the cerebellum (**Table 2**; **Figure 3**).

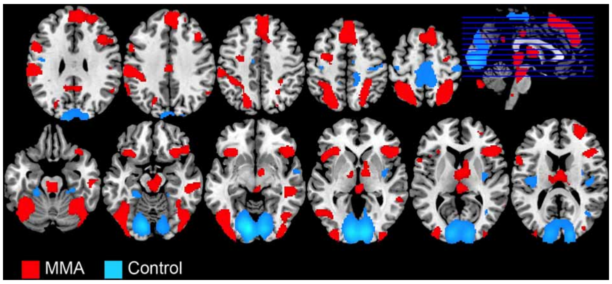


**FIGURE 1 | Whole brain analysis showing BOLD signal changes in 49 male participants while watching videos depicting a full contact, Mixed Martial Arts (MMA, in red) vs. control videos depicting Capoeira (Control, in blue).**

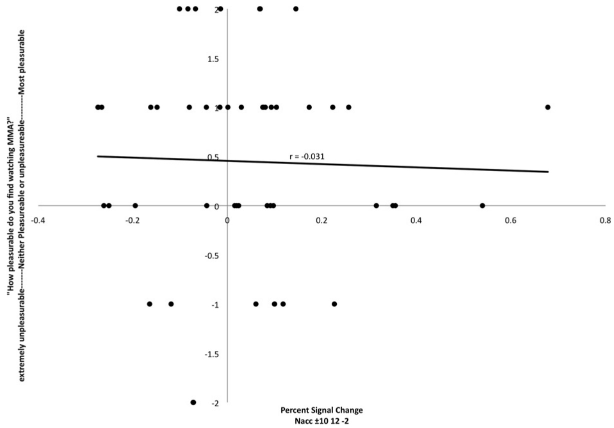


**FIGURE 2 | Relationship between participant's reports of pleasure from watching MMA and percent signal change in the nucleus accumbens (MNI coordinates: ±10, 12, −2).** Each point corresponds to a single subject's percent signal change on the *X* axis and self-reported pleasure or displeasure rating on the *Y* axis.

## DISCUSSION

This study was designed to investigate how the neural response and shift in FC seeded from the NAcc are modulated by the relative pleasure individuals experience when watching violent videos. The violent videos selected were MMA, chosen based of their intensity, pervasiveness, and popularity, as well as the wide range of variability regarding how pleasurable subjects perceived the stimuli.

The whole brain analysis level, contrasting MMA vs. the control condition, identified a pattern of brain response consistent with a large literature that has previously documented the neural response to observing or imagining pain and injury in others. Regions recruited included the midcingulate cortex, SMA, PAG, AI, and somatosensory cortex (Akitsuki and Decety, 2009; Decety

**Table 2 | Regions that show greater effective functional connectivity seeded in NAcc.**

| Brain region | MNI (*x*, *y*, *z*) | | | *k* | *t*-value |
|---|---|---|---|---|---|
| **GREATER FUNCTIONAL CONNECTIVITY WITH NAcc, PLEASURE** | | | | | |
| Left anterior insula/subgenual cingulate cortex | −24 | 24 | −6 | 363 | 3.94 |
| **GREATER FUNCTIONAL CONNECTIVITY WITH NAcc, DISPLEASURE** | | | | | |
| Left dorsolateral prefrontal cortex | −48 | 12 | 44 | 303 | 6.67 |
| Right superior parietal | 30 | −72 | 46 | 185 | 4.2 |
| Left cerebellum | −36 | −58 | −26 | 243 | 4.96 |

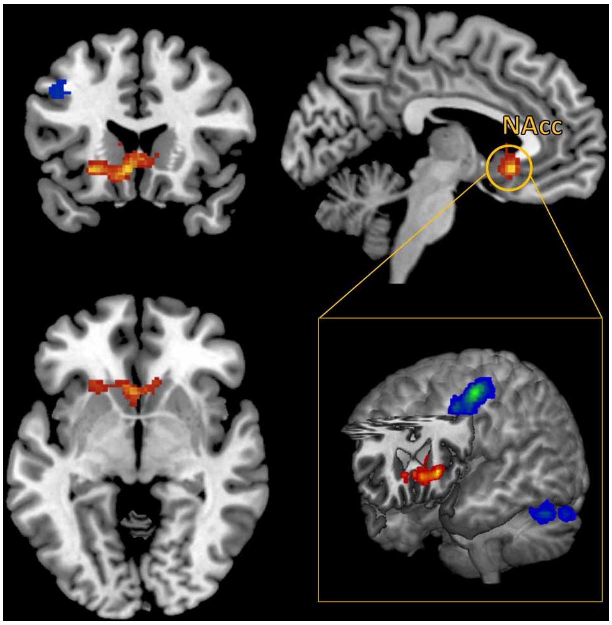


**FIGURE 3 | Functional connectivity seeded in the nucleus accumbens (MNI coordinates: ±10, 12, −2) associated with pleasurable ratings (in red) and connectivity associated with unpleasurable ratings (in blue).**

et al., 2009; Decety and Michalska, 2010; Decety and Porges, 2011; Lamm et al., 2011). In addition, we observed the involvement of regions implicated in theory of mind, goal oriented attention, and executive function (Corbetta et al., 2008), including a parietal cluster that encompassed the TPJ (Decety and Lamm, 2007) and a cluster over the dorsolateral cortex (Meyer et al., 2012). Regions implicated in autonomic arousal, such as the brainstem and mid-cingulate cortex (Critchley, 2009), were also activated. Finally, the orbitofrontal cortex, which plays a critical role in emotion and decision-making, but is also responsive to both positive and aversive stimuli (Rolls et al., 2003), and the temporal pole, a region that binds visceral emotional information to perceptual information (Olson et al., 2007), were also detected at the whole brain level, these areas have previously been reported to be responsive to strongly valenced stimuli, such as unpleasant images (Aldhafeeri et al., 2012) from the International Affective Picture System (Lang et al., 1997). Compared to previous neuroimaging research that employed violent, though social acceptable, entertainment with non-violent entertainment (e.g., Murray et al., 2006a), our study found many of the same regions of increased hemodynamic activity including in visual areas, thalamus, caudate, and precentral gyrus. One notable difference, however, is the lack of amygdala activation in response to MMA. While this structure that plays a primary role in the detection and appraisal of biologically relevant stimuli (not restricted to fear) has been widely reported as being responsive to the observation of violence (e.g., Mathiak and Weber, 2006; Decety et al., 2012), or even imaging harming another person (Decety and Porges, 2011), it is important to note that our study used very well matched control stimuli (Capoeira) that contained a large amount of extremely intense and similar visual information, with the main difference from MMA being the lack injury or intent to injure. This is in contrast to the previous literature (e.g., Murray et al., 2006a) which utilized control conditions with dramatically less visual activity, such a National Geographic program made for a child audience. In our study, both MMA and Capoeira, evaluated relative to a fixation cross, produced robust activation in the amygdala.

The stimuli used in the Capoeira control condition were, as in the MMA condition, extracted from commercial videos. Unavoidably, this resulted in some variability between the stimulus classes, with Capoeira stimuli containing a slightly more variable range of environments than MMA, which was always conducted in an environment with consistent features (an octagon shaped cage). In addition, the MMA stimuli tended to have a tighter focus on the participants and slightly fewer kicks and punches than the Capoeira stimuli. These differences, may account for the regions that were more active during the control condition relative to MMA. For example, the posterior region of the parahippocampal gyrus, that we report being more active in this condition than in MMA, has been previously reported to be preferentially responsive to novel visual scenes (Epstein et al., 1999; Epstein, 2011), of which the Capoeira stimuli contained more than the MMA stimuli. Additionally, greater activations during the Capoeira stimuli were observed in the precentral gyrus, a region that, while recruited during motor generation, is also sensitive to the observation of human movements (Servos, 2002; Grosbras et al., 2012). Finally the posterior insula also showed greater activation to the control condition. Though the activation of this region has been widely reported to be involved in nociception, but not the observation of pain in others (Lamm et al., 2011; Garcia-Larrea, 2012), in this study is best accounted for by its role supporting eye movements in complex visual scenes (Blurton et al., 2012; Indovina et al., 2013), a possible consequence of the increased visual motion in the control condition. When activity at the whole brain analysis was covaried with subjects' reported pleasure when watching MMA (data corrected for multiple comparisons), no cluster survived.

Given our predictions regarding the role of the NAcc in response to the perception of highly salient stimuli that have survival implications, an ROI analysis extracted from a bilateral NAcc mask for the MMA condition was performed. The extracted values were then covaried with subjects reported pleasure values from watching MMA. At this level of inquiry, no relationship between the HRF in the NAcc and reported pleasure scores was found (**Figure 2**). This finding mirrors published studies reporting that this region has a comparable response to the anticipation of both positive and negative outcomes (Cooper and Knutson, 2008; Carter et al., 2009). There is some electrophysiological evidence in animal models that there is a degree of functional segregation for appetitive and aversive responses in the NAcc (Reynolds and Berridge, 2002; Anselme et al., 2013), but perhaps due to limits of spatial resolution in fMRI, this has not yet been demonstrated in humans.

While pleasure ratings were not correlated with brain activity at the whole brain level, they did predict FC seeded in the NAccs. More interestingly, using reported pleasure from MMA as a covariate in the FC analysis revealed a unique pattern of brain activation. Pleasure, while watching MMA, predicted increased connectivity in the subgenual region of the ACC as well as the left AI. The subgenual region has extensive dopaminergic projections from the mesolimbic pathway (Gaspar et al., 1989) and has been previously reported to be responsive to viewing one's offspring (Bartels and Zeki, 2004), making charitable donations (Moll et al., 2006), and has been shown to reflect guilt (Zahn et al., 2009; Green et al., 2012). More generally, it has been associated in animal models with social attachment and pair bonding (Insel and Young, 2001). In patients with major depressive disorder, hypoactivity has been reported in this region (Drevets et al., 1997), and remission of symptoms have been associated with a normalizing of this region's function (Ressler and Mayberg, 2007). The increased FC between the NAcc and this subgenual prefrontal cortex, predicted by self-reported pleasure when watching MMA, may be evidence of other oriented pleasure and perhaps simultaneously guilt around enjoyment derived from the pain of another.

The cluster in the subgenual cingulate cortex extended to and had a second distinct peak in the left AIC. Recent work has begun to integrate the wide range of conditions under which activation of the AIC has been reported in functional neuroimaging studies. This work suggests that the AIC is central to understanding and representing one's own physical and somatic state (Craig, 2009). Furthermore, in our study, the connectivity with the AIC was found only in the left hemisphere, which has been associated with positive affect, appetitive behavior, and group affiliated emotion (Craig, 2005).

Conversely, displeasure while watching MMA predicted increased connectivity in the left dorsolateral cortex and the right superior parietal lobule. Interestingly, these two regions are recruited when physicians watch a painful medical procedure taking place, and this has been interpreted as a sign of top-down regulation and modulation of the aversive aspects of the procedure (Cheng et al., 2007). In addition, supporting the interpretation that recruitment of these regions supports executive function, these regions belong the dorsal frontoparietal, goal directed attention network (Corbetta and Shulman, 2002; Corbetta et al., 2008). It seems that participants, who dislike watching MMA either engage in top-down emotion regulation or, at the least, experience a more effortful appraisal of the stimuli.

In conclusion, when individuals watch violent videos, the overall pattern of brain response is not the most informative source of information in predicting the extent to which one derives pleasure or displeasure. Rather, the connectivity seeded in the NAcc, a structure known to be responsive to salient information in the environment (Smith et al., 2011), demonstrates clear and distinct responses depending on the relationship of the perceiver to the stimuli.

## FUTURE DIRECTIONS

Since an objective of this study was to use ecologically valid stimuli, it was necessary to use videos extracted from commercial broadcast in conjunction with longer stimuli durations. In the future, an event-related design, where temporal dynamics of the response could be better established would be even more informative, and provide additional information. For example, based on work that has attempted to decouple the responses of the ventral striatum to both the anticipation of positive and negative outcomes as well as the outcome itself (Cooper and Knutson, 2008), one would predict that pleasure would have little impact on the response to the anticipation of violence, but it would have significant impact on during sustained observation of violence. Our study utilized stimuli containing a dyad involved in a violent interaction and the influence of these stimuli on the FC of anatomy in a dopaminergic pathway. The response of this neuroanatomy to social conflict, has previously been demonstrated to be influenced by a participants' implicit need for power (Schultheiss and Schiepe-Tiska, 2013). The inclusion of a measure of need for power would likely have accounted for some of the variability in the responses to violent stimuli. Furthermore, the explicit measure of participant's pleasure derived from watching MMA used in the current study could be complemented by an implicit measure of need for power. Finally, our study intentionally restricted the sample to male participants within a narrow age range and geographic region, future studies would benefit from including female subjects and participants from different cultures representing a broader age range.

## ACKNOWLEDGMENTS

This study was supported by a grant (# BCS-0718480) from the National Science Foundation and a grant from NIH/NIMH (# MH84934-01A1) to Dr. Jean Decety. We are extremely grateful for the time, effort, and patience of Karen E. Smith and Janique Santos dedicated toward this project in their assistance in scheduling participants and collecting data.

## REFERENCES

Akitsuki, Y., and Decety, J. (2009). Social context and perceived agency affects empathy for pain: an event-related fMRI investigation. *Neuroimage* 47, 722–734.

Aldhafeeri, F. M., Mackenzie, I., Kay, T., Alghamdi, J., and Sluming, V. (2012). Regional brain responses to pleasant and unpleasant IAPS pictures: different networks. *Neurosci. Lett.* 512, 94–98. doi: 10.1016/j.neulet.2012.01.064

Anselme, P., Robinson, M. J. F., and Berridge, K. C. (2013). Reward uncertainty enhances incentive salience attribution as sign-tracking. *Behav. Brain Res.* 238, 53–61. doi: 10.1016/j.bbr.2012.10.006

Badrinarayan, A., Wescott, S. A., Vander Weele, C. M., Saunders, B. T., Couturier, B. E., Maren, S., et al. (2012). Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J. Neurosci.* 32, 15779–15790.

Bartels, A., and Zeki, S. (2004). The neural correlates of maternal and romantic love. *Neuroimage* 21, 1155–1166.

Behzadi, Y., Restom, K., Liau, J., and Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37, 90–101.

Benuzzi, F., Lui, F., Duzzi, D., Nichelli, P. F., and Porro, C. A. (2008). Does it look painful or disgusting. Ask your parietal and cingulate cortex. *J. Neurosci.* 28, 923–931.

Blurton, S. P., Raabe, M., and Greenlee, M. W. (2012). Differential cortical activation during saccadic adaptation. *J. Neurophysiol.* 107, 1738–1747. doi: 10.1152/jn.00682.2011

Carter, R. M., Macinnes, J. J., Huettel, S. A., and Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Front. Behav. Neurosci.* 3:21. doi: 10.3389/neuro.08.021.2009

Cheng, Y., Lin, C.-P., Liu, H.-L., Hsu, Y.-Y., Lim, K.-E., Hung, D., et al. (2007). Expertise modulates the perception of pain in others. *Curr. Biol.* 17, 1708–1713. doi: 10.1016/j.cub.2007.09.020

Cheng, Y., Yang, C.-Y., Lin, C.-P., Lee, P.-L., and Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: a magnetoencephalography study. *Neuroimage* 40, 1833–1840.

Cooper, J. C., and Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *Neuroimage* 39, 538–547.

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324.

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Craig, A. D. B. (2005). Forebrain emotional asymmetry: a neuroanatomical basis. *Trends Cogn. Sci.* 9, 566–571. doi: 10.1016/j.tics.2005.10.005

Craig, A. D. B. (2009). How do you feel–now. The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Critchley, H. D. (2009). Psychophysiology of neural, cognitive and affective integration: fMRI and autonomic indicants. *Int. J. Psychophysiol.* 73, 88–94. doi: 10.1016/j.ijpsycho.2009.01.012

Damasio, A., and Damásio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain. Penguin Books*. New York, NY: G. P. Putnam's sons.

Decety, J., and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist* 13, 580–593. doi: 10.1177/1073858407304654

Decety, J., and Michalska, K. J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Dev. Sci.* 13, 886–899.

Decety, J., Michalska, K. J., Akitsuki, Y., and Lahey, B. B. (2009). Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. *Biol. Psychol.* 80, 203–211. doi: 10.1016/j.biopsycho.2008.09.004

Decety, J., Michalska, K. J., and Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cereb. Cortex* 22, 209–220. doi: 10.1093/cercor/bhr111

Decety, J., and Porges, E. C. (2011). Imagining being the agent of actions that carry different moral consequences: an fMRI study. *Neuropsychologia* 49, 2994–3001. doi: 10.1016/j.neuropsychologia.2011.06.024

Drevets, W. C., Price, J. L., Simpson, J. R., Todd, R. D., Reich, T., Vannier, M., et al. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 386, 824–827.

Epstein, R. A. (2011). Cognitive neuroscience: scene layout from vision and touch. *Curr. Biol.* 21, R437–R438. doi: 10.1016/j.cub.2011.04.037

Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding. *Neuron* 23, 115–125.

Garcia-Larrea, L. (2012). The posterior insular-opercular region and the search of a primary cortex for pain. *Clin. Neurophysiol.* 42, 299–313. doi: 10.1016/j.neucli.2012.06.001

Gaspar, P., Berger, B., Febvret, A., Vigny, A., and Henry, J. P. (1989). Catecholamine innervation of the human cerebral cortex as revealed by comparative immunohistochemistry of tyrosine hydroxylase and dopamine-beta-hydroxylase. *J. Comp. Neurol.* 279, 249–271. doi: 10.1002/cne.902790208

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.

Gläscher, J. (2009). Visualization of group inference data in functional neuroimaging. *Neuroinformatics* 7, 73–82. doi: 10.1007/s12021-008-9042-x

Green, S., Lambon Ralph, M. A., Moll, J., Deakin, J. F. W., and Zahn, R. (2012). Guilt-selective functional disconnection of anterior temporal and subgenual cortices in major depressive disorder. *Arch. Gen. Psychiatry* 69, 1014–1021.

Gregory, S., and Osborne, J. (2009). White vs. Fedor: Ultimate fighting's cold war gets hotter - TIME. *Time Magazine*. Available online at: http://www.time.com/time/arts/article/0,8599,1909703,00.html

Grosbras, M.-H., Beaton, S., and Eickhoff, S. B. (2012). Brain regions involved in human movement perception: a quantitative voxel-based meta-analysis. *Hum. Brain Mapp.* 33, 431–454. doi: 10.1002/hbm.21222

Gu, X., and Han, S. (2007). Attention and reality constraints on the neural processes of empathy for pain. *Neuroimage* 36, 256–267.

Indovina, I., Maffei, V., Pauwels, K., Macaluso, E., Orban, G. A., and Lacquaniti, F. (2013). Simulated self-motion in a visual gravity field: sensitivity to vertical and horizontal heading in the human brain. *Neuroimage* 71, 114–124. doi: 10.1016/j.neuroimage.2013.01.005

Insel, T. R., and Young, L. J. (2001). The neurobiology of attachment. *Nat. Rev. Neurosci.* 2, 129–136. doi: 10.1038/35053579

Jackson, P. L., Brunet, E., Meltzoff, A. N., and Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* 44, 752–761. doi: 10.1016/j.neuropsychologia.2005.07.015

Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005a). How do we perceive the pain of others. a window into the neural processes involved in empathy*Neuroimage* 24, 771–779.

Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005b). How do we perceive the pain of others. A window into the neural processes involved in empathy. *Neuroimage* 24, 771–779.

Jensen, J., McIntosh, A. R., Crawley, A. P., Mikulis, D. J., Remington, G., and Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257.

Lamm, C., Batson, C. D., and Decety, J. (2007a). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58.

Lamm, C., Nusbaum, H. C., Meltzoff, A. N., and Decety, J. (2007b). What are you feeling. Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE* 2:e1292. doi: 10.1371/journal.pone.0001292

Lamm, C., and Decety, J. (2008). Is the extrastriate body area (EBA) sensitive to the perception of pain in others? *Cereb. Cortex* 18, 2369–2373. doi: 10.1093/cercor/bhn006

Lamm, C., Decety, J., and Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* 54, 2492–2502. doi: 10.1016/j.neuroimage.2010.10.014

Lamm, C., Porges, E. C., Cacioppo, J. T., and Decety, J. (2008). Perspective taking is associated with specific facial responses during empathy for pain. *Brain Res.* 1227, 153–161.

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1997). *International Affective Picture System (IAPS): Technical manual and affective ratings*. NIMH Center for the Study of Emotion and Attention.

Leknes, S., and Tracey, I. (2008). A common neurobiology for pain and pleasure. *Nat. Rev. Neurosci.* 9, 314–320. doi: 10.1038/nrn2333

Luijten, M., Veltman, D. J., Hester, R., Smits, M., Pepplinkhuizen, L., and Franken, I. H. A. (2012). Brain activation associated with attentional bias in smokers is modulated by a dopamine antagonist. *Neuropsychopharmacology* 37, 2772–2779. doi: 10.1038/npp.2012.143

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., and Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19, 1233–1239.

Mathiak, K., and Weber, R. (2006). Toward brain correlates of natural behavior: fMRI during violent video-games. *Hum. Brain Mapp.* 27, 948–956. doi: 10.1002/hbm.20234

Meyer, M. L., Spunt, R. P., Berkman, E. T., Taylor, S. E., and Lieberman, M. D. (2012). Evidence for social working memory from a parametric functional MRI study. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1883–1888. doi: 10.1073/pnas.1121077109

Michalska, K. J., Kinzler, K. D., and Decety, J. (2013). Age-related sex differences in explicit measures of empathy do not predict brain responses across childhood and adolescence. *Dev. Cogn. Neurosci.* 3, 22–32. doi: 10.1016/j.dcn.2012.08.001

Moll, J., De Oliveira-Souza, R., Garrido, G. J., Bramati, I. E., Caparelli-Daquer, E. M. A., Paiva, M. L. M. F., et al. (2007). The self as a moral agent: linking the neural bases of social agency and moral sensitivity. *Soc. Neurosci.* 2, 336–352. doi: 10.1080/17470910701392024

Moll, J., Krueger, F., Zahn, R., Pardini, M., De Oliveira-Souza, R., and Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15623–15628. doi: 10.1073/pnas.0604475103

Moriguchi, Y., Decety, J., Ohnishi, T., Maeda, M., Mori, T., Nemoto, K., et al. (2007). Empathy and judging other's pain: an fMRI study of alexithymia. *Cereb. Cortex* 17, 2223–2234.

Morrison, I., Lloyd, D., Di Pellegrino, G., and Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue. *Cogn. Affect. Behav. Neurosci.* 4, 270–278. doi: 10.3758/CABN.4.2.270

Murray, J. P., Liotti, M., Ingmundson, P. T., Mayberg, H. S., Pu, Y., Zamarripa, F., et al. (2006a). Children's brain activations while viewing televised violence revealed by fMRI. *Media Psychol.* 8, 25–37.

Ogino, Y., Nemoto, H., Inui, K., Saito, S., Kakigi, R., and Goto, F. (2007). Inner experience of pain: imagination of pain while viewing images showing painful events forms subjective pain representation in human brain. *Cereb. Cortex* 17, 1139–1146.

Olson, I. R., Plotzker, A., and Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* 130, 1718–1731.

Ressler, K. J., and Mayberg, H. S. (2007). Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nat. Neurosci.* 10, 1116–1124. doi: 10.1038/nn1944

Reynolds, S. M., and Berridge, K. C. (2002). Positive and negative motivation in nucleus accumbens shell: bivalent rostrocaudal gradients for GABA-elicited eating, taste "liking"/"disliking" reactions, place preference/avoidance, and fear. *J. Neurosci.* 22, 7308–7320.

Reynolds, S. M., and Berridge, K. C. (2008). Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nat. Neurosci.* 11, 423–425. doi: 10.1038/nn2061

Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., and McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cereb. Cortex* 13, 308–317.

Schechter, H. (2005). *Savage Pastimes?: a Cultural History of Violent Entertainment*. *1st CN* -. New York, NY: St. Martin's Press.

Schultheiss, O. C., and Schiepe-Tiska, A. (2013). The role of the dorsoanterior striatum in implicit motivation: the case of the need for power. *Front. Hum. Neurosci.* 7:141. doi: 10.13389/fnhum.2013.00141

Servos, P. (2002). The neural substrates of biological motion perception: an fMRI study. *Cereb. Cortex* 12, 772–782. doi: 10.1093/cercor/12.7.772

Smith, K. S., Berridge, K. C., and Aldridge, J. W. (2011). Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. *Proc. Natl. Acad. Sci. U.S.A.* 108, E255–E264.

Whitehead, N. L. (2004). *Violence. 1st CN* -. Santa Fe, NM: School of American Research Press.

Whitfield-Gabrieli, S., and Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* 2, 125–141. doi: 10.1089/brain.2012.0073

Wood, R., Gallese, V., and Cattaneo, L. (2010). Visuotactile empathy within the primary somatosensory cortex revealed by short-latency afferent inhibition. *Neurosci. Lett.* 473, 28–31. doi: 10.1016/j.neulet.2010.02.012

Zahn, R., De Oliveira-Souza, R., Bramati, I., Garrido, G., and Moll, J. (2009). Subgenual cingulate activity reflects individual differences in empathic concern. *Neurosci. Lett.* 457, 107–110. doi: 10.1016/j.neulet.2009.03.090

Zaki, J., Ochsner, K. N., Hanelin, J., Wager, T. D., and Mackey, S. C. (2007). Different circuits for different pain: patterns of functional connectivity reveal distinct networks for processing pain in self and others. *Soc. Neurosci.* 2, 276–291. doi: 10.1080/17470910701401973

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 05 April 2013; accepted: 19 July 2013; published online: 13 August 2013.*

*Citation: Porges EC and Decety J (2013) Violence as a source of pleasure or displeasure is associated with specific functional connectivity with the nucleus accumbens. Front. Hum. Neurosci. 7:447. doi: 10.3389/fnhum.2013.00447*

frontiers in
HUMAN NEUROSCIENCE

OPINION ARTICLE
published: 18 June 2013
doi: 10.3389/fnhum.2013.00287

# Fluid intelligence, social cognition, and perspective changing abilities as pointers of psychosocial adaptation

**David Huepe**[1]* **and Natalia Salas**[2]

[1] *Laboratory of Cognitive and Social Neuroscience (LaNCyS), UDP-INECO Foundation Core on Neuroscience (UIFCoN), Universidad Diego Portales, Santiago de Chile*
[2] *Faculty of Education, Cognitive Development Center, Universidad Diego Portales, Santiago de Chile*
**Correspondence: david.huepe@mail.udp.cl*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*Cecilia I. Calero, Universidad de Buenos Aires, Argentina*

The prefrontal cortex in human brain is the main area that is related to the capacity for establishing relationships to others and with your environment. It is also the core for the development of superior mental functions such as plan and motor outcome, cognitive, affective, and social behavior across time (Kolb et al., 2012). Some cognitive functions related to this lobule include fluid intelligence (FI), social cognition (SC), and perspective changing abilities (PCA), which are necessary for adaptation to social contexts and solving problems in new situations (Barkley, 2001; Crisp and Meleady, 2012). These abilities, in turn, appear to be dependent on contextual keys, thus requiring flexibility, which is associated with frontal lobe functioning (Nestor et al., 2013; Pfeifer and Peake, 2012), particularly, in the case of PCA some areas are specifically related to the prefrontal cortex such as Brodmann area 10 (BA10) (Buckner and Carroll, 2007). In this opinion paper, we propose a model that integrates these components (FI, SC, and PCA) as indicators of psychosocial adaptation (PSA) in contexts of social vulnerability or diminished social/cultural conditions, in contrast to contributions in neurosciences made from evidence of patients with brain damage or psychiatric disorders.

PSA, defined as the quality of social life and subjective well-being of an individual in context (Bishop et al., 2008; Cox et al., 2010), is relevant for proper development. Research suggests that the prefrontal cortex plays a major role in adaptation, given its involvement in behavioral flexibility, executive functions, FI, and SC (Van Horn et al., 2012; Waters-Wood et al., 2012). Similarly, SC tasks, FI (Duncan et al., 1995; Roca et al., 2010) and cognitive flexibility have been associated with this area (Shamay-Tsoory et al., 2009; Larquet et al., 2010). Damage or alterations in the frontal lobe, have a direct impact on these functions, mainly resulting in maladaptive behaviors. Likewise, deficits in maturation or development of the cortex have been associated with social behavior disorders (Schore, 2000; Kolb et al., 2012).

FI has been defined as the ability to think logically and to solve problems in new situations, regardless of the acquisition of knowledge (Cattell, 1963). This reflects the ability to reason and to think abstractly in contrast to what is called crystallized intelligence (Cattell, 1967), which depends on cultural and academic learning. From a neuroanatomical viewpoint, FI has been associated with frontal lobe functions (Duncan et al., 1995). Injuries in this area affect the performance of these cognitive abilities (Roca et al., 2010; Woolgar et al., 2010). Besides, neuroimaging studies of FI have shown activation of frontal areas (Duncan et al., 2000; Bishop et al., 2008). There is consensus that the frontal lobe represents the neural basis of FI, but its association with cognitive flexibility and social behavior has not been studied until fairly recently. Support for the association between FI and cognitive flexibility and social behavior also comes from studies on frontal lobe lesions (Hynes et al., 2006; Shamay-Tsoory et al., 2009; Larquet et al., 2010).

In brief, the relationship between FI, PCA and dysfunction of executive functions, has been extensively studied (Bechara et al., 2000; Duncan, 2010; Dumontheil et al., 2011). In addition, the link between FI and abstract reasoning (Bunting, 2006; Perfetti et al., 2009) has also been established. However, research that associates this set of variables with social behavior in contexts of interaction with non-pathological samples, are very scarce, even more when it comes to PSA (Roca et al., 2010; Huepe et al., 2011).

The BA10 area could have a role in the ability of self-projection and in PCA, which are needed for other social skills (Buckner and Carroll, 2007). This area plays a central role in the so-called "default network" (Default Mode Network-DMN-) (Buckner and Vincent, 2007; Chen et al., 2008; Ko et al., 2011) and there is now abundant evidence that the DMN has an atypical configuration in subjects exhibiting some form of mental or psychiatric disorder such as attention deficit-hyperactivity disorder-ADHD, depression, Alzheimer's disease, schizophrenia, bipolar disorder, autism spectrum disorders (ASD), among others (Broyd et al., 2009; Minshew and Keller, 2010; Ongür et al., 2010; Pomarol-Clotet et al., 2011; Meda et al., 2012). Several of these disorders represent extreme points within an adaptation continuum. It is therefore possible to hypothesize that individuals belonging to vulnerable contexts, and exhibiting maladaptive behaviors, could present affectation of prefrontal cortex's (FI, SC, and PCA) main functions. In effect, the literature shows that people with problems associated to impulse control, violent behavior, decision taking, morality, empathy, FI, among others, are related with malfunction of prefrontal areas (Raine, 2002; Bechara and Van Der Linden, 2005; Seitz et al., 2006). Similarly, but more radically, we can see similar behaviors in patients with brain injuries in prefrontal cortex. In short, PCA and FI could be linked to aspects of SC such as theory of mind (ToM), handling multiple tasks (multitasking) and frontal functions (Torralva et al., 2007, 2009). Yet, the question still remains whether individuals belonging to vulnerable contexts—with maladaptive behaviors—would show decline of main prefrontal cortex functions and whether the three cognitive components discussed here can predict PSA in this population.

## A PRELIMINARY MODEL OF FI, SC, AND PCA AS PREDICTORS OF PSA

PSA includes multiple dimensions such as social behavior, emotional regulation, and the development of social habits (Bishop et al., 2005; Cox et al., 2010). Psychosocial functioning represents an ecological approach to everyday adaptation as well as a theoretical approach which integrates cognition and emotion (Wilson, 2008). We propose that variables such as FI, SC, and PCA would influence PSA processes, based on the evidence from injured patients, briefly summarized here. We would also expect that PCA will be altered given their link to complex cognitive functions and high-level cognitive skills such as mental state attribution, empathy, and understanding of social contextual cues (Buckner and Carroll, 2007; Ibañez and Manes, 2012). This last issue could be expected from evidence of researches about the role BA10 plays in the understanding of contextual cues. Patients with damages in this area have exhibited significant difficulties to read correctly certain social meanings such as the ability to infer feelings, thoughts and other complex set of functions associated to ToM (Gilbert et al., 2006, 2007).

**Figure 1** illustrates the relationships among FI, SC, and PCA as a framework to assess PSA individual differences. Arrows indicate the direction of the relationship: FI, SC, and PCA are explained by the prefrontal activation and maturation. In turn, FI and SC predict PSA. Finally, PCA shows an indirect effect mediated by the effect of SC on PSA.

We hypothesize that the degree of PSA that people show in vulnerable social contexts would be partly explained by the level of performance exhibited on FI and SC tasks. SC, in turn, would depend on PCA levels. Previous research supports each of these relationships separately (Gilbert et al., 2006, 2007; Torralva et al., 2007, 2009; Roca et al., 2010, 2011; Huepe et al., 2011). However, how these relationships could predict PSA is a matter of further research.

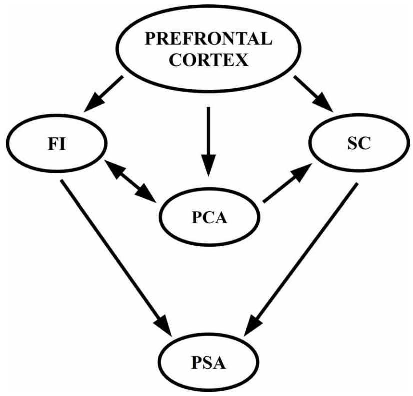


**FIGURE 1 | Tentative model of a framework for the relationship between FI, PCA, SC, and PSA.** Prefrontal cortex, FI, and SC (which includes PCA) predict PSA.

This framework would help to design empirical models of individual differences of these variables in vulnerable social contexts. Preliminary evidence supports this assertion. For instance, children under vulnerability with better social adaptation have high levels of FI (Huepe et al., 2011). It is also known that people under unfavorable living conditions, have a better PSA when they have high level cognitive skills, facilitating a better social adaptation (Flores et al., 2005; Cicchetti and Blender, 2006).

According to this tentative model, we specifically expect to find that higher levels of FI are associated with PSA. Moreover, a good performance on PCA and SC, positively linked to PSA, is also expected. Indirect evidence from Roca et al. (2011) has suggested that FI predicts the performance on different tests (executive functions and SC) in a group of frontal patients with lesions in BA10 (associated with PCA) when compared to a group of frontal patients (non-BA10), and a group of healthy controls. This evidence shows the importance of BA10 in SC. Patients with lesions in BA10 exhibited lower performance in SC tests and not in ecologically executive functions. Hence, it can be established that these components would be networks that, anatomically and neurally, to some extent are independent, although data is not conclusive on the extent of the kind of association among these components. Thus, we suggest a particular pattern of effects of these variables in our tentative model.

Complex modern societies demand a strong capacity for social adaptation. Bullying and violence, addictive behavior, mental health impairments, and other social behaviors are strongly linked to quality of life. Current agenda includes discovering the processes whereby individuals at high risk do not develop maladaptive or pathological behaviors. Resilience, defined as the accomplishment of competences regardless of significant adversity (Cicchetti and Blender, 2006), could be an example of the implications of social adaptation. Therefore, the framework presented would help to better clarify the psychosocial factors related to how resilience works, by evaluating cognitive protective factors that could be contributing to social adaptation in vulnerable populations.

Our model proposes a straightforward association between levels of FI, PCA, and SC regarding PSA. To prove that certain cognitive abilities have a major impact on PSA is crucial both from a scientific and a political point of view, especially in countries where inequality limits the opportunities of development and academic performance (Lutha and Cicchetti, 2000; Salas et al., 2010). The influence of cognitive functions related to social adaptation may represent a contribution by orienting changes in policy regarding the possibilities of development and intervention. This would be especially relevant in populations suffering from poverty, drug abuse, violence, among others, offering a background to design cognitive interventions that are socially and contextually focused.

In brief, assessing the effects of FI, PCA, and SC would be crucial for understanding the different levels of PSA in vulnerable contexts. This kind of studies would favor a multiple-level-analysis viewpoint in order to design and evaluate interventions that aim at recognizing outcomes related to resilience, cognitive changes, and social adaptation, in persons facing significant adversity (Lutha and Cicchetti, 2000). Neurocognitive markers would help to enlighten the impact of cognitive functions on SC (Ibáñez et al., 2009; Ibañez et al., 2012). Further empirical development of this framework would promote possible future forms of social intervention based on the theoretical and empirical co-construction of tools provided by social neuroscience, neuropsychology, and social psychology.

## ACKNOWLEDGMENTS

This work was supported by grant FONDECYT (1130920).

## REFERENCES

Barkley, R. A. (2001). The executive functions and self-regulation: an evolutionary neuropsychological perspective. *Neuropsychol. Rev.* 11, 1–29. doi: 10.1023/A:1009085417776

Bechara, A., Tranel, D., and Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 123(Pt 11), 2189–2202. doi: 10.1093/brain/123.11.2189

Bechara, A., and Van Der Linden, M. (2005). Decision-making and impulse control after frontal lobe injuries. *Curr. Opin. Neurol.* 18, 734–739. doi: 10.1097/01.wco.0000194141.56429.3c

Bishop, A. J., Marteau, T. M., Hall, S., Kitchener, H., and Hajek, P. (2005). Increasing women's intentions to stop smoking following an abnormal cervical smear test result. *Prev. Med.* 41, 179–185. doi: 10.1016/j.ypmed.2004.09.046

Bishop, S. J., Fossella, J., Croucher, C. J., and Duncan, J. (2008). COMT val158met genotype affects recruitment of neural mechanisms supporting fluid intelligence. *Cereb. Cortex* 18, 2132–2140. doi: 10.1093/cercor/bhm240

Broyd, S. J., Demanuele, C., Debener, S., Helps, S. K., James, C. J., and Sonuga-Barke, E. J. S. (2009). Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci. Biobehav. Rev.* 33, 279–296. doi: 10.1016/j.neubiorev.2008.09.002

Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004

Buckner, R. L., and Vincent, J. L. (2007). Unrest at rest: default activity and spontaneous network correlations. *Neuroimage* 37, 1091–1096. Discussion: 1097–1099. doi: 10.1016/j.neuroimage.2007.01.010

Bunting, M. (2006). Proactive interference and item similarity in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 183–196. doi: 10.1037/0278-7393.32.2.183

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychol.* 54, 1–22. doi: 10.1037/h0046743

Cattell, R. B. (1967). The theory of fluid and crystallized general intelligence checked at the 5-6 year-old level. *Br. J. Educ. Psychol.* 37, 209–224. doi: 10.1111/j.2044-8279.1967.tb01930.x

Chen, A. C. N., Feng, W., Zhao, H., Yin, Y., and Wang, P. (2008). EEG default mode network in the human brain: spectral regional field powers. *Neuroimage* 41, 561–574. doi: 10.1016/j.neuroimage.2007.12.064

Cicchetti, D., and Blender, J. A. (2006). A multiple-levels-of-analysis perspective on resilience: implications for the developing brain, neural plasticity, and preventive interventions. *Ann. N.Y. Acad. Sci.* 1094, 248–258. doi: 10.1196/annals.1376.029

Cox, K. S., Wilt, J., Olson, B., and McAdams, D. P. (2010). Generativity, the big five, and psychosocial adaptation in midlife adults. *J. Pers.* 78, 1185–1208. doi: 10.1111/j.1467-6494.2010.00647.x

Crisp, R. J., and Meleady, R. (2012). Adapting to a multicultural future. *Science* 336, 853–855. doi: 10.1126/science.1219009

Dumontheil, I., Thompson, R., and Duncan, J. (2011). Assembly and use of new task rules in fronto-parietal cortex. *J. Cogn. Neurosci.* 23, 168–182. doi: 10.1162/jocn.2010.21439

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* 14, 172–179. doi: 10.1016/j.tics.2010.01.004

Duncan, J., Burgess, P., and Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia* 33, 261–268. doi: 10.1016/0028-3932(94)00124-8

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science* 289, 457–460. doi: 10.1126/science.289.5478.457

Flores, E., Cicchetti, D., and Rogosch, F. A. (2005). Predictors of resilience in maltreated and nonmaltreated Latino children. *Dev. Psychol.* 41, 338–351. doi: 10.1037/0012-1649.41.2.338

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., et al. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J. Cogn. Neurosci.* 18, 932–948. doi: 10.1162/jocn.2006.18.6.932

Gilbert, S. J., Williamson, I. D. M., Dumontheil, I., Simons, J. S., Frith, C. D., and Burgess, P. W. (2007). Distinct regions of medial rostral prefrontal cortex supporting social and nonsocial functions. *Soc. Cogn. Affect. Neurosci.* 2, 217–226. doi: 10.1093/scan/nsm014

Huepe, D., Roca, M., Salas, N., Canales-Johnson, A., Rivera-Rei, Á. A., Zamorano, L., et al. (2011). Fluid intelligence and psychosocial outcome: from logical problem solving to social adaptation. *PLoS ONE* 6:e24858. doi: 10.1371/journal.pone.0024858

Hynes, C. A., Baird, A. A., and Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia* 44, 374–383. doi: 10.1016/j.neuropsychologia.2005.06.011

Ibáñez, A., Haye, A., González, R., Hurtado, E., and Henríquez, R. (2009). Multi-level analysis of cultural phenomena: the role of ERPs approach to prejudice. *J. Theory Soc. Behav.* 39, 81–110. doi: 10.1111/j.1468-5914.2008.00391.x

Ibañez, A., Melloni, M., Huepe, D., Helgiu, E., Rivera-Rei, A., Canales-Johnson, A., et al. (2012). What event-related potentials (ERPs) bring to social neuroscience? *Soc. Neurosci.* 7, 632–649. doi: 10.1080/17470919.2012.691078

Ibañez, A., and Manes, F. (2012). Contextual social cognition and the behavioral variant of frontotemporal dementia. *Neurology* 78, 1354–1362. doi: 10.1212/WNL.0b013e3182518375

Ko, A. L., Darvas, F., Poliakov, A., Ojemann, J., and Sorensen, L. B. (2011). Quasi-periodic fluctuations in default mode network electrophysiology. *J. Neurosci.* 31, 11728–11732. doi: 10.1523/JNEUROSCI.5730-10.2011

Kolb, B., Mychasiuk, R., Muhammad, A., Li, Y., Frost, D. O., and Gibb, R. (2012). Experience and the developing prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109(Suppl. 2), 17186–17193. doi: 10.1073/pnas.1121251109

Larquet, M., Coricelli, G., Opolczynski, G., and Thibaut, F. (2010). Impaired decision making in schizophrenia and orbitofrontal cortex lesion patients. *Schizophr. Res.* 116, 266–273. doi: 10.1016/j.schres.2009.11.010

Lutha, S. S., and Cicchetti, D. (2000). The construct of resilience: implications for interventions and social policies. *Dev. Psychopathol.* 12, 857–885. doi: 10.1017/S0954579400004156

Meda, S. A., Gill, A., Stevens, M. C., Lorenzoni, R. P., Glahn, D. C., Calhoun, V. D., et al. (2012). Differences in resting-state functional magnetic resonance imaging functional network connectivity between schizophrenia and psychotic bipolar probands and their unaffected first-degree relatives. *Biol. Psychiatry* 71, 881–889. doi: 10.1016/j.biopsych.2012.01.025

Minshew, N. J., and Keller, T. A. (2010). The nature of brain dysfunction in autism: functional brain imaging studies. *Curr. Opin. Neurol.* 23, 124–130. doi: 10.1097/WCO.0b013e32833782d4

Nestor, P. G., Nakamura, M., Niznikiewicz, M., Thompson, E., Levitt, J. J., Choate, V., et al. (2013). In search of the functional neuroanatomy of sociality: MRI subdivisions of orbital frontal cortex and social cognition. *Soc. Cogn. Affect. Neurosci.* 8, 460–467. doi: 10.1093/scan/nss018

Ongür, D., Lundy, M., Greenhouse, I., Shinn, A. K., Menon, V., Cohen, B. M., et al. (2010). Default mode network abnormalities in bipolar disorder and schizophrenia. *Psychiatry Res.* 183, 59–68. doi: 10.1016/j.pscychresns.2010.04.008

Perfetti, B., Saggino, A., Ferretti, A., Caulo, M., Romani, G. L., and Onofrj, M. (2009). Differential patterns of cortical activation as a function of fluid reasoning complexity. *Hum. Brain Mapp.* 30, 497–510. doi: 10.1002/hbm.20519

Pfeifer, J. H., and Peake, S. J. (2012). Self-development: integrating cognitive, socioemotional, and neuroimaging perspectives. *Dev. Cogn. Neurosci.* 2, 55–69. doi: 10.1016/j.dcn.2011.07.012

Pomarol-Clotet, E., Moro, N., Sarró, S., Goikolea, J. M., Vieta, E., Amann, B., et al. (2011). Failure of de-activation in the medial frontal cortex in mania: evidence for default mode network dysfunction in the disorder. *World J. Biol. Psychiatry* 13, 616–626. doi: 10.3109/15622975.2011.573808

Raine, A. (2002). Annotation: the role of prefrontal deficits, low autonomic arousal, and early health factors in the development of antisocial and aggressive behavior in children. *J. Child Psychol. Psychiatry* 43, 417–434. doi: 10.1111/1469-7610.00034

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., et al. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain* 133, 234–247. doi: 10.1093/brain/awp269

Roca, M., Torralva, T., Gleichgerrcht, E., Woolgar, A., Thompson, R., Duncan, J., et al. (2011).

The role of Area 10 (BA10) in human multitasking and in soc cognition: a lesion study. *Neuropsychologia* 49, 3525–3531. doi: 10.1016/j.neuropsychologia.2011.09.003

Salas, N., Assael, C., Huepe, D., Pérez, T., González, F., Morales, A., et al. (2010). Application of IE-basic program to promote cognitive and affective development in preschoolers: a chilean study. *J. Cogn. Educ. Psychol.* 9, 285–297. doi: 10.1891/1945-8959.9.3.285

Schore, A. N. (2000). Attachment and the regulation of the right brain. *Attach. Hum. Dev.* 2, 23–47. doi: 10.1080/146167300361309

Seitz, R. J., Nickel, J., and Azari, N. P. (2006). Functional modularity of the medial prefrontal cortex: involvement in human empathy. *Neuropsychology* 20, 743–751. doi: 10.1037/0894-4105.20.6.743

Shamay-Tsoory, S. G., Aharon-Peretz, J., and Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. doi: 10.1093/brain/awn279

Torralva, T., Kipps, C. M., Hodges, J. R., Clark, L., Bekinschtein, T., Roca, M., et al. (2007). The relationship between affective decision-making and theory of mind in the frontal variant of frontotemporal dementia. *Neuropsychologia* 45, 342–349. doi: 10.1016/j.neuropsychologia.2006.05.031

Torralva, T., Roca, M., Gleichgerrcht, E., Bekinschtein, T., and Manes, F. (2009). A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. *Brain* 132, 1299–1309. doi: 10.1093/brain/awp041

Van Horn, J. D., Irimia, A., Torgerson, C. M., Chambers, M. C., Kikinis, R., and Toga, A. W. (2012). Mapping connectivity damage in the case of phineas gage. *PLoS ONE* 7:e37454. doi: 10.1371/journal.pone.0037454

Waters-Wood, S. M., Xiao, L., Denburg, N. L., Hernandez, M., and Bechara, A. (2012). Failure to learn from repeated mistakes: persistent decision-making impairment as measured by the iowa gambling task in patients with ventromedial prefrontal cortex lesions. *J. Int. Neuropsychol. Soc.* 18, 927–930. doi: 10.1017/S135561771200063X

Wilson, B. A. (2008). Neuropsychological rehabilitation. *Annu. Rev. Clin. Psychol.* 4, 141–162. doi: 10.1146/annurev.clinpsy.4.022007.141212

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., et al. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14899–14902. doi: 10.1073/pnas.1007928107

*Received: 16 April 2013; accepted: 31 May 2013; published online: 18 June 2013.*

*Citation: Huepe D and Salas N (2013) Fluid intelligence, social cognition, and perspective changing abilities as pointers of psychosocial adaptation. Front. Hum. Neurosci.* ***7****:287. doi: 10.3389/fnhum.2013.00287*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 26 September 2013
doi: 10.3389/fnhum.2013.00551

# The implicit processing of categorical and dimensional strategies: an fMRI study of facial emotion perception

***Yoshi-Taka Matsuda*[1,2,3] **, *Tomomi Fujimura*[1,2], *Kentaro Katahira*[1,2,4], *Masato Okada*[1,5], *Kenichi Ueno*[6], *Kang Cheng*[6] and *Kazuo Okanoya*[1,2,4] ****

[1] *Okanoya Emotional Information Project, Exploratory Research for Advanced Technology (ERATO), Japan Science and Technology Agency (JST), Wako, Japan*
[2] *Emotional Information Joint Research Laboratory, RIKEN Brain Science Institute, Saitama, Japan*
[3] *Center for Baby Science, Doshisha University, Kyoto, Japan*
[4] *Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan*
[5] *Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan*
[6] *Support Unit for Functional MRI, RIKEN Brain Science Institute, Wako, Japan*

***Edited by:***
*Maria Ruz, Universidad de Granada, Spain*

***Reviewed by:***
*Adam K. Anderson, University of Toronto, Canada*
*Andrew C. Connolly, Dartmouth College, USA*

****Correspondence:***
*Yoshi-Taka Matsuda and Kazuo Okanoya. Okanoya Emotional Information Project, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
*e-mail: matsuda@brain.riken.jp; cokanoya@mail.ecc.u-tokyo.ac.jp*

Our understanding of facial emotion perception has been dominated by two seemingly opposing theories: the categorical and dimensional theories. However, we have recently demonstrated that hybrid processing involving both categorical and dimensional perception can be induced in an implicit manner (Fujimura et al., 2012). The underlying neural mechanisms of this hybrid processing remain unknown. In this study, we tested the hypothesis that separate neural loci might intrinsically encode categorical and dimensional processing functions that serve as a basis for hybrid processing. We used functional magnetic resonance imaging to measure neural correlates while subjects passively viewed emotional faces and performed tasks that were unrelated to facial emotion processing. Activity in the right fusiform face area (FFA) increased in response to psychologically obvious emotions and decreased in response to ambiguous expressions, demonstrating the role of the FFA in categorical processing. The amygdala, insula and medial prefrontal cortex exhibited evidence of dimensional (linear) processing that correlated with physical changes in the emotional face stimuli. The occipital face area and superior temporal sulcus did not respond to these changes in the presented stimuli. Our results indicated that distinct neural loci process the physical and psychological aspects of facial emotion perception in a region-specific and implicit manner.

**Keywords: individual differences, fMRI, facial expressions, categorical processes, implicit**

## INTRODUCTION

Two seemingly opposing theories dominate the field of facial emotion perception: the categorical theory and the dimensional theory (**Figure 1**). The categorical theory posits the existence of six basic, distinct and universal emotions (Ekman et al., 1969; Ekman and Friesen, 1971, 1976; Ekman, 1992): happiness, anger, sadness, surprise, disgust, and fear (Tomkins and McCarter, 1964; Ekman and Friesen, 1971; Johnson-Laird and Oatley, 1992). The dimensional theory posits the existence of two fundamental dimensions of emotional space: valence and arousal (Russell, 1980; Russell and Bullock, 1985). Valence represents hedonic tone or position on a pleasantness–unpleasantness continuum, whereas arousal (Russell, 1980; Russell and Bullock, 1985) or tension (Schlosberg, 1954) refers to the level of energy of an affective experience. Evidence has been accumulated to support the occurrence of both categorical (Etcoff and Magee, 1992; Calder et al., 1996; Young et al., 1997) and dimensional (Katsikitis, 1997; Takehara and Suzuki, 1997, 2001) processing during facial emotion perception. In response to this apparent conflict, certain investigators have proposed a hybrid theory (Russell, 2003; Christie and Friedman, 2004; Panayiotou, 2008; cf. Rosch and Mervis, 1975) in which categorical and dimensional perception operate simultaneously to decode facial expressions. However, during the determination of facial emotions, it remains unknown which of these types of perception is relatively dominant and how these types of perceptions interact.

In a recent study (Fujimura et al., 2012), we used a stimulus set of morphed continua and two different sets of experimental instructions (detailing categorical vs. dimensional strategies) to investigate the relative dominance of categorical or dimensional perception during the decoding of facial expressions. Category boundaries were observed for not only identification and discrimination tasks, which require a categorical strategy, but also the Affect Grid task (Russell et al., 1989), which requires a dimensional strategy. Thus, our results indicate that categorical processing is dominant relative to dimensional processing and that categorical processing can be induced either explicitly or implicitly during facial emotion perception. Interestingly, this dominance of categorical processing was observed in the valence dimension but not the arousal dimension (Fujimura et al., 2012, Exp. 1). Despite this psychological evidence for the dominance of categorical processing, the neural mechanisms underlying this phenomenon remain unclear.

Facial emotion perception has been most frequently studied through the use of neuroimaging as a component of a general

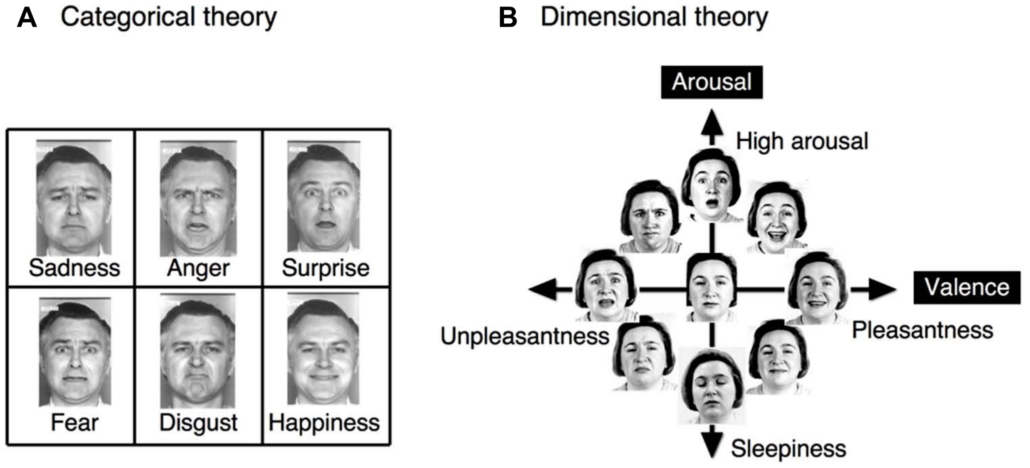


**FIGURE 1 | The two different theories regarding facial emotion perception. (A)** The categorical theory proposes the existence of six basic, distinct and universal emotions, with clear boundaries separating emotional states. **(B)** The dimensional theory proposes the existence of the two fundamental dimensions of valence and arousal, which form an emotional space. In the dimensional theory, each emotional state may be continuously represented as a linear combination of values in these two dimensions. The original faces were drawn from Ekman and Friesen (1976), and Russell (1997).

face perception study; this approach tends to indicate that a distributed neural system mediates both face processing and other types of cognitive processing (Sergent et al., 1992; Courtney et al., 1996; Haxby et al., 2000; Ishai et al., 2004, 2005; Calder and Young, 2005; Atkinson and Adolphs, 2011). The cortical network for face perception includes the inferior occipital gyrus (IOG) and lateral fusiform gyrus (FG; Kanwisher et al., 1997; Grill-Spector et al., 2004; Rotshtein et al., 2005); the superior temporal sulcus (STS; Puce et al., 1998; Hoffman and Haxby, 2000; Calder et al., 2007); the amygdala and insula (Breiter et al., 1996; Morris et al., 1996; Phillips et al., 1997; Vuilleumier et al., 2001; Ishai et al., 2004); the inferior frontal gyrus (IFG; Leveroni et al., 2000; Ishai et al., 2000); and the nucleus accumbens and orbitofrontal cortex (OFC; Aharon et al., 2001; O'Doherty et al., 2003; Kranz and Ishai, 2006). This distributed representation of face processing may reflect the fact that although the recognition of facial identity is based on invariant facial features, animated aspects of the face, such as speech-related movement and emotional expression, contribute to social communication (Ishai, 2008). This idea of a neural dichotomy between perceptions of facial identity and expressions within the "face network" represents an influential model in the field (Haxby et al., 2000); however, this proposal remains debated, as it has been demonstrated that certain brain regions are involved in perceiving both facial identities and expressions (Gorno-Tempini et al., 2001; Vuilleumier et al., 2001; Kaufmann and Schweinberger, 2004; Winston et al., 2004; Calder and Young, 2005; Ganel et al., 2005; Ishai et al., 2005; Fox and Barton, 2007; Palermo and Rhodes, 2007; Fox et al., 2008). However, to an extent, there exists consensus that specific neural processing mechanisms are associated with particular emotional facial expressions; for instance, disgust is processed in the insula, whereas fear is processed in the amygdala (Fusar-Poli et al., 2009). Three studies have utilized morphed faces to investigate the neural representation of categorical processing in facial emotion perception (Furl et al., 2007; Fox et al., 2009b; Harris et al., 2012); however, these studies have reported different neural bases for this phenomenon. Furl et al. (2007) used adaptation-induced after effects to identify the neural basis of categorical perception in the anterior temporal lobe. In contrast, Fox et al. (2009b) and Harris et al. (2012) employed a functional magnetic resonance imaging (fMRI) adaptation method to address this topic, although these fMRI studies produced inconsistent results. In particular, Fox et al. (2009b) found that the fusiform face area (FFA) and STS were responsive to the categorical differences in facial emotion continua, whereas Harris et al. (2012) demonstrated that categorical processing occurred in the amygdala. Similarly, studies of dimensional processing in facial emotion perception have provided inconsistent findings regarding the neural bases of this type of processing (Gerber et al., 2008; Fox et al., 2009b; Harris et al., 2012).

In this study, we sought to disentangle this question using our facial emotion stimulus sets, which were designed to simultaneously investigate not only categorical and dimensional processing but also valence and arousal, the two fundamental dimensions in emotional space (Fujimura et al., 2012). Furthermore, we sought to investigate which loci within the face network underlie the implicit categorical processing that we previously determined was dominant over dimensional processing in the valence dimension but not the arousal dimension.

We hypothesized that in emotional perception, categorical processing (i.e., the processing of psychological changes) and dimensional processing (i.e., the processing of physical changes) are encoded by separate neural loci of the face network. In particular, we hypothesized that there exists a functional dichotomy between the cortical and subcortical systems of the face network; in other words, this dichotomy exists between psychological processing in the cortical system and physical processing in the subcortical/limbic system. To test this hypothesis, we performed experiments using fMRI. Subjects were asked to perform irrelevant tasks to generate passive and implicit viewing of each face stimulus.

Each stimulus was randomly presented from two morphed continua, which were created using the physical features of emotional faces (in particular, the happiness–fear continuum and the anger–disgust continuum were employed to examine the valence and arousal dimensions, respectively). We also asked subjects to perform psychological experiments that we have previously described (Fujimura et al., 2012), allowing us to identify individual differences in the categorical boundaries of the morphed continua. Individual categorical boundaries were then utilized to realign the response curves obtained from each subject's blood oxygen level-dependent (BOLD) signal.

## MATERIALS AND METHODS

### STUDY PARTICIPANTS

A total of 22 Japanese adults participated in this study (12 males and 10 females; the mean ± SD of subjects' ages was 27.7 ± 5.3 years). All of the subjects were right-handed and neurologically normal. Each subject provided written informed consent in accordance with procedures approved by the RIKEN Brain Science Institute Ethics Committee and Functional MRI Safety and Ethics Committee (Wako, Japan). Data obtained from four additional subjects were excluded from the analysis due to excessive head movement (>1 mm) and/or the termination of the experiment at the subject's request.

### FACIAL STIMULI

To simultaneously investigate both categorical and dimensional processing, facial stimuli were chosen from the six basic emotions, and a morphing technique was used to create intermediate stimuli among these emotions that could be interpreted from both categorical and dimensional perspectives. This set of visual stimuli, which was utilized in our prior research (Fujimura et al., 2012), was selected from the Facial Expressions of Emotion Stimuli and Tests (FEEST) devised by Young et al. (2002). The stimuli utilized in our research were designed to dissociate the two fundamental dimensions of valence and arousal in the emotional space (**Figure 2**). In particular, our dimensional strategy attempted to dissociate the valence and arousal dimensions by creating two orthogonal continua that lay parallel to these two dimensions of the emotional space. To satisfy this constraint and the requirement that the emotional stimuli must be chosen from among the six basic emotions, fear and happiness were selected to create the valence-related continuum. This selection was motivated by the reasoning that these two emotions are the only pair of facial emotions that remain largely parallel to the two sides of the valence axis in the emotional space (with fear and happiness associated with negative and positive valences, respectively). Faces conveying disgust and anger were chosen to satisfy the aforementioned constraint for the arousal dimension (**Figure 1B**). However, there exist cross-cultural differences in facial emotion perception (e.g., Jack et al., 2012), particularly in the arousal dimension. Japanese individuals tend to perceive the disgust-anger continuum to be largely parallel to the arousal axis (Fujimura et al., 2012), whereas Caucasian individuals perceive the fear-sadness or surprise-sadness continua to be parallel to this axis (Russell, 1980). Expressions of fear and happiness were utilized as the endpoints for the creation

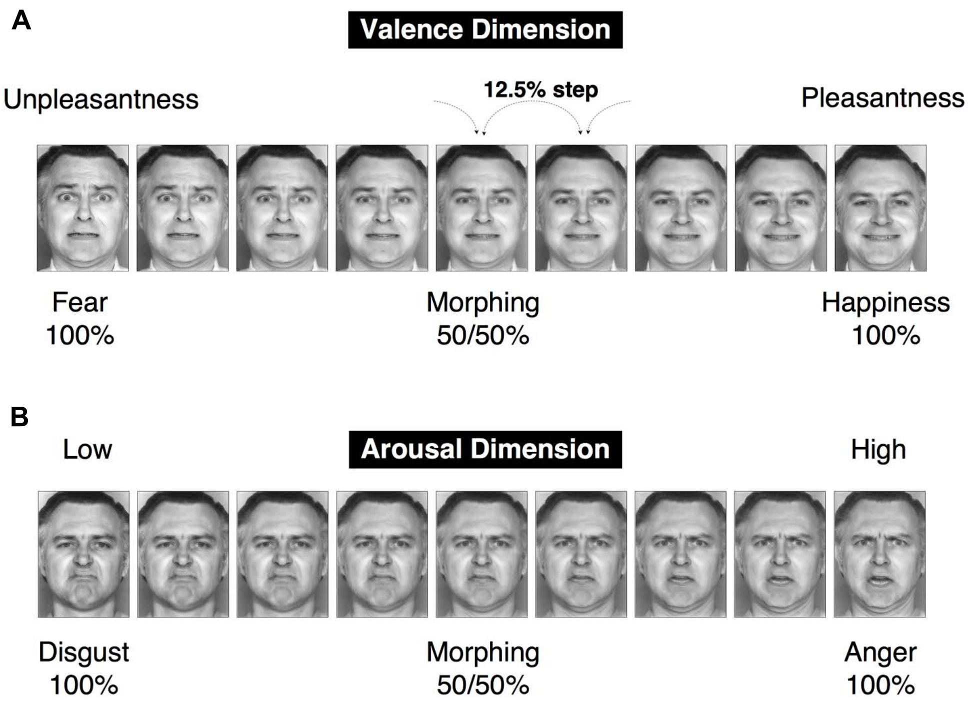


**FIGURE 2 | A set of facial stimuli.** Two morphed continua were created, with a 12.5% step between a morphed image and the subsequent image in a continuum. The happiness–fear continuum represents change in the pleasantness–unpleasantness continuum (the valence dimension). The anger–disgust continuum represents change in the arousal–sleepiness continuum (the arousal dimension).

of a morphed continuum in the valence dimension; in particular, the faces representing these two emotions were morphed to create seven intermediate images separated by steps of 12.5% (**Figure 2**). Using the same procedure, disgust and anger were employed to create a morphed continuum in the arousal dimension (**Figure 2**). We validated the stimuli to ensure their effectiveness for the study participants (see **Figure 3** in the section Results of this study).

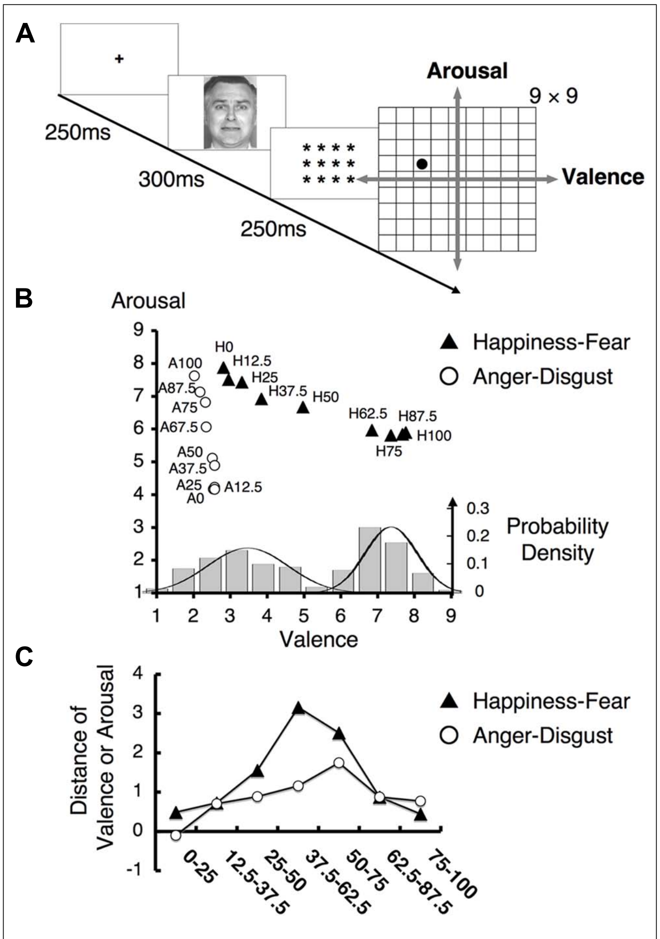


**FIGURE 3 | Behavioral results for the Affect Grid task. (A)** The experimental design of the Affect Grid task. Subjects were asked to rate each facial emotion by selecting the appropriate location on a two-dimensional square grid representing the emotional space. This grid represents nine levels of both the valence and arousal dimensions; in particular, subjects rated each emotion from "1" = "unpleasant" to "9" = "pleasant" in the valence dimension and from "1" = "low" to "9" = "high" in the arousal dimension. **(B)** The mean ratings on the Affect Grid task. The labels "H12.5" and "A75" indicate the 12.5% happiness and 75% anger facial stimuli, respectively. The ratings for the happiness − fear continuum were nearly parallel to the valence axis, whereas the ratings for the anger–disgust continuum were approximately parallel to the arousal axis. (Bottom) The histogram indicates how frequently each facial stimulus on the happiness–fear continuum was assigned to each level of the valence scale on the grid representing the emotional space. The displayed data were obtained from every study participant and were averaged for each type of stimulus across all participants. The solid lines indicate normal distributions. **(C)** The mean differences in valence and arousal ratings between faces separated by two steps along the happiness–fear continuum and the anger–disgust continuum, respectively. The label "75–100" refers to the difference in valence (arousal) rating between the 75% happiness (anger) image and the 100% happiness (anger) image.

We also created movie stimuli of the four emotional expressions (happiness, fear, anger, and disgust) and four mosaic faces. These stimuli were generated using the same faces of the same models (a male and a female) that were used for the main experimental sessions. These movie stimuli were used in a localizer task to effectively detect the "face network" in individual subjects by accounting for individual differences in brain structure (see the section fMRI DESIGN of this study). It has been reported that greater activation of the face network (e.g., the, FFA and IOG) is produced by emotional expression movies (which are known as "dynamic expressions") than by static images of the same expressions (Kilts et al., 2003; LaBar et al., 2003; Sato et al., 2004; Fox et al., 2009a; Foley et al., 2012). Thus, we used these movies in a functional localizer task to define regions of interest (ROIs) in individual subjects (Fox et al., 2009a). Importantly, the same ROIs were obtained regardless of whether dynamic expressions or conventional static stimuli were used to define face-selective regions (Fox et al., 2009a; Pitcher et al., 2011a). This consistency was also reported in the ROIs defined using dynamic and static body postures as stimuli (Russell, 1997). We created the aforementioned movie stimuli in the following way. First, 24 intermediate images between the neutral (0%) and emotional (100%) expressions, separated by steps of 4%, were generated using computer-based morphing techniques. To create a movie clip, the resulting 26 images (the neutral image, the 24 intermediate images, and the final, 100% emotional image) were presented in succession. Each intermediate image was presented for 40 ms, and the first and last images were each presented for 520 ms; thus, the duration of each movie clip was 2,000 ms. It has previously been demonstrated that this presentation speed adequately reflects natural changes that occur in dynamic facial expressions of fear and happiness (Sato et al., 2004). A control condition was established by generating dynamic mosaic images from the same images used for the experimental condition (Sato et al., 2004). All of the aforementioned face images were arranged on a 12 × 18 grid and randomly reordered using a constant algorithm, causing each image to become unrecognizable as a face. A set of 26 of these mosaic images was serially presented as a movie clip. The presentation speed for these mosaic images was identical to the presentation speed utilized for the dynamic expression stimuli. These manipulations caused the dynamic mosaic images to be nearly equivalent to their corresponding original dynamic expression stimuli with respect to size, brightness and dynamic information.

## PSYCHOLOGICAL EXPERIMENT

Subjects were asked to perform a psychological experiment in a laboratory using the aforementioned facial stimuli. Events in the psychological experiment were controlled using a program written in Inquisit 3.0 (Millisecond, Seattle, USA) and implemented on a computer (Vostro 420, Dell) that used the Windows XP operating system (Microsoft). Static image stimuli were presented on a 19-inch LCD monitor (E1902S, Iiyama; 1024 × 768 pixels, 75 Hz refresh rate) and subtended a visual angle of approximately 10.0° × 7.3°. Participants rated the same facial stimuli using three types of tasks: the Affect Grid task, the identification task, and the

ABX discrimination task. No time restrictions were applied. Four training trials were conducted prior to the performance of each task. The order of the three tasks was counterbalanced across the study participants.

### *The affect grid task (dimension task)*

The 9 × 9 Affect Grid assesses affect along the dimensions of valence and arousal (Russell et al., 1989). Study participants were asked to rate the emotions expressed on the faces they viewed by utilizing a computer mouse to select the appropriate location on a two-dimensional square that represented the emotional space. Each trial began with the presentation of a fixation point for 250 ms, which was followed by the presentation of a blank screen for 250 ms and a facial stimulus for 300 ms. Following a mask stimulus of asterisks that lasted 250 ms, the Affect Grid was displayed until the participant responded. Each facial stimulus was presented twice in random order, producing a total of 72 trials that were divided into two blocks based on the two models who were utilized. The order of the blocks was counterbalanced across the participants.

### *The identification task (categorization task 1)*

Participants were asked to identify depicted facial expressions by choosing between the two emotions on the endpoints of the continuum containing each depiction. For instance, participants were asked to identify the 87.5% happiness stimulus as either "happiness" or "fear." Each trial began with the presentation of a fixation point for 250 ms, which was followed by the presentation of a blank screen for 250 ms and a facial stimulus for 300 ms. Subsequently, a mask consisting of a cluster of asterisks was displayed for 250 ms; two emotional words were then presented ("happiness? fear?" for the valence dimension and "anger? disgust?" for the arousal dimension), and participants were asked to select one of these two words by pressing the assigned button. Each face was presented eight times in random order, producing a total of 288 trials. The two models and two continua (happiness–fear and anger–disgust) of this study were used to divide these trials into four blocks. The order of the four blocks was randomly determined for each study participant.

### *The ABX discrimination task (categorization task 2)*

The ABX discrimination task required participants to discriminate between faces from a continuum. Each trial began with the presentation of a fixation point for 250 ms, which was followed by the presentation of a blank screen for 250 ms and three successive images of faces. The first (A) and second (B) faces were presented for 300 ms each, and the third (X) face was presented for 1,000 ms. The blank intervals between the presentation of A and B, between the presentation of B and X, and after X were 250, 1,000, and 250 ms, respectively. Participants were asked to press a response button to indicate whether X matched A or B. In each trial, the stimuli A and B differed by two steps along one of the examined continua; thus, there was a 25% gap between the paired faces (e.g., happiness 100% and happiness 75%), and there were seven potential pairs for each continuum. The third face, X, was always identical to either A or B. Four presentation orders were possible: (ABX) = (ABA), (ABB), (BAA) or (BAB). The same order was presented twice for each pair, resulting in a total of 56 trials for each continuum. One block consisted of pairs from one continuum; thus, there were a total of four blocks. The order of the trials within a block and the blocks themselves were randomized across the study participants.

Chronologically, the psychological experiment was performed after the fMRI experiment (discussed below) to avoid any habituation/adaptation effects on the BOLD signal caused by excessive repetition of exposure to the same stimuli.

## fMRI DESIGN

Stimuli were displayed via a back-projection screen placed at the head of the scanner bore (Avotec Inc., Stuart, FL, USA; resolution: 800 × 600; refresh rate: 60 Hz), which was viewed by each subject via a mirror attached to the table near the subject's head. A pair of plastic glasses was used to correct each subject's vision to normal levels. All visual stimuli (260 × 360 pixels, gray images) were restricted to a maximum of 30° of visual angle. Manual responses were recorded using an MRI-compatible single-button box. An fMRI experiment consisted of three separate sessions (one functional localizer session and two main sessions).

In the localizer task of this study, dynamic expression stimuli (2,000 ms) were used to effectively define face network ROIs in individual subjects (Fox et al., 2009a; Pitcher et al., 2011a). Dynamic expressions and dynamic mosaics for the four types of examined expressions (happiness, fear, anger, and disgust) were randomly presented in turn, with each stimulus presented five times for each of the two model faces (a male and a female). Between trials, a central fixation cross was displayed for a jittered inter-trial interval of 3–6 s (mean = 4.5 s). In the localizer session, subjects were required to monitor a stream of dynamic expressions and dynamic mosaics and rapidly press a button when they detected a target "house" stimulus (100 ms), which was also randomly presented. This irrelevant task was performed to attract the attention of subjects, causing study participants to passively and implicitly view the presented facial expressions. Chronologically, the localizer session was performed after the main experiments (discussed below) to avoid any habituation/adaptation effects on the BOLD signal caused by excessive repetition of stimuli involving the same models' faces.

In the main sessions, subjects were required to monitor a stream of "static" stimuli and rapidly press a button when they detected an upside-down "target" face. This irrelevant task was performed to attract the attention of subjects, causing study participants to passively and implicitly view the presented facial expressions. The targets were neutral faces that differed from the main stimuli but involved the same individuals. Each of the 18 expression images in the main stimulus set, which included the nine images of the morphed continuum for each of the two dimensions of emotional space (valence and arousal), was displayed (for 300 ms each) in quasi-random order. To avoid habituation/adaptation effects, there was no consecutive presentation of the same image in a raw. Between trials, a central fixation cross was displayed for a jittered inter-trial interval of 4–6 s (mean = 5 s). Different models (a male and a female with the same expressions) were used for each main session, and the order of the two main sessions was counterbalanced across the study participants. Each stimulus was presented six times for each model face.

## fMRI SCANNING PROCEDURE

Functional magnetic resonance imaging experiments were performed on a 4 T Agilent whole-body MRI system (Agilent Inc., Santa Clara, CA, USA) with a circularly polarized quadrature birdcage radio frequency coil as a transmitter and four-channel receiver surface coils as receivers (Nova Medical Inc., Wilmington, MA, USA). A total of 40 axial slices (24 cm field of view (FOV), 64 × 64 matrix, 3 mm thickness, 0 mm gap) with 30° forward rotation from the anterior commissure–posterior commissure (AC–PC) plane were acquired using a two-shot echo-planar imaging (EPI) pulse sequence [volume TR (repetition time) 4.4 s, TE (echo time) 25 ms, flip angle 78°] for the three functional runs (two main runs and one localizer run), each of which consisted of 156 volumes. After TSENSE (sensitivity encoding incorporating temporal filtering; Pruessmann et al., 1999; Kellman et al., 2001) reconstruction (with an acceleration factor of 2), the sampling frequency was doubled, causing the effective volume TR to become 2.2 s. Prior to and between the functional runs, a set of high-resolution (1 $mm^3$) and low-resolution (1.72 $mm^3$) whole-brain anatomical images were acquired using a magnetization-prepared 3D FLASH (fast low-angle shot) pulse sequence.

## ANALYSES OF fMRI DATA

### Preprocessing

After EPI image reconstruction, intensity alternation between the odd- and even-numbered volumes produced by TSENSE reconstruction was removed from each functional run (odd and even volumes were averaged on a voxel-by-voxel basis, and we calculated the multiplication factor between the volumes). Cardiac and respiratory fluctuations were also removed using a retrospective estimation and correction method that we have previously described (Chen et al., 2001). The data were then preprocessed and analyzed using the BrainVoyager QX software package (Brain Innovation, Maastricht, Netherlands).

Data from the two main runs were serially connected in the time course direction. Preprocessing included slice scan time correction (using sinc interpolation), linear trend removal, temporal high-pass filtering to remove low-frequency non-linear drifts (0.00505 Hz) and 3D motion correction to detect and correct for small head movements by spatially aligning all volumes to a target volume via rigid-body transformations. A relatively small spatial smoothing isotropic Gaussian kernel that was 4 mm at full-width half-maximum (FWHM) was applied to the resulting volumes to investigate finer structures in the subcortical regions (e.g., the amygdala subregions). This spatial smoothing was applied to each subject's native space (as a default component of preprocessing with Brain Voyager QX). Functional images were then coregistered to the anatomical volume, using both position parameters obtained from the scanner and manual adjustments to achieve the optimal fit. We omitted data from the four subjects who exceeded 1 mm in head movement (as estimated by the motion correction algorithm) and/or performed the tasks with less than 90% accuracy. Data from the remaining 22 adult subjects were used for the following analyses.

### Functional localizers

After the transformation of each subject's registered functional images into Talairach space (Talairach and Tournoux, 1988), whole-brain activation maps were obtained using a standard voxelwise general linear model (GLM) at the single-subject level. For the functional localizer task, regressors encoding the perceptual processing of the four types of dynamic expressions that were examined (fear, happiness, disgust, and anger) and four types of dynamic mosaics were convolved using a theoretical two-gamma hemodynamic response function (HRF) and regressed against the observed BOLD data. To determine individual face-network ROIs for each subject, we compared the activity associated with all dynamic expression stimuli with the activity associated with all dynamic mosaics. For each subject, flexible thresholds [from uncorrected $p < 0.001$ to $p < 0.00001$, cluster size >4 voxels (36 $mm^3$) and ≤10 voxels (90 $mm^3$)] were employed to assess the results of this comparison and identify and isolate cluster of face-sensitive voxels in ROIs, particularly for subcortical regions, such as amygdala subregions. We adjusted $p$ value for each ROI in each subject, so that the cluster size of each ROI fit within the definition (5∼10 voxels) as small as possible. The $t$-statistics produced by these comparisons were computed to detect activation levels in each ROI. The three face-related regions comprising the "core" system of face perception were defined in the following manner (Fox et al., 2009a). Face-related voxels located on the lateral temporal portion of the FG were designated as the FFA, whereas voxels located on the lateral surface of the IOG were designated as the occipital face area (OFA). Face-related voxels located on the posterior segment of the STS were designated as the posterior portion of the superior temporal sulcus (pSTS). In addition to these core face-processing ROIs, regions comprising the "extended" system of face perception were also defined. Face-related voxels within the amygdala and insula were identified, and face-related voxels within the medial region of the prefrontal cortex were designated as the medial prefrontal cortex (mPFC).

### ROI analyses

We then applied GLM analysis to each ROI of each subject to extract the enhanced BOLD signals involved in the perceptual processing of each emotional image (which included 18 emotional expressions in total from the valence and arousal continua, irrespective of the model who was examined). We generated 18 multiple regressors by convolving a boxcar representation of each type of stimulus-presentation period with the HRF. We then extracted the parameter estimates from this subject-specific (i.e., first-level) analysis. The parameter estimate for each ROI was extracted from the mean of all voxels in the cluster. Subsequent statistical correlational analyses of the generated linear and $U$-shaped models were performed using SPSS (IBM Corporation, Chicago, USA). All reported results were subjected to Greenhouse-Geisser correction for non-sphericity. To analyze the effect of the experimental conditions on BOLD signals in each ROI of the face network, we performed two types of analysis: a stimulus-based analysis and a perception-based analysis. Because fMRI results during a face/object detection task have indicated that the measured effects were driven more strongly by a subject's perception than by stimulus category (Grill-Spector et al., 2004), it is possible that categorical effects observed in this study may be influenced less by the morph characteristics of the image than by whether the subject perceives the image as a specific expression of emotion. In fact, our previous study demonstrated that individual variations were observed when subjects were asked to identify categorical boundaries in morphed continua of facial expressions (Fujimura et al., 2012). To address this issue, we first performed a stimulus-based analysis using our *a priori* stimuli of morphed continua and aligned 50:50 morphed images in the centers of the tested continua ("5" in the valence/arousal score). We then calculated the BOLD response averages to each valence/arousal score across subjects (producing the "Stimulus-based" results, see fMRI RESULTS). Our second analysis realigned the stimulus images based on each subject's category boundary, which was determined by the behavioral results from the categorization tasks (**Table 1**). We performed this realignment operation by implementing a parallel shift of each individual's response curve from the stimulus-based analysis that adjusted subjective category boundaries to the center point for each dimension ("5" in the valence/arousal score). We then calculated the BOLD response averages to each valence/arousal score across subjects (producing the "Perception-based" results, see fMRI RESULTS).

A second-stage random-effect analysis (RFX) was performed solely to determine the coordinates of group-wise face network ROIs (uncorrected $p < 0.001$), which are indicated in **Table 2**.

### *Validation with subject-wise analysis*

We considered the issue that dimensional processing could be reflected in the linear relationships observed between BOLD signals and degrees of morphing. To analyze group-wise profiles, we calculated the $R$-values (the Pearson product-moment correlation coefficients) between the ordered integers 1 through 9 and the average BOLD signals across subjects for the nine stimuli along each dimension of the emotional space (valence or arousal, see fMRI RESULTS). We confirmed the group-wise profiles by performing a validation using the power of the random variable (the subjects) in the following manner. First, for each ROI and each subject, we calculated the Pearson $R$-value, which was subsequently converted to a $z$-value using Fisher's $z$-transform. For each ROI, the 22 $z$-values from the study subjects were subjected to a one-sample $t$-test against zero. The average correction value (averaging the $R$-values across subjects) and the $t$-values with their associated $p$-values have been reported in **Table 3**.

## RESULTS

### BEHAVIORAL RESULTS

We first confirmed that our stimuli affected the subjects of this study in a manner consistent with the effects observed in our previous study (Fujimura et al., 2012). In particular, we administered the Affect Grid task (**Figure 3A**), which used a dimensional strategy to confirm our previous findings regarding hybrid categorical and dimensional processing (Fujimura et al., 2012).

The mean scores on the Affect Grid task are illustrated in **Figure 3B**. Ratings of facial stimuli along the happiness–fear and anger–disgust continua shifted in accordance with changes in the valence or arousal dimensions, respectively, and were consistent with physical changes in the stimuli. These results indicated the existence of dimensional perception of the tested stimuli. The happiness–fear and anger–disgust continua were nearly orthogonal and remained relatively parallel to the valence and arousal dimensions, respectively, in the emotional space (**Figure 3B**). Although the happiness–fear continuum appeared to demonstrate a slight linear increase in arousal ratings as stimuli shifted from happiness to fear, this increase produced no significant difference between happiness and fear with respect to arousal ratings.

Notably, ratings on the happiness–fear continuum appeared to demonstrate a gap between the 62.5 and 50% stimuli, indicating the likelihood of a category boundary between these stimuli. However, no clear boundary was observed for the anger–disgust continuum; these findings were consistent with the results of our

**Table 1 | Individual category boundaries for each subject.**

| Subject | Category boundary (1–9) | |
|---|---|---|
| | Valence dimension | Arousal dimension |
| | (happiness–fear continuum) | (anger–disgust continuum) |
| ID 1 | 5 | 5 |
| ID 2 | 5 | 6 |
| ID 3 | 4 | 5 |
| ID 4 | 5 | 6 |
| ID 5 | 6 | 6 |
| ID 6 | 5 | 5 |
| ID 7 | 5 | 6 |
| ID 8 | 5 | 6 |
| ID 9 | 5 | 5 |
| ID 10 | 6 | 6 |
| ID 11 | 5 | 6 |
| ID 12 | 5 | 6 |
| ID 13 | 5 | 5 |
| ID 14 | 6 | 6 |
| ID 15 | 6 | 6 |
| ID 16 | 5 | 5 |
| ID 17 | 6 | 5 |
| ID 18 | 6 | 6 |
| ID 19 | 6 | 6 |
| ID 20 | 6 | 6 |
| ID 21 | 5 | 6 |
| ID 22 | 8 | 5 |

*The individual category boundaries for each subject were defined as the facial stimulus with an identification rate closest to 50%. Each number corresponded to a morphed stimulus; as indicated in* ***Figure 2****, these stimuli ranged from "1" = "0% happiness" to "9" = "100% happiness" in the valence dimension and from "1" = "0% anger" to "9" = "100% anger" in the arousal dimension. "5" denotes faces with a 50% contribution from each of the two opposing emotions of a continuum.*

**Table 2 | The ROIs defined by the functional localizer task of facial emotion perception.**

| Region-of-interest | *t*-value of peak voxel | Talairach coordinates | | |
|---|---|---|---|---|
| | | *x* | *y* | *z* |
| FFA (R) | 4.82 ± 0.31 | 41 ± 7 | −50 ± 9 | −18 ± 6 |
| OFA (R) | 5.13 ± 0.47 | 44 ± 8 | −65 ± 11 | −11 ± 8 |
| Amygdala (L) | 4.47 ± 0.14 | −25 ± 6 | −2 ± 2 | −12 ± 4 |
| Insula (L) | 4.86 ± 0.28 | −31 ± 6 | 6 ± 8 | −9 ± 7 |
| mPFC | 5.10 ± 0.41 | −2 ± 2 | 55 ± 10 | 18 ± 8 |
| pSTS (R) | 5.51 ± 0.36 | 40 ± 8 | −50 ± 11 | 3 ± 5 |

*The data reported above represent the average results from the functional localizer. The t-value of the peak voxel and the Talairach coordinates are indicated for each ROI (and expressed in terms of mean ± SD). The mean Talairach coordinates are reported solely to facilitate comparisons with the findings of other studies. FFA, fusiform face area; OFA, occipital face area; mPFC, medial prefrontal cortex; pSTS, posterior portion of the superior temporal sulcus; L, left hemisphere; R, right hemisphere.*

prior investigation (Fujimura et al., 2012). To visualize data distributions for the Affect Grid task, a frequency histogram for the happiness–fear continuum is provided in **Figure 3B** (bottom), indicating the ratings of the tested facial stimuli along the valence dimension. The rating data for each morphed stimulus were averaged for each subject. A Gaussian mixture model (McLachlan and Basford, 1988), which uses probability models to account for clustering in data distributions, was utilized for analysis. The Bayesian information criterion (BIC) was used to evaluate the fitness of the model; smaller BIC values indicate greater model suitability. We found two normal distributions in the data for the happiness–fear continuum (single Gaussian distribution: BIC = 971.07; two Gaussian distributions: BIC = 881.92). Thus, this study confirmed our previously reported finding (Fujimura et al., 2012) that valence may be divided into two clusters of rating scores; in other words, the study results support the existence of implicit categorical processing of emotion perceptions. To verify the occurrence of categorical perception in the presence of ratings based on continua in emotional space, we calculated the difference in valence or arousal ratings between facial stimuli differing by two steps on each examined continuum. The mean difference in valence rating between the elements of each pair of facial stimuli on the happiness–fear continuum and the mean difference in arousal rating between the elements of each pair of facial stimuli on the anger–disgust continuum are depicted in **Figure 3C**. We conducted an analysis identical to the approach used in the ABX discrimination task to confirm that the largest difference in valence rating among stimulus pairs on the happiness–fear continuum occurred between the stimuli with 67.5 and 37.5% happiness. This maximum difference in valence was compared with the combined mean valence differences of all other stimulus pairs on the happinessfear continuum, and a *t*-test demonstrated that the difference in valence between the 62.5 and 37.5% happiness images was significantly greater than the average valence difference between these other stimulus pairs ($t_{21} = 7.20$, $p < 0.001$). These results confirmed the validity of our stimuli and reproduced our previous findings regarding hybrid categorical and dimensional processing within the dimensional strategy (Fujimura et al., 2012).

We then performed categorization tasks to identify individual category boundaries for each subject, which were used in the *post hoc* analysis to realign individual fMRI data (to the "Perception-based analysis" results, as described in the section Methods of this study).

The study subjects were asked to perform the two different categorization tasks of identification (**Figure 4A**) and ABX discrimination (**Figure 5A**). The results from the identification task are depicted in **Figures 4B,C**, which indicate the group-wise mean percentages for the two identified emotions on each continuum (happiness or fear in **Figure 4B** and anger or disgust in **Figure 4C**). Visual inspection of these figures reveals that the identification rates exhibit sigmoidal non-linear distributions, indicating the presence of an abrupt category shift within each continuum. On the happiness–fear continuum, the category boundary between happiness and fear appeared to be located

**Table 3 | Analyses of dimensional (linear) processing in each ROI, after subject-wise validation.**

| ROI | Valence dimension | | | | | | Arousal dimension | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Stimulus-based | | | Perception-based | | | Stimulus-based | | | Perception-based | | |
| | Ave. *R* | $t_{21}$ | *p* | Ave. *R* | $t_{21}$ | *p* | Ave. *R* | $t_{21}$ | *p* | Ave. *R* | $t_{21}$ | *p* |
| FFA (R) | −0.18 | −2.03 | 0.06 | −0.11 | −0.93 | 0.36 | 0.10 | 1.68 | 0.11 | 0.13 | 2.08 | <0.05* |
| Insula (L) | −0.15 | −2.09 | <0.05* | −0.13 | −1.22 | 0.24 | 0.07 | 0.99 | 0.33 | 0.18 | 2.76 | <0.05* |
| Amygdala (L) | −0.14 | −2.11 | <0.05* | −0.10 | −1.26 | 0.22 | 0.06 | 0.51 | 0.62 | 0.08 | 0.77 | 0.45 |
| mPFC | 0.19 | 3.18 | <0.01** | 0.23 | 3.89 | <0.001*** | −0.03 | −0.47 | 0.65 | −0.01 | −0.07 | 0.95 |
| OFA (R) | −0.06 | −0.79 | 0.44 | −0.02 | −0.21 | 0.84 | 0.12 | 1.58 | 0.13 | 0.10 | 1.30 | 0.21 |
| pSTS (R) | 0.02 | 0.41 | 0.68 | 0.01 | 0.11 | 0.92 | −0.08 | −0.84 | 0.41 | −0.07 | −0.77 | 0.45 |

*The group-wise profile for dimensional processing in each ROI (**Figures 6** and **7**, as well as **Table 2**) was validated after the subject-wise validation (see the section Methods of this study). Average correction values (Ave. R, representing the mean of the R values from the 22 study subjects) and t-values with associated p-values are reported. *, ** and *** denote statistically significant correlations with p-values of less than 0.05, 0.01 and 0.001, respectively.*

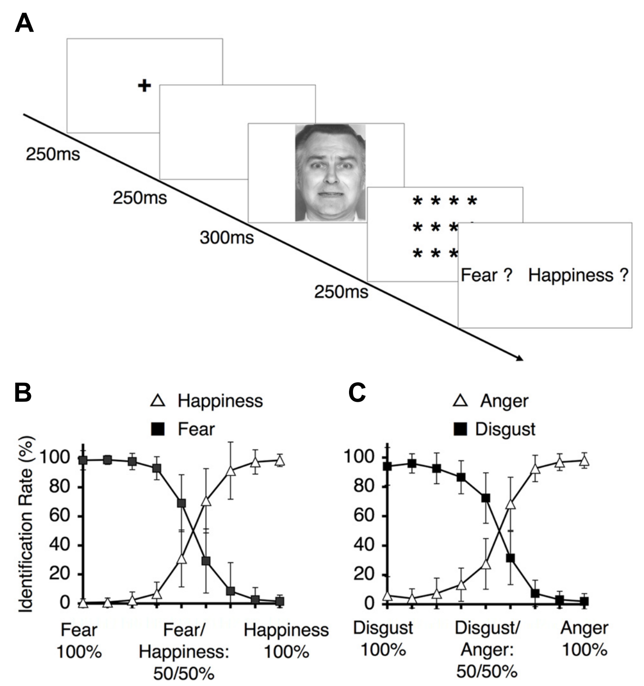


**FIGURE 4 | Behavioral results for the identification task. (A)** The experimental design for the identification task. Subjects were asked to identify a depicted facial expression by choosing between the two emotions at the endpoints of the continuum to which the depiction belonged (i.e., fear vs. happiness, or disgust vs. anger). **(B)** The identification rates for the happiness−fear continuum. These rates indicate the frequencies at which happiness and fear were identified for the morphed faces depicted in **Figure 2A**. **(C)** The identification rates for the anger–disgust continuum. These rates indicate the frequencies at which anger or disgust were identified for each of the morphed faces depicted in **Figure 2B** (as indicated by the labels along the *x*-axis).

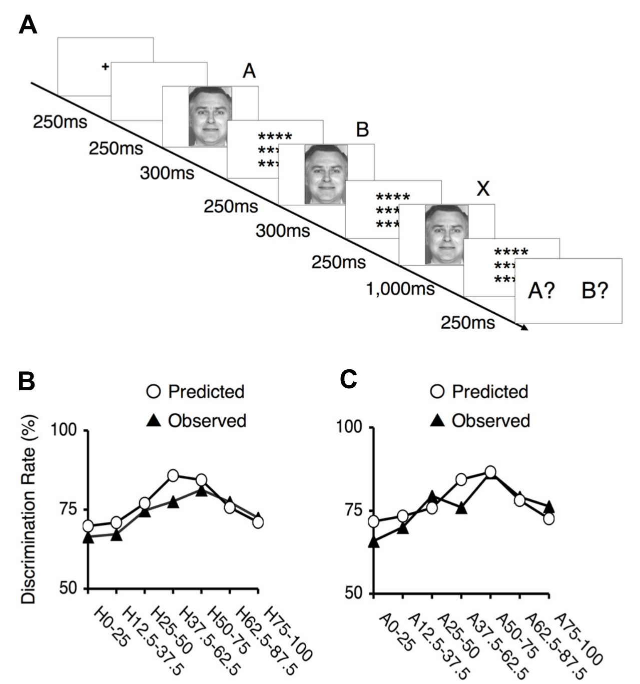


**FIGURE 5 | Behavioral results for the ABX task. (A)** The experimental design of the ABX task. Subjects were asked to indicate whether the third face "X" matched the initial face "A" or the second face "B". In each trial, the facial stimuli A and B differed by two steps on one of the examined continua; thus, there was always a 25% gap between this pair of faces (e.g., 100% happiness and 75% happiness), and seven potential pairs existed along each continuum. The third face, "X," was always identical to either "A" or "B." **(B)** The mean values of the observed discrimination rates and the predicted data for the happiness − fear continuum. The depicted rates were based on the frequency of correct responses in the ABX discrimination task. The labels indicate which pair of facial stimuli correspond to each point. For example, "H25–50" indicates a trial in which 25 and 50% happiness were presented as images A and B. **(C)** The mean values of the observed discrimination rates and the predicted data for the anger–disgust continuum. The labels indicate which pair of facial stimuli correspond to each point. For example, "A25−50" indicates a trial in which 25 and 50% anger were presented as images A and B.

between the stimuli with 62.5 and 50% happiness. The identification rates for the anger–disgust continuum also exhibited patterns similar to the patterns observed for the happiness–fear continuum.

To assess the occurrence of categorical perception, we applied a method used in previous studies (Calder et al., 1996; Young et al., 1997). First, we predicted subjects' performance in the ABX discrimination task based on identification and ABX discrimination data. This approach assumes that two factors determine the ability to discriminate between two facial expressions: the physical differences between pairs of facial stimuli, independent of the expressions involved, and contributions from categorical perceptions of facial expressions. To estimate the first of these two factors, we utilized the mean of the discrimination rates for the pairs at the endpoints of each continuum. Categorical perception did not significantly contribute to these results because these stimuli were similar to prototypical facial expressions. To assess the second factor, we calculated the differences between the identification rates for the two relevant stimuli in each pair and multiplied this difference by 0.25 (a constant). By totaling the estimates for these two factors, we obtained performance predictions for the discrimination task. If these predicted values correlated with the observed ABX discrimination data, we could conclude that categorical perception occurred within the examined continuum.

The predicted rates of correct responses and the mean actual rates of correct responses for the discrimination task are indicated in **Figures 5B,C**. The observed and predicted curves appear to largely coincide, and the correlations between the observed and predicted results were significant for each continuum (for the happiness–fear continuum: $R = 0.85$, $t_5 = 3.56$, $p < 0.05$; for the anger–disgust continuum: $R = 0.76$, $t_5 = 2.58$, $p < 0.05$). Thus, categorical perception contributed to the observed responses to the facial stimuli within each continuum. If categorical perception occurs for each continuum, participants should more readily discriminate between a pair of facial stimuli that are on opposite sides of a categorical boundary than between a pair of facial stimuli that fall within the same category. To confirm this hypothesis, the peak correct discrimination rate for each continuum was compared with the mean of the correct discrimination rates for all other stimulus pairs along the continuum in question. A $t$-test revealed that the correct discrimination rate for the stimulus pair involving images of 75 and 50% happiness was significantly higher than the correct discrimination rates for other pairs along the happiness–fear continuum ($t_{21} = 3.34$, $p < 0.01$). On the anger–disgust continuum, significantly superior discrimination performance was observed for the stimulus pair involving images of 75 and 50% anger than for the other examined pairs ($t_{21} = 4.67$, $p < 0.001$). These results were consistent with our previous reports that 62.5% happiness and 62.5% anger may represent group-wise means that constitute category boundaries for the happiness–fear and anger–disgust continua, respectively, (Fujimura et al., 2012). Individual category boundaries for each subject were defined as the facial stimulus level with an identification rate that was closest to 50% (**Table 1**). If multiple facial stimuli exhibited identification rates that were approximately equivalent relative to this standard, the facial stimulus closest to the center of the continuum (that is, the stimulus that was closest to 50% for each of the two opposing emotions of a continuum) was defined as the category boundary for the subject in question.

### fMRI RESULTS

Given that our psychological results exhibited hybrid categorical and dimensional processing in the valence dimension but not the arousal dimension (**Figure 3**), we sought to identify the neural basis of this phenomenon. To test our hypothesis that categorical and dimensional processing are intrinsically encoded by separate neural loci, we induced study subjects to passively view each emotional stimulus by asking these subjects to perform irrelevant tasks. The fMRI experiments involving passive viewing of the tested emotional stimuli were conducted prior to the aforementioned behavioral experiments to minimize any neural adaptation effects that might be induced by the repetitive presentation of the same stimuli in both sets of experiments. We obtained data from limited ROIs defined by a subsequent functional localizer task (as described in the section Methods of this study).

In accordance with the findings of numerous other reports, the functional localizer experiment conducted in this study revealed the existence of a distributed neural network for emotional face processing (**Table 2**). We depicted the response properties along the happiness–fear and anger–disgust continua in each ROI and classified the observed responses into three different groups with respect to facial emotion processing: categorical, dimensional (linear) and constant (uniform) processing. The valence and arousal dimensions were separately analyzed using the happiness–fear and anger–disgust continua, respectively.

#### *Categorical processing in the valence dimension*

The right FFA exhibited non-linear processing instead of a linear or uniform response in the happiness–fear continuum (**Figure 6A**, "Stimulus-based" results); as responses in the right FFA did not demonstrate a statistically linear correlation with the valence scores of facial emotions after the subject-wise validation (group-wise $R = -0.19$; after the subject-wise validation: $t_{21} = -2.03$, $p = 0.06$; see the section Methods of this study for additional details). Visual inspection indicates a $U$-shaped response curve in this ROI, with greater signal for both the happier and more fearful faces but lower signal for faces that are approximately evenly divided between these two emotions. This phenomenon suggests that the right FFA contributes to categorical processing via the detection of unambiguous faces (at both ends of the happiness–fear continuum) but filters out ambiguous faces near the category boundary. To confirm this abrupt change in response property for approximately evenly divided faces, we statistically compared the average responses of the right FFA to the three faces on each end of the happiness–fear continuum (i.e., the 100-0, 87.5-12.5 and 75-25 faces at both the happiness and fear ends of this continuum) with the three faces at the center of the continuum (the 37.5-62.5, 50-50 and 62.5-37.5 faces). This statistical comparison confirmed that the right FFA exhibited significantly different responses to the faces at the ends of the happiness–fear continuum than to the faces in the center of this continuum ($t_{21} = 2.51$, $p < 0.05$). The group-wise response property of categorical processing became more evident when the fMRI data from each study subject were realigned in accordance with the subject's category boundary, as determined by the behavioral results from the categorization tasks (**Table 1** and the "Perception-based" results in **Figure 6A**). We performed this realignment by implementing a parallel translation of each individual's response curve that adjusted the category boundary to the center for each dimension (as represented by a valence or arousal score of "5"). For the study participants, neural responses in the right FFA to faces that were perceived to be unambiguous (the three faces at each end of a subject's perceived happiness–fear continuum) were not only relatively constant but also significantly greater than the neural responses in this ROI for faces that were perceived to be ambiguous (the central three faces of a subject's perceived happiness–fear continuum). In particular, in the comparison of average responses for the three faces on each end of the continuum with average responses for the three central faces, $t_{21} = 2.17$, with $p < 0.05$. This response property of categorical processing differed from the typical sigmoidal curve obtained in the behavioral results (**Figure 4B**). Instead, the observed response, which appears to approximate the first derivative of a sigmoid function rather than a simple $U$-shaped curve, exhibited categorical processing with abrupt changes in response near a perception-based categorical boundary and with constant processing near both ends. We statistically compared response curves of the FFA between stimulus-based and perception-based analyses to find a better fit by the first derivative of sigmoid function. There was a trend that a perception-based response curve was better fitted by the first derivative of sigmoid function ($t_{21} = 1.89$, $p = 0.07$). This result indicated that realignment of each individual's response curve by using individual's category boundary was essential for the better fit, and hence implicated that the FFA exhibited categorical processing. These results indicated that the FFA not only engages in categorical processing but also detects an individual's categorical boundary in the valence dimension and decreases its response accordingly.

Although the observed atypical "$U$-shaped" curve for the BOLD response as a function of morphing may not constitute compelling evidence for categorical processing, the right FFA was

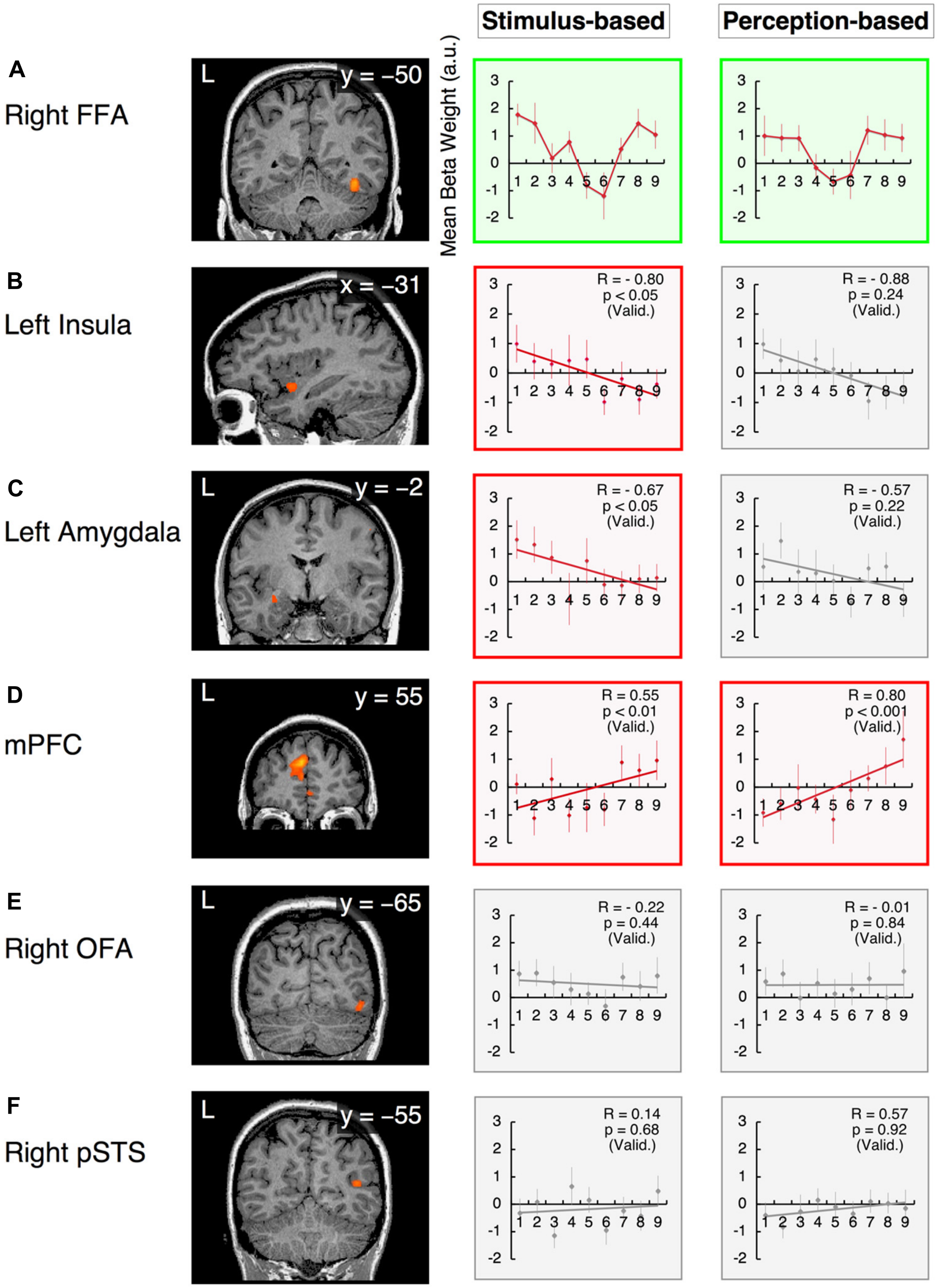


**FIGURE 6 | Functional magnetic resonance imaging results in the valence dimension.** The x-values represent morphed faces on the happiness – fear continuum depicted in **Figure 2A**, ranging from "1" = "0% happiness" to "9" = "100% happiness." The *y*-values represent the BOLD signals [with arbitrary units (a.u.)] obtained from each ROI. Two graphs are provided for each ROI to indicate the results from the stimulus-based and perception-based analyses; in the latter analysis, each individual's fMRI data are realigned in accordance with the individual's behavioral data with respect to category boundaries (**Table 1**). The graphs in green squares represent *U*-shaped or categorical processing that differs from linear or constant processing. Graphs in red squares represent dimensional (linear) processing with statistically significant correlations after the subject-wise validation (**Table 3**). Graphs in gray squares represent constant processing without statistically significant correlations. Error bars denote SEM. R: the group-wise correlation coefficient. p (Valid.): the probability after the subject-wise validation (**Table 3**). **(A)** The right fusiform face area (FFA); **(B)** the left insula; **(C)** the left amygdala; **(D)** the mPFC; **(E)** the right OFA; and **(F)** the posterior STS (pSTS) in the right hemisphere.

the only region in the face network to exhibit categorical processing in the valence dimension that clearly differed from linear or constant processing (as discussed below).

In contrast, in the arousal dimension, the right FFA exhibited dimensional processing (as discussed below) rather than categorical processing. This phenomenon was consistent with our psychological results, which indicated that implicit categorical processing occurred in the valence dimension but not the arousal dimension. Thus, the right FFA is a candidate ROI for the neural basis of psychological processing in facial emotion perception (see the section Discussion of this study).

#### *Dimensional (linear) processing*

The response in the left insula exhibited a negative correlation with the valence scores of facial emotions (**Figure 6B**). Activity in this region increased in response to increasingly fearful, "unpleasant" faces but decreased in response to increasingly happy, "pleasant" faces. The linear response property of this ROI was statistically significant in stimulus-based analyses but not perception-based analyses (stimulus-based results: group-wise $R = -0.80$, with $t_{21} = -2.09$ and $p < 0.05$ after the subject-wise validation; perception-based results: group-wise $R = -0.88$, with $t_{21} = -1.22$ and $p = 0.24$ after the subject-wise validation; **Figure 6B**, **Table 3**). The left amygdala also exhibited a negative correlation with valence scores that was significant only in stimulus-based analyses (stimulus-based results: group-wise $R = -0.67$, with $t_{21} = -2.11$ and $p < 0.05$ after the subject-wise validation; perception-based results: group-wise $R = -0.57$, with $t_{21} = -1.26$ and $p = 0.22$ after the subject-wise validation; **Figure 6C**, **Table 3**). The mPFC exhibited a significant positive correlation with valence scores in both the stimulus-based and perception-based analyses (stimulus-based results: group-wise $R = 0.55$, with $t_{21} = 3.18$ and $p < 0.01$ after the subject-wise validation; perception-based results: group-wise $R = 0.80$, with $t_{21} = 3.89$ and $p < 0.001$ after the subject-wise validation; **Figure 6D**, **Table 3**).

However, in the arousal dimension, the right FFA exhibited a positive correlation with arousal scores of facial emotions, although this correlation was significant only in perception-based analyses (stimulus-based results: group-wise $R = 0.65$, with $t_{21} = 1.68$ and $p = 0.11$ after the subject-wise validation; perception-based results: group-wise $R = 0.80$, with $t_{21} = 2.08$ and $p < 0.05$ after the subject-wise validation; **Figure 7A**, **Table 3**). The left insula also exhibited a significant positive correlation with arousal scores in perception-based analyses but not stimulus-based analyses (stimulus-based results: group-wise $R = 0.67$, with $t_{21} = 0.99$ and $p = 0.33$ after the subject-wise validation; perception-based results: group-wise $R = 0.60$, with $t_{21} = 2.76$ and $p < 0.05$ after the subject-wise validation; **Figure 7B**, **Table 3**).

This linear-fit analysis appeared to render linear processing and categorical "sigmoid-curve" processing indistinguishable because both of these processing types should exhibit high correlation coefficients (monotonically increasing functions). This confounding problem may be inevitable given that we measured BOLD signals with large per-trial variability; furthermore, this issue may have been exacerbated in this investigation because we used a small number of trials (12 for each emotional face, regardless of the model involved) to prevent adaptation effects. Indeed, there were no statistical differences between sigmoidal- and linear-fit models to explain our neural response function ($t_{21} < 1.29$ and $p > 0.21$ for all ROIs). Thus, we only categorized the remarkable case of the right FFA, which involved a clear boundary as an instance of categorical processing (as discussed above).

We excluded the possibility that data from the aforementioned areas that demonstrated dimensional processing could also be explained by $U$-shaped categorical processing. For these areas, there were no statistically significant differences between the neural responses to unambiguous faces (the three faces at each end of a continuum) and the neural responses to ambiguous faces (the three faces in the center of the continuum in question) for either stimulus-based or perception-based analyses (in particular, for all cases, the comparison of the average responses of the three faces at each end of a continuum with the three faces at the center of the curriculum produced results of $t_{21} < 1.70$ and $p > 0.1$).

#### *Constant processing*

The left amygdala and mPFC demonstrated relatively constant (or uniform) processing in the arousal dimension in both stimulus-based and perception-based analyses [stimulus-based results for the left amygdala: group-wise $R = 0.30$, with $t_{21} = 0.51$ and $p = 0.62$ after the subject-wise validation; perception-based results for the left amygdala: group-wise $R = 0.18$, with $t_{21} = 0.77$ and $p = 0.45$ after the subject-wise validation (**Figure 7C** and **Table 3**); stimulus-based results for the mPFC: group-wise $R = -0.15$, with $t_{21} = -0.47$ and $p = 0.65$ after the subject-wise validation; perception-based results for the mPFC: group-wise $R = -0.28$, with $t_{21} = -0.07$ and $p = 0.95$ after the subject-wise validation (**Figure 7D** and **Table 3**)]. The right OFA and right pSTS also demonstrated constant processing in both the valence and arousal dimensions (for all cases, group-wise $R < 0.65$, with $p > 0.2$ after the subject-wise validation; **Figures 6E,F** and **7E,F** as well as **Table 3**). However, the OFA and STS differed in activity levels. The right OFA exhibited significant activation relative to baseline levels ($t_{21} = 2.68$, $p < 0.05$), whereas no significant activation was observed for the STSs (for both the left and right STS, $t_{21} < 0.36$ and $p > 0.70$). Notably, we defined functional ROIs using dynamic-expression stimuli of emotional faces in an implicit manner (**Table 2**) but investigated categorical/dimensional processing using static facial images (as described in the section Methods of this study). Our results indicated that in the absence of explicit attention, the right OFA responds to both static and dynamic emotional face stimuli, whereas the right STS responds to dynamic but not static images.

We excluded the possibility that the data from areas that exhibited constant processing could be explained by U-shaped categorical processing. For these areas, there were no statistically significant differences between the neural responses to unambiguous faces (the three faces at each end of a continuum) and the neural responses to ambiguous faces (the three faces in the center of the continuum in question) for either stimulus-based or perception-based analyses (in particular, for all cases, the comparison of the average responses of the three faces at each end of a continuum with the three faces at the center of the curriculum produced results of $t_{21} < 1.30$ and $p > 0.2$).

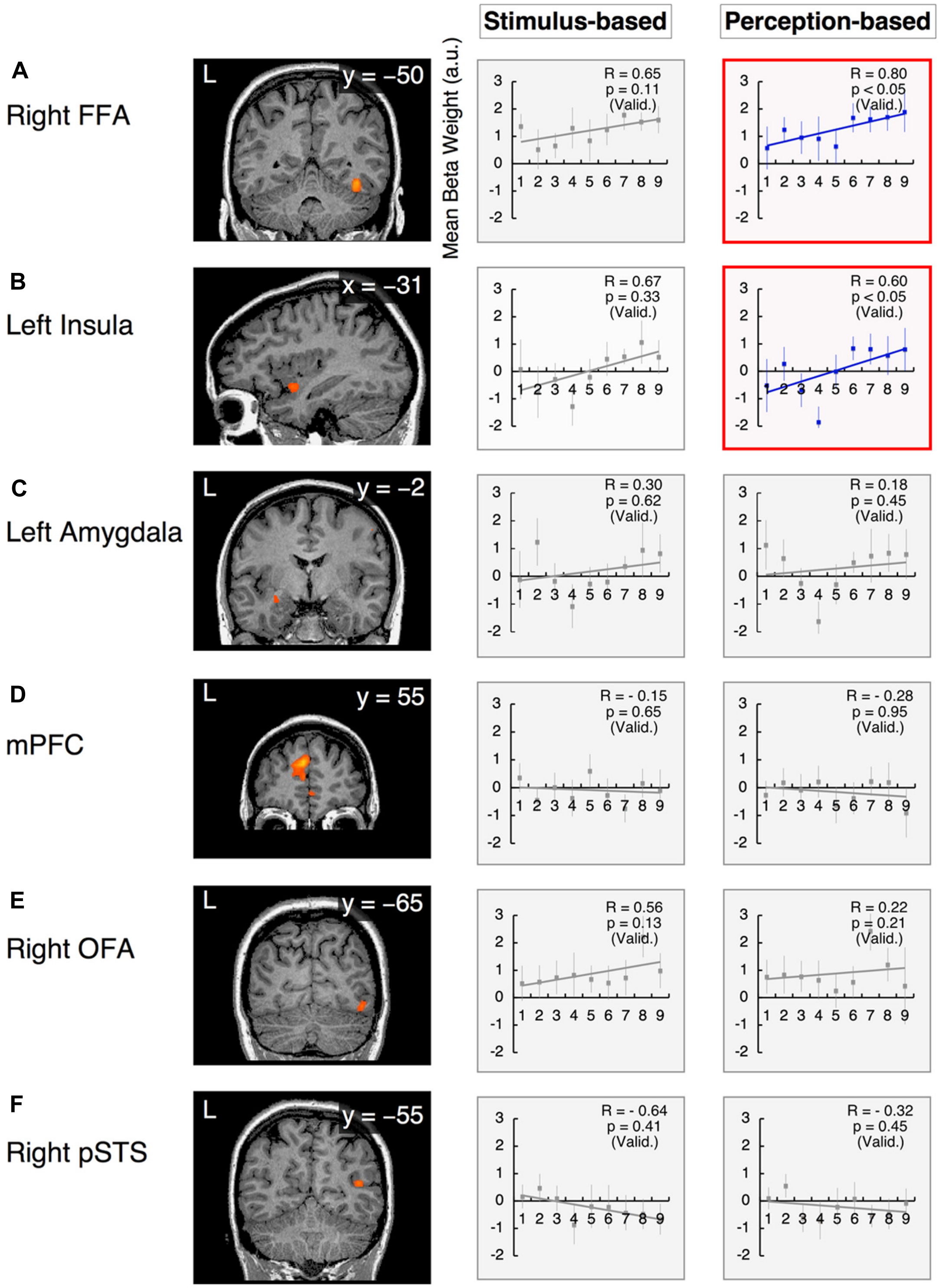


**FIGURE 7 | Functional magnetic resonance imaging results in the arousal dimension.** The *x*-values represent morphed faces on the anger–disgust continuum depicted in **Figure 2B**, ranging from "1" = "0% anger" to "9" = "100% anger." The *y*-axes represent the BOLD signals [with arbitrary units (a.u.)]. Two graphs are provided for each ROI to indicate the results from the stimulus-based and perception-based analyses. Graphs in red squares represent dimensional (linear) processing with statistically significant correlations after the subject-wise validation (**Table 3**). Graphs in gray squares represent constant processing without statistically significant correlations. Error bars denote SEM. R: the group-wise correlation coefficient. p (Valid.): probability after the subject-wise validation (**Table 3**). **(A)** The right FFA; **(B)** the left insula; **(C)** the left amygdala; **(D)** the mPFC; **(E)** the right OFA; and **(F)** the right pSTS.

## DISCUSSION

The current investigation is the first study to simultaneously investigate the implicit processing of both categorical and dimensional aspects of facial emotion perception. Our experimental design dissociates the two fundamental dimensions of valence and arousal in the emotional space by establishing two orthogonal continua of morphed stimuli that parallel these dimensions. Because the four endpoints of the morphed continua represented four basic emotions (happiness, fear, anger, and disgust), this design enabled the interpretation of morphed stimuli from both categorical and dimensional perspectives. The results obtained from an fMRI study involving the passive viewing of emotional stimuli along the examined continua confirmed our hypothesis that the implicit encoding of categorical and dimensional (linear) processing in separate neural loci serves as a basis for our psychological observations of hybrid processing. In the valence dimension (the happiness–fear continuum), the right FFA exhibited categorical processing, whereas the left insula, left amygdala and mPFC exhibited linear processing. However, in the arousal dimension (the anger–disgust continuum), the right FFA and left insula demonstrated linear processing, and none of the examined loci demonstrated clearly categorical processing. Notably, the right FFA was the only area to reflect our psychological results, which indicated that categorical processing occurred in the valence dimension and linear processing occurred in the arousal dimension in an implicit manner (**Figure 3**; Fujimura et al., 2012).

### CATEGORICAL PROCESSING IN THE VALENCE DIMENSION

Categorical processing in the right FFA became more obvious in the group-wise profile after the fMRI data from individual subjects were realigned to conduct perception-based analyses based on the category boundaries established for each subject from the behavioral results (**Figure 6A**). This phenomenon indicates that implicit categorical processing in the right FFA is likely to reflect the explicit behavioral strategy of categorical processing. Given that category boundaries vary by individual (**Table 1**), the right FFA may implicitly process subjective values of category boundaries in the facial continuum of the valence dimension. This conjecture is consistent with the known role of the FFA in processing psychological (i.e., subjective) rather than physical aspects of face stimuli. Moreover, in morphed face experiments involving individual identification tasks, Rotshtein et al. (2005) and Fox et al. (2009b) have determined that the FFA is sensitive to perceived similarity but not physical similarity. It has also been reported that psychological processing in the FFA exhibits the face-inversion effect, which involves a higher response to upright faces than to inverted faces; this difference has been correlated with behavioral effects across subjects (Yovel and Kanwisher, 2005). The findings from the current investigation, in combination with the aforementioned results from published studies and other accumulating evidence supporting the role of the FFA in emotion processing (for a review, see Fusar-Poli et al., 2009), suggest that the FFA processes not only the psychological (i.e., the categorical) aspects of face identification but also contributes to emotion recognition in face perception; this conjecture is consistent with previously reported results (Fox et al., 2009b). In this investigation, we have also confirmed that the implicit categorical processing of facial emotion perception occurs in the valence dimension (the happiness–fear continuum) but not the arousal dimension (the anger–disgust continuum). The right FFA was the only examined brain area to reflect our psychological results indicating the existence of categorical processing in the valence dimension and linear processing in the arousal dimension in an implicit manner (**Figure 3**; Fujimura et al., 2012).

### DIMENSIONAL PROCESSING IN THE VALENCE DIMENSION

Activity in the left insula and left amygdala was negatively correlated with the valence scores of facial emotions; thus, in these regions, increasingly unpleasant faces produced greater neural responses, whereas increasingly pleasant faces produced smaller neural responses (**Figures 6B,C**). Previous attempts to investigate how the response of the insula relates to the valence of a stimulus have generated conflicting results; one study has reported that these traits are negatively correlated (Cunningham et al., 2004), several investigations have stated that these traits are positively correlated (Heinzel et al., 2005; Posner et al., 2009), another study has claimed that both negative and positive correlations between these traits could be observed depending on which insula subregion is examined (Lewis et al., 2007), and one investigation has identified an inverted U-shaped relationship between stimulus valence and insula response (Viinikainen et al., 2010). The discrepancies among these findings might reflect the differing stimulus sets used in the aforementioned investigations. Although conventional stimulus sets [e.g., the International Affective Picture System (IAPS); Lang et al., 2005] span a wide range of emotional space, each stimulus within a conventional set consists of a completely different face or scene. Furthermore, in these stimulus sets, valence and dimensional scores tend to be correlated; in particular, both highly pleasant and highly unpleasant stimuli appear to elicit high arousal. Thus, valence or arousal scores in the emotional space cannot be systematically altered using these stimulus sets, rendering it difficult to functionally dissociate the valence and arousal dimensions in a research investigation. In contrast, our morphed stimulus set utilizes faces with the same identities and has been designed to enable either the valence or arousal score alone to change as the remaining parameter remains fixed. This stimulus set design can ideally dissociate the valence and arousal dimensions of the emotional space, resulting in data sets suitable for subsequent analysis. It has been demonstrated that the insula plays a role in a number of unpleasant affective states, such as anger (Damasio et al., 2000); disgust (Calder et al., 2000); physical pain and social distress (Eisenberger et al., 2003); empathy for others' pain (Singer et al., 2004); and guilt (Shin et al., 2000; for a review, see Craig, 2009). These studies support our results, which indicate a negative correlation between insula response and the valence scores of facial emotions. However, further studies are required to elucidate how insula activity relates to valence scores; in particular, these future investigations could examine each subregion of the insula and/or asymmetry between the left and right hemispheres (Craig, 2009).

Although it has been well documented that the amygdala is activated during the presentation of emotional stimuli (LaBar et al., 1995, 1998; Zald, 2003), the precise role of the amygdala in emotional processing and the direction of amygdala activation remain the subject of considerable controversy (for a review, see Costafreda et al., 2008). For example, the amygdala has been implicated in mediating emotional responses not only to discrete emotions, such as fear (LaBar et al., 1995), but also to the properties of emotional cues, such as valence (Morris et al., 1998; Garavan et al., 2001; Killgore and Yurgelun-Todd, 2004; Fusar-Poli et al., 2009; Gerdes et al., 2010; Viinikainen et al., 2010), intensity (i.e., arousal; Anderson et al., 2003; Small et al., 2003; Glascher and Adolphs, 2003; Anders et al., 2004; Cunningham et al., 2004; Kensinger and Corkin, 2004; Kensinger and Schacter, 2006; Lewis et al., 2007; Gerber et al., 2008) or a combination of both valence and intensity (Anders et al., 2004; Kensinger and Corkin, 2004; Winston et al., 2005). The discrepancies among these findings might reflect the insufficient resolution of conventional fMRI for investigations of the human amygdaloid complex. In fact, recent studies have sought to dissociate the different functional subregions of the amygdala through the use of high-resolution fMRI (Davis et al., 2010; Gamer et al., 2010). Our findings from the current study indicating that the amygdala is preferentially involved in processing valence rather than arousal are consistent with the results of these recent studies, which have demonstrated greater activity in the lateral subregion of the amygdala in response to unpleasant faces than in response to pleasant faces (Davis et al., 2010; Gamer et al., 2010).

The mPFC exhibited positive correlations with valence scores for facial emotions, with increased responses to pleasant faces and decreased responses to unpleasant faces (**Figure 6D**, "Perception-based" results). The mPFC has a well-established role in emotional processing, particularly with respect to reward mechanisms (Knutson et al., 2003; Phillips et al., 2003, 2008; Sabatinelli et al., 2007). The mPFC is more consistently activated when a subject perceives or attempts to modulate a subjective feeling during the perception of an emotionally evocative stimulus than when these processes occur in response to an emotionally neutral stimulus (Lane et al., 1997; Reiman et al., 1997; Ochsner et al., 2002; Phan et al., 2002). Interestingly, the opposing correlations with valence observed in this study for mPFC activity (positively correlated with valence) and amygdala activity (negatively correlated with valence; **Figures 6C,D**) are consistent with the findings from animal models and human imaging studies, which indicate a reciprocal relationship between prefrontal and amygdala activities (Davidson, 2002; Quirk et al., 2003; Likhtik et al., 2005). Our findings suggest that higher levels of neural activity in the mPFC in response to stimuli with greater positive valences are accompanied by concurrent declines in amygdala activity. This result is consistent with the findings of previous studies of functional connectivity that have demonstrated inverse correlations between PFC activity and amygdala activity (Zald et al., 1998; Kim et al., 2003; Quirk et al., 2003). Several investigators have suggested that this reciprocal relationship may represent a regulatory or feedback system that serves to modulate and dampen affective responses that would otherwise be excessive (Garcia et al., 1999; Ochsner et al., 2002).

### DIMENSIONAL PROCESSING IN THE AROUSAL DIMENSION

The right FFA and left insula exhibited positive correlations with the arousal scores of facial emotions, with increased responses to faces with higher arousal levels (**Figures 7A,B**). Previous studies have also proposed that the arousal of facial emotions affects FFA activity (Glascher et al., 2004; Kanwisher and Yovel, 2006; Brassen et al., 2010), although our results extend these findings by suggesting that valence (pleasant–unpleasant) also affects FFA activity (as discussed above). We found that the right FFA differentially processes the arousal and valence dimensions, engaging in dimensional (linear) processing for arousal and categorical processing for valence (**Figures 6A** and **7A**). However, simple arousal effects alone cannot explain our FFA activity results. For example, the arousal score of the 100% disgust face (i.e., the 0% anger face) was lower than the arousal score of the 50% happiness face (i.e., the 50% fear face; **Figure 3B**); however, there was a greater FFA response to the 100% disgust face than to the 50% happiness face ($t_{22} = 3.14$, $p < 0.01$). This phenomenon indicate the integrated nature of arousal and valence processing in the right FFA, which enhances its activity in response to higher-arousal faces and filters out emotionally ambiguous faces in the valence dimension.

However, the left insula demonstrated linear processing in both the arousal and valence dimensions (**Figures 6A** and **7A**). In particular, enhanced left insula activity was observed for faces with greater arousal and greater unpleasantness. These results are consistent with previous reports demonstrating that negative words with higher arousal activated the middle subregion of the left insula (Lewis et al., 2007), which is located in the vicinity of the functional ROI that we identified in the left insula in the current investigation.

### CONSTANT PROCESSING IN THE VALENCE AND AROUSAL DIMENSIONS

The right OFA and right pSTS exhibited constant (or uniform) processing in both the valence and arousal dimensions (**Figures 6E,F** and **7E,F**). The right OFA is known to play an essential role in face recognition, as lesions in the right OFA induce prosopagnosia (Rossion et al., 2003; Sorger et al., 2007). It has also been suggested that the OFA preferentially represents certain facial components, including the eyes, nose, and mouth, at an early stage of visual perception (Pitcher et al., 2011b). In fact, it has been hypothesized that by representing these facial components, the OFA functions as the first stage in a hierarchical face perception network in which increasingly complex facial features are subsequently processed by higher face-selective cortical regions (Haxby et al., 2000). Consistent with these studies, our investigation demonstrated the constant activation of the right OFA to any morphed stimulus of emotional faces, irrespective of the valence and arousal scores of these faces. This phenomenon might reflect the fact that our stimuli involved the same facial identities (i.e., all of the examined emotional expressions utilized the same models); thus, the right OFA might respond similarly to the identical facial components within these stimuli.

However, the right pSTS demonstrated virtually no response to any stimulus in the valence and arousal dimensions when static facial expressions were presented in an implicit manner (**Figures 6F** and **7F**), in contrast to the results observed with explicit processing (Puce et al., 1998; Hoffman and Haxby, 2000; Calder et al., 2007). Interestingly, the right pSTS was responsive when dynamic facial expressions were used as functional localizer stimuli (**Table 2**). One possible explanation for this finding is that even in an implicit context, relative to the traits of static faces, the characteristics of dynamic faces might produce enhanced visual motion analysis by recruiting more attentional resources and inducing stronger pSTS activation (Kilts et al., 2003; LaBar et al., 2003; Sato et al., 2004; Fox et al., 2009a; Foley et al., 2012). In studies of face and body perception, the pSTS has also been associated with the processing of the biological motions of changeable stimulus components (Puce et al., 1998; Allison et al., 2000; Haxby et al., 2000; Grossman and Blake, 2002; Pelphrey et al., 2003). These biological motions of stimuli (e.g., an individual's gaze or changes in facial muscles) apparently convey relatively complex social cues that are critical for adequate social communication and are therefore likely to evoke strong neural activation patterns.

### INDIVIDUAL DIFFERENCES IN EMOTION PROCESSING

When the individuals' fMRI data were realigned based on the obtained behavioral data, certain brain areas demonstrated enhanced group-wise profiles for categorical/dimensional processing, whereas other regions exhibited no enhancement (**Figures 6** and **7**). In particular, the processing profiles of the right FFA and mPFC became more distinct, whereas the processing profile of the left amygdala became more obscure, and the processing profile of the left insula was enhanced in one dimension of the emotional space but obscured in the other dimension. These differences may reflect differences in the hierarchical organization of processing stages in the face-processing network. As discussed above, it is known that the FFA processes the psychological rather than the physical aspects of faces (Rotshtein et al., 2005; Fox et al., 2009b) and that the mPFC is activated when a subject perceives or attempts to regulate the effects of an emotionally evocative stimulus (Lane et al., 1997; Reiman et al., 1997; Ochsner et al., 2002; Phan et al., 2002). These findings suggest that the FFA and mPFC may participate in higher-order stages of face processing, implying that these areas may reflect explicit behavioral strategies. In contrast, the amygdala mediates implicit learning even for emotional stimuli that are not consciously perceived (Morris et al., 1998). The amygdala exhibits anatomically diverse connections to both lower- and higher-order brain areas and is therefore subject to both top-down and bottom-up modulations of its responses (Pessoa and Adolphs, 2010). Our implicit paradigm of facial emotion processing might involve reduced top-down modulation and enhanced bottom-up processing in the amygdala, resulting in amygdala activity that is significantly correlated with the physical changes of facial stimuli (as indicated by the "Stimulus-based" results in **Figure 6C**) but not the psychological changes of these stimuli (as indicated by the "Perception-based" results in **Figure 6C**). An anatomically and functionally graded representation of facial stimuli is thought to form within the structure of the insula, resulting in; a posterior-to-mid-to-anterior pattern for the integration of interoceptive information in this region (Craig, 2009). In particular, the integration of subjective (self-related) and objective (other-related) information in the anterior-/mid-insula may support our finding that there exists a relatively stable yet complex relationship between insula activity and the physical and psychological changes in facial stimuli, with changes in insula activity dependent on alterations in the dimensions of these stimuli (**Figures 6B** and **7B**).

The implicit hybrid processing that we observed might be achieved through the integrated function of separate neural loci. In particular, both types of processing involved in perceiving facial emotions may occur in the cortex for the "psychological" aspects of facial stimuli and in the subcortical/limbic system for the "physical" aspects of these stimuli; these processing mechanisms could function synergistically to produce the hybrid categorical/dimensional processing of facial emotion perceptions.

### FUTURE WORK

We examined the functional dissociation between the two fundamental dimensions of valence and arousal in the emotional space by creating two orthogonal continua of morphed stimuli that were parallel to these dimensions. Although we employed these morphed faces to generate inferences regarding dimensional and categorical perspectives (Fujimura et al., 2012), greater caution might be required before utilizing the morphed stimuli of this study as direct, independent representations of valence and arousal for drawing conclusions regarding neural representations.

(1) The affect grid ratings indicating that the happiness–fear continuum is categorical may simply reflect the fact that happiness is the most obvious positive emotion among the largely negative emotions that comprise the set of basic emotions (with the potential exception of surprise, which tends to be relatively neutral). This phenomenon could indirectly relate to the concept of the valence scale as bivariate scales indicating degrees of positivity and negativity (Cacioppo and Berntson, 1994), which may cause valence to be categorically distinguished more readily than arousal.

(2) Anger and disgust do not span arousal in a valence-neutral manner; instead, both emotions represent negative-valence forms of arousal. An independent measure of arousal would include both positive- and negative-valence expressions. In addition, it has been demonstrated that anger and disgust are structurally similar (Susskind and Anderson, 2008); thus, these two emotions might readily be perceived in a more similar fashion than the distinct opposites of fear and happiness. (However, the distinct categorical separation of disgust and anger may have been emphasized in the identification task of this study by the fact that only two options were offered.)

Thus, it might not be evident that the morphed-face stimuli of this study represent continua of valence and arousal. Instead, it may be more appropriate to clarify that these morphed-face stimuli represent indices of categorical vs. dimensional perception rather than clearly defined delineations. The definitive region-based coding of the valence and arousal dimensions and specific facial expression categories would require the use of a broader range of controlled stimuli.

## ACKNOWLEDGMENTS

This work is supported by funding from the Japan Science and Technology Agency, Exploratory Research for Advanced Technology, Okanoya Emotional Information Project. We thank T.Asamizuya and C. Suzuki for help in data acquisition.

## REFERENCES

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., and Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32, 537–551. doi: 10.1016/S0896-6273(01)00491-3

Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278. doi: 10.1016/S1364-6613(00)01501-1

Anders, S., Lotze, M., Erb, M., Grodd, W., and Birbaumer, N. (2004). Brain activity underlying emotional valence and arousal: a response-related fMRI study. *Hum. Brain Mapp.* 23, 200–209. doi: 10.1002/hbm.20048

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* 6, 196–202. doi: 10.1038/nn1001

Atkinson, A. P., and Adolphs, R. (2011). The neuropsychology of face perception: beyond simple dissociations and functional selectivity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1726–1738. doi: 10.1098/rstb.2010.0349

Brassen, S., Gamer, M., Rose, M., and Buchel, C. (2010). The influence of directed covert attention on emotional face processing. *Neuroimaging* 50, 545–551. doi: 10.1016/j.neuroimage.2009.12.073

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., et al. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 17, 875–887. doi: 10.1016/S0896-6273(00)80219-6

Cacioppo, J. T., and Berntson, G. G. (1994). Relationship between attitudes and evaluative space: a critical review: with emphasis on the separability of positive and negative substrates. *Psychol. Bull.* 115, 401–423. doi: 10.1037/0033-2909.115.3.401

Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., et al. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Curr. Biol.* 17, 20–25. doi: 10.1016/j.cub.2006.10.052

Calder, A. J., Keane, J., Manes, F., Antoun, N., and Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nat. Neurosci.* 3, 1077–1078. doi: 10.1038/80586

Calder, A. J., and Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nat. Rev. Neurosci.* 6, 641–651. doi: 10.1038/nrn1724

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., and Rowland, D. (1996). Categorical perception of morphed facial expression. *Vis. Cogn.* 3, 81–117. doi: 10.1080/713756735

Chen, K., Waggoner, R. A., and Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron* 32, 359–374. doi: 10.1016/S0896-6273(01)00477-9

Christie, I. C., and Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *Int. J. Psychophysiol.* 51, 143–153. doi: 10.1016/j.ijpsycho.2003.08.002

Costafreda, S. G., Brammer, M. J., David, A. S., and Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 pet and fMRI studies. *Brain Res. Rev.* 58, 57–70. doi: 10.1016/j.brainresrev.2007.10.012

Courtney, S. M., Ungerleider, L. G., Keil, K., and Haxby, J. V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb. Cortex* 6, 39–49. doi: 10.1093/cercor/6.1.39

Craig, A. D. (2009). How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Cunningham, W. A., Raye, C. L., and Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *J. Cogn. Neurosci.* 16, 1717–1729. doi: 10.1162/0898929042947919

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* 3, 1049–1056. doi: 10.1038/79871

Davidson, R. J. (2002). Anxiety and affective style: role of prefrontal cortex and amygdala. *Biol. Psychiatry* 51, 68–80. doi: 10.1016/S0006-3223(01)01328-2

Davis, F. C., Johnstone, T., Mazzulla, E. C., Oler, J. A., and Whalen, P. J. (2010). Regional response differences across the human amygdaloid complex during social conditioning. *Cereb. Cortex* 20, 612–621. doi: 10.1093/cercor/bhp126

de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., and Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neurosci. Biobehav. Rev.* 34, 513–527. doi: 10.1016/j.neubiorev.2009.10.008

Eisenberger, N. I., Lieberman, M. D., and Williams, K. D. (2003). Does rejection hurt? An FMRI study of social exclusion. *Science* 302, 290–292. doi: 10.1126/science.1089134

Ekman, P. (1992). Argument for basic emotions. *Cogn. Emot.* 6, 169–200. doi: 10.1080/02699939208411068

Ekman, P., and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* 17, 124–129. doi: 10.1037/h0030377

Ekman, P., and Friesen, W. V. (1976). *Pictures of Facial Affect.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Sorenson, E. R., and Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotions. *Science* 164, 86–88. doi: 10.1126/science.164.3875.86

Etcoff, N. L., and Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition* 44, 227–240. doi: 10.1016/0010-0277(92)90002-Y

Foley, E., Rippon, G., Thai, N. J., Longe, O., and Senior, C. (2012). Dynamic facial expressions evoke distinct activation in the face perception network: a connectivity analysis study. *J. Cogn. Neurosci.* 24, 507–520. doi: 10.1162/jocn_a_00120

Fox, C. J., and Barton, J. J. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Res.* 1127, 80–89. doi: 10.1016/j.brainres.2006.09.104

Fox, C. J., Iaria, G., and Barton, J. J. (2008). Disconnection in prosopagnosia and face processing. *Cortex* 44, 996–1009. doi: 10.1016/j.cortex.2008.04.003

Fox, C. J., Iaria, G., and Barton, J. J. S. (2009a). Defining the face processing network: optimization of the functional localizer in fMRI. *Hum. Brain Mapp.* 30, 1637–1651. doi: 10.1002/hbm.20630

Fox, C. J., Moon, S. Y., Iaria, G., and Barton, J. J. S. (2009b). The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. *NeuroImage* 44, 569–580. doi: 10.1016/j.neuroimage.2008.09.011

Fujimura, T., Matsuda, Y.-T., Katahira, K., Okada, M., and Okanoya, K. (2012). Categorical and dimensional perceptions in decoding emotional facial expressions. *Cogn. Emot.* 26, 587–601. doi: 10.1080/02699931.2011.595391

Furl, N., van Rijsbergen, N. J., Treves, A., and Dolan, R. J. (2007). Face adaptation aftereffects reveal anterior medial temporal cortex role in high level category representation. *Neuroimage* 37, 300–310. doi: 10.1016/j.neuroimage.2007.04.057

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., et al. (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.* 34, 418–432.

Gamer, M., Zurowski, B., and Buchel, C. (2010). Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. *Proc. Nat. Acad. Sci. U.S.A.* 107, 9400–9405. doi: 10.1073/pnas.1000985107

Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., and Goodale, M. A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia* 43, 1645–1654. doi: 10.1016/j.neuropsychologia.2005.01.012

Garavan, H., Pendergrass, J. C., Ross, T. J., Stein, E. A., and Risinger, R. C. (2001). Amygdala response to both positively and negatively valenced stimuli. *Neuroreport* 12, 2779–2783. doi: 10.1097/00001756-200108280-00036

Garcia, R., Vouimba, R. M., Baudry, M., and Thompson, R. F. (1999). The amygdala modulates prefrontal cortex activity relative to conditioned fear. *Nature* 402, 294–296. doi: 10.1038/46286

Gerber, A. J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., et al. (2008). An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia* 46, 2129–2139. doi: 10.1016/j.neuropsychologia.2008.02.032

Gerdes, A. B., Wieser, M. J., Mulberger, A., Weyers, P., Alpers, G. W., Plichta, M. M., et al. (2010). Brain activations to emotional pictures are differentially associated with valence and arousal ratings. *Front. Hum. Neurosci.* 4:175. doi: 10.3389/fnhum.2010.00175

Glascher, J., and Adolphs, R. (2003). Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *J. Neurosci.* 23, 10274–10282.

Glascher, J., Tuscher, O., Weiller, C., and Buchel, C. (2004). Elevated responses to constant facial emotions in different faces in the human amygdala: an fMRI study of facial identity and expression. *BMC Neurosci.* 5:45. doi: 10.1186/1471-2202-5-45

Gorno-Tempini, M. L., Pradelli, S., Serafini, M., Pagnoni, G., Baraldi, P., Porro, C., et al. (2001). Explicit and incidental facial expression processing: an fMRI study. *Neuroimage* 14, 465–473. doi: 10.1006/nimg.2001.0811

Grill-Spector, K., Knouf, N., and Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7, 555–562. doi: 10.1038/nn1224

Grossman, E. D., and Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron* 35, 1167–1175. doi: 10.1016/S0896-6273(02)00897-8

Harris, R. J., Young, A. W., and Andrews, T. J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proc. Nat. Acad. Sci. U.S.A.* 109, 21164–21169. doi: 10.1073/pnas.1212207110

Haxby, J. V., Hoffman, E. A., and Gobbini, I. M. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0

Heinzel, A., Bermpohl, F., Niese, R., Pfennig, A., Pascual-Leone, A., Schlaug, G., et al. (2005). How do we modulate our emotions? Parametric fMRI reveals cortical midline structures as regions specifically involved in the processing of emotional valences. *Brain Res. Cogn. Brain Res.* 25, 348–358. doi: 10.1016/j.cogbrainres.2005.06.009

Hoffman, E. A., and Haxby, J. V. (2000). Distinct representation of eye gaze and identity in the distributed human neural system for face perception. *Nat. Neurosci.* 3, 80–84. doi: 10.1038/71152

Ishai, A. (2008). Let's face it: it's a cortical network. *Neuroimage* 40, 415–419. doi: 10.1016/j.neuroimage.2007.10.040

Ishai, A., Pessoa, L., Bikle, P. C., and Ungerleider, L. G. (2004). Repetition suppression of faces is modulated by emotion. *Proc. Nat. Acad. Sci. U.S.A,* 101, 9827–9832. doi: 10.1073/pnas.0403559101

Ishai, A., Schmidt, C. F., and Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Res. Bull.* 67, 87–93. doi: 10.1016/j.brainresbull.2005.05.027

Ishai, A., Ungerleider, L. G., and Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990. doi: 10.1016/S0896-6273(00)00168-9

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., and Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proc. Nat. Acad. Sci. U.S.A.* 109, 7241–7244. doi: 10.1073/pnas.1200155109

Johnson-Laird, P. N., and Oatley, K. (1992). Basic emotions, rationality, and folk theory. *Cogn. Emot.* 6, 201–223. doi: 10.1080/02699939208411069

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 2109–2128. doi: 10.1098/rstb.2006.1934

Katsikitis, M. (1997). The classification of facial expressions of emotion: a multidimensional-scaling approach. *Perception* 26, 613–626. doi: 10.1068/p260613

Kaufmann, J. M., and Schweinberger, S. R. (2004). Expression influences the recognition of familiar faces. *Perception* 33, 399–408. doi: 10.1068/p5083

Kellman, P., Epstein, F. H., and McVeigh, E. R. (2001). Adaptive sensitivity encoding incorporating temporal filtering (TSENSE). *Magn. Reson. Med.* 45, 846–852. doi: 10.1002/mrm.1113

Kensinger, E. A., and Corkin, S. (2004). Two routes to emotional memory: distinct neural processes for valence and arousal. *Proc. Nat. Acad. Sci. U.S.A.* 101, 3310–3315. doi: 10.1073/pnas.0306408101

Kensinger, E. A., and Schacter, D. L. (2006). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *J. Neurosci.* 26, 2564–2570. doi: 10.1523/JNEUROSCI.5241-05.2006

Killgore, W. D., and Yurgelun-Todd, D. A. (2004). Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. *Neuroimage* 21, 1215–1223. doi: 10.1016/j.neuroimage.2003.12.033

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., and Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage* 18, 156–168. doi: 10.1006/nimg.2002.1323

Kim, H., Somerville, L. H., Johnstone, T., Alexander, A. L., and Whalen, P. J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *Neuroreport* 14, 2317–2322. doi: 10.1097/00001756-200312190-00006

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., and Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage* 18, 263–272. doi: 10.1016/S1053-8119(02)00057-5

Kranz, F., and Ishai, A. (2006). Face perception is modulated by sexual preference. *Curr. Biol.* 16, 63–68. doi: 10.1016/j.cub.2005.10.070

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., and McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cereb. Cortex* 13, 1023–1033. doi: 10.1093/cercor/13.10.1023

LaBar, K. S., LeDoux, J. E., Spencer, D. D., and Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *J. Neurosci.* 15, 6846–6855.

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., and Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* 20, 937–945. doi: 10.1016/S0896-6273(00)80475-4

Lane, R. D., Reiman, E. M., Ahern, G. L., Schwartz, G. E., and Davidson, R. J. (1997). Neuroanatomical correlates of happiness, sadness, and disgust. *Am. J. Psychiatry* 154, 926–933.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., et al. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia* 35, 1437–1444. doi: 10.1016/S0028-3932(97)00070-5

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report a-6.* Gainesville, FL: University of Florida.

Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., and Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *J. Neurosci.* 20, 878–886.

Lewis, P. A., Critchley, H. D., Rotshtein, P., and Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cereb. Cortex* 17, 742–748. doi: 10.1093/cercor/bhk024

Likhtik, E., Pelletier, J. G., Paz, R., and Pare, D. (2005). Prefrontal control of the amygdala. *J. Neurosci.* 25, 7429–7437. doi: 10.1523/JNEUROSCI.2314-05.2005

McLachlan, G. J., and Basford, K. E. (1988). *Mixture Models: Inference and Applications to Clustering.* New York, NY: Marcel Dekker.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., and Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383, 812–815. doi: 10.1038/383812a0

Morris, J. S., Ohman, A., and Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature* 393, 467–470. doi: 10.1038/30976

Ochsner, K. N., Bunge, S. A., Gross, J. J., and Gabrieli, J. D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229. doi: 10.1162/089892902760807212

O'Doherty, J., Winston, J., Critchley, H. D., Perrett, D., Burt, D. M., and Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41, 147–155. doi: 10.1016/S0028-3932(02)00145-8

Palermo, R., and Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia* 45, 75–92. doi: 10.1016/j.neuropsychologia.2006.04.025

Panayiotou, G. (2008). Emotional dimensions reflected in ratings of affective scripts. *Pers. Indiv. Diff.* 44, 1795–1806. doi: 10.1016/j.paid.2008.02.006

Pelphrey, K. A., Singerman, J. D., Allison, T., and McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia* 41, 156–170. doi: 10.1016/S0028-3932(02)00146-X

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* 11, 773–783. doi: 10.1038/nrn2920

Phan, K. L., Wager, T., Taylor, S. F., and Liberzon, I. (2002). Functional neuroanatomy of emotion:

a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16, 331–348. doi: 10.1006/nimg.2002.1087

Phillips, M. L., Drevets, W. C., Rauch, S. L., and Lane, R. (2003). Neurobiology of emotion perception: I. The neural basis of normal emotion perception. *Biol. Psychiatry* 54, 504–514. doi: 10.1016/S0006-3223(03)00168-9

Phillips, M. L., Ladouceur, C. D., and Drevets, W. C. (2008). A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol. Psychiatry* 13, 829–857. doi: 10.1038/mp.2008.82

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature* 389, 495–498. doi: 10.1038/39051

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., and Kanwisher, N. (2011a). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage* 56, 2356–2363. doi: 10.1016/j.neuroimage.2011.03.067

Pitcher, D., Walsh, V., and Duchaine, B. (2011b). The role of the occipital face area in the cortical face perception network. *Exp. Brain Res.* 209, 481–493. doi: 10.1007/s00221-011-2579-1

Posner, J., Russell, J. A., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., et al. (2009). The neurophysiological bases of emotion: an fMRI study of the affective circumplex using emotion-denoting words. *Hum. Brain Mapp.* 30, 883–895. doi: 10.1002/hbm.20553

Pruessmann, K. P., Weiger, M., Scheidegger, M. B., and Boesiger, P. (1999). SENSE: sensitivity encoding for fast MRI. *Magn. Reson. Med.* 42, 952–962. doi: 10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S

Puce, A., Allison, T., Bentin, S., Gore, J. C., and McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* 18, 2188–2199.

Quirk, G. J., Likhtik, E., Pelletier, J. G., and Pare, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *J. Neurosci.* 23, 8800–8807.

Reiman, E. M., Lane, R. D., Ahern, G. L., Schwartz, G. E., Davidson, R. J., Friston, K. J., et al. (1997). Neuroanatomical correlates of externally and internally generated human emotion. *Am. J. Psychiatry* 154, 918–925.

Rosch, E., and Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cogn. Psychol.* 7, 573–605. doi: 10.1016/0010-0285(75)90024-9

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., and Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126, 2381–2395. doi: 10.1093/brain/awg241

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., and Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat. Neurosci.* 8, 107–113. doi: 10.1038/nn1370

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Russell, J. A. (1997). "Reading emotions from and into faces," in *The Psychology Of Facial Expression*, eds J. A. Russell and J. M. Fernandez-Dols (Paris, France: Cambridge University Press), 295–320.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295X.110.1.145

Russell, J. A., and Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *J. Pers. Soc. Psychol.* 48, 1290–1298. doi: 10.1037/0022-3514.48.5.1290

Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. *J. Pers. Soc. Psychol.* 57, 493–502. doi: 10.1037/0022-3514.57.3.493

Sabatinelli, D., Bradley, M. M., Lang, P. J., Costa, V. D., and Versace, F. (2007). Pleasure rather than salience activates human nucleus accumbens and medial prefrontal cortex. *J. Neurophysiol.* 98, 1374–1379. doi: 10.1152/jn.00230.2007

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., and Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cogn. Brain Res.* 20, 81–91. doi: 10.1016/j.cogbrainres.2004.01.008

Schlosberg, H. (1954). The dimensions of emotions. *Psychol. Rev.* 61, 81–88. doi: 10.1037/h0054570

Sergent, J., Ohta, S., and MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* 115, 15–36. doi: 10.1093/brain/115.1.15

Shin, L. M., Dougherty, D. D., Orr, S. P., Pitman, R. K., Lasko, M., Macklin, M. L., et al. (2000). Activation of anterior par- alimbic structures during guilt-related script-driven imagery. *Biol. Psychiatry* 48, 43–50. doi: 10.1016/S0006-3223(00)00251-1

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162. doi: 10.1126/science.1093535

Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., and Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39, 701–711. doi: 10.1016/S0896-6273(03)00467-7

Sorger, B., Goebel, R., Schiltz, C., and Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *Neuroimage* 35, 836–852. doi: 10.1016/j.neuroimage.2006.09.051

Susskind, J. M., and Anderson, A. K. (2008). Facial expression form and function. *Commun. Integr. Biol.* 1, 148–149. doi: 10.4161/cib.1.2.6999

Takehara, T., and Suzuki, N. (1997). Morphed images of basic emotional expressions: ratings on Russell's bipolar field. *Percept. Mot. Skills* 85, 1003–1010. doi: 10.2466/pms.1997.85.3.1003

Takehara, T., and Suzuki, N. (2001). Differential processes of emotion space over time. *North Am. J. Psychol.* 3, 217–228.

Talairach, J., and Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain.* New York: Thieme Medical Publishers.

Tomkins, S. S., and McCarter, R. (1964). What and where are the primary affect? Some evidence for a theory. *Percept. Mot. Skills* 18, 119–158. doi: 10.2466/pms.1964.18.1.119

Viinikainen, M., Jaaskelainen, I. P., Alexandrov, Y., Balk, M. H., Autti, T., and Sams, M. (2010). Nonlinear relationship between emotional valence and brain activity: evidence of separate negative and positive valence dimensions. *Hum. Brain Mapp.* 31, 1030–1040. doi: 10.1002/hbm.20915

Vuilleumier, P., Armony, J. L., Driver, J., and Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841. doi: 10.1016/S0896-6273(01)00328-2

Winston, J. S., Gottfried, J. A., Kilner, J. M., and Dolan, R. J. (2005). Integrated neural representations of odor intensity and affective valence in human amygdala. *J. Neurosci.* 25, 8903–8907. doi: 10.1523/JNEUROSCI.1569-05.2005

Winston, J. S., Henson, R. N., Fine-Goulden, M. R., and Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *J. Neurophysiol.* 92, 1830–1839. doi: 10.1152/jn.00155.2004

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., and Ekman, P. (2002). *Facial Expressions of Emotion Stimuli and Tests (FEEST).* Bury St Edmunds, England: Thames Valley Test Company.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., and Perrett, D. I. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition* 63, 271–313. doi: 10.1016/S0010-0277(97)00003-6

Yovel, G., and Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Curr. Biol.* 15, 2256–2262. doi: 10.1016/j.cub.2005.10.072

Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res. Brain Res. Rev.* 41, 88–123. doi: 10.1016/S0165-0173(02)00248-5

Zald, D. H., Donndelinger, M. J., and Pardo, J. V. (1998). Elucidating dynamic brain interactions with across-subjects correlational analyses of positron emission tomographic data: the functional connectivity of the amygdala and orbitofrontal cortex during olfactory tasks. *J. Cereb. Blood Flow Metab.* 18, 896–905. doi: 10.1097/00004647-199808000-00010

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 April 2013; accepted: 21 August 2013; published online: 26 September 2013.*

*Citation: Matsuda YT, Fujimura T, Katahira K, Okada M, Ueno K, Cheng K and Okanoya K (2013) The implicit processing of categorical and dimensional strategies: an fMRI study of facial emotion perception. Front. Hum. Neurosci.* **7***:551. doi: 10.3389/fnhum.2013.00551*

*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 26 September 2013
doi: 10.3389/fnhum.2013.00626

# Valence of emotions and moral decision-making: increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults

***Martina Carmona-Perera*[1], *Celia Martí-García*[2], *Miguel Pérez-García*[1,3,4] and *Antonio Verdejo-García*[1,5,6,7]*****

[1] *Department of Personality, Assessment and Psychological Treatment, University of Granada, Granada, Spain*
[2] *School of Health Sciences, University of Granada, Granada, Spain*
[3] *Centro de Investigación Mente, Cerebro y Comportamiento, University of Granada, Granada, Spain*
[4] *Centro de Investigación Biomédica en Red de Salud Mental, University of Granada, Granada, Spain*
[5] *Institute of Neuroscience Federico Olóriz, University of Granada, Armilla, Spain*
[6] *Red de Trastornos Adictivos, Instituto Carlos III, University of Granada, Spain*
[7] *School of Psychology and Psychiatry, Monash University, Melbourne, VIC, Australia*

***Edited by:***
*Maria Ruz, Universidad de Granada, Spain*

***Reviewed by:***
*Maria Ruz, Universidad de Granada, Spain*
*Florian Cova, Swiss Centre for Affective Sciences, University of Geneva, Switzerland*

****Correspondence:***
*Antonio Verdejo-García, School of Psychology and Psychiatry, Monash University, 3800 Wellington Road, Melbourne, VIC, Australia*
*e-mail: antonio.verdejo@monash.edu*

Moral decision-making is a key asset for humans' integration in social contexts, and the way we decide about moral issues seems to be strongly influenced by emotions. For example, individuals with deficits in emotional processing tend to deliver more utilitarian choices (accepting an emotionally aversive action in favor of communitarian well-being). However, little is known about the association between emotional experience and moral-related patterns of choice. We investigated whether subjective reactivity to emotional stimuli, in terms of valence, arousal, and dominance, is associated with moral decision-making in 95 healthy adults. They answered to a set of moral and non-moral dilemmas and assessed emotional experience in valence, arousal and dominance dimensions in response to neutral, pleasant, unpleasant non-moral, and unpleasant moral pictures. Results showed significant correlations between less unpleasantness to negative stimuli, more pleasantness to positive stimuli and higher proportion of utilitarian choices. We also found a positive association between higher arousal ratings to negative moral laden pictures and more utilitarian choices. Low dominance was associated with greater perceived difficulty over moral judgment. These behavioral results are in fitting with the proposed role of emotional experience in moral choice.

**Keywords: moral-decision making, utilitarian choices, moral emotions, valence, arousal**

## INTRODUCTION

Moral decision-making is an essential asset for humans' integration in social contexts. Emotional processes contribute to moral judgment by assigning affective value to the moral decision-making scenarios, thus guiding the distinction between acceptable and inacceptable behaviors (Haidt, 2001). The presentation of hypothetical scenarios involving moral violations typically generate subjective unpleasantness and increased arousal, which are thought to guide subsequent moral appraisals and decisions (Moll et al., 2002a; Ostrosky-Solís et al., 2003; Vélez-García et al., 2003; Harenski and Hamann, 2006). Moreover, the presentation of different types of moral stimuli, including moral-laden pictures (Moll et al., 2002a; Harenski and Hamann, 2006), moral statements (Moll et al., 2002b) or moral dilemmas (Greene et al., 2001, 2004; Heekeren et al., 2003, 2005; Blair, 2007) evoke significant changes in brain networks specialized in emotional processing, such as the ventromedial prefrontal cortex. Conversely, individuals with ventromedial prefrontal dysfunction (by virtue of psychopathology or brain lesions) and emotion processing deficits are typically more prone to endorse utilitarian choices, which maximize the aggregate welfare at the expense of the emotional implications of harming an innocent person (Greene et al., 2001; Koenigs et al., 2007; Carmona-Perera et al., 2012; Young et al., 2012).

According to the dual process theory (Greene, 2007; Greene et al., 2008) utilitarian choices are associated with higher order cognitive control, as illustrated by the impact of cognitive biasing factors on this type of judgments, including reasoning styles (Amit and Greene, 2012), cognitive load (Greene et al., 2008; Moore et al., 2008), priming reflection (Paxton et al., 2011), or attentional bias (Van Dillen et al., 2012). By contrast, deontological choices are preferentially supported by aversive emotional processing (Greene et al., 2001, 2004; Koenigs et al., 2007; Moretto et al., 2010; Carmona-Perera et al., 2013a,b). Recent studies have demonstrated that transient manipulation of specific emotions can bias moral decision-making toward utilitarian or deontological choices in response to moral dilemmas. Specifically, several studies have demonstrated that the induction of positively valenced emotions (e.g., happiness, humorous) favors the tendency to endorse utilitarian choices, whereas the induction of negatively valenced emotions (e.g, sadness, disgust) favors the tendency to endorse deontological choices (Wheatley and Haidt, 2005; Valdesolo and DeSteno, 2006; Schnall et al., 2008; Pastötter et al., 2012; Van Dillen et al., 2012). Complementarily, several studies have shown that the motivational tendency primed by the specific emotion induced is significantly associated with utilitarian vs. deontological choices in moral dilemmas. Specifically, the induction of approach-related emotions (e.g., anger) fosters the tendency to endorse utilitarian choices, whereas the induction of avoidance-related emotions (e.g., disgust) fosters the tendency to endorse deontological choices (Harlé and Sanfey, 2010; Ugazio et al., 2012). Although these studies elegantly show how transient manipulations of particular emotions can bias moral decision-making in different directions, considerably less is known about how more stable individual differences in emotional experience (in response to a range of emotionally competent stimuli) are associated with decision-making patterns in moral vs. non-moral scenarios. The Lang bio-informational model of emotion assumes that individual differences in emotional experience can be reliably and efficiently tracked using the subjective responses to emotional stimuli on three relevant aspects of emotion: valence (pleasantness/unpleasantness of the experience), arousal (activation generated by the experience), and dominance (degree of control that one is able to exert over the emotional experience induced; Greenwald et al., 1989; Lang et al., 1993). In this dimensional system, categorical emotions are quantitatively represented; for example, anger would be linked to high unpleasantness, high arousal and high dominance, whereas fear would be linked to high unpleasantness, high arousal but low dominance.

In this study we aimed to investigate whether individual differences in emotional experience, based on the Lang model, are associated with individual differences in moral decision-making patterns, as measured by a battery of moral (and non-moral) dilemmas (Greene et al., 2001). Specifically, we examined whether individual differences in subjective reactivity to affective stimuli is specifically associated with moral (vs. non-moral) decision-making in healthy adults, and whether individual variations in the valence, arousal and dominance subjective emotional ratings are associated with specific utilitarian vs. deontological choice patterns. Based on the previous literature, we hypothesize that (1) individual differences in emotional experience will be specifically correlated with decision-making in moral but not in non-moral scenarios; (2) subjective ratings indexing greater unpleasantness, high arousal and low dominance emotional experience will be associated with predominantly deontological choice patterns, whereas subjective ratings indexing lower experience of unpleasantness, low arousal and high dominance will be associated with predominantly utilitarian choice patterns.

## MATERIALS AND METHODS

### PARTICIPANTS

The sample consisted of 95 healthy adults (49 males and 46 females). All participants were of European-Caucasian origin and were recruited from local community and recreational centers during the first semester of 2011 through flyers-based advertisement and word-of-mouth communication. Eligibility criteria were defined as follows: (i) to be literate enough to ensure reading comprehension, and correctly complete the tests; (ii) not having lifetime use of illegal drugs in more than five occasions or current or past diagnoses of substance dependence (with the exception of nicotine); (iii) not having history of head injury or neurological disorders; (iv) and not having clinically significant psychiatric symptoms. The Interview for Research on Addictive Behaviour (IRAB; Verdejo-García et al., 2005) was used to assess compliance with the absence of drug use/dependence criterion, and the Symptom Checklist-90-Revised (SCL-90-R; Derogatis, 1977) was used to assess compliance with the absence of significant psychiatric symptoms criterion. The sample had a mean (standard deviation) of 49 years-old (10.67) and 18 years of education (2.38). Socioeconomic status was assessed through occupation prestige and mean family income (through self-reports). We classified the participants into three socioeconomic categories: low level (17.9%), average level (63.2%) and high level (17.9%). None of these demographic variables affected moral decision-making (all $p > 0.05$).

### INSTRUMENTS

#### *Emotional experience task*

We used a set of 40 picture stimuli extracted from the International Affective Picture System (IAPS; Lang et al., 1988) and other sources such as the internet. Based on the IAPS norms (Lang et al., 1988), we defined four picture categories or conditions of interest: (i) neutral (10 pictures displaying landscapes, household objects), (ii) pleasant (10 pictures displaying sexual and radical sports scenes), (iii) unpleasant non-moral laden (10 pictures displaying accident-related casualties or mutilations), and (iv) unpleasant moral laden (10 pictures displaying poverty or one to one violence scenes). Since moral content is not addressed in the IAPS norms, we conducted a pilot study ($n = 83$ undergraduate students) to evaluate "perceived moral content" in an initial pool of 22 images with suitable contents for the unpleasant moral laden category. The 10 images with higher "perceived moral content" ratings (>7.5 in a 1–10 range) were finally included in this (iv) category. As a further check the 40 images included in the emotional experience task were also evaluated for "perceived moral content" by the study sample ($n = 95$), and we confirmed that the 10 images included in this category significantly differed from those included in the other categories on "perceived moral content" [$F(3, 282)$ 721.20, $p < 0.001$]. The main dependent measure for each of the picture categories were the subjective ratings of valence (from 1 –unpleasant– to 9 –pleasant–), arousal (from 1 –relaxed– to 9 –aroused–), and dominance (from 1 –dominant– to 9 –dominated–). The responses were recorded using the Self-Assessment Manikin (SAM; Lang, 1980). As dependent variables we used the mean of valence, arousal and dominance scores for each of the four categories of pictures.

#### *Moral decision-making task*

We used a subset of 32 hypothetical dilemmas extracted from Greene battery (Greene et al., 2001). The original Greene battery was adapted to Spanish language through a back-translation process. The ensuing items were evaluated using Rasch analysis to obtain a briefer construct-valid measure of moral decision-making (Carmona-Perera et al., under review). We used the calibration and item fit tests to remove redundant and low quality items impacted by commonly confounding variables outside moral decision-making (e.g., socio-demographic factors). We also excluded those moral dilemmas that fell at the tails (>95%) of the deontological or utilitarian response distributions, since they are less likely to constitute an actual decision dilemma. The final 32-item Spanish version has demonstrated adequate psychometric properties (Cronbach's alpha = 0.78, Spearman Brown coefficient = 0.76; Carmona-Perera et al., under review). This task is composed by eight non-moral dilemmas involving a rational decision without moral content (e.g., to travel by train or bus given certain time constraints, or to buy a new camera or to have your old camera repaired for the same price), and 24 moral dilemmas which concern the appropriateness of moral violations for a higher benefit (e.g., smothering a baby to save a group of people, or throwing a dying person into the sea to keep a lifeboat of survivors afloat). These moral dilemmas involve different degrees of emotional salience based on the extent of personal involvement and the ensuing severity of harm (Greene et al., 2001, 2004). Therefore, the task included both Personal dilemmas (16 items) which involve higher emotional salience and Impersonal dilemmas (8 items) which involve lower emotional salience. Participants were asked to provide "choice" (affirmative vs. negative) and "perceived difficulty" responses (from 1 –low– to 10 –high–). For moral dilemmas affirmative answers were considered "utilitarian" (e.g., to kill someone to save a group of people), and negative answers "deontological" (e.g., to refuse the harmful action regardless the aggregate well-being). For non-moral dilemmas affirmative answers were considered "efficient" (e.g., to travel by the fastest transport to arrive on time), and negative answers "non-efficient" (e.g., to travel by the preferred transport despite off to arrive late). The proportion of affirmative choices and the mean of perceived level of difficulty for moral and non-moral scenarios were computed as main dependent variables.

### PROCEDURE

This study was approved by the Ethics Committee for Human Research of the University of Granada. Before testing, all participants were informed about the study protocols and they signed a written informed consent to certificate their voluntary collaboration. The information sheet included the following information: "We are interested in exploring how you make decisions in relation to a set of moral and non-moral hypothetical scenarios, and how you experience emotions in relation to a set of affective pictures. We will ask you to decide whether you would accept or refuse to take a proposed action concerning moral and non-moral scenarios. In a separate task, we will ask you to report your subjective emotional experience in response to both pleasant and unpleasant stimuli." To describe each SAM scale we used the standardized guidelines of Lang et al. (2001). Participants were assessed individually in a single session that lasted approximately 90 min. The emotional experience and the moral decision-making tasks were administered in computerized format using two different orders, such that half of the sample performed first the dilemmas and then the pictures and half of the sample did it in the reverse sequence. In the emotional experience task categories were presented in a counter-balanced order across participants. In all cases, each picture was presented during 6 s, followed by a 2 s black screen with a fixation cross. Participants were instructed to stare at the picture and rate their emotional experience using the SAM scales of valence, arousal and dominance, with no time limits established. In the moral dilemmas task the different subsets of dilemmas (moral personal, moral impersonal, non-moral) were also presented in a counter-balanced order. Each was presented through three subsequent computer screens. The first screen described the dilemma (presentation); the second screen presented the response options and requested the choice (decision-making), and the third screen presented the difficulty scale and requested the perceived difficulty rating. Each screen continued with no time limit as the participants read and responded to the dilemmas.

### STATISTICAL ANALYSIS

We used repeated-measures ANOVAs to test the main effects of picture categories on subjective valence, arousal and dominance ratings in the emotional experience task, and of type of dilemma on the affirmative choices and perceived difficulty in the moral decision-making task. Pairwise Bonferrroni *post-hoc* tests were used to examine specific effects driven by the different picture categories and types of dilemmas. To test our main assumptions, we conducted Pearson product-moment correlation analyses between the valence, arousal, and dominance ratings to the images included in the different picture categories, and the choices and difficulty ratings to the moral and non-moral dilemmas. The Bonferroni correction was used to adjust the significance levels of correlation coefficients for multiple comparisons (Curtin and Schulz, 1998). Results are presented reporting the corrected $p$ values.

## RESULTS

### SUBJECTIVE REACTIVITY TO EMOTIONAL STIMULI

Results showed the expected significant differences between the valence, arousal and dominance ratings evoked by the images grouped in the different picture categories (see **Table 1**). Pairwise comparisons showed significant effects in all contrasts, with the exception of the contrast between unpleasant non-moral laden and unpleasant moral laden categories, which had similar valence, arousal and dominance ratings. Moreover, the unpleasant pictures (non-moral and moral) yielded higher unpleasantness and arousing ratings, and lower dominance ratings, than all other conditions.

### DECISION-MAKING AND DIFFICULTY RATINGS TO MORAL AND NON-MORAL DILEMMAS

ANOVA analyses showed significant differences between moral and non-moral dilemmas in terms of affirmative choices [$F(1, 94) = 593.82$, $p < 0.001$], and difficulty [$F(1, 94)$ 346.34, $p < 0.001$]. Moral dilemmas yielded less affirmative choices ($M = 58.72$, SD = 13.78) and higher perceived difficulty ($M = 4.08$, SD = 1.38) than non-moral dilemmas ($M = 97.24$, SD = 5.82, and $M = 1.58$, SD = 0.64, respectively). We also found significant differences between personal and impersonal moral dilemmas on utilitarian choices [$F(1, 94)$ 765.56, $p < 0.001$] and difficulty ratings [$F(1, 94)$ 51.81, $p < 0.001$]. Personal dilemmas yielded less utilitarian choices (personal: $M = 28.62$, SD = 20.01; impersonal: $M = 88.82$, SD = 14.29) and higher perceived

**Table 1 | Descriptive scores, ANOVA and *post-hoc* comparisons for emotional valence, arousal, and dominance.**

| Subjective reactivity | Picture ctegories, Mean (SD) | | | | *F* | ***Post-hoc*** Non-significant (*p*) |
|---|---|---|---|---|---|---|
| | **Neutral (1)** | **Pleasant (2)** | **Unpleasant non-moral laden (3)** | **Unpleasant moral laden (4)** | | |
| Valence | 5.65 (0.86) | 7.47 (0.90) | 1.94 (0.72) | 1.96 (0.73) | 143.97* | $3 = 4$ ($p = 1.000$) |
| Arousal | 3.97 (1.22) | 5.82 (1.60) | 7.46 (0.86) | 7.60 (0.83) | 213.18* | $3 = 4$ ($p = 0.511$) |
| Dominance | 8.42 (0.90) | 7.49 (1.43) | 5.26 (2.27) | 5.07 (2.35) | 148.67* | $3 = 4$ ($p = 1.000$) |

*F value associated to the tests of within-subjects effects (Huynh-Feldt correction); *p value <0.01*

difficulty (personal: $M = 4.69$, SD = 1.70; impersonal: $M = 3.48$, SD = 1.50).

### ASSOCIATION BETWEEN SUBJECTIVE REACTIVITY TO EMOTIONAL STIMULI AND UTILITARIAN CHOICES AND DIFFICULTY RATINGS TO DILEMMAS

Results showed that the subjective ratings evoked by the emotional stimuli were specifically associated with decision-making in moral, but not in non-moral, scenarios. The proportion of affirmative (utilitarian) choices in moral dilemmas (merging both personal and impersonal dilemmas) correlated with both valence and arousal ratings (see **Figure 1**). However, we found non-significant correlations between the proportion of affirmative (efficient) choices in non-moral dilemmas and the subjective ratings of valence, arousal or dominance (all $p \geq 0.170$). Separate correlations between personal or impersonal moral dilemmas and the emotional experience task failed to show significant effects in all picture categories for valence ($p \geq 0.181$), arousal ($p \geq 0.096$) and dominance ratings ($p \geq 0.235$).

Significant correlations between moral (personal and impersonal) decisions and emotional experience indicated that moral choices were associated with valence ratings to both unpleasant (moral: $r = -0.29$, $p = 0.016$; non-moral: $r = -0.26$, $p = 0.043$) and pleasant images ($r = 0.26$, $p = 0.047$); experiencing less unpleasantness in response to unpleasant images (both moral and non-moral), and more pleasantness in response to pleasant images were associated with more utilitarian choices (see **Figure 1A**). Moral choices were also associated with arousal ratings to unpleasant moral laden images ($r = 0.34$, $p = 0.004$); higher arousal responses correlated with more utilitarian choices (see **Figure 1B**). The perceived difficulty ratings to the moral dilemmas were negatively correlated with dominance ratings across the moral and non-moral negative picture categories (unpleasant non-moral, $r = -0.26$; $p = 0.043$; and unpleasant moral, $r = -0.29$; $p = 0.016$). Conversely, moral choices or difficulty ratings failed to correlate with the perceived moral content of the images.

## DISCUSSION

The main findings of this study are the following: (1) individual differences in self-reported emotional experience correlate with decision-making in moral scenarios, but not in non-moral scenarios; (2) lower experience of unpleasantness to both moral and non-moral unpleasant images and higher experience of pleasantness to pleasant images are associated with utilitarian choice patterns; (3) higher experience of arousal (specifically in response to moral laden images) are associated with more utilitarian

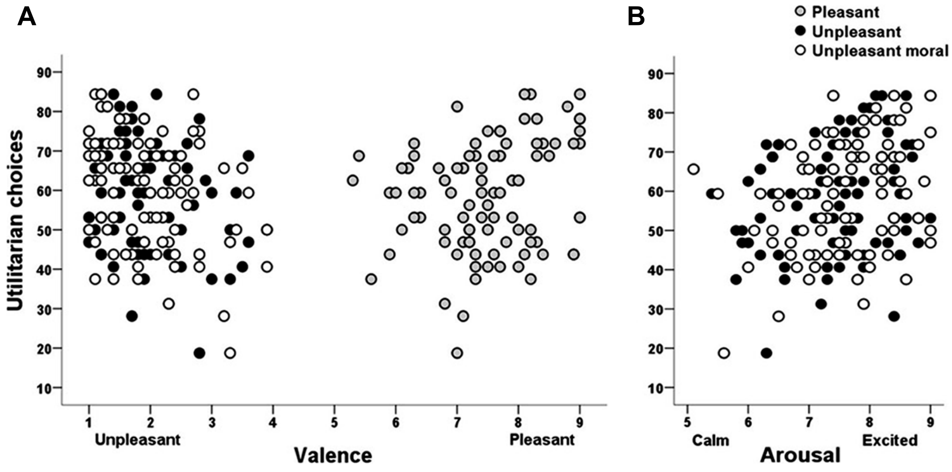


**FIGURE 1 | Dispersion graph of correlation between proportion of utilitarian choices in moral dilemmas and subjective reactivity of valence (A) and arousal (B) to affective stimuli.**

choices; and (4) lower dominance over emotions is significantly associated with higher perceived difficulty to make decisions in moral scenarios. In agreement with our initial hypotheses, these findings support the specific association between emotional experience and moral decision-making, and support the notion that diminished experience of unpleasantness favors utilitarian choice patterns. The association between higher arousal to unpleasant moral laden pictures and utilitarian choices, and between low dominance and higher moral difficulty were not originally predicted and may warrant further research.

In agreement with previous findings, we showed that decision-making in healthy populations is sensitive to the impact of moral vs. non-moral content scenarios (Moll et al., 2001; Harenski and Hamann, 2006; Tassy et al., 2013; Van Bavel et al., 2013), and to the impact of personal vs. impersonal involvement within these moral scenarios (Greene et al., 2001, 2004; Moretto et al., 2010; Koenigs et al., 2007; Carmona-Perera et al., 2013a). Difficulty of judgment may also contribute to describe the emotional weight attached to these choices. For example to push a button to kill someone (low emotional salience) is considered easier than to push a person to the train tracks (high emotional salience). Therefore, participants demonstrated sensitivity to moral content, and to the degree of emotional salience associated with this content.

Correlation analyses showed that moral-related patterns of choice (including both personal and impersonal dilemmas) correlate with subjective emotional experience, at difference with non-moral related decisions. Because separate consideration of personal and impersonal dilemmas did not result in significant correlations with emotional experience, our results can only speak of the association between moral-related decisions and emotional experience. In this respect, our findings are in agreement with those of previous studies that have demonstrated associations between the processing of moral (vs. non-moral) contents and emotional reactivity (Moll et al., 2001; Harenski and Hamann, 2006; Tassy et al., 2013; Van Bavel et al., 2013). The direction of the significant correlations between higher subjective valence ratings and higher proportion of utilitarian choices in moral dilemmas are in agreement with the specific role of emotional processes in moral decision-making (Greene et al., 2001; Haidt, 2001). The dual process model of moral judgment posits that a decreased sensitivity to the negative emotional input attached to moral violations may foster utilitarian choice patterns (Greene, 2007; Greene et al., 2008). Therefore, it is expectable that those individuals with less ability to experience unpleasantness are more prone to endorse utilitarian choices. The findings can also be theoretically accounted by the "undoing hypothesis," which proposes that positive moods can "undo" the cognitive and physiological effects of negative emotions, thus decreasing experience of unpleasantness and increasing utilitarian biases (Fredrickson et al., 2000; Fredrickson and Branigan, 2005). These findings are also in agreement with a plethora of previous evidence demonstrating that induction of positive emotions reliably bias moral decision-making toward utilitarian patterns (Pastötter et al., 2012; Valdesolo and DeSteno, 2006).

In partial disagreement with our initial hypothesis (lower arousal associated with utilitarian choices) we found a positive correlation between higher arousal ratings to unpleasant moral laden pictures and higher proportion of utilitarian choices. These findings can be accounted by the inverted U-shaped association between arousal and decision-making, whereby moderate levels of arousal are optimal to process the emotional input that is relevant for decision-making, but too much or too little arousal become disrupting (Miu et al., 2008). Specifically, it has been demonstrated that high levels of arousal are associated with reduced ability to detect the relevant aspects of emotional input in the context of emotion regulation for dilemmas-solving (Blair et al., 2012). Therefore, we tentatively suggest that higher arousal sensitivity may be associated with greater influence of the emotional information that is irrelevant to address the moral dilemmas. Alternatively, these findings could be interpreted in the context of attentional control models of emotions, which postulate an attentional interference effects due to a higher arousal levels (Schimmack and Derryberry, 2005; Jefferies et al., 2008). Decreased attentional control has been recently linked to utilitarian choices (Van Dillen et al., 2012). Therefore, attentional process may also account for these findings, playing a moderator role between emotional experience and utilitarian choices.

An additional interesting finding was the association between lower dominance over emotions and higher perceived difficulty to decide about the moral dilemmas. Previous cognitive neuroscience studies have associated individual differences in emotional regulation with moral decision-making, identifying that lower emotional control increase the difficulty to decide, since the individual is driven into a more exhaustive appraisal process (Harenski et al., 2009; Bartels and Rips, 2010; Koven, 2011). Lower dominance ratings are associated with lower emotional control in the perceived situation (Bradley and Lang, 1994), such that our results agree with the notion that lower emotional control associates with more complex (more highly difficult-perceived) appraisal processes.

In summary, we provide novel evidence about the association between subjective emotional experience and moral decision-making in a community sample. The strengths of the study include the use of a representative community sample from the healthy population, the use of well-validated quantitative measures of emotional experience and moral decision-making, and the potential relevance of our findings for clinical implications. Because we show that variations in emotional experience, but not in subjective perceptions of moral content, are associated with utilitarian biases, we reason that the interventions for individuals with moral judgment problems should focus on training and shaping emotional response, rather than working on the "rules" characterizing moral violations. These type of emotional interventions may be useful to restore social decision-making in patients with acquired brain injuries (Koenigs et al., 2007; Moretto et al., 2010), psychopathy (Blair, 2007; Young et al., 2012) or drug addictions (Carmona-Perera et al., 2012; Khemiri et al., 2012). Our results should be also interpreted in the context of its relevant limitations. First, personal and impersonal moral dilemmas (differing on emotional salience) did not differentially correlate with emotional experience. Hence, future studies are warranted to explore whether our findings can be replicated in more heterogeneous samples allowing further variance within personal and impersonal categories. Second, since emotional input impacts not only on moral choice but also on a range of other decision-making processes (Paulus and Yu, 2012) future studies are also warranted to determine whether reported associations apply only to utilitarian vs. deontological moral decision-making choices, or to a wider spectrum of decision-making scenarios. An additional limitation is the choice to base the emotional measurement only on subjective responses, disregarding complementary physiological or external behavioral indices (Lang et al., 1993) that should be included in future studies; and the non-measurement of some potential moderators of the link between emotion and moral decision-making – e.g., cognitive processes (Greene et al., 2008; Moore et al., 2008; Paxton et al., 2011; Amit and Greene, 2012; Van Dillen et al., 2012), personality traits (Bartels and Pizarro, 2011), or desirability to social and experimental demands (Lumer, 2010; Hess and Kotter-Grühn, 2011; Caravita et al., 2012).

## ACKNOWLEDGMENTS

This research was supported by the "Red de Trastornos Adictivos," RETICS Program, Instituto de Salud Carlos III, Spanish Ministry of Health (PI: AVG) and the Junta de Andalucía under the Research Project P07.HUM 03089 (PI: MPG). MCP is funded by FPU predoctoral research grant (AP 2008-01848) from Spanish Ministry of Education and Science. We would like to thank all participants involved in the study for his collaboration.

## REFERENCES

Amit, E., and Greene, J. D. (2012). You see, the ends don't justify the means: visual imagery and moral judgment. *Psychol. Sci.* 23, 861–868. doi: 10.1177/0956797611434965

Bartels, D. M., and Pizarro, D. A. (2011). The mismeasure of morals: antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition* 121, 154–161. doi: 10.1016/j.cognition.2011.05.010

Bartels, D. M., and Rips, L. J. (2010). Psychological connectedness and intertemporal choice. *J. Exp. Psychol. Gen.* 139, 49–69. doi: 10.1037/a0018062

Blair, K. S., Geraci, M., Smith, B. W., Hollon, N., DeVido, J., Otero, M., et al. (2012). Reduced dorsal anterior cingulate cortical activity during emotional regulation and top-down attentional control in generalized social phobia, generalized anxiety disorder, and comorbid generalized social phobia/generalized anxiety disorder. *Biol. psychiatry* 72, 476–482. doi: 10.1016/j.biopsych.2012.04.013

Blair, R. J. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends Cogn. Sci.* 11, 387–392. doi: 10.1016/j.tics.2007.07.003

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment Manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Caravita, S. C., Gini, G., and Pozzoli, T. (2012). Main and moderated effects of moral cognition and status on bullying and defending. *Aggress. Behav.* 38, 456–468. doi: 10.1002/ab.21447

Carmona-Perera, M., Clark, L., Young, L., Pérez-García, M., and Verdejo-García, A. (2013a). Impaired decoding of fear and disgust predicts utilitarian moral judgment in alcohol-dependent individuals. *Alcohol. Clin. Exp. Res.* (in press).

Carmona-Perera, M., Reyes del Paso, G. A., Pérez-García, M., and Verdejo-García, A. (2013b). Heart rate correlates of utilitarian moral decision-making in alcoholism. *Drug Alcohol Depend.* doi: 10.1016/j.drugalcdep.2013.06.023 [Epub ahead of print].

Carmona-Perera, M., Verdejo-García, A., Young, L., Molina-Fernández, A., and Pérez-García, M. (2012). Moral decision-making in polysubstance dependent individuals. *Drug Alcohol Depend.* 126, 389–392. doi: 10.1016/j.drugalcdep.2012.05.038

Curtin, F., and Schulz, P. (1998). Multiple correlations and Bonferroni's correction. *Biol. Psychiatry* 44, 775–777. doi: 10.1016/S0006-3223(98)00043-2

Derogatis, L. R. (1977). *SCL-90–R: Administration, Scoring and Procedures Manual I for the Revised Version of Other Instruments of the Psychopathology Rating Scale Series.* Baltimore: John Hopkins University.

Fredrickson, B. L., and Branigan, C. A. (2005). Positive emotions broaden the scope of attention and thought–action repertoires. *Cogn. Emot.* 19, 313–332. doi: 10.1080/02699930441000238

Fredrickson, B. L., Mancuso, R. A., Branigan, C., and Tugade, M. M. (2000). The undoing effect of positive emotions. *Motiv. Emot.* 24, 237–258. doi: 10.1023/A:1010796329158

Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends Cogn. Sci.* 11, 322–323; author reply 323–324. doi: 10.1016/j.tics.2007.06.004

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., and Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition* 107, 1144–1154. doi: 10.1016/j.cognition.2007.11.004

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., and Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400. doi: 10.1016/j.neuron.2004.09.027

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108. doi: 10.1126/science.1062872

Greenwald, M. K., Cook, E. W., and Lang, P. J. (1989). Affective judgment and psychophysiological response: dimensional covariation in the evaluation of pictorial stimuli. *J. Psychophysiol.* 3, 51–64

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295X.108.4.814

Harenski, C. L., and Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. *Neuroimage* 30, 313–324. doi: 10.1016/j.neuroimage.2005.09.034

Harenski, C. L., Kim, S. H., and Hamann, S. (2009). Neuroticism and psychopathy predict brain activation during moral and non-moral emotion regulation. *Cogn. Affect. Behav. Neurosci.* 9, 1–15. doi: 10.3758/CABN.9.1.1

Harlé, K. M., and Sanfey, A. G. (2010). Effects of approach and withdrawal motivation on interactive economic decisions. *Cogn. Emot.* 24, 1456–1465. doi: 10.1080/02699930903510220

Heekeren, H. R., Wartenburger, I., Schmidt, H., Prehn, K., Schwintowski, H. P., and Villringer, A. (2005). Influence of bodily harm on neural correlates of semantic and moral decision-making. *Neuroimage* 24, 887–897. doi: 10.1016/j.neuroimage.2004.09.026

Heekeren, H. R., Wartenburger, I., Schmidt, H., Schwintowski, H.-P., and Villringer, A. (2003). An fMRI study of simple ethical decision-making. *Neuroreport* 14, 1215–1219. doi: 10.1097/00001756-200307010-00005

Hess, T. M., and Kotter-Grühn, D. (2011). Social knowledge and goal-based influences on social information processing in adulthood. *Psychol. Aging* 26, 792–802. doi: 10.1037/a0023775

Jefferies, L. N., Smilek, D., Eich, E., and Enns, J. T. (2008). Emotional valence and arousal interact in attentional control. *Psychol. Sci.* 19, 290–295. doi: 10.1111/j.1467-9280.2008.02082.x

Khemiri, L., Guterstam, J., Franck, J., and Jayaram-Lindström, N. (2012). Alcohol dependence associated with increased utilitarian moral judgment: a case control study. *PLoS ONE* 7:e39882. doi: 10.1371/journal.pone.0039882

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature* 446, 908–911. doi: 10.1038/nature05631

Koven, N. S. (2011). Specificity of meta-emotion effects on moral decision-making. *Emotion* 11, 1255–1261. doi: 10.1037/a0025616

Lang, P. J. (1980). "Behavioral treatment and bio-behavioral assessment: computer applications," in *Technology in Mental Health Care Delivery*, ed. T. A. Williams (Norwood, NY: Ablex), 119–137.

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2001). *International Affective Picture System (IAPS): Instruction Manual and Affective Ratings. Technical Report A-5, The Center for Emotion and Motivation.* Gainesville, FL: University of Florida, Center for Research in Psychophysiology.

Lang, P. J., Greenwald, M. K., Bradley, M. M., and Hamm, A. O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30, 261–273. doi: 10.1111/j.1469-8986.1993.tb03352.x

Lang, P. J., Öhman, A., and Vaitl, D. (1988). *The International Affective Picture System [Photographic Slides].* Gainesville, FL: University of Florida, Center for Research in Psychophysiology.

Lumer, C. (2010). Moral desirability and rational decision. *Ethical*

*Theory Moral Pract.* 13, 561–584. doi: 10.1007/s10677-010-9227-x

Miu, A. C., Heilman, R. M., and Houser, D. (2008). Anxiety impairs decision-making: psychophysiological evidence from an Iowa Gambling Task. *Biol. Psychol.* 77, 353–358. doi: 10.1016/j.biopsycho.2007.11.010

Moll, J., De Oliveira-Souza, R., Bramati, I. E., and Grafman, J. (2002b). Functional networks in emotional moral and non-moral social judgments. *Neuroimage* 16, 696–703. doi: 10.1006/nimg.2002.1118

Moll, J., De Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., et al. (2002a). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.

Moll, J., Eslinger, P. J., and De Oliveira-Souza, R. (2001). Frontopolar and anterior temporal cortex activation in a moral judgment task: preliminary functional MRI results in normal subjects. *Arq. Neuropsiquiatr.* 59, 657–664. doi: 10.1590/S0004-282X2001000500001

Moore, A. B., Clark, B. A., and Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychol. Sci.* 19, 549–557. doi: 10.1111/j.1467-9280.2008.02122.x

Moretto, G., Làdavas, E., Mattioli, F., and Di Pellegrino, G. (2010). A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *J. Cogn. Neurosci.* 22, 1888–1899. doi: 10.1162/jocn.2009.21367

Ostrosky-Solís, F., Chayo-Dichy, R., Castillo-Parra, G., Vélez-García, A. E., and Arias-García, N. (2003). Valencia, activación, dominancia y contenido moral, ante estímulos visuales con contenido emocional y moral: un estudio en población mexicana. *Rev. Esp. Neuropsicol.* 5, 213–225.

Pastötter, B., Gleixner, S., Neuhauser, T., and Bäuml, K. H. T. (2012). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition* 126, 373–377. doi: 10.1016/j.cognition.2012.11.003

Paulus, M. P., and Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn. Sci.* 16, 476–483. doi: 10.1016/j.tics.2012.07.009

Paxton, J. M., Ungar, L., and Greene, J. D. (2011). Reflection and reasoning in moral judgment. *Cogn. Sci.* 36, 163–177. doi: 10.1111/j.1551-6709.2011.01210.x

Schimmack, U., and Derryberry, D. (2005). Attentional interference effects of emotional pictures: threat, negativity, or arousal? *Emotion* 5, 55–66. doi: 10.1037/1528-3542.5.1.55

Schnall, S., Haidt, J., Clore, G. L., and Jordan, A. H. (2008). Disgust as embodied moral judgment. *Pers. Soc. Psychol. Bull.* 34, 1096–1109. doi: 10.1177/0146167208317771

Tassy, S., Deruelle, C., Mancini, J., Leistedt, S., and Wicker, B. (2013). High levels of psychopathic traits alters moral choice but not moral judgment. *Front. Hum. Neurosci.* 7:229. doi: 10.3389/fnhum.2013.00229

Ugazio, G., Lamm, C., and Singer, T. (2012). The role of emotions for moral judgments depends on the type of emotion and moral scenario. *Emotion* 12, 579–590. doi: 10.1037/a0024611

Valdesolo, P., and DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychol. Sci.* 17, 476–477. doi: 10.1111/j.1467-9280.2006.01731.x

Van Bavel, J. J., Packer, D. J., Haas, I. J., and Cunningham, W. A. (2013). The importance of moral construal: moral versus non-moral construal elicits faster, more extreme, universal evaluations of the same actions. *PLoS ONE* 7:e48693. doi: 10.1371/journal.pone.0048693

Van Dillen, L. F., Van der Wal, R. C., and Van den Bos, K. (2012). On the role of attention and emotion in morality: attentional control modulates unrelated disgust in moral judgments. *Pers. Soc. Psychol. Bull.* 38, 1222–1231. doi: 10.1177/0146167212448485

Vélez-García, A. E., Chayo-Dichy, R., García, N. A., Castillo-Parra, G., and Ostrosky-Solís, F. (2003). Emociones morales, una batería para su medición. *Rev. Neuropsicol. Neuropsiquiatr. Neurosci.* 5, 189–199.

Verdejo-García, A., López-Torrecillas, F., Aguilar de Arcos, F., and Pérez-García, M. (2005). Differential effects of MDMA, cocaine, and cannabis use severity on distinctive components of the executive functions in polysubstance users: a multiple regression analysis. *Addict. Behav.* 30, 89–101. doi: 10.1016/j.addbeh.2004.04.015

Wheatley, T., and Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychol. Sci.* 16, 780–784. doi: 10.1111/j.1467-9280.2005.01614.x

Young, L., Koenigs, M., Kruepke, M., and Newman, J. P. (2012). Psychopathy increases perceived moral permissibility of accidents. *J. Abnorm. Psychol.* 121, 659–667. doi: 10.1037/a0027489

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 30 April 2013; accepted: 10 September 2013; published online: 26 September 2013.*

*Citation: Carmona-Perera M, Martí-García C, Pérez-García M and Verdejo-García A (2013) Valence of emotions and moral decision-making: increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults. Front. Hum. Neurosci. **7**:626. doi: 10.3389/fnhum.2013.00626*

*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 12 November 2013
doi: 10.3389/fnhum.2013.00745

# Emotional reactions of peers influence decisions about fairness in adolescence

**Eduard T. Klapwijk[1,2,3] *, Sabine Peters[2,3] , Robert R. J. M. Vermeiren[1,3] and Gert-Jan Lelieveld[2,3]**

[1] Department of Child and Adolescent Psychiatry, Curium – Leiden University Medical Centre, Leiden, Netherlands
[2] Institute of Psychology, Leiden University, Leiden, Netherlands
[3] Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands

**Edited by:**
*Maria Ruz, Universidad de Granada, Spain*

**Reviewed by:**
*Maria Ruz, Universidad de Granada, Spain*
*Nora Vetter, Technische Universität Dresden, Germany*

***Correspondence:**
*Eduard T. Klapwijk, Department of Child and Adolescent Psychiatry, Curium – Leiden University Medical Centre, Endegeesterstraatweg 27, 2342 AK, Oegstgeest, Netherlands*
*email: e.t.klapwijk@curium.nl*

During adolescence, peers take on increasing importance, while social skills are still developing. However, how emotions of peers influence social decisions during that age period is insufficiently known. We therefore examined the effects of three different emotional responses (anger, disappointment, happiness) on decisions about fairness in a sample of 156 adolescents aged 12–17 years. Participants received written emotional responses from peers in a version of the Dictator Game to a previous unfair offer. Adolescents reacted with more generous offers after disappointed reactions compared to angry and happy reactions. Furthermore, we found preliminary evidence for developmental differences over adolescence, since older adolescents differentiated more between the three emotions than younger adolescents. In addition, individual differences in social value orientation played a role in decisions after happy reactions of peers to a previous unfair offer, such that participants with a "proself" orientation made more unfair offers to happy peers than "prosocial" participants. Taken together, our findings demonstrate that adolescents take emotions of peers into account when making social decisions, while individual differences in social value orientation affect these decisions, and age seems to influence the nature of the reaction.

**Keywords: interpersonal effects of emotions, adolescence, social value orientation, dictator game, social interactions, development**

## INTRODUCTION

Emotions play a pivotal role in social interactions, particularly during adolescence, a life stage of significant social development (Blakemore, 2008). For instance, research in adults showed that being in a happy mood themselves makes people rely more on available cues and schemas about their interaction partner (Lount, 2010) and makes them more cooperative in negotiations (e.g., Forgas, 1998). In addition, emotions expressed by others also affect the perceiver's behavior in (further) interactions with the expresser (Parkinson, 1996; Van Kleef et al., 2010). These interpersonal effects of others' emotions have not yet been studied in adolescence and are therefore the focus of the current study.

Social functional accounts of emotions (e.g., Frijda, 1986; Parkinson, 1996; Keltner and Haidt, 1999; Morris and Keltner, 2000; Elfenbein, 2007; Van Kleef, 2009; Van Kleef et al., 2010) posit that one important function of emotion expression is communication to influence the behavior of others. Expressed emotions – either facial, verbal, or postural – may trigger affective reactions and inferences about the other person that influence subsequent social interactions with this person (Van Kleef, 2009).

In adults, interpersonal effects of emotions have been fruitfully investigated in negotiation settings and allocation games (e.g., Van Kleef et al., 2004; Kopelman et al., 2006; Lelieveld et al., 2012; for a review, see Van Kleef et al., 2010). Generally, these studies have demonstrated that bargainers react differently to distinct emotional expressions when responding to others. However, the effects of emotions may depend on the situational context and can be influenced by individual personality differences such as social value orientation.

### INTERPERSONAL EFFECTS OF ANGER, DISAPPOINTMENT, AND HAPPINESS

The emotions used in the current study include anger, disappointment and happiness. These three emotions were selected because they enable us to compare both the effects of negative and positive communicated emotions, as well as the effects of different types of negative emotions (anger vs. disappointment). Studying discrete emotional expressions is important since they carry more information than the more vague and diffuse (negative vs. positive) category of moods (Frijda, 1986; Van Kleef et al., 2010). Another reason to include anger, disappointment, and happiness in this study is that we wanted to know if results for adolescents would differ from the results of adults in a previous study by Lelieveld et al. (2013a).

When being confronted with anger, the perceiver's affective reactions may cause reciprocal anger (leading to competition) or complementary fear (leading to concessions; Lelieveld et al., 2012). Furthermore, someone else's anger can also lead to the conclusion that it might be more beneficial to concede before the situation will escalate. Indeed people make higher concessions to angry compared to happy opponents, because of the toughness communicated by the angry person (Van Kleef et al., 2004). However, communicating anger from a low power position (i.e., when having a low influence on the other's outcomes) might backfire; in these situations the angry recipient gets offered less (Van Dijk et al., 2008; Lelieveld et al., 2013a). Although disappointment – like anger – is a negative emotion, its interpersonal effects differ from the effects of anger. The few studies on the interpersonal effects of disappointment have shown that communicated disappointment invokes feelings of guilt (Lelieveld et al., 2013b), leading to higher offers compared to communicated anger (Van Kleef et al., 2006; Lelieveld et al., 2011, 2012). Finally, happy reactions in negotiations may lead to increased liking and higher subsequent offers in cooperative settings. In more competitive settings people will infer that the happy other is satisfied, which will take away the need to concede and encourage lower offers (Van Kleef et al., 2004, 2010).

## INDIVIDUAL DIFFERENCES

The interpersonal effects of emotions may partly depend on individual personality differences. One important personality trait that has consistently been shown to influence one's bargaining behavior (e.g., McClintock and Liebrand, 1988; Parks, 1994; Van Lange et al., 2007), and, more specifically, the impact of emotions of others on one's bargaining position (e.g., Van Kleef and Van Lange, 2008), is a person's social value orientation (SVO; Messick and McClintock, 1968). Social value orientation is the dispositional preference for distributions of outcomes for the benefit of either self or others. Persons with a prosocial SVO try to maximize both one's own outcomes and the other's outcomes and to minimize differences in outcomes for themselves and others. Proselfs are more interested in maximizing differences between themselves and others (i.e., competition) or they are solely concerned about their own outcomes with little or no interest in others' outcomes (i.e., individualism; Van Lange et al., 1997). One previous study found that proselfs were more likely to concede than prosocials after a disappointed reaction from the other player in a multi-round negotiation, likely due to strategic motivations (Van Kleef and Van Lange, 2008). In order to look at the direct effects of the communicated emotions on prosocials and proselfs, we will investigate the interpersonal effects of emotions in a situation without strategic considerations (i.e., in the case of the Dictator Game). In such situations one might expect prosocials to be more affected by other's emotions, because they are more empathic than proselfs (Declerck and Bogaert, 2008). To test this hypothesis, we decided to consider the effects of social value orientation in the current study.

## INTERPERSONAL EFFECTS OF EMOTIONS IN ADOLESCENCE

Research with allocation games in the domain of interpersonal effects of emotions has mainly focused on adults. To our knowledge, no previous study has focused on the interpersonal effects of emotions in adolescence. However, studying this topic during adolescence is relevant for several reasons. First, notable social changes are seen during this life stage. There is an increased focus on peer relationships and an improvement in social skills that are used to form more complex social relations (Steinberg and Morris, 2001). Second, some studies suggest that the capacity to recognize facial emotions of all six basic emotions (i.e., happiness, sadness, anger, fear, disgust, and surprise) is still developing throughout adolescence and into adulthood (e.g., McGivern et al., 2002; Herba and Phillips, 2004; Wade et al., 2006; Thomas et al., 2007). Hence, it might be harder for (younger) adolescents to differentiate between emotions. Third, adolescents increasingly take the situational context, such as the perspective of their interaction partner, into account during social interactions. Changes in fairness views continue to impact sharing in allocation games throughout adolescence (Crone, 2013). For example, when allowed to distribute jointly earned resources, older adolescents share more with peers that earned more (Almås et al., 2010). Older adolescents also make more use of knowledge about the intentions of others when offering money or when evaluating offers made by others (Güroğlu et al., 2009). Moreover, an increase in social cognitive abilities with adolescence has also been shown in other complex perspective-taking tasks (Dumontheil et al., 2010) and the capacity to attribute mental states to others (i.e., theory of mind) has been found to improve during and beyond adolescence. For example, older adolescents are faster at attributing emotional mental states to others than younger adolescents (Keulers et al., 2010) and adolescents make more errors than adults when evaluating complex theory of mind stories (Vetter et al., 2012). In order to bring these findings together, we studied how adolescents use emotional information from peers during allocation games. This can be an important novel approach to characterize the effects of emotions of peers during social interactions in adolescence.

## THE CURRENT STUDY

In the current study, we therefore investigated interpersonal effects of emotions on allocations in adolescence. We used a procedure developed by Lelieveld et al. (2013a), in which we examined participants' choices in a Dictator Game after receiving verbal emotional reactions from a peer (depicting disappointment, anger, or happiness) to a previous unfair offer. In the Dictator Game (Kahneman et al., 1986), one player divides an amount of money between oneself and another player. The other player is forced to accept this – the dictator's – offer. The Dictator Game allows one to study the interpersonal effects of emotions in a clear and controlled setting. Allocators do not need to consider whether a low offer will be rejected (as opposed to the Ultimatum Game, where the other player can reject the offer), which minimizes the interference of strategic motivations.

This study will test the following hypotheses. First, in line with the results from Lelieveld et al. (2013a), we hypothesized that angry reactions from peers to a previous unfair offer would lead to more unfair offers compared to receiving happy statements in response to identical unfair offers (Van Dijk et al., 2008; Lelieveld et al., 2013a). In addition, we expected less unfair offers in reaction to disappointed compared to angry reactions because disappointment leads to a concern for the outcomes of others (Lelieveld et al., 2011, 2012, 2013a). Second, we explored age differences in the amount of unfair offers for the three distinct emotions. Given the increasing incorporation of the situational context with age (Güroğlu et al., 2009; Almås et al., 2010; Dumontheil et al., 2010) and adolescents' heightened susceptibility to peer influence (Gardner and Steinberg, 2005), we explored if older adolescents would differentiate more between the three emotions than younger adolescents. Third, we investigated effects of individual differences in

SVO (i.e., prosocials vs. proselfs). Previous research has shown that the effects of disappointment depend on a person's SVO (Van Kleef and Van Lange, 2008), which we extend by examining the effects of SVO on anger, disappointment as well as happiness. We expected participants with a proself orientation to make more unfair offers compared to participants with a prosocial orientation and to differentiate less between the emotional expressions of others (cf. Van Lange et al., 1997).

## MATERIALS AND METHODS

### PARTICIPANTS

The final sample of participants included 156 adolescents (76 girls) aged 12.41–17.75 years ($M = 15.17$, SD = 1.22) who were recruited from local secondary schools in The Netherlands. All had then experienced a social change from primary to secondary school. Participants were recruited from schools whose populations have common Dutch ethnicity. Self-report indicated that 97.4% of the participants was born in The Netherlands and that 79.6% had two parents with a Dutch background; 22.4% had at least one parent that was born in another country. No information about socioeconomic status and (history of) psychiatric disorders was collected. In total, 209 eligible participants initially took part in the study, of which 19 were excluded because they did not choose an unfair option in the first phase of the experiment (see below). Of the remaining 190 participants, there was incomplete data for 28 participants due to computer problems, who were therefore excluded from further analyses. A 10-min time limited version of the Standard Raven's Progressive Matrices (Raven et al., 1998) was administered to assess fluid reasoning skills in order to obtain an estimate of general intellectual ability. Since we used a time limited version of the Raven, raw test scores were used. Participants who scored more than 2 SD from the mean on this measure were excluded from further analysis ($N = 6$), because we wanted our participants to be of comparable intelligence. Informed consent was obtained from all participants and from primary caregivers.

### EXPERIMENTAL TASK

Using a procedure developed by Lelieveld et al. (2013a), we examined participants' choices in a Dictator Game after receiving emotional reactions to a previous unfair offer from others.

#### *Phase one*

In the first phase of the experiment, participants read a scenario presented on paper in which they negotiated for a company. They were instructed to divide 10 tokens between themselves and another person. Participants chose between two predetermined distributions of the 10 tokens in two different scenarios. First, they had to choose between a 6–4 distribution in favor of themselves and a 5–5 equal distribution. Second, participants chose between a 6–4 distribution in favor of themselves and a 4–6 distribution in favor of the other. The negotiation scenario was meant to create a business setting, in order to assure that most participants chose the 6–4 option in this phase of the study. We wrote in the scenario that we are researchers interested in monetary negotiations by companies that try to maximize their profits, and that we are interested in how youths think about these negotiations. Adolescents were instructed to imagine which option they would choose if they were part of a company and had to make a profit.

Indeed, most participants (190 participants out of 209) chose a 6–4 distribution in at least one of the two scenarios. To ensure credibility of the second phase, only these 190 participants took part in phase two of the experiment.

#### *Phase two*

The second phase took place 1 week after the first phase, to support the credibility of the researchers collecting reactions from others to the offer. In this phase, participants were told that their unfair offer (the 6–4 distribution) was presented to 60 same-aged peers who were given the opportunity to write out their reaction upon receiving the offer. In reality, the 60 reactions were pre-programmed and rated for emotional significance in a pilot-study with adult participants (see Lelieveld et al., 2013a). On each trial, participants read one of the reactions to the unfair offer, presented on a computer screen. Subsequently, they played a version of the Dictator Game with the person who provided the emotional reaction, in which participants were again allocators in a Dictator Game and divided 10 tokens. Participants could now choose between a 7–3 distribution (i.e., 7 tokens for themselves and 3 for the other) and a 5–5 distribution. We did not include a 6–4 distribution, to ensure that a desire to be consistent with the participants' first offer did not influence the results. We emphasized that the (simulated) recipients did not know that their written reactions would be sent back to the participant to ensure that participants trusted the emotional reactions to be non-strategic (i.e., not aimed to influence the participant's new offer). There was no business setting in this phase anymore; participants now played as individuals. They were told that they would play for real money and that after the experiment, three participants from every class would be selected by a computer to get some rounds actually paid out. Since participants were playing for real money (at least some of their rounds could get paid out), we have no reason to assume participants did not believe the cover story. Additionally, participants in the Lelieveld et al. (2013a) did not express any doubts about the cover story either.

On each trial, participants were paired with a different player, whose first name (indicating gender) was provided. Emotional reactions were either angry, disappointed, or happy. Examples of emotional reactions depicting anger were "I feel really angry after receiving this offer," "This annoying person really pisses me off," and "I am starting to get really furious right now." Examples of reactions depicting disappointment were "This really disappoints me," "I expected more from the other person," and "I am really disappointed in the other person." Examples of reactions depicting happiness were "I am really happy with this offer," "The other person made my day," and "This is perfect, I am really satisfied" (see also Lelieveld et al., 2013a).

### QUESTIONNAIRES

#### *Social value orientation*

We measured social value orientation (Messick and McClintock, 1968) with a nine-item questionnaire developed by Van

Lange et al. (1997). Participants can choose from three different options to allocate valuable points between themselves and a recipient. Responses can be categorized in accordance with three different social value orientations: "prosocials," who maximize outcomes for themselves and the recipient and minimize differences between these outcomes; "individualists," who maximize their own outcomes without regard for the recipient's outcomes; and "competitors" who maximize their own outcomes relative to the recipient's outcomes. Participants were classified into one of these categories when they made six or more consistent choices. Participants who did not meet this requirement were excluded from further analyses that investigated relations with SVO. Similar to several previous studies (e.g., De Cremer and Van Vugt, 1999; Olekalns and Smith, 1999; Joireman and Duell, 2005; Van Dijk and De Cremer, 2006; Van Kleef and Van Lange, 2008), we chose to combine individualists and competitors into one category, the "proselfs," in order to compare self-interest and collective interest. In our sample, 35% of the participants were classified as prosocials, 50% as proself, and 15% could not be classified.

## RESULTS

For phase two of the experimental task, several participants informed us that the experiment was too long and that toward the end of the task, it became difficult to still concentrate on the emotional reactions. To ensure that we were analyzing meaningful results, we decided before any analyses were performed to limit our analyses to the first 30 trials. We made this decision based on a trade-off between statistical power and motivation of participants. By analyzing only half the trials, we ensured in the best possible way that participants were sufficiently motivated for all trials while still retaining statistical power. Our analyses on these trials showed that when collapsing all types of emotional reactions together, participants chose an unfair 7–3 distribution in a mean of 49% of the trials (SD = 19%).

To check for differences in unfair choices between the three emotional reactions, we performed a repeated-measures analysis of variance (ANOVA) with emotion (anger vs. disappointment vs. happiness) as a within-subjects variable and percentage of unfair choices as the dependent variable. This analysis yielded a main effect of emotion, $F(2{,}310) = 4.58$, $p = 0.02$, $\eta^2 = 0.03$. Least significant difference (LSD) *post hoc* tests showed that participants chose the unfair option more often when dealing with angry recipients ($M = 51\%$, SD = 33%, $p < 0.001$) or happy recipients ($M = 53\%$, SD = 31%, $p = 0.01$) than when dealing with disappointed recipients ($M = 43\%$, SD = 31%). In other words, disappointed reactions of a peer to a previous unfair offer led to more generous offers than angry or happy reactions. There was no difference in the amount of unfair offers for angry and happy recipients ($p = 0.53$). Although we did not expect any sex differences we explored an effect of sex, which was not found, $F(2{,}308) = 0.52$, $p = 0.52$, $\eta^2 = 0.003$.

To investigate the time course of the responses for the different emotions, we compared the percentage of unfair offers per emotion for the first trial and the last trial (i.e., the 10th trial for each emotion). A repeated-measures ANOVA with emotion and trial number as between-subjects variables indicated that there was a slight interaction between emotion and trial number, $F(2{,}310) = 3.06$, $p = 0.051$. This indicates that there was a marginally significant effect of trial number, i.e., that there were differences in the percentage of unfair offers per emotion over the course of the task. Because this effect is quite small and we averaged over 10 responses per emotion, we expect that the effect of trial number is relatively small.

### AGE EFFECTS

Collapsed over all emotions, no correlation was found between the total amount of unfair distributions and age ($r = 0.10$, $p = 0.21$). We also checked for effects of age for the three emotions separately, by performing a repeated-measures ANOVA with emotion (anger vs. disappointment vs. happiness) as a within-subjects variable and percentage of unfair choices as the dependent variable, with age as a covariate. No effects of age were found, $F(2{,}308) = 1.34$, $p = 0.26$, $\eta^2 = 0.01$. We also divided our sample in three similar sized age groups: young adolescents ($M = 13.75$ years, SD = 0.65 years), mid adolescents ($M = 15.29$ years, SD = 0.32 years), and late adolescents ($M = 16.49$ years, SD = 0.40 years). See **Table 1** for details about the age groups. There was no significant difference between the age groups for sex, $\chi^2 = 2.26$, $p = 0.32$, SVO, $\chi^2 = 0.07$, $p = 0.97$, and Raven scores, $F(2{,}132) = 1.01$, $p = 0.37$, $\eta^2 = 0.02$. No significant interaction was found for the three emotions and age group, $F(4{,}306) = 1.11$, $p = 0.34$, $\eta^2 = 0.01$ (see **Figure 1**). However, based on our expectations that younger adolescents would differentiate less between the different emotions than older adolescents, we looked at the interpersonal effects of emotions for each age group separately. We conducted a repeated-measures ANOVA with emotion (anger vs. disappointment vs. happiness) as a within-subjects variable and percentage of unfair choices as the dependent variable separately for each age group. We found no main effect of emotion in young adolescents, $F(2{,}102) = 0.02$, $p = 0.92$, $\eta^2 < 0.001$, but we did find a trend for mid adolescents, $F(2{,}104) = 3.22$, $p = 0.06$, $\eta^2 = 0.06$, and a significant effect for late adolescents, $F(2{,}100) = 3.57$, $p = 0.05$, $\eta^2 = 0.07$. LSD *post hoc* tests indicated that for young adolescents none of the effects of the emotions differed from each other, that for mid adolescents, disappointment differed from anger ($p = 0.001$) and happiness ($p = 0.04$), and that for late adolescents, disappointment also differed from anger ($p = 0.005$) and happiness ($p = 0.03$). The effects of anger and happiness did not differ from each other in any age group (all *p*s > 0.05). **Figure 1** depicts the means for each age group, and shows that with increasing age, adolescents seem the differentiate more between the three emotions. That is, the emotions seem to affect the adolescents differently with increasing age.

### EFFECTS OF SOCIAL VALUE ORIENTATION

Collapsed over all emotions, prosocials made less unfair offers than proselfs, although this effect is slightly above significance level, $t(130) = 1.92$, $p = 0.057$, $d = 0.34$. Furthermore, a repeated-measures ANOVA with emotion as a within-subjects variable and social value orientation (proself vs. prosocial) as a between-subjects variable yielded a significant interaction effect, $F(2{,}260) = 4.37$, $p = 0.03$, $\eta^2 = 0.03$ (see **Figure 2**). Follow-up *t*-test comparisons indicated that proselfs proposed more unfair

**Table 1 | Demographic characteristics of the three age groups.**

| Age group | Young adolescents | Mid adolescents | Late adolescents |
|---|---|---|---|
| *N* | 52 | 53 | 51 |
| Age range | 12.41–14.69 | 14.78–15.89 | 15.92–17.75 |
| Mean age (SD) | 13.75 (0.65) | 15.29 (0.33) | 16.49 (0.40) |
| *N* females | 22 | 25 | 29 |
| *N* proself SVO | 26 | 26 | 26 |
| *N* prosocial SVO | 17 | 19 | 18 |
| Mean % unfair (SD) | 47.37 (18.29) | 50.13 (20.58) | 49.74 (17.28) |
| Mean raven (SD) | 34.06 (6.13) | 35.62 (5.50) | 35.57 (6.43) |

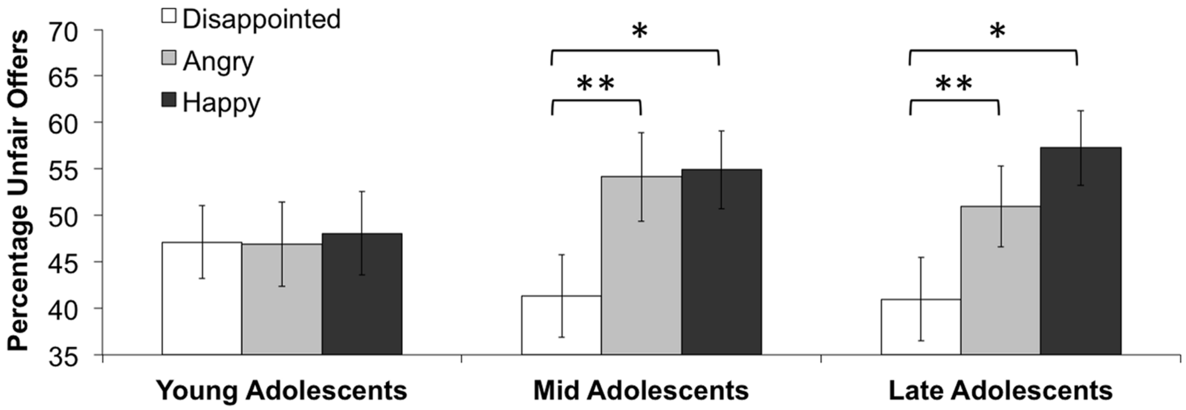


**FIGURE 1 | Percentage of unfair offers for the three emotions, separate for three age groups.** Single asterisk (*) indicates $p < 0.05$; double asterisk (**) indicates $p < 0.01$.

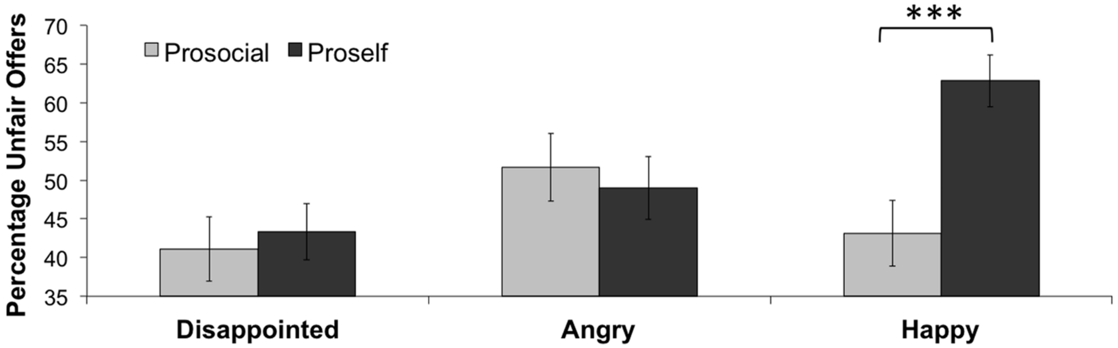


**FIGURE 2 | Percentage of unfair offers for the three emotions, separate for prosocial and proself SVO.** Triple asterisk (***) indicates $p < 0.001$.

offers than prosocials, but only after happy emotional reactions, $t(130) = 3.68$, $p < 0.001$, $d = 0.67$. We found no age effects for this analysis, that is, no further interactions of SVO and age were found. We also performed these analyses separate for the three age groups. We only found an interaction between emotion and SVO in the late adolescent group, $F(2,84) = 4.61$, $p = 0.03$, $\eta^2 = 0.10$. Follow up $t$-test comparisons showed that this effect is again due to proselfs proposing more unfair offers than prosocials, but only after happy reactions, $t(42) = 3.04$, $p = 0.004$, $d = 0.9$.

## DISCUSSION

In this study we investigated adolescents' choices in a Dictator Game after receiving angry, disappointed, and happy emotional reactions from others to an unfair offer. We found that adolescents took information about emotional reactions into account when they were allocating resources. Even though the participant had a lower payoff for the fair than for the unfair option, disappointed reactions led to more generous offers than angry or happy reactions. Happy reactions were more likely to evoke a tendency to act in a self-interested manner, especially amongst participants with a "proself" social value orientation. Young adolescents seemed to differentiate less between the three emotions compared to late adolescents.

### ADOLESCENTS VS. ADULTS

Based on previous research with adult participants (Van Dijk et al., 2008; Van Kleef et al., 2010; Lelieveld et al., 2013a), we hypothesized that anger would lead to more unfair offers than happiness and disappointment. Angry expressions indeed led to more unfair offers than disappointed expressions, which emphasizes the importance of studying discrete emotional expressions instead of valence alone. Both anger and disappointment are reactions to undesirable behavior of others, and both communicate a wish for a behavioral change in the other (Van Dijk and Zeelenberg, 2002). However, a difference between disappointment and anger is that disappointment might lead to more fair offers via a feeling of guilt, whereas anger might instead lead to more unfair offers when communicated from a low-power position (which was the case in our study) because of reciprocal anger (Lelieveld et al., 2011, 2012). Disappointment also differs from anger and happiness because it is regarded more complex than the basic emotions of anger and happiness. However, research suggests that disappointment and anger are two of the most frequently expressed negative emotions (Van Dijk and Zeelenberg, 2002; Lelieveld et al., 2011). We were interested how different negative emotions that signal dissatisfaction would impact offers by adolescents. Future studies may include other basic emotions such as sadness to contrast different basic emotions.

The current results from an adolescent sample differ partly from previous findings using the same paradigm with adults (Lelieveld et al., 2013a). The adolescents in our study made more unfair offers in response to happy and angry reactions compared to disappointed reactions, while the adults in the Lelieveld et al. (2013a) study only made lower offers in response to angry (and not happy) reactions. When reading happy reactions the adolescents possibly inferred that the happy other was satisfied with the previous unfair offer, and accordingly chose more unfair offers (Van Kleef et al., 2010). This is in line with the idea that the general function of positive emotions is to serve as a cue to continue the current course of action (Cacioppo and Gardner, 1999). Previous studies with adults showed this pattern of utilization of happiness as well (Van Kleef et al., 2004). However, the effect of happiness was mainly driven by the proself (SVO) participants in the oldest age group in our sample (see below).

### AGE EFFECTS

Our analyses did not reveal any effects of age on the total amount of unfair offers. Previous studies also found that fairness preferences (as measured with the Dictator Game) do not seem to change substantially after childhood (Gummerum et al., 2008; Güroğlu et al., 2009; Steinbeis et al., 2012). However, since older adolescents are better at considering consequences of their decisions for others (Crone, 2013) and we expected the adolescents to increasingly take the situational context (i.e., different emotional reactions) into account when growing older, we analyzed effects of emotional reactions separately for three age groups. Although we did not find an interaction between age and the offers after the different emotions, we did find differences when we compared the interpersonal effects of the discrete emotions within the separate age groups. This analysis suggested that the youngest adolescents did not differentiate between the three emotions, while older adolescents did. Only the older adolescents made less unfair offers in response to disappointed expressions compared to angry and happy expressions.

Although future studies are needed to confirm these results, the finding that younger adolescents seem to make less use of information conveyed in others' emotions than older adolescents is in line with previous research on emotion recognition and perspective-taking. Possibly, younger adolescents are less able at distinguishing the written emotional expressions, as previously observed for facial expressions of the six basic emotions (i.e., happiness, sadness, anger, fear, disgust, and surprise; Wade et al., 2006) and anger and fear specifically (Thomas et al., 2007). However, we have no reason to believe that adolescents did not perceive the correct emotion from the different reactions, since most disappointed, angry, and happy statements contained the words disappointed, angry, or happy respectively (or any clear synonym). Verbally presented emotions might also be processed differently with increasing age, as was found in younger children (Jenkins and Ball, 2000). In addition, more mixed feelings in response to social emotion scenarios have been linked to pubertal development in young adolescents (Burnett et al., 2011b), suggesting more complex understanding of verbally described emotions with advancing puberty. Hence, the difference between the youngest and oldest adolescents in the current study seems to reflect an improvement with age in understanding emotions and incorporating emotional information into decision-making.

In addition to behavioral evidence, several neuroimaging studies have revealed that brain regions important for social decisions are changing during adolescence (reviewed in Burnett et al., 2011a; Crone and Dahl, 2012). For example, age-related increases in temporo-parietal junction (TPJ) activation were observed when adolescents played the Trust Game, in which it is important to take the perspective of the other player into account (Van den Bos et al., 2010). In adults, heightened TPJ activation in response to happy reactions was found using the paradigm employed in the current study (Lelieveld et al., 2013a). An interesting but speculative direction for future research would be to investigate if the increased amount of unfair offers in response to happiness is related to lower TPJ activity in adolescents compared to adults.

### SOCIAL VALUE ORIENTATION

In order to evaluate individual personality differences in responses to others' emotions we measured social value orientation. The hypothesis that proselfs would make more unfair offers in general than prosocials was confirmed, although the effect was only marginally significant. Additionally, as considered above, the higher amount of unfair offers in response to happy reactions in our study can be attributed to proself (SVO) participants who made more unfair offers than prosocial participants in this condition (see **Figure 2**). It is possible that proselfs perceived the other's happiness with a previous unfair offer as a signal of satisfaction that could be answered with more unfairness. Prosocials may have liked the happy others more and therefore may have made less unfair offers. Previous studies have demonstrated that proselfs tend to be more concerned about their own outcomes in allocation games than prosocials (Parks, 1994; Balliet et al., 2009). In the current study, proselfs tended to be somewhat more selfish in general, but not univocally egocentric. Proselfs did not differ from prosocials in their choices after disappointed and angry expressions. It appears that proselfs only act more selfishly in certain situations (Ketelaar and Au, 2003; Van Kleef and Van Lange, 2008), such as acting upon the perceived satisfaction that was signaled by the happy others in our study. Future research could investigate these underlying mechanisms, and further explore whether SVO also modulates other positive emotions such as gratitude and hope.

### LIMITATIONS AND FUTURE DIRECTIONS

There are several limitations in the current study. First, the small effect of age that we found within our sample should be treated with caution. Our findings require replication, preferably in studies that directly compare children, adolescents, and adults. Second, the emotions in our study were verbal expressions presented on a computer screen, which may lead to concerns about ecological validity. We used verbal expressions since there is no typical facial expression for disappointment. Since emotions can be expressed in a variety of other (non-verbal) ways, future research is needed to confirm whether these effects are generalizable to more natural settings. Previous research suggests that findings with verbally expressed emotions (Van Kleef et al., 2004) are comparable with those of studies using face-to-face negotiations (Kopelman et al., 2006; Sinaceur and Tiedens, 2006) or facial emotional expressions (Pietroni et al., 2008). Furthermore, computer-mediated communication is well-established among adolescents, since the vast majority are Internet and social network users (Madden et al., 2013). However, these online interactions naturally differ from the experimental paradigm that participants engaged in for this study. Future research could investigate if similar results are found with face-to-face interactions for this particular paradigm. Third, participants reported to lose interest in the experiment toward the end of the task. The 60 trials in which participants had to choose between the same distributions may have been too repetitive. Future studies could avoid this problem by alternating between different options for fair and unfair distributions or by simply presenting less than 60 trials (e.g., 30 trials as in the current analyses). In the current study, we chose to discard the second half of the trials before performing any analyses, to avoid the inclusion of trials that are confounded by lower participant motivation and/or concentration. Finally, with the current design, we were not able to study the underlying mechanisms of participants' fairness decisions. Future research may investigate interpersonal effects of emotions in adolescence focused on the underlying (affective) reactions; this could be realized by adding a post-questionnaire to assess feelings of, for example, guilt and anger of participants.

So far, only two studies (with adult participants) have used neuroimaging to elucidate the neural mechanisms underlying the effects of emotions in interactions (Ruz and Tudela, 2011; Lelieveld et al., 2013a). Future studies might use similar experiments to shed light on the developing brain in social interactions. Furthermore, it would be interesting to see whether psychiatric populations known for problems in empathy and emotion recognition might process interpersonal emotions in a different way, such as autism spectrum disorders or psychopathy (Blair, 2005; Kennedy and Adolphs, 2012). In addition, future work may include other discrete emotions such as sadness in order to find out whether there are other emotions that will influence younger adolescents' fairness decisions differently. Furthermore, studies can be designed to examine the effects of emotions in situations beyond allocation games such as cooperative settings.

## CONCLUSION

The current study aimed to elucidate the role of interpersonal emotions in the context of social interaction in adolescence. In general, adolescents reacted with more generous offers after disappointed reactions from peers compared to angry or happy reactions. We also provide preliminary evidence for a developmental increase of incorporating emotional information in social decisions within adolescence. Furthermore, individual differences in SVO were related to reactions upon happiness: proselfs made more unfair offers to happy others than prosocials. Our results emphasize the importance of distinguishing between different types of emotions during social interactions in adolescence and the role of social value orientation. We hope that this study provides a fruitful starting point for the investigation of the interpersonal effects of emotions in adolescence.

## ACKNOWLEDGMENTS

We thank participating schools for their cooperation, and Lisette Burgmeijer, Nina van der Mast, Laura Oosterveer, Rebecca Piët, Malou Stoop, and Marie-Fleur Westenburg for their help with data collection. We also thank Olivier Colins and Geert-Jan Will for comments on an earlier version of this manuscript. This study was supported by the Netherlands Organization for Scientific Research (NWO) Grant No. 056-23-011.

## REFERENCES

Almås, I., Cappelen, A. W., Sorensen, E. O., and Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science* 328, 1176–1178. doi: 10.1126/science.1187300

Balliet, D., Parks, C., and Jeff Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: a meta-analysis. *Group Process. Intergroup Relat.* 12, 533–547. doi: 10.1177/1368430209105040

Blair, R. J. R. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718. doi: 10.1016/j.concog.2005.06.004

Blakemore S. J. (2008). The social brain in adolescence. *Nat. Rev. Neurosci.* 9, 267–277. doi: 10.1038/nrn2353

Burnett, S., Sebastian, C., Cohen Kadosh, K., and Blakemore, S.-J. (2011a). The social brain in adolescence: evidence from functional magnetic resonance imaging and behavioural studies. *Neurosci. Biobehav. Rev.* 35, 1654–1664. doi: 10.1016/j.neubiorev.2010.10.011

Burnett, S., Thompson, S., Bird, G., and Blakemore, S.-J. (2011b). Pubertal development of the understanding of social emotions: implications for education. *Learn. Individ. Differ.* 21, 681–689. doi: 10.1016/j.lindif.2010.05.007

Cacioppo, J. T., and Gardner, W. L. (1999). Emotion. *Annu. Rev. Psychol.* 50, 191–214. doi: 10.1146/annurev.psych.50.1.191

Crone, E. A. (2013). Considerations of fairness in the adolescent brain. *Child Dev. Perspect.* 7, 97–103. doi: 10.1111/cdep.12022

Crone, E. A., and Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nat. Rev. Neurosci.* 13, 636–650. doi: 10.1038/nrn3313

Declerck, C. H., and Bogaert, S. (2008). Social value orientation: related to empathy and the ability to read the mind in the eyes. *J. Soc. Psychol.* 148, 711–726. doi: 10.3200/SOCP.148.6.711-726

De Cremer, D., and Van Vugt, M. (1999). Social identification effects in social dilemmas: a transformation of motives. *Eur. J. Soc. Psychol.* 29, 871–893. doi: 10.1002/(SICI)1099-0992(199911)29:7<871::AID-EJSP962>3.0.CO;2-I

Dumontheil, I., Apperly, I. A., and Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Dev. Sci.* 13, 331–338. doi: 10.1111/j.1467-7687.2009.00888.x

Elfenbein, H. A. (2007). Emotion in organizations: a review and theoretical integration. *Acad. Manag. Ann.* 1, 315–386. doi: 10.1080/078559812

Forgas, J. P. (1998). On feeling good and getting your way: mood effects on negotiator cognition and bargaining strategies. *J. Pers. Soc. Psychol.* 22, 31–42.

Frijda, N. H. (1986). *The Emotions.* Cambridge: Cambridge University Press.

Gardner, M., and Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Dev. Psychol.* 41, 625–635. doi: 10.1037/0012-1649.41.4.625

Gummerum, M., Keller, M., Takezawa, M., and Mata, J. (2008). To give or not to give: children's and adolescents' sharing and moral negotiations in economic decision situations. *Child Dev.* 79, 562–576. doi: 10.1111/j.1467-8624.2008.01143.x

Güroğlu, B., Van den Bos, W., and Crone, E. A. (2009). Fairness considerations: increasing understanding of intentionality in adolescence. *J. Exp. Child Psychol.* 104, 398–409. doi: 10.1016/j.jecp.2009.07.002

Herba, C., and Phillips, M. (2004). Annotation: development of facial expression recognition from childhood to adolescence: behavioural and neurological perspectives. *J. Child Psychol. Psychiatry* 45, 1185–1198. doi: 10.1111/j.1469-7610.2004.00316.x

Jenkins, J. M., and Ball, S. (2000). Distinguishing between negative emotions: children's understanding of the social-regulatory aspects of emotion. *Cogn. Emot.* 14, 261–282. doi: 10.1080/026999300378969

Joireman, J., and Duell, B. (2005). Mother Teresa versus Ebenezer Scrooge: mortality salience leads proselfs to endorse self-transcendent values (unless proselfs are reassured). *Pers. Soc. Psychol. Bull.* 31, 307–320. doi: 10.1177/0146167204271593

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1986). Fairness and the assumptions of economics. *J. Bus.* 59, S285–S300. doi: 10.1086/296367

Keltner, D., and Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cogn. Emot.* 13, 505–521. doi: 10.1080/026999399379168

Kennedy, D. P., and Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends Cogn. Sci. (Regul. Ed.)* 16, 559–572. doi: 10.1016/j.tics.2012.09.006

Ketelaar, T., and Au, W. T. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: an affect-as-information interpretation of the role of emotion in social interaction. *Cogn. Emot.* 17, 429–453. doi: 10.1080/02699930143000662

Keulers, E. H. H., Evers, E. A. T., Stiers, P., and Jolles, J. (2010). Age, sex, and pubertal phase influence mentalizing about emotions and actions in adolescents. *Dev. Neuropsychol.* 35, 555–569. doi: 10.1080/87565641.2010.494920

Kopelman, S., Rosette, A. S., and Thompson, L. (2006). The three faces of Eve: an examination of the strategic display of positive, negative, and neutral emotions in negotiations. *Organ. Behav. Hum. Decis. Process.* 99, 81–101. doi: 10.1016/j.obhdp.2005.08.003

Lelieveld, G.-J., Van Dijk, E., Güroğlu, B., Van Beest, I., Van Kleef, G. A., Rombouts, S. A. R. B., et al. (2013a). Behavioral and neural reactions to emotions of others in the distribution of resources. *Soc. Neurosci.* 8, 52–62. doi: 10.1080/17470919.2012.735621

Lelieveld, G.-J., Van Dijk, E., Van Beest, I., and Van Kleef, G. A. (2013b). Does communicating disappointment in negotiations help or hurt? Solving an apparent inconsistency in the social-functional approach to emotions. *J. Pers. Soc. Psychol.* 105, 605–620. doi: 10.1037/a0033345

Lelieveld, G.-J., Van Dijk, E., Van Beest, I., Steinel, W., and Van Kleef, G. A. (2011). Disappointed in you, angry about your offer: distinct negative emotions induce concessions via different mechanisms. *J. Exp. Soc. Psychol.* 47, 635–641. doi: 10.1016/j.jesp.2010.12.015

Lelieveld, G.-J., Van Dijk, E., Van Beest, I., and Van Kleef, G. A. (2012). Why anger and disappointment affect other's bargaining behavior differently: the moderating role of power and the mediating role of reciprocal and complementary emotions. *Pers. Soc. Psychol. Bull.* 38, 1209–1221. doi: 10.1177/0146167212446938

Lount, R. B. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *J. Pers. Soc. Psychol.* 98, 420–433. doi: 10.1037/a0017344

Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., Smith, A., et al. (2013). *Teens, Social Media, and Privacy.* Washington, DC: Pew Research Center's Internet & American Life Project. Available at: http://pewinternet.org/Reports/2013/Teens-Social-Media-And-Privacy.aspx (accessed August 6, 2013).

McClintock, C. G., and Liebrand, W. B. G. (1988). The role of interdependence structure, individual value orientation and other's strategy in social decision making: a transformational analysis. *J. Pers. Soc. Psychol.* 55, 396–409. doi: 10.1037/0022-3514.55.3.396

McGivern, R. F., Andersen, J., Byrd, D., Mutter, K. L., and Reilly, J. (2002). Cognitive efficiency on a match to sample task decreases at the onset of puberty in children. *Brain Cogn.* 50, 73–89. doi: 10.1016/S0278-2626(02)00012-X

Messick, D. M., and McClintock, C. G. (1968). Motivational bases of choice in experimental games. *J. Exp. Soc. Psychol.* 4, 1–25. doi: 10.1016/0022-1031(68)90046-2

Morris, M. W., and Keltner, D. (2000). How emotions work: an analysis of the social functions of emotional expression in negotiations. *Res. Organ. Behav.* 22, 1–50. doi: 10.1016/S0191-3085(00)22002-9

Olekalns, M., and Smith, P. L. (1999). Social value orientations and strategy choices in competitive negotiations. *Pers. Soc. Psychol. Bull.* 25, 657–668. doi: 10.1177/0146167299025006002

Parkinson, B. (1996). Emotions are social. *Br. J. Psychol.* 87, 663–683. doi: 10.1111/j.2044-8295.1996.tb02615.x

Parks, C. (1994). The predictive ability of social values in resource dilemmas and public goods games. *Pers. Soc. Psychol. Bull.* 20, 431–438. doi: 10.1177/0146167294204010

Pietroni, D., Van Kleef, G. A., De Dreu, C. K. W., and Pagliaro, S. (2008). Emotions as strategic information: effects of other's emotional expressions on fixed-pie perception, demands, and integrative behavior in negotiation. *J. Exp. Soc. Psychol.* 44, 1444–1454. doi: 10.1016/j.jesp.2008.06.007

Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview.* San Antonia, TX: Harcourt Assessment.

Ruz, M., and Tudela, P. (2011). Emotional conflict in interpersonal interactions. *Neuroimage* 54, 1685–1691. doi: 10.1016/j.neuroimage.2010.08.039

Sinaceur, M., and Tiedens, L. Z. (2006). Get mad and get more than even: when and why anger expression is effective in negotiations. *J. Exp. Soc. Psychol.* 42, 314–322. doi: 10.1016/j.jesp.2005.05.002

Steinbeis, N., Bernhardt, B. C., and Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron* 73, 1040–1051. doi: 10.1016/j.neuron.2011.12.027

Steinberg, L., and Morris, A. S. (2001). Adolescent development. *Annu. Rev. Psychol.* 52, 83–110. doi: 10.1146/annurev.psych.52.1.83

Thomas, L. A., De Bellis, M. D., Graham, R., and LaBar, K. (2007). Development of emotional face recognition in late childhood and adolescence. *Dev. Sci.* 10, 547–558. doi: 10.1111/j.1467-7687.2007.00614.x

Van den Bos, W., Van Dijk, E., Westenberg, M., Rombouts, S. A. R. B., and Crone, E. A. (2010). Changing brains, changing perspectives: the neurocognitive development of reciprocity. *Psychol. Sci.* 22, 60–70. doi: 10.1177/0956797610391102

Van Dijk, E., and De Cremer, D. (2006). Self-benefiting in the allocation of scarce resources: leader-follower effects and the moderating effect of social value orientations. *Pers. Soc. Psychol. Bull.* 32, 1352–1361. doi: 10.1177/0146167206290338

Van Dijk, E., Van Kleef, G. A., Steinel, W., and Van Beest, I. (2008). A social functional approach to emotions in bargaining: when communicating anger pays and when it backfires. *J. Pers. Soc. Psychol.* 94, 600–614. doi: 10.1037/0022-3514.94.4.600

Van Dijk, W. W., and Zeelenberg, M. (2002). What do we talk about when we talk about disappointment? Distinguishing outcome-related disappointment from person-related disappointment. *Cogn. Emot.* 16, 787–807. doi: 10.1080/02699930143000563

Van Kleef, G. A. (2009). How emotions regulate social life: the emotions as social information (EASI) model. *Curr. Dir. Psychol. Sci.* 18, 184–188. doi: 10.1111/j.1467-8721.2009.01633.x

Van Kleef, G. A., De Dreu, C. K. W., and Manstead, A. S. R. (2004). The interpersonal effects of anger and happiness in negotiations. *J. Pers. Soc. Psychol.* 86, 57–76. doi: 10.1037/0022-3514.86.1.57

Van Kleef, G. A., De Dreu, C. K. W., and Manstead, A. S. R. (2006). Supplication and appeasement in conflict and negotiation: the interpersonal effects of disappointment, worry, guilt, and regret. *J. Pers. Soc. Psychol.* 91, 124–142. doi: 10.1037/0022-3514.91.1.124

Van Kleef, G. A., De Dreu, C. K. W., and Manstead, A. S. R. (2010). An interpersonal approach to emotion in social decision making: the emotions as social information model. *Adv. Exp. Soc. Psychol.* 42, 45–96. doi: 10.1016/S0065-2601(10)42002-X

Van Kleef, G. A., and Van Lange, P. A. M. (2008). What other's disappointment may do to selfish people: emotion and social value orientation in a negotiation context. *Pers. Soc. Psychol. Bull.* 34, 1084–1095. doi: 10.1177/0146167208318402

Van Lange, P. A. M., De Cremer, D., and Van Dijk, E., and Van Vugt, M. (2007). "Self-interest and beyond: basic principles of social interaction," in *Social Psychology: Handbook of Basic Principles,* 2nd Edn, eds A. W. Kruglanski and E. T. Higgins (New York: Guilford), 540–561.

Van Lange, P. A. M., Otten, W., De Bruin, E. N. M., and Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *J. Pers. Soc. Psychol.* 73, 733–746. doi: 10.1037/0022-3514.73.4.733

Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., and Altgassen, M. (2012). Ongoing development of social cognition in adolescence. *Child Neuropsychol.* doi: 10.1080/09297049.2012.718324 [Epub ahead of print].

Wade, A. M., Lawrence, K., Mandy, W., and Skuse, D. (2006). Charting the development of emotion recognition from 6 years of age. *J. Appl. Stat.* 33, 297–315. doi: 10.1080/02664760500445756

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 28 May 2013; accepted: 18 October 2013; published online: 12 November 2013.*

*Citation: Klapwijk ET, Peters S, Vermeiren RRJM and Lelieveld G-J (2013) Emotional reactions of peers influence decisions about fairness in adolescence. Front. Hum. Neurosci.* ***7****:745. doi: 10.3389/fnhum.2013.00745*

*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 26 July 2013
doi: 10.3389/fnhum.2013.00406

# Unfair offers, unfair offenders? Fairness considerations in incarcerated individuals with and without psychopathy

**Sina Radke[1]*, Inti A. Brazil[1,2], Inge Scheper[3], Berend H. Bulten[2] and Ellen R. A. de Bruijn[4]**

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands
[2] Pompestichting, Nijmegen, Netherlands
[3] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands
[4] Department of Clinical, Health and Neuropsychology, Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands

**Edited by:**
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

**Reviewed by:**
*Steve W. C. Chang, Duke University, USA*
*Yin Wu, University of Cambridge, UK*
*Claudio Lavin, Universidad Diego Portales, Chile*

***Correspondence:**
*Sina Radke, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, PO Box 9104, 6500 HE Nijmegen, Netherlands*
*e-mail: s.radke@donders.ru.nl*

Offenders with psychopathy have often committed crimes violating social norms, which may suggest a biased moral reasoning in psychopathy. Yet, as findings on utilitarian decisions remain conflicting, the current study investigated different aspects of fairness considerations in offenders with psychopathy, offenders without psychopathy and healthy individuals ($N = 18/14/18$, respectively). Unfair offers in a modified Ultimatum Game (UG) were paired with different unselected alternatives, thereby establishing the context of a proposal, and made under opposing intentionality constraints (intentional vs. unintentional). As in previous studies, unfair offers were most often rejected when the alternative was fair and when the offer was made intentionally. Importantly, however, offenders with psychopathy demonstrated a similar rejection pattern to that of healthy individuals, i.e., taking the unselected alternative into account. In contrast, delinquents without psychopathy did not adjust their decision behavior to the alternatives to an offer, suggesting stronger impairments in social decision-making. Crucially, the mechanisms and processes underlying rejection decisions might differ, particularly with regard to cognitive vs. emotional competencies. While preserved cognitive perspective-taking could drive seemingly intact decision patterns in psychopathy, emotional empathy is likely to be compromised.

**Keywords: psychopathy, social decision-making, ultimatum game, fairness, antisocial**

## INTRODUCTION

Social deficits are evident in various psychiatric disorders with their expression ranging from withdrawal in e.g., social phobia to antisocial behavior and even social predation as observed in psychopathy. Offenders with psychopathy often show a history of serious violent crimes committed against another person (i.e., murder, rape) and increased recidivism of criminal behavior (D'Silva et al., 2004) that exceeds the relapse rate of offenders without psychopathy by a factor of up to four (Harris et al., 1991; Hemphill et al., 1998).

The initial concept of psychopaths as "moral imbeciles" (Maudsley, 1895) attributed their deviations to a decreased ability for moral reasoning. Indeed, core traits of psychopathy such as manipulative behavior, callousness, and lack of guilt/remorse have been associated with overlooking moral principles for non-moral incentives such as money as well as a negative appreciation of the moral values of fairness and harm prevention (Glenn et al., 2009). In order to target the cognitive component of morality, hypothetical moral dilemmas that assess decisive judgments, e.g., how "appropriate" an action in the given situation is or whether one would execute that action, are frequently used (Greene et al., 2001). Although offenders with psychopathy tend to maximize overall benefit in these scenarios, i.e., demonstrate utilitarian choice patterns (Koenigs et al., 2012), there is no consistent evidence that individuals with psychopathy differ from healthy groups in explicit moral judgments (Blair et al., 1995; Cima et al., 2010; Aharoni et al., 2012). Findings on differences in moral reasoning between incarcerated populations with and without psychopathy are similarly inconsistent (Cima et al., 2010; Koenigs et al., 2012) and a recent meta-analysis reported a negative relation between moral development and recidivism for offenders in general, irrespective of psychopathic traits (Van Vugt et al., 2011). It therefore remains important to compare individuals with psychopathy not only to a healthy, but also to another forensic reference group which has also been convicted for serious offenses that essentially violate social and moral norms.

In contrast to hypothetical scenarios, an association between psychopathic traits and an increased focus on self-interest has been derived from social decision-making paradigms (Rilling et al., 2007; Mokros et al., 2008; Koenigs et al., 2010; Osumi and Ohira, 2010). Economic games, such as the Ultimatum Game (UG; Güth et al., 1982), are frequently used to capture strategies in interpersonal settings that involve weighting self-interest and other-interest. Here, the first player proposes a split of a resource, which can be either accepted or rejected by the second player (responder). Acceptance implements the proposal, but rejection leaves both players with nothing. Instead of "rationally" maximizing their payoff by accepting anything, responders frequently reject unfair offers, which has been attributed to fairness considerations (Güth et al., 1982).

In individuals with psychopathic traits, the observed disregard for fairness norms (Glenn et al., 2009; Aharoni et al., 2011) is mirrored in altered responder behavior in the UG, although the findings remain conflicting. On the one hand, students scoring high on psychopathic traits displayed lower rejection rates of unfair offers, interpreted as favoring self-interest (Osumi and Ohira, 2010). On the other hand, incarcerated patients with psychopathy showed the opposite pattern: individuals with primary psychopathy, i.e., psychopathy with low trait anxiety, rejected more unfair offers relative to individuals with secondary psychopathy or without psychopathy, which the authors relate to deficits in regulating anger and frustration (Koenigs et al., 2010). A recently published study found similar rejection behavior in participants with high and low psychopathic tendencies and suggests different underlying decision mechanisms, i.e., rejection as a reaction to frustration in individuals scoring high on psychopathic traits (Vieira et al., 2013).

As the classic UG assesses outcome-based fairness considerations, i.e., a comparison of outcomes of the self and the other (Radke et al., 2012), without an explicit normative reference point, it remains unresolved which factors underlie the deviations in social decision-making. Along these lines, previous results on psychopathy and UG decisions (Koenigs et al., 2010; Osumi and Ohira, 2010) can only be interpreted on the basis of outcome-driven judgments, but not in terms of social dynamics. In contrast, information derived from context and perceived intentionality guide not only social interactions in our daily lives, but also influence UG decisions (Blount, 1995; Falk et al., 2003, 2008; Güroğlu et al., 2009; Radke et al., 2012).

Interestingly, a recent finding revealed that offenders with psychopathy rate accidents as more morally permissible than delinquents without psychopathy (Young et al., 2012). This inclination suggests that they might weight the intention behind an action greater than its (harmful) outcome, stemming from the deficit of generating an emotional response to the victim's suffering (Young et al., 2012). It remains open, however, in how far this partiality in moral judgments might also apply to imbalanced decision-making. Of note, in laboratory settings, moral judgments are made from a detached perspective as the situation to be judged remains hypothetical, even when probed by a "would you do … in order to …?" question. Therefore, both the implementation of one's choice and the absence vs. presence of self-interest are important methodological distinctions between the use of hypothetical scenarios and socioeconomic games. Whereas the former usually depict vignettes or actions that do not affect oneself, economic games traditionally involve real, to-be-paid-out stakes and thus outcomes relevant to oneself and one's interaction partner.

Using a modified UG enables us to investigate how social decisions involving fairness considerations are resolved. In this version, information is provided about an unselected alternative, thereby establishing the "context" in which an offer is selected, and about the intentionality of an offer. From a fixed set of two allocations of 10 coins, either the first player (proposer) himself or the computer randomly chooses one. The fixed set allows to manipulate the reference point ("context") of the proposal (Falk et al., 2003; Güroğlu et al., 2009; Radke et al., 2012), whereas the agency of the proposer constitutes the manipulation of intentionality, i.e., whether the offer was selected by proposers themselves vs. by the computer (Radke et al., 2012). This setup allows to investigate perspective-taking from the side of responders. Here, of particular interest are unfair proposals (8 coins for the proposer and 2 coins for the responder) that are contrasted against either fair, hyperfair, or hyperunfair alternatives, or no alternative at all. Previous findings show that unfair offers are more often rejected when the alternative was fair compared to all other three alternatives (Falk et al., 2003; Güroğlu et al., 2009; Radke et al., 2012), which has also been associated with developmental advances in cognitive perspective-taking abilities (Güroğlu et al., 2009). In adults, a similar rejection pattern was evident for intentional and unintentional decisions (Radke et al., 2012). However, intentionality played a crucial role in the decision process when an unfair treatment was made explicit and salient, i.e., when paired with a fair alternative: These unfair offers were more often rejected when the offer was selected by proposers themselves than when selected by the computer, underlining the importance of punishing intentional social norm violations (Radke et al., 2012).

The current study is the first to investigate social decision-making based on different aspects of fairness considerations and their (social-) cognitive demands (Radke et al., 2012) in a forensic sample. The behavior of offenders with psychopathy was compared to a group of offenders without psychopathy and a group of healthy controls. For healthy individuals, we expected to replicate previous findings on the effects of context referring to the manipulation of alternative offers as in (Güroğlu et al., 2009; Radke et al., 2012) and intentionality (Radke et al., 2012). As these features of social decision-making have not been assessed in our populations of interest until now, it is difficult to predict the rejection patterns in the offender groups.

Based on studies pointing to a relative integrity of cognitive functioning in psychopathy as opposed to non-psychopathy (Morgan and Lilienfeld, 2000; Gao and Raine, 2009; Brazil et al., 2012), one might expect the group with psychopathy to take the context or the intentionality of an offer into account. On the contrary, offenders without psychopathy might show a more impulsive behavioral pattern, not differentiating on the basis of additional information. In sum, we aimed to investigate to what extent the alterations in moral judgments in psychopathy and non-psychopathy translate to decisions with not only moral, but also utilitarian outcomes.

## MATERIALS AND METHODS

### PARTICIPANTS

The offender groups were recruited from the patient population of the Pompestichting Forensic Psychiatric Institute in Nijmegen, The Netherlands[1]. The study was approved by the local medical ethics committee and in accordance with the Declaration of

---

[1] The Pompestichting is a "TBS-clinic" located in Nijmegen. TBS is a disposal to be treated, on behalf of the state, for people who committed serious criminal offences in connection with having a mental disorder. TBS is not a punishment, but an entrustment act for mentally disordered offenders (diminished responsibility). These court orders are an alternative to either long-term imprisonment or confinement in psychiatric hospital, with the goal to strike a balance between security, treatment, and protection.

Helsinki. All participants received written information about the experiment and gave written informed consent.

All participants were male. The group with psychopathy consisted of 18 offenders, the forensic group without psychopathy consisted of 14 offenders and the control group comprised 18 healthy volunteers without criminal records or a history of psychiatric disorders who were recruited through advertisements and matched with the delinquents on age and intelligence (see **Table 1** for characteristics of the study population). The Psychopathy Checklist-Revised (PCL-R; Hare, 2003) was used to assess psychopathy. The PCL-R is an instrument that allows the assessment of psychopathy through a semi structured interview and information on criminal history (Hare, 2003), and is regarded as the golden standard for the assessment of clinical psychopathy. The instrument consists of 20 items capturing behavioral correlates of core aspects of psychopathy, which are coded as either not present (0), moderately present (1), or certainly present (2). Certified psychologists administered the PCL-R after placement in the Dutch forensic mental health system and for the present study the PCL-R scores were retrieved from offenders' files. As common in European countries, participants with a score of 26 or more were included in the group with psychopathy and participants with a score below 26 were assigned to the non-psychopathy group (Hildebrand et al., 2004). As healthy controls did not have criminal records, is was not possible to obtain reliable PCL-R scores in this group. Exclusion criteria were assessed with the Dutch version of Mini International Neuropsychiatric Interview Plus 5.0.0. (Van Vliet et al., 2000) and Structured Clinical Interview for DSM-IV Axis II Personality Disorders (Weertman et al., 2000) and included all major Axis I and Axis II disorders (except antisocial personality disorder in the offender groups) or any CNS injuries (Brazil et al., 2009, 2011, 2012). Additional information was retrieved from each offender's clinical files. An estimation of intelligence level was made for all participants by using the Dutch version of the Adult Reading Test (NLV; Schmand et al., 1991). All assessments were carried out by trained psychologists.

## DESIGN

Participants were responders in a computerized version of the modified UG with two within-subject factors: Intentionality and Context. Intentionality had two levels based on *who* selects the offer: the human player (i.e., the proposer) him/herself (intentional) or the computer (unintentional). Intentionality was thus manipulated in a binary fashion. Context had four levels based on alternatives to an unfair distribution (8:2): a fair-alternative (5:5 vs. 8:2), a hyperfair-alternative (2:8 vs. 8:2), a hyperunfair-alternative (10:0 vs. 8:2), and no-alternative (8:2 vs. 8:2). Hence, the factor Context pertains to the alternative outcome that had *not* been chosen. Pitting an unfair offer (8:2) against a fair alternative (5:5) can be seen as an explicit version of the classic UG in which decision-making is generally based on comparing any offer to a potential equal split. The resulting 8 conditions were presented 16 times each (counterbalanced for proposers' gender and position of the unfair offer). As the no-alternative condition entails an 8:2 offer for either alternative, an unfair offer (8:2) was presented in 5 of the 8 conditions, equivalent to 80 trials. The three genuine alternative offers (i.e., 5:5, 2:8, or 10:0) were selected on 48 trials, yielding 128 trials in total. Contrary to subjects' belief, all choices were computer-generated.

**Table 1 | Sociodemographic and clinical characteristics of study participants (mean [*SD*]).**

| | PP (*N* = 18) | Non-PP (*N* = 14) | HC (*N* = 18) | *p*-value |
|---|---|---|---|---|
| Age in years | 42.5 (6.7) | 39.7 (7.7) | 37.4 (8.8) | 0.15 |
| PCL-R | 31.0 (3.6) | 15.8 (5.1) | n/a | 0.00* |

*PP, offenders with psychopathy; non-PP, offenders without psychopathy; HC, healthy controls. *significant difference between PP and non-PP.*

## MATERIAL

**Figure 1** depicts the timeline of a trial in the intentional fair-alternative condition. Each round started with a fixation cross (1000 ms), followed by the presentation of the two available options (1000 ms). Next, the selected offer was surrounded by a red square (1000 ms). Subsequently, "Yes" and "No" buttons were presented while the selection remained visible. As the task was self-paced, participants had unlimited amount of time to respond via pressing one of two buttons using the keyboard. Participants' response remained on the screen for 2000 ms before the next round started.

## PROCEDURE

Participants were led to believe that they were coupled with data from others who had previously participated as proposers and that they would play every trial with a new partner (Güroğlu et al., 2009; Radke et al., 2012). They were told that on some trials the other players would make an offer themselves and on other trials the computer would take over and randomly select one of the two options. Participants' task was to decide whether to accept or reject an offer. If accepted, the coins were distributed as proposed; if rejected, neither player received anything. Participants were informed that at the end of the experiment, a random number of rounds would be selected to determine their payoff. This was done to assure participants' motivation and to strengthen the concept of a one-shot game as every round could influence their financial outcome. Moreover, it was emphasized that participants' decisions also affected the other players' outcome because their payoff would be determined by participants' response, irrespective of who made the proposal in a particular round (i.e., themselves vs. computer). Proposers would be paid after all data from responders had been collected. The payoff was set around 5 Euro (± 5 cent) to manage an equal payment for all participants, but simultaneously minimize suspicion.

## STATISTICAL ANALYSES

Rejection rates of unfair offers were entered into a repeated measures ANOVA with Intentionality (two levels: intentional vs. unintentional) and Context (four levels: fair vs. hyperfair vs. hyperunfair vs. no alternative) as within-subject factors and Group (three levels: offenders with psychopathy, offenders without psychopathy, healthy controls) as a between-subject factor. In case of interactions involving the factor Group, separate ANOVAs

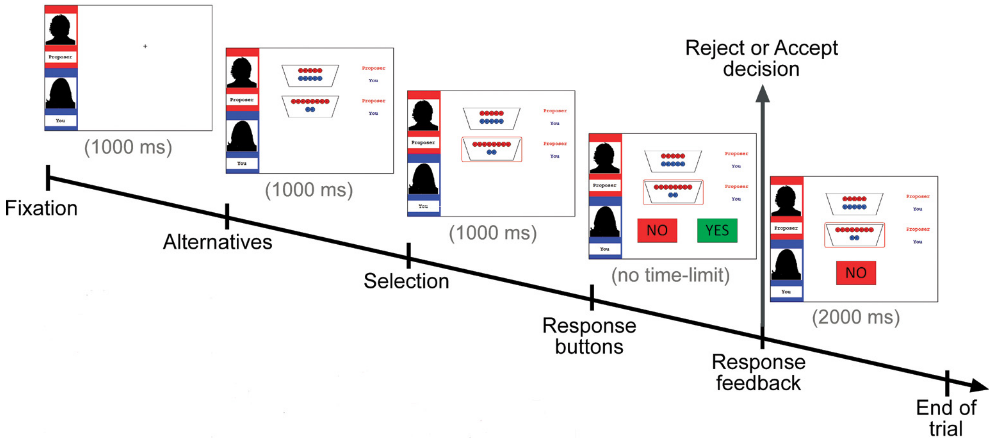


**FIGURE 1 | Display of a trial in the intentional fair-alternative condition.** The name of the proposer is shown at the top (here "Proposer") and the name of the participant is shown underneath (here "You"). The two potential distributions are specified by red and blue coins (red for proposer, blue for responder), with the offer selected by the proposer encircled in red. Participants have to indicate via button press whether to accept ("Yes") or reject ("No") the offer. Note that for unintentional offers (not shown here), the otherwise black silhouette of the proposer was purple with a banner displaying "Computer chooses," which was also depicted instead of the proposer's name.

for the three different groups are conducted with the above mentioned within-subject factors.

In order to test for replicating the results of Radke et al. (2012), i.e., higher rejection rates for unfair offers paired with a fair alternative when the offer was selected by proposers themselves than when selected by the computer, the effect of intentionality will be tested in the fair-alternative context on the whole group level by means of a repeated measures ANOVA with Intentionality (two levels: intentional vs. unintentional) as a within-subject factor.

## RESULTS

There were significant main effects of Context, $F_{(3,\ 141)} = 16.49$, $p < 0.001$, partial $\eta^2 = 0.26$, and Intentionality, $F_{(1,\ 47)} = 4.95$, $p = 0.03$, partial $\eta^2 = 0.10$. Moreover, the interaction between Context and Group was significant, $F_{(6,\ 141)} = 3.58$, $p = 0.01$, partial $\eta^2 = 0.13$. None of the interactions involving the within-subject factor Intentionality (Intentionality × Context, Intentionality × Group, Intentionality × Context × Group) was significant (all *ps* > 0.27) nor was the main effect of Group ($p = 0.98$).

Pairwise comparisons revealed that rejection rates were highest for the fair-alternative condition (63.3%) compared to the more disadvantageous alternatives (no-alternative: 41.8%, $p = 0.001$; hyperunfair: 39.1%, $p < 0.001$). Rejection rates for the hyperfair alternative condition (53.8%) were higher than for the hyperunfair and no-alternative condition (both *ps* = 0.001). The latter two conditions did not differ significantly ($p = 1$). With respect to the main effect of intentionality, rejection rates were higher when the offer was selected intentionally (52.2%) than when selected unintentionally/by the computer (49.2%).

To investigate the Context × Group interaction, separate analyses for the three different groups were conducted (see **Figures 2**, **3**). There was a significant effect of Context in healthy controls, $F_{(3,\ 51)} = 14.03$, $p < 0.001$, partial $\eta^2 = 0.45$, as well as in the forensic sample with psychopathy, $F_{(3,\ 51)} = 3.96$, $p = 0.039$, partial $\eta^2 = 0.19$, but not in offenders without psychopathy, $F_{(3,\ 39)} = 1.63$, $p = 0.27$, partial $\eta^2 = 0.10$. For the healthy controls, the same pattern as on the whole-group level was evident: rejection rates were highest for the fair-alternative condition (72.2%) compared to the more disadvantageous alternatives (no-alternative: 36.7%, $p < 0.001$; hyperunfair: 28.5%, $p = 0.001$). Rejection rates for the hyperfair alternative condition (67.7%) were higher than for the hyperunfair and no-alternative condition (both ps < 0.002), with the latter two not differing significantly ($p = 0.15$). Reactions to the fair and hyperfair alternative conditions did not differ ($p = 0.48$). For the forensic sample with psychopathy, rejection rates were highest for the fair-alternative condition (60.7%) compared to the more disadvantageous alternatives (no-alternative: 45.1%, $p = 0.03$; hyperunfair: 43.9%, $p = 0.02$). The remaining pairwise comparisons did not yield significant differences ($p > 0.08$).

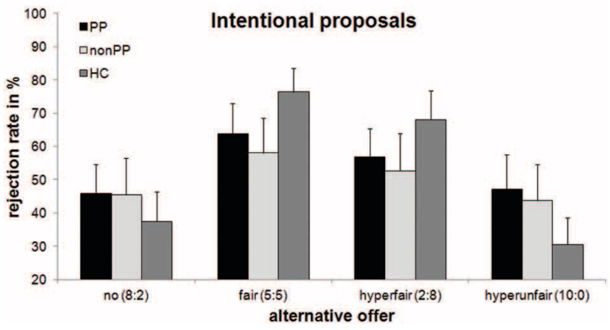


**FIGURE 2 | Rejection rates of intentional unfair offers with regard to alternative offers and group.** Mean percentage and standard errors of rejection of 8:2-offers are displayed. PP, offenders with psychopathy; non-PP, offenders without psychopathy; HC, healthy controls.

In order to directly test for replication of Radke et al. (2012), analysis of the fair alternative condition indicated higher rejection rates for intentional (66.8%) vs. unintentional/computer offers (59.8%), $F_{(1,\ 47)} = 6.78$, $p = 0.01$, partial $\eta^2 = 0.13$. In contrast, the effect of Intentionality did not reach significance for the other three contexts (all $F$s $< 0.69$, all $p$s $> 0.41$; see also **Figure 4**).

## DISCUSSION

In the current study, social decision-making based on different aspects of fairness considerations was investigated in a forensic sample and a matched healthy control group. In particular, we sought to explore in how far altered moral judgments in psychopathy apply to decisions when not only moral values, but also self-relevant outcomes are at stake.

In addition to replicating previous findings on context and intentionality (Falk et al., 2003; Güroğlu et al., 2009; Radke et al., 2012), group differences in context sensitivity were evident. Essentially, offenders with psychopathy displayed a similar pattern of rejection behavior to that of healthy individuals, i.e., an effect of context. In contrast, the decisions in delinquents without psychopathy were not influenced by the alternative offer to an unfair proposal.

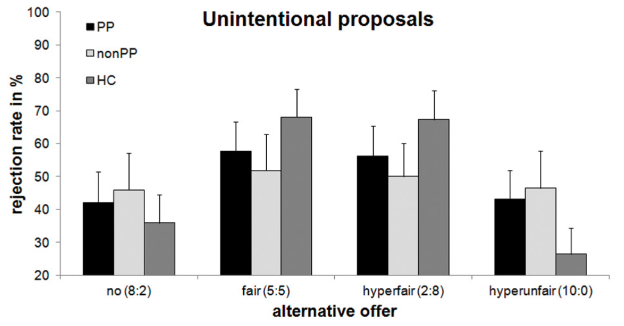


**FIGURE 3 | Rejection rates of unintentional unfair offers with regard to alternative offers and group.** Mean percentage and standard errors of rejection of 8:2-offers are displayed. PP, offenders with psychopathy; non-PP, offenders without psychopathy; HC, healthy controls.

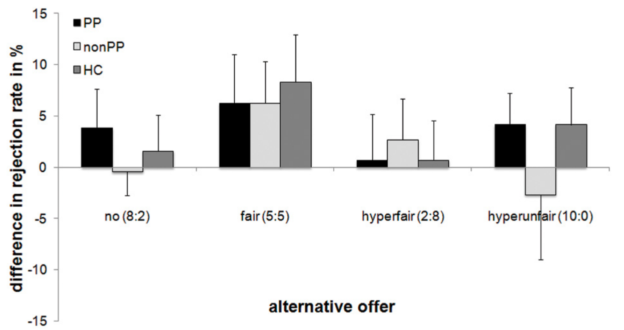


**FIGURE 4 | Difference between intentional and unintentional unfair offers with regard to alternative offers and group.** Mean percentage and standard errors of rejection of 8:2-offers are displayed. PP, offenders with psychopathy; non-PP, offenders without psychopathy; HC, healthy controls.

Recently, there has been some disagreement on which processes in fairness considerations are targeted by the context manipulation, i.e., higher-order social functions like perspective-taking (Falk et al., 2003; Güroğlu et al., 2009) or straightforward outcome comparisons (Brandts and Sola, 2001; Sandbu, 2007). This discussion (see also Radke et al., 2012) is crucial with regard to drawing inferences about possible impaired and preserved (social-) cognitive abilities in psychiatric/forensic populations. Although both processes rely on counterfactual thinking, i.e., representations of alternatives to past events (Roese, 1997), the representation of another, social agent is not necessary for comparing the outcomes for the self-resulting from the chosen and the unchosen alternative. Such a quantitative evaluation can be achieved—quite parsimoniously—without taking the perspective of another person. Our design resolves this disagreement by making the social dimension explicit, i.e., contrasting intentional and unintentional offers. The processes of outcome comparisons and intentionality considerations can thereby easily be disentangled. Importantly, the current data replicates previous findings on their relative contribution to social decision-making (Radke et al., 2012), which provides a solid basis for the investigation of intergroup variations.

As delinquents without psychopathy did not adjust their behavior to the context or intentionality of an offer, their decisions might be dominated by rather basic motives not directly targeted by these manipulations. Given that their overall mean rejection rate did not differ from the other groups either, the offenders without psychopathy seem to be guided by the magnitude of the proposal's intrinsic distribution (what the proposer gets vs. what the responder gets) with a dislike of unequal outcomes, i.e., inequity aversion. Fairness is determined by payoffs available in the here and now, which may reflect the preference for immediate options, hinting at a hyperactivity of the impulsive system (Buckholtz et al., 2010; Dean et al., 2013).

Offenders with psychopathy, on the other hand, behaved similarly to healthy volunteers. The lack of differences in overall rejection rates compared to both healthy and incarcerated individuals is in line with the behavioral results of Vieira et al. (2013), but at odds with other earlier, yet inconsistent, findings on altered responder behavior in the UG reporting higher (Koenigs et al., 2010), and lower rejection rates (Osumi and Ohira, 2010), respectively. These studies, however, also diverge in sample characteristics, with the group being either very small [$N = 6$ for high anxious psychopaths, (Koenigs et al., 2010)] or consisting of students (Osumi and Ohira, 2010) or a community sample (Vieira et al., 2013), warranting caution for generalization to a large forensic population. The main focus of the current study, however, lay not on rejection rates as such, but on its modulation by context and intentionality.

Most interestingly therefore, in the current sample, individuals with psychopathy showed an analogous sensitivity to the alternatives to a given outcome, which converges with findings on intact moral judgments (Blair et al., 1995; Cima et al., 2010; Aharoni et al., 2012), Theory of Mind (Blair et al., 1996; Richell et al., 2003) and other aspects of cognitive functioning (Blair et al., 2006; Brazil et al., 2012). Despite the behavioral similarity with healthy controls, in psychopathy the underlying mechanisms might differ and reflect a distinct motivation. For instance, in healthy individuals with psychopathic traits, rejection of unfair offers was associated with increased activation in the anterior cingulate cortex and ventromedial prefrontal cortex (Vieira et al., 2013). The recruitment of these clusters might indicate impairments in automatic emotion regulation, leading to anger-motivated instead of fairness-motivated rejection (Vieira et al., 2013).

Conversely, we observed behavioral differences in the context-dependency of rejection decisions between offenders with psychopathy compared to offenders without psychopathy. This differentiation between subgroups of violent offenders might be attributable to the cognitive nature of the task, i.e., its assessment of cognitive perspective-taking (Güroğlu et al., 2009). Previous research suggests relatively intact cognitive functioning in psychopathy, but shortages in non-psychopathy (Morgan and Lilienfeld, 2000; Gao and Raine, 2009; Brazil et al., 2012). Moreover, in contrast to emotional aspects of empathy, making inferences about others' mental states, i.e., mentalizing or cognitive inferences, does not seem to be compromised in clinical psychopathy (Blair et al., 1996; Cima et al., 2010; Shamay-Tsoory et al., 2010).

Along these lines, other (than financial) self-serving motivations supported by mentalizing might underlie the different decision patterns between offenders with psychopathy and those without. While both groups show an increased risk for frustration (Blair, 2010), individuals with psychopathy might be more successful in cognitively regulating impulsive tendencies and possibly even use a "fair" disguise instrumentally in order to obtain unobtrusive advantages, e.g., the appreciation of the experimenter. One might also speculate in how far the initial emotional responses to unfairness differ between the forensic groups. Both might react equally fervently to potentially frustrating unfair offers, but base their decisions on other features, leading to the distinct behavioral patterns. Future studies assessing physiological indicators of emotional reactivity, such as skin conductance, would be useful to explore in how far initial, affective reactions might be restrained by cognitive mechanisms of control or impression management. More tailored paradigms could also identify effects of impulsivity or serial decisions.

With regard to the relative weighting of the intention behind an action and its outcome in offenders with psychopathy (Young et al., 2012), the current data does not allow for drawing firm conclusions on the influence of intentionality for the subgroup of offenders with psychopathy. Its effects manifested only on the whole-group level, which precludes further investigation for the groups separately. Likewise, in contrast to the study by Young et al. (2012) who used hypothetical scenarios in which negative outcomes meant harm or death of another person, even the worst consequences in the current design were, naturally due to the implementation of choices, much less severe. Besides, they did not imply positive punishment, i.e., harm, but negative punishment, i.e., the withholding of coins in the case of an rejection decision and thereby forgoing potential gain. Despite the methodological strength of executing the choices in an interactive setting, this approach is less likely to trigger empathetic reactions, also since in UG settings, the most pronounced emotions arise in responders facing unfairness (Pillutla and Murnighan, 1996; Xiao and Houser, 2005).

In sum, our findings indicate discrepancies between the two offender samples: On the one hand, offenders without psychopathy seem to neglect aspects of fairness considerations that go beyond the comparisons based on payoffs, whereas, on the other hand, the decisions of offenders with psychopathy did not differ from those of healthy individuals. Distinct processes in cognition and affect might underlie these behavioral similarities. Importantly, central features of psychopathy, i.e., manipulating or deceiving others, require certain knowledge about social rules as well as cognitively taking the perspective of others, so that offenders with psychopathy might succeed in an environment where all possible outcome variants, intentions, self- and other-interests are explicitly stated. In contrast, real-life interactions with others are not only more complex and subtle, but also require emotional skills, such as generating empathic responses, regulating one's emotions, and adequately reacting to others' feelings, that are likely to be impaired in psychopathy, as evident in their antisocial lifestyles and violent crimes.

## ACKNOWLEDGMENTS

Ellen R. A. de Bruijn was supported by a VENI grant from the Netherlands Organization for Scientific Research (NWO; 451-07-022). Inti A. Brazil was supported by a Mosaic grant from the NWO (240-00-244). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## REFERENCES

Aharoni, E., Antonenko, O., and Kiehl, K. A. (2011). Disparities in the moral intuitions of criminal offenders: the role of psychopathy. *J. Res. Pers.* 45, 322–327. doi: 10.1016/j.jrp.2011.02.005

Aharoni, E., Sinnott-Armstrong, W., and Kiehl, K. A. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *J. Abnorm. Psychol.* 121, 484–497. doi: 10.1037/a0024796

Blair, K. S., Newman, C., Mitchell, D. G. V., Richell, R. A., Leonard, A., Morton, J., et al. (2006). Differentiating among prefrontal substrates in psychopathy: neuropsychological test findings. *Neuropsychology* 20, 153–165. doi: 10.1037/0894-4105.20.2.153

Blair, R. J. R. (2010). Psychopathy, frustration, and reactive aggression: the role of ventromedial prefrontal cortex. *Br. J. Psychol.* 101, 383–399. doi: 10.1348/000712609X418480

Blair, R. J. R., Jones, L., Clark, F., and Smith, M. (1995). Is the Psychopath "morally insane". *Pers. Individ. Dif.* 19, 741–752. doi: 10.1016/0191-886900087-M

Blair, R. J. R., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., et al. (1996). Theory of mind in the psychopath. *J. Forensic Psychiatry* 7, 15–25. doi: 10.1080/09585189608409914

Blount, S. (1995). When social outcomes aren't fair—the effect of causal attributions on preferences. *Organ. Behav. Hum. Decis. Process.* 63, 131–144. doi: 10.1006/obhd.1995.1068

Brandts, J., and Sola, C. (2001). Reference points and negative reciprocity in simple sequential games. *Games Econ. Behav.* 36, 138–157. doi: 10.1006/game.2000.0818

Brazil, I. A., De Bruijn, E. R. A., Bulten, B. H., Von Borries, A. K. L., Van Lankveld, J., Buitelaar, J. K., et al. (2009). Early and late components of error monitoring in violent offenders with psychopathy. *Biol. Psychiatry* 65, 137–143. doi: 10.1016/j.biopsych.2008.08.011

Brazil, I. A., Mars, R. B., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., and De Bruijn, E. R. A. (2011). A neurophysiological dissociation between monitoring one's own and others' actions in psychopathy.

*Biol. Psychiatry* 69, 693–699. doi: 10.1016/j.biopsych.2010.11.013

Brazil, I. A., Verkes, R. J., Brouns, B. H. J., Buitelaar, J. K., Bulten, B. H., and De Bruijn, E. R. A. (2012). Differentiating psychopathy from general antisociality using the P3 as a psychophysiological correlate of attentional allocation. *PLoS ONE* 7:e50339. doi: 10.1371/journal.pone.0050339

Buckholtz, J. W., Treadway, M. T., Cowan, R. L., Woodward, N. D., Benning, S. D., Li, R., et al. (2010). Mesolimbic dopamine reward system hypersensitivity in individuals with psychopathic traits. *Nat. Neurosci.* 13, 419–421. doi: 10.1038/nn.2510

Cima, M., Tonnaer, F., and Hauser, M. D. (2010). Psychopaths know right from wrong but don't care. *Soc. Cogn. Affect. Neurosci.* 5, 59–67. doi: 10.1093/scan/nsp051

Dean, A. C., Altstein, L. L., Berman, M. E., Constans, J. I., Sugar, C. A., and McCloskey, M. S. (2013). Secondary psychopathy, but not primary psychopathy, is associated with risky decision-making in noninstitutionalized young adults. *Pers. Individ. Dif.* 54, 272–277. doi: 10.1016/j.paid.2012.09.009

D'Silva, K., Duggan, C., and McCarthy, L. (2004). Does treatment really make psychopaths worse. A review of evidence. *J. Pers. Disord.* 18, 163–177.

Falk, A., Fehr, E., and Fischbacher, U. (2003). On the nature of fair behavior. *Econ. Inq.* 41, 20–26. doi: 10.1093/ei/41.1.20

Falk, A., Fehr, E., and Fischbacher, U. (2008). Testing theories of fairness—intentions matter. *Games Econ. Behav.* 62, 287–303. doi: 10.1016/j.geb.2007.06.001

Gao, Y., and Raine, A. (2009). P3 event-related potential impairments in antisocial and psychopathic individuals: a meta-analysis. *Biol. Psychol.* 82, 199–210. doi: 10.1016/j.biopsycho.2009.06.006

Glenn, A. L., Iyer, R., Graham, J., Koleva, S., and Haidt, J. (2009). Are all types of morality compromised in psychopathy. *J. Pers. Disord.* 23, 384–398. doi: 10.1521/pedi.2009.23.4.384

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108. doi: 10.1126/science.1062872

Güroğlu, B., Van Den Bos, W., and Crone, E. A. (2009). Fairness considerations: increasing understanding of intentionality during adolescence. *J. Exp. Child Psychol.* 104, 398–409. doi: 10.1016/j.jecp.2009.07.002

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental-analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388. doi: 10.1016/0167-268190011-7

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist, 2nd Edn.* Toronto, ON: Multi-Health Systems.

Harris, G. T., Rice, M. E., and Cormier, C. A. (1991). Psychopathy and violent recidivism. *Law Hum. Behav.* 15, 625–637. doi: 10.1007/BF01065856

Hemphill, J. F., Hare, R. D., and Wong, S. (1998). Psychopathy and recidivism: a review. *Leg. Criminol. Psychol.* 3, 141–172. doi: 10.1111/j.2044-8333.1998.tb00355.x

Hildebrand, M., De Ruiter, C., and Nijman, H. (2004). PCL-R psychopathy predicts disruptive behavior among male offenders in a dutch forensic psychiatric hospital. *J. Interpers. Violence* 19, 13–29. doi: 10.1177/088626050 3259047

Koenigs, M., Kruepke, M., and Newman, J. P. (2010). Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia* 48, 2198–2204. doi: 10.1016/j.neuropsychologia.2010.04.012

Koenigs, M., Kruepke, M., Zeier, J., and Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Soc. Cogn. Affect. Neurosci.* 7, 708–714. doi: 10.1093/scan/nsr048

Maudsley, H. (1895). *The Pathology of Mind*. London: Macmillan.

Mokros, A., Menner, B., Eisenbarth, H., Alpers, G. W., Lange, K. W., and Osterheider, M. (2008). Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *J. Abnorm. Psychol.* 117, 406–413. doi: 10.1037/0021-843X.117.2.406

Morgan, A. B., and Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clin. Psychol. Rev.* 20, 113–136. doi: 10.1016/S0272-735800096-8

Osumi, T., and Ohira, H. (2010). The positive side of psychopathy: emotional detachment in psychopathy and rational decision-making in the ultimatum game. *Pers. Individ. Dif.* 49, 451–456. doi: 10.1016/j.paid.2010.04.016

Pillutla, M. M., and Murnighan, J. K. (1996). Unfairness, anger, and spite: emotional rejections of ultimatum offers. *Organ. Behav. Hum. Decis. Process.* 68, 208–224. doi: 10.1006/obhd.1996.0100

Radke, S., Güroğlu, B., and De Bruijn, E. R. A. (2012). There's something about a fair split: intentionality moderates context-based fairness considerations in social decision-making. *PLoS ONE* 7:e31491. doi: 10.1371/journal.pone.0031491

Richell, R. A., Mitchell, D. G., Newman, C., Leonard, A., Baron-Cohen, S., and Blair, R. J. (2003). Theory of mind and psychopathy: can psychopathic individuals read the 'language of the eyes'. *Neuropsychologia* 41, 523–526. doi: 10.1016/S0028-393200175-6

Rilling, J. K., Glenn, A. L., Jairam, M. R., Pagnoni, G., Goldsmith, D. R., Elfenbein, H. A., et al. (2007). Neural correlates of social cooperation and Non-cooperation as a function of psychopathy. *Biol. Psychiatry* 61, 1260–1271. doi: 10.1016/j.biopsych.2006.07.021

Roese, N. J. (1997). Counterfactual thinking. *Psychol. Bull.* 121, 133–148. doi: 10.1037/0033-2909.121.1.133

Sandbu, M. E. (2007). Fairness and the roads not taken: an experimental test of non-reciprocal set-dependence in distributive preferences. *Games Econ. Behav.* 61, 113–130. doi: 10.1016/j.geb.2006.11.003

Schmand, B. A., Bakker, D., Saan, R. J., and Louman, J. (1991). The dutch reading test for adults: a measure of premorbid intelligence level. *Tijdschr. Gerontol. Geriatr.* 22, 15–19.

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., and Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex* 46, 668–677. doi: 10.1016/j.cortex.2009.04.008

Van Vliet, I. M., Leroy, H., and Van Megen, H. J. M. (2000). *De MINI-Internationaal Neuropsychiatrisch Interview: een Kort Gestructureerd Diagnostisch Interview voor DSM-IV en ICD-10 Psychiatrische Stoornissen [in Dutch].* Leiden, The Netherlands: LUMC.

Van Vugt, E., Gibbs, J., Stams, G. J., Bijleveld, C., Hendriks, J., and Van Der Laan, P. (2011). Moral development and recidivism. *Int. J. Offender Ther. Comp. Criminol.* 55, 1234–1250. doi: 10.1177/0306624X11396441

Vieira, J. B., Almeida, P. R., Ferreira-Santos, F., Barbosa, F., Marques-Teixeira, J. O., and Marsh, A. A. (2013). Distinct neural activation patterns underlie economic decisions in high and low psychopathy scorers. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst1093. [Epub ahead of print].

Weertman, A., Arntz, A., and Kerkhofs, M. (2000). *Gerstructureerd Diagnostisch Interview voor DSM-IV Persoonlijkheidsstoornissen (SCID II; in Dutch).* Lisse, The Netherlands: Swets Test Publisher.

Xiao, E., and Houser, D. (2005). Emotion expression in human punishment behavior. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7398–7401. doi: 10.1073/pnas.0502399102

Young, L., Koenigs, M., Kruepke, M., and Newman, J. P. (2012). Psychopathy increases perceived moral permissibility of accidents. *J. Abnorm. Psychol.* 121, 659–667. doi: 10.1037/a0027489

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 01 June 2013; accepted: 10 July 2013; published online: 26 July 2013.*

*Citation: Radke S, Brazil IA, Scheper I, Bulten BH and de Bruijn ERA (2013) Unfair offers, unfair offenders? Fairness considerations in incarcerated individuals with and without psychopathy. Front. Hum. Neurosci. 7:406. doi: 10.3389/fnhum.2013.00406*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 16 August 2013
doi: 10.3389/fnhum.2013.00467

# Structural neuroimaging of social cognition in progressive non-fluent aphasia and behavioral variant of frontotemporal dementia

**Blas Couto**[1,2,3†], **Facundo Manes**[1,2†], **Patricia Montañés**[4], **Diana Matallana**[4], **Pablo Reyes**[4], **Marcela Velasquez**[1,2], **Adrián Yoris**[1,2], **Sandra Baez**[1,2,3] **and Agustin Ibáñez**[1,2,3,5]*

[1] Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive Neurology (INECO), Favaloro University, Buenos Aires, Argentina
[2] Institute of Neuroscience, Favaloro University, Buenos Aires, Argentina
[3] National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
[4] Intellectus Memory and Cognition Center, Mental Health and Psychiatry Department, San Ignacio Hospital, Aging Institute, Pontifical Javeriana University, Bogotá, Colombia
[5] Laboratory of Cognitive and Social Neuroscience (LaNCyS), UDP-INECO Foundation Core on Neuroscience (UIFCoN), Diego Portales University, Santiago, Chile

***Edited by:***
*Maria Ruz, Universidad de Granada, Spain*

***Reviewed by:***
*Ingrid R. Olson, Temple University, USA*
*Mario Parra, University of Edinburgh, UK*

****Correspondence:***
*Agustín Ibáñez, Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive Neurology (INECO) and CONICET. Pacheco de Melo 1860, Buenos Aires, Argentina*
*e-mail: aibanez@ineco.org.ar*
*† First Authors.*

Social cognition impairments are pervasive in the frontotemporal dementias (FTD). These deficits would be triggered by (a) basic emotion and face recognition processes as well as by (b) higher level social cognition (e.g., theory of mind, ToM). Both emotional processing and social cognition impairments have been previously reported in the behavioral variant of FTD (bvFTD) and also in other versions of FTDs, including primary progressive aphasia. However, no neuroanatomic comparison between different FTD variants has been performed. We report selective behavioral impairments of face recognition, emotion recognition, and ToM in patients with bvFTD and progressive non-fluent aphasia (PNFA) when compared to controls. Voxel-based morphometry (VBM) shows a classical impairment of mainly orbitofrontal (OFC), anterior cingulate (ACC), insula and lateral temporal cortices in patients. Comparative analysis of regional gray matter related to social cognition deficits (VBM) reveals a differential pattern of fronto-insulo-temporal atrophy in bvFTD and an insulo-temporal involvement in PNFA group. Results suggest that in spite of similar social cognition impairments reported in bvFTD and PNFA, the former represents an inherent ToM affectation whereas in the PNFA these deficits could be related to more basic processes of face and emotion recognition. These results are interpreted in the frame of the fronto-insulo-temporal social context network model (SCNM).

**Keywords: PNFA, bvFTD, theory of mind, voxel-based morphometry, social context network model, fronto-insulo-temporal network**

## INTRODUCTION

Frontotemporal dementia (FTD) is a neurodegenerative disease commonly characterized by behavioral changes, as well as emotional and cognitive impairments (Neary et al., 1998; Perry and Miller, 2001; Ritchie and Lovestone, 2002). These symptoms present differently in each variant of FTD, namely the behavioral variant FTD (bvFTD), semantic dementia (SD), and progressive non-fluent aphasia (PNFA). Here we compare the structural neuroimaging of social cognition impairments in the bvFTD and PNFA variants.

Facial information (Baron-Cohen et al., 1997; Flombaum and Santos, 2005; Itier and Batty, 2009) and emotion recognition (Sollberger et al., 2010; Shany-Ur et al., 2012; Shany-Ur and Rankin, 2011) are crucial for social cognition (Ibanez et al., 2013a,b). Together with agency perception, they allow for the emergence of inferences about others' affective and cognitive mental states (theory of mind, ToM; Baron-Cohen et al., 1997).

Impaired social cognition and ToM have been shown in bvFTD (Rankin et al., 2003; Torralva et al., 2007, 2009; Manes et al., 2011). Voxel-based morphometry (VBM) studies in bvFTD patients have related those deficits to specific cortical and subcortical atrophy (Eslinger et al., 2007, 2011; Rankin et al., 2009). Furthermore, bvFTD is generally associated with atrophy of the frontal lobe (medial prefrontal, gyrus rectus, orbitofrontal/subgeneual cortices), amygdale, insula, right temporal pole, and white matter tracts, including the anterior corpus callosum, uncinate and arcuate fasiculus, and superior and inferior longitudinal fasciculi (Rosen et al., 2002a; Seeley, 2008; Whitwell et al., 2009; Hornberger et al., 2011; Agosta et al., 2012; Mendez and Shapira, 2013). These areas are engaged in emotion and affective states (Piguet et al., 2011) and also in the social cognition networks (Ibáñez and Manes, 2012; Couto et al., 2013a).

Emotion processing deficits have been also reported in SD (Rosen et al., 2004) and patients suffering from linguistic variants of FTD appear to have deficits in negative emotions (Rosen et al., 2004; Omar et al., 2011). However, no VBM analysis of ToM assessment in PNFA has been reported. This temporal variant of FTD is characterized by left perisylvian, insular and temporal atrophy (Rosen et al., 2002b; Gorno-Tempini et al., 2004; Whitwell et al., 2005; Mesulam et al., 2009).

There are some studies on the morphometry of PNFA related to deficits in emotion recognition (Kumfor et al., 2011; Rohrer et al., 2012; Zhang et al., 2013). However, no previous report has compared the specific atrophy of bvFTD and PNFA related to ToM and none has shown the differences between these samples in the domains of face recognition, basic emotion processing and their interaction with social cognition. Consequently, more research needs to be conducted in this area.

In this study, we describe the differential patterns of brain atrophy that are associated with face recognition, emotional processing, and ToM deficits in a sample of PNFA in comparison with bvFTD patients and control subjects. First, we assessed the profiles of behavioral performance on social cognition. We also compared the atrophy patterns in both FTD groups regarding controls. In addition, a VBM analysis was done on *a priori* selected regions of interest associated with social cognition impairments in FTD (Ibáñez and Manes, 2012), including the orbitofrontal (OFC), insular, amygdaline areas, and temporal pole. The identification of possible deterioration of this network in PNFA would help to clarify whether similar neural correlates for ToM are shared in both FTD groups.

## MATERIALS AND METHODS

### PARTICIPANTS

We recruited 22 patients who fulfilled criteria for diagnoses of behavioral variant of FTD ($n = 12$, 45% male; mean age = 69.81, $SD = 7.35$; mean years of education = 16, $SD = 6.51$) and PNFA ($n = 10$, 55% male; mean age = 64.9, $SD = 8.68$; mean years of education = 12.3, $SD = 4.49$). All patients were in early (mild) stages of the disease.

Patients were assessed and diagnosis was initially made by two experts in FTD (Patricia Montañés and Facundo Manes). Each patient was reviewed in the context of a multidisciplinary clinical meeting, where cognitive neurologists, psychiatrists, and neuropsychologists discussed each patient's case under the current criteria (Gorno-Tempini et al., 2011; Rascovsky et al., 2011). All FTD patients were recruited as part of a broader ongoing study on fronto-temporal dementia and underwent a standard examination battery including neurological, neuropsychiatric, and neuropsychological examinations, and in a separate session, a MRI. They all showed frontal or temporal atrophy on MRI, hence, they belong to probable FTD under the new diagnostic criteria (Gorno-Tempini et al., 2011). Inter-rater reliability for diagnosis was assessed (Cohen's $\kappa = 0.91$). The patients described in the present study presented with prominent changes in personality social behavior verified by a caregiver and did not meet criteria for any other psychiatric disorder.

Two groups of controls were assessed. Eighteen healthy controls were recruited for the behavioral assessment and were matched in age (mean, 6.24; *SD*, 7.24), gender (70% male) and years of education (mean, 14.5; *SD*, 3.71; see **Table 1**). They were recruited from a large pool of volunteers who did not have a history of drug abuse or a family history of neurodegenerative or psychiatric disorders. A second group of 12 healthy controls, age (mean, 60.63; *SD*, 4.59) and education (mean, 15.5; *SD*, 3.13) matched were scanned with a structural MRI to be compared with patients (See **Table 1**). All participants provided written informed consent in agreement with the declaration of Helsinki and the institution's ethics committee approved the study.

### BEHAVIORAL ASSESSMENT

Patients and behavioral controls' sample received a series of tasks previously reported by Torralva et al. (2009) that were designed to assess face recognition, facial emotion recognition, and ToM (Reading the mind in the eyes test, RMET; Baron-Cohen et al., 1997).

#### *Face recognition*

This test assesses the ability to discriminate permanent facial features and a person's identity. The task consists of 10 sheets, each one with 7 black and white pictures of faces of people of different age and gender, who are not facing the camera. The target picture is located at the top of each sheet and is repeated among other six options located at the bottom of the sheet. Participants are asked to match the target face with the option that corresponds to the same person. Performance was expressed as the number of correct responses out of the total, 10.

#### *Facial emotion recognition*

The task consists of a sheet with cartoons showing the face of a child with 8 different expressions including 5 different basic emotions (anger, sadness, surprise, fear, and happiness). The three remaining expressions were envy, rage, and crying. Participants

**Table 1 | Demographic data and results of cognitive state and language assessments.**

| | bvFTD (*n = 12*) | PNFA (*n = 10*) | Controls (*n = 18*) | bvFTD vs. Controls | PNFA vs. Controls | bvFTD vs. PNFA |
|---|---|---|---|---|---|---|
| **DEMOGRAPHIC DATA** | | | | | | |
| Age (years) | 69.8 (7.3) | 64.9 (8.6) | 69.8 (7.3) | N.S | N.S | N.S |
| Gender (F:M) | (5:7) | (4:6) | (6:12) | N.S | N.S | N.S |
| Education (years) | 16.0 (6.5) | 12.3 (4.4) | 14.7 (3.7) | N.S | N.S | N.S |
| **GENERAL COGNITIVE STATE AND LANGUAGE** | | | | | | |
| MMSE | 21.7 (8.2) | 21.1 (6.2) | 29.4 (0.6) | 0.002 | 0.001 | N.S |
| Naming | 48.7 (7.3) | 29.5 (19.1) | 57.6 (3.6) | 0.05 | 0.0001 | 0.02 |
| Phonological fluency | 6.4 (3.6) | 7.8 (2.4) | 14.5 (4.9) | 0.001 | 0.004 | N.S |
| Semantic fluency | 5.5 (3.9) | 8.8 (3.9) | 16.3 (3.0) | 0.0001 | 0.0002 | N.S |

*Note: The results are shown as the mean (SD). MMSE, Mini-mental state examination; bvFTD, behavioral variant FTD; PNFA, progressive non-fluent aphasia.*

were asked to identify which one of these pictures corresponded to the expression that the examiner indicated. Negative emotions were grouped in a single global score for VBM correlations. The total score (8) was derived from the number of correct responses.

#### *Reading the mind in the eyes test (RMET)*

This test (Baron-Cohen et al., 1997) assesses the ToM's emotional inference. This is a computerized test of 17 pictures of the eye region. Participants were asked to choose which of four words best described what the person was thinking or feeling in each photograph. The total score (17) was derived from the number of correct responses.

#### *Imaging recordings*

Both groups of patients (bvFTD and PNFA) and 12 controls participants were scanned in a 1.5 T Phillips Intera scanner with a standard head coil. A T1-weighted spin echo sequence was used to generate 120 contiguous axial slices ($TR = 2300$ ms; $TE = 13$ ms; flip angle = 68°; FOV = rectangular 256 mm; matrix size $= 256 \times 240 \times 120$; slice thickness = 1 mm) which covered all the brain surface and tissue.

#### *Voxel-based morphometry (VBM)*

Images were preprocessed for VBM analysis using DARTEL Toolbox and following procedures previously described (Ashburner and Friston, 2000). Following, modulated, 12 mm full-width half-maximum kernel smoothed as suggested in other reports (Good et al., 2001) and normalized to MNI space, images were analyzed within general linear models in SPM-8 2nd level analyses (http://www.fil.ion.ucl.ac.uk/spm/software/spm8).

First, a two-sample *t*-test between controls and the whole FTD group was performed in order to account for global atrophy pattern in patients, correcting by total intracranial volume. Second, we performed region of interest (ROI) analyses based on a fronto-insulo-temporal network to describe differential patterns of atrophy on these areas related to the social-context network (Ibáñez and Manes, 2012). The differential patterns of fronto-insulo-temporal atrophy associated with each specific behavioral impairment (face recognition, emotion, and ToM scores separately) in PNFA relative to bvFTD, were evaluated by testing the interaction terms in three different SPM regression designs. This has been shown to be powerful for assessing between-group differences in brain-behavior association slopes when adjusting for interaction terms which in our case, was intracranial volume (O'Brien et al., 2011). Following previous reports, statistical threshold was set at the $p < 0.05$ voxel level, besides small-volume corrected (Grossman et al., 2004) and cluster size corrected (Forman et al., 1995). The GM ROIs were defined *a priori* using the WFU-Pick Atlas (http://www.nitrc.org/projects/wfupickatlas/) in SPM8. Selected ROIs were lateral and medial OFC (Brodmann Area, BA 47/11), gyrus rectus (BA 11), fusiform gyrus (BA 37), bilateral temporal pole (BA 37–38), bilateral amygdale and bilateral insula (BA 13). It should be noted that correlation scatter plots of all tests reported here were assessed for the presence of outliers which may have affected the results, which revealed no outliers.

## RESULTS

### GENERAL COGNITIVE STATUS AND LANGUAGE ASSESSMENT

Compared with behavioral control sample, patients had no significant differences in age [$F_{(2,\ 36)} = 1.16$, $p = 0.32$], gender [$X^2_{(2)}$ Kruskal–Wallis $= 1.29$, $p = 0.25$], or education [$F_{(2,\ 36)} = 1.36$, $p = 0.27$]. Similarly, regarding MRI control sample, there were no significant differences in age [$F_{(2,\ 36)} = 0.27$, $p = 0.76$], gender [$X^2_{(2)}$ Kruskal–Wallis $= 6.41$, $p = 0.72$], nor in education [$F_{(2,\ 36)} = 2.10$, $p = 0.14$]. See **Table 1**.

As expected, differences among bvFTD and PNFA groups [$F_{(2,\ 36)} = 10.19$, $p < 0.01$] were observed in MMSE total score. *Post-hoc* bivariate comparisons (Tukey HSD, $MS = 30.74$, $df = 36$) showed that both bvFTD ($p < 0.01$) and PNFA ($p < 0.01$) patients had lower scores than controls. Furthermore, the performance on the naming task differed among groups [$F_{(2,\ 36)} = 21.15$, $p < 0.01$]. *Post-hoc* analysis (Tukey HSD, $MS = 103.85$, $df = 36$) evidenced that both patient groups, bvFTD ($p < 0.01$) and PNFA ($p < 0.01$) showed impairments compared to controls, and that PNFA scored lower than the bvFTD group ($p < 0.01$). Differences among groups were also observed in the phonological fluency task [$F_{(2,\ 36)} = 17$, $p < 0.01$]. According to *post-hoc* analysis (Tukey HSD, $MS = 17.50$, $df = 0$ 36), both bvFTD ($p < 0.01$) and PNFA ($p < 0.01$) performed lower than controls' group. Finally, groups also differed on the semantic fluency task performance [$F_{(2,\ 36)} = 33.14$, $p < 0.01$]. *Post-hoc* comparisons (Tukey HSD, $MS = 12.31$, $df = 36$) revealed that both bvFTD ($p < 0.01$) and PNFA ($p < 0.01$) patients exhibited lower fluency than controls (see **Table 1**).

### BEHAVIORAL RESULTS

Compared to controls, the bvFTD were not impaired on the face recognition task ($t = 1.18$, $p = 0.24$). However, bvFTD showed impairment in total emotion recognition score ($t = 2.28$, $p = 0.03$), and ToM ($t = 2.93$ $p < 0.01$) in comparison to normal subjects. The PNFA group showed significant impairments in face recognition ($t = 2.69$, $p = 0.01$). Furthermore, they also showed a trend to misrecognize emotions ($t = 1.93$, $p = 0.06$) and a significant deficit in ToM ($t = 4.80$, $p < 0.001$) when compared to healthy subjects (See **Figure 1** and **Table 2**).

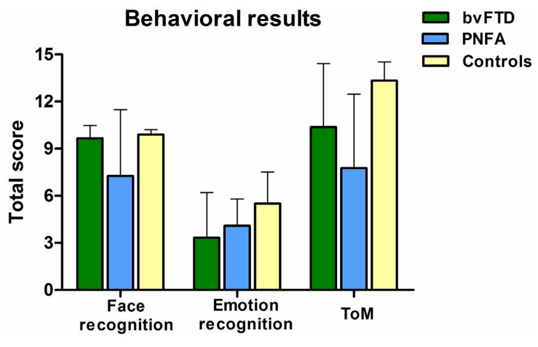


**FIGURE 1 | Scores of three behavioral tasks of face and emotion recognition and ToM.** PNFA patients in light blue, bvFTD patients in green, and controls in beige.

**Table 2 | Behavioral results descriptive and difference statistics.**

| | bvFTD (*n* = *12*) | PNFA (*n* = *10*) | Controls (*n* = *18*) | bvFTD vs. Controls | PNFA vs. Controls | bvFTD vs. PNFA |
|---|---|---|---|---|---|---|
| Face recognition | 9.6 (0.8) | 7.2 (4.2) | 9.8 (0.3) | N.S | 0.01 | N.S |
| Emotion recognition | 3.3 (2.8) | 4 (1.7) | 5.5 (2.0) | 0.03 | 0.06 | N.S |
| Theory of mind | 10.3 (4.0) | 7.7 (4.7) | 13.3 (1.1) | 0.006 | < 0.001 | N.S |

*The results are shown as the mean (SD). bvFTD, behavioral variant FTD; PNFA, progressive non-fluent aphasia. VBM, voxel-based morphometry, bvFTD, behavioral variant FTD; PNFA, progressive non-fluent aphasia; BA, Brodman Area; IFG, inferior frontal gyrus; OFC, orbitofrontal cortex.*

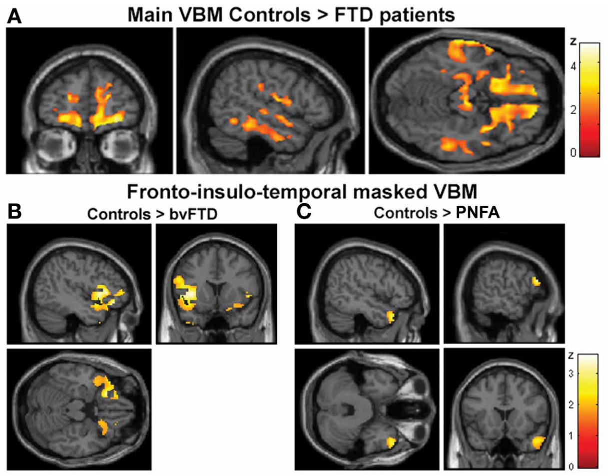


**FIGURE 2 | (A)** Main VBM results, atrophy of all FTD patients compared to controls, $p < 0.05$, $\alpha = 0.2$ FDR corrected. **(B)** Atrophy of bvFTD group compared to controls in the frontal, insular, and temporal regions, $p < 0.05$, small volume corrected. **(C)** Atrophy of PNFA group compared to controls in the frontal, insular, and temporal regions, $p < 0.05$, small volume corrected.

## VBM RESULTS

### *Global atrophy of patients compared to controls*

The VBM analysis revealed a pattern of global atrophy in the FTD group (both versions) in frontal and temporal lobe structures, as expected from and reported in previous studies (Rankin et al., 2006; Seeley et al., 2009; Garibotto et al., 2011). These included orbital (OFC) and motor cortices, right superior and mid temporal gyri, bilateral insula, right anterior cingulate cortex (ACC) and left parietal cortex ($p < 0.05$, $\alpha = 0.2$ FDR corrected. See **Figure 2A** and **Table 3A**).

Furthermore, regional analysis of the fronto-insulo-temporal sites in the bvFTD group relative to controls showed atrophy in bilateral anterior insula, lateral OFC bilaterally, left inferior frontal gyrus (IFG), left superior and right mid temporal pole regions ($p < 0.05$, small volume corrected, see **Figure 2B**). On the other hand, the same procedure revealed diminished GM in the PNFA group in right mid temporal pole and left IFG ($p < 0.05$, small volume corrected. See **Figure 2C** and **Table 3B**).

In sum, main GM loss was evidenced for FTD patients mainly in OFC, ACC, insula, and lateral temporal cortices. In addition, regional small volume corrected atrophy was greater for bvFTD, in frontal inferior, bilateral insula, and right temporal pole than for PNFA, where only left IFG and temporal pole atrophy was observed.

### *Relative atrophy of PNFA to bvFTD associated to face recognition, emotion, and social cognition*

In a second stage, the specific engagement of the fronto-insulo-temporal network on emotions and social cognition was assessed through ROI analyses. Face recognition in bvFTD was associated with GM *decreasing* in bilateral OFC ($r = 0.45$; $p < 0.05$), IFG ($r = 0.19$; $p < 0.05$), right gyrus rectus ($r = 0.36$; $p < 0.05$), and right insula ($r = 0.45$; $p < 0.05$). Bilateral fusiform gyrus involvement ($r = 0.56$; $p < 0.05$), temporal pole ($r = 0.55$; $p < 0.05$), insula ($r = 0.35$; $p < 0.05$) and IFG ($r = 0.46$; $p < 0.05$) were observed in PNFA (see **Figures 3A–F** and **Table 4A**).

Emotion recognition in bvFTD correlated with bilateral OFC ($r = 0.74$; $p < 0.05$), right gyrus rectus ($r = 0.46$; $p < 0.05$), and right insula ($r = 0.26$; $p < 0.05$). In PNFA, emotion recognition was associated with atrophy of bilateral insula ($r = 0.38$; $p < 0.05$), temporal pole ($r = 0.72$; $p < 0.05$) bilateral amygdala ($r = 0.23$; $p < 0.05$) and left rolandic opercula ($r = 0.62$; $p < 0.05$; see **Figures 4A–F**; **Table 4B**). The recognition of negative emotions correlated with bilateral OFC ($r = 0.51$; $p < 0.05$), and fronto-insular cortex (FIC, $r = 0.66$; $p < 0.05$) and right gyrus rectus ($r = 0.44$; $p < 0.05$) in bvFTD; and in PNFA with right insula ($r = 0.72$; $p < 0.05$) and right temporal pole ($r = 0.50$; $p < 0.05$).

Finally, ToM correlated with gyrus rectus ($r = 0.49$; $p < 0.05$) in the bvFTD group; and with bilateral insula ($r = 0.54$; $p < 0.05$), temporal pole ($r = 0.36$; $p < 0.05$) and amygdale ($r = 0.36$; $p < 0.05$) in the PNFA group (see **Figures 5A–E** and **Table 4C**).

Summarizing, face recognition was associated in bvFTD with atrophy in bilateral orbitofrontal regions, and right rectus gyrus, while in PNFA face recognition was related to right insula and bilateral fusiform gyrus. Emotion deficits were associated to relative atrophy of FIC (bvFTD) and of right temporal pole plus bilateral insula (PNFA), whereas ToM impairments correlated with fronto-insular areas (bvFTD) and with right insula and temporal pole (PNFA; See **Figures 3–5**; and **Table 4**).

## DISCUSSION

This work compared the structural neuroimaging signatures associated with face recognition, emotional processing and ToM impairments in bvFTD and PNFA. Compared to controls, impairments in face recognition (PNFA), emotion recognition and ToM (PNFA and bvFTD) were observed. VBM analysis showed the expected atrophy of orbital, medial, and lateral frontal structures, insula, temporal, and parietal cortices in both FTD versions compared to age and gender matched healthy controls. Furthermore, behavioral deficits were associated with different

**Table 3 | VBM results of whole brain atrophy of patients.**

| Region | | BA | Coordinates | | | *Z*-peak |
|---|---|---|---|---|---|---|
| | | | *x* | *y* | *z* | |
| **(A) VBM MAIN EFFECT*** | | | | | | |
| Superior Frontal Gyrus L | | 6 | −22.62 | 0.73 | 60.02 | 1.69 |
| Inferior Frontal Gyrus R | | 44 | 56.24 | 7.19 | 21 | 3.42 |
| Precentral gyrus L | | 4 | −62 | −7.5 | 22.5 | 1.93 |
| Supplementary Motor Area R | | 6/9 | 66 | 9 | 13.5 | 3.14 |
| Supplementary Motor Area L | | 6/9 | −15 | 0.73 | 63.9 | 2.81 |
| Anterior Cingulate R | | 32 | 5.82 | 39.5 | 0 | 3.01 |
| Frontal Superior Medial R | | 8 | 7.11 | 45.94 | 45.80 | 2.49 |
| Frontal Inferior Triangular L | | 8 | −9 | 36.90 | 57.43 | 2.37 |
| Frontal Superior Orbital R | | 11/47 | 13.57 | 42.06 | −18.85 | 2.65 |
| Frontal Superior Orbital L | | 11/47 | −34.26 | 30.79 | −21.43 | 1.79 |
| Rectus gyrus R | | 11 | 3 | 22.5 | −22.5 | 2.34 |
| Rectus gyrus L | | 11 | −4.5 | 49.5 | −22.5 | 2.07 |
| Insula R | | 13 | 40.7 | 18 | −11 | 2.14 |
| Insula L | | 13 | −31.67 | 21.8 | 3 | 1.83 |
| Superior Temporal Gyrus R | | 22 | 67.5 | −21 | −1.5 | 2.93 |
| Temporal Mid gyrus R | | 21 | 51 | 25.10 | −4.63 | 1.73 |
| Inferior Temporal Gyrus R | | 20 | 65.29 | −36.73 | −24 | 2.07 |
| Parahippocampal Gyrus L | | – | −29 | −25.10 | −16.50 | 2.03 |
| Fusiform Gyrus L | | 37 | −37.5 | −69 | −18 | 4.78 |
| Postcentral gyrus L | | 3. 1. 2 | −60 | −2252 | 28 | 2.56 |
| Supramarginal gyrus L | | 40 | −60 | −21.23 | 39.33 | 3.32 |
| **(B) SPECIFIC bvFTD AND PNFA ATROPHY**** | | | | | | |
| **bvFTD** | Anterior insula R | 13 | 45.90 | 7.19 | −2.04 | 2.58 |
| | Anterior insula L | 13 | −40.73 | 9.77 | −4.63 | 2.69 |
| | Lateral OFC R | 47 | 45.90 | 36.90 | −12.38 | 1.85 |
| | Lateral OFC L | 47 | −49.77 | 27.85 | −12.38 | 2.56 |
| | Temporal pole sup L | 38 | 47.19 | 11.06 | −43.41 | 1.97 |
| | Temporal pole mid R | 38 | −37.17 | 10.09 | −43.09 | 2.35 |
| **PNFA** | Temporal pole mid R | 38 | 52.36 | 18.81 | −29.19 | 2.73 |
| | IFG L | 45/46 | −56.24 | 26.56 | 16.06 | 2.68 |

*$^{*}p < 0.05$, FDR corrected $\alpha = 0.2$;*

*$^{**}p < 0.05$, Small volume corrected.*

patterns of atrophy in each FTD variant. ROI analysis looking at the fronto-insulo-temporal network in bvFTD revealed a pattern of bilateral orbitofrontal and gyrus rectus, right fronto-insula and insula associated with emotion and social cognition deficits respectively, whereas in PNFA these impairments were related to right insula and right temporal pole atrophy. This is the first report unraveling the structural correlates of face recognition, emotion, and ToM in both variants of FTD. We discuss these findings in the light of relatively differential (low level) emotion-face recognition processes vs. (higher level) social cognition impairments in these disorders.

### ATROPHY PATTERNS IN PNFA AND bvFTD AND THEIR RELATION WITH EMOTION/SOCIAL COGNITION PERFORMANCE

The general pattern of atrophy found in our FTD patients described above is in coincidence with previous morphometric evidence and distinct methodological techniques (Seeley et al., 2008; Zamboni et al., 2008; Rohrer et al., 2010). Moreover, functional and structural connectivity studies have shown atrophy in similar regions (Seeley, 2008; Zhou et al., 2010; Agosta et al., 2012; Mendez and Shapira, 2013) although the structural correlates of face recognition, emotion recognition, and ToM in both groups reveal different levels of social cognition impairments.

#### *Face recognition*

When analyzing the bvFTD fronto-insulo-temporal pattern of atrophy, face recognition performance correlated with orbitofrontal cortex. This region is associated with covert face recognition in prosopagnosia (Valdes-Sosa et al., 2011), with visual encoding of face stimuli (Frey and Petrides, 2003; Henson et al., 2003) and with familiar faces (Taylor et al., 2009). Hence, in bvFTD, OFC can be traced as a region subserving face perception.

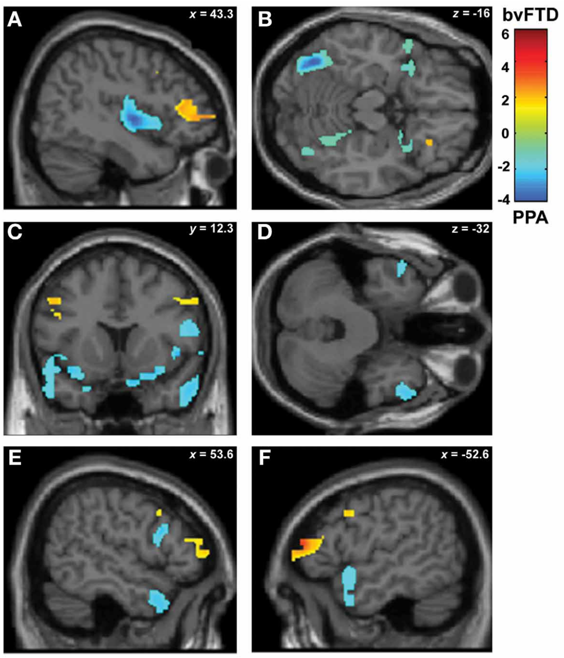


**FIGURE 3 | Contrast vectors of bvFTD (red-yellow) and PNFA (light blue) atrophy that correlates with face recognition.** $p < 0.05$, small volume corrected. **(A)** Right insular (PNFA) and IFG (bvFTD) correlations; **(B)** Bilateral posterior fusiform and temporal pole correlations for PNFA; **(C)** Bilateral insular and temporal pole (PNFA) and mid frontal gyrus (bvFTD) correlations; **(D)** Bilateral temporal pole in PNFA; **(E)** and **(F)**. Most lateral right **(E)** and left **(F)** frontal and temporal pole associations for PNFA and bvFTD. PNFA, progressive non-fluent aphasia, bvFTD, behavioral variant FTD; IFG, inferior frontal gyrus.

On the other hand, face recognition impairments in PNFA were related to atrophy of bilateral posterior fusiform gyrus, bilateral insular cortex, and anterior temporal lobe (ATL). Specifically, atrophy of the posterior fusiform is not an unexpected result, since the Fusiform Face Area (FFA) selectively engaged on early stages of face recognition, was initially described by Kanwisher et al. (1997) less than 10 mm further from our PNFA atrophy peak (MNI $x, y, z$ coordinates, left: −35.35 −64.35 −15.63; right: 40.40 −56.11 −15.16, see **Table 4** to compare with atrophy coordinates). On the other hand, an influential macaque study (Freiwald and Tsao, 2010) describes a network for face recognition with its ATL patches located in ventral and superior temporal pole as well as in the anterior bank of the STS, which has been recently confirmed in humans by a combination of fMRI meta-analytic and empirical results (Von Der Heide et al., 2013). In addition, a recent review by Gainotti (2007) shows that patients with right temporal pole damage are more prone to familiar face recognition deficits and poorer naming from facial (visual) stimuli than those who have left temporal pole lesions. This work also points to models of continuity between multimodal perceptual features and conceptual activities, leading to the emergence of familiarity feelings (Bruce and Young, 1986). In line with this wealth of evidence, we found a pattern of posterior fusiform gyrus and right temporal pole atrophy associated to face recognition scores in PNFA that would suggest an engagement of both early discriminative and person-specific stages of face recognition and supports their role in indexing semantic/biographical knowledge (Zahn et al., 2007; Mion et al., 2010; Ross and Olson, 2010; Simmons et al., 2010). Nonetheless, the process of face recognition includes the extraction of emotional expression (Haxby et al., 2000), which contributes to familiarity feelings and person perception (Young and Bruce, 2011). In the classical Bruce and Young model (1986), semantic processing is an integral part of the face structural processing which indexes the attribution of meaning, valence, and salience to facial expressions. With respect to this, we found bilateral insular atrophy in PNFA associated to face recognition, which we speculate may be accounting for the conveyed emotional component of this task (Nakamura et al., 2000; Josephs et al., 2008; Nielson et al., 2010). Furthermore, the structural connectivity of anterior insula, temporal pole, and orbitofrontal regions points to a dynamic interaction among these areas in ascribing emotional salience to perceived stimuli (Cloutman et al., 2012). Hence, we propose that the neuroanatomical correlates of PNFA's face recognition would engage both basic face recognition and emotional processing areas.

### *Emotion recognition*

In bvFTD, the fronto-insulo-temporal pattern of atrophy that correlated with emotion recognition included orbitofrontal, gyrus rectus and right posterior insula atrophy, regions which have been known to participate in assigning emotional valence to facial stimuli (Gobbini and Haxby, 2007). In addition, these emotion recognition deficits in PNFA involved some of the same regions related to face recognition deficits (insula, right temporal pole, and rolandic opercula). These are structures engaged in the binding of perception and visceral emotional responses (Olson et al., 2013; Visser et al., 2012). Additionally, we found that emotion recognition deficits in PNFA also related to specific atrophy in the left amygdale, a structure classically engaged in emotional processing (Kennedy and Adolphs, 2012) and previously related to emotion deficits in PNFA and SD (Whitwell et al., 2005; Garibotto et al., 2011; Yang et al., 2012). This corresponds to the temporo-insular (but not frontal) atrophy associated with ToM observed in PNFA, and it leads us to posit that lower level recognition of basic emotions in PNFA together with the face recognition impairments might represent basic deficits which could be triggering ToM deficits in these patients.

### *ToM*

Similarly, in bvFTD the ToM deficit was mainly associated with two core regions for social skills such as OFC and gyrus rectus bilaterally. These areas index high level social cognition processes (Viskontas et al., 2007; Nestor et al., 2012). Moreover, previous morphometric reports related OFC atrophy to ToM deficits in bvFTD (Mesulam et al., 2009; Eslinger et al., 2011).

Second, we found that the ToM deficit in PNFA was associated with temporal pole and insular cortex degeneration. These regions are engaged in semantic knowledge and in the integration of emotional body states and external milieu information respectively, two processes closely related to social cognition (Olson et al., 2007; Craig, 2009; Ibanez et al., 2010; Couto et al., 2013b;

**Table 4 | Patterns of trophy correlated with task performance in bvFTD and PNFA.**

| Regional GM-task | Region | BA | Coordinates | | | *Z*-peak |
|---|---|---|---|---|---|---|
| | | | *x* | *y* | *z* | |
| **(A) FACE RECOGNITION** | | | | | | |
| bvFTD > PNFA | Inferior frontal (IFG) R | 45 | 39 | 33 | 12 | 3.22 |
| | Orbitofrontal | 11/47 | 21.33 | 16.67 | −21.43 | 2.17 |
| | | 11/47 | −34.26 | 30.79 | −21.43 | 1.79 |
| PNFA > bvFTD | Insula R | 13 | 49.77 | 1.26 | 0 | 3.22 |
| | | | 40.7 | 18 | −11 | 2.14 |
| | Insula L | | −31.67 | 21.8 | 3 | 1.83 |
| | Fusiform L | 37 | −37.5 | −69 | −18 | 4.78 |
| | Fusiform R | 37 | 42 | −66 | −18 | 2.25 |
| | Temporal Pole L | 38 | −53.65 | 12.35 | −16.26 | 1.85 |
| | Temporal Pole R | 38 | 56.24 | 11 | −29.19 | 2.32 |
| | Inferior frontal (IFG) R | 45 | 55.5 | 12.35 | 17.35 | 2.10 |
| **(B) EMOTION RECOGNITION** | | | | | | |
| bvFTD > PNFA | Rectus | 11 | −4.5 | 22.5 | 22.5 | 2.34 |
| | | 11 | 3 | 49.5 | −22.5 | 2.07 |
| | Orbitofrontal R | 11/47 | 15 | 57 | −13.5 | 3.28 |
| | Insula R | 13 | 45.9 | −17.35 | 16 | 1.82 |
| PNFA > bvFTD | Insula | 13 | 43 | −12 | 7 | 3.07 |
| | | 13 | −38.14 | 7.19 | −2.04 | 4.67 |
| | Temporal Pole | 38 | −31.68 | 5.9 | −22.72 | 2.15 |
| | | 38 | 35.55 | 6 | −22.5 | 2.95 |
| | Amygdale | – | −25.21 | 2.02 | −18.85 | 2.95 |
| | | – | 35.55 | 6 | −22.5 | 2.95 |
| | Rolandic Opercula L | 46 | −47.19 | 4.6 | 12.18 | 1.92 |
| **(C) ToM** | | | | | | |
| bvFTD > PNFA | Rectus R | 11 | 5 | 50 | −22 | 1.98 |
| PNFA > bvFTD | Insula | 13 | 40.7 | 18 | −11 | 1.83 |
| | | 13 | −38.14 | 11.06 | −4.5 | 2.42 |
| | Amygdale | – | −26.5 | 4.6 | −18.85 | 2.86 |
| | | – | 25.21 | 0.73 | −19.5 | 2.44 |
| | Temporal pole | 38 | −40.73 | 3.31 | −17.55 | 2.31 |
| | | 38 | 34 | 11 | 8 | 1.99 |

*VBM, voxel-based morphometry; BA, Brodmann area; PNFA, progressive non-fluent aphasia; bvFTD, behavioral variant-FTD; ToM, theory of mind; R, rigth; L, left; GM, gray matter.*

Melloni et al., 2013). As mentioned above, biographical and emotional knowledge contribute to the recognition of faces. Further evidence suggests that temporal pole atrophy in SD could be related to ToM deficits (Duval et al., 2012) and right temporal pole atrophy has been associated to prosopagnosia in SD (Josephs et al., 2008). Therefore, right temporal pole and insular cortex atrophy in PNFA could be related to ToM deficits as they are associated with specific emotional and face recognition processes crucial for this social domain. This is convergent with the anatomical correlations of face and emotion recognition deficits described above. In other words, we suggest that the lower level emotional and face recognition deficits associated with the right temporal pole and insular atrophy could be the roots of subtle ToM impairments in PNFA, rather than a *sui generis* frontal involvement as is the case of bvFTD (Olson et al., 2007, 2013; Visser et al., 2012).

Although on the basis of these results we cannot rule out the possibility that concomitant loss of executive functions, language, and semantic memory may be the cause of social skills deployment in our PNFA sample, the hypothesis outlined above is also consistent with the multifactorial nature of emotional deficits proposed in FTD (Miller et al., 2012). Indeed, the fact that these recognition impairments impact the processing of social meaningful features on PNFA could be considered an extension of the semantic phenotype to the social and emotional domains. This parallels what has been shown to occur in bvFTD, in which the same ToM would be associated with other cognitive domains such as executive functions (Torralva et al., 2009). This has two

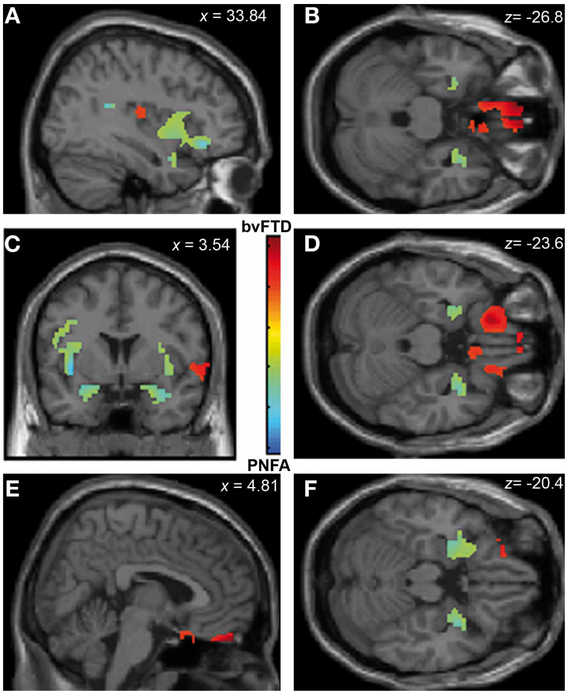


**FIGURE 4 | Contrast vectors of bvFTD (red-yellow) and PNFA (light blue) atrophy correlated with emotion recognition.** $p < 0.05$, small volume corrected. **(A)** Right insular and temporal correlations; **(B,D,F)** Orbitofrontal and mid temporal regions; **(C)** Bilateral insular correlations of PNFA emotions; **(E)** Right gyrus rectus for bvFTD. PNFA, progressive non-fluid aphasia, bvFTD, behavioral variant FTD.

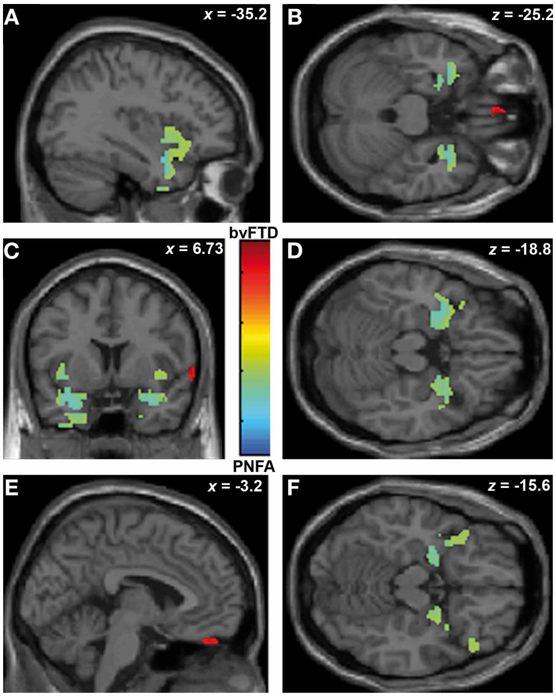


**FIGURE 5 | Contrast vectors of bvFTD (red-yellow) and PNFA (light blue) atrophy that correlates with ToM.** $p < 0.05$, small volume corrected. **(A)** Right insulo-temporal correlations for PNFA; **(B,D,F)** Temporal pole and amygdaline correlations for PNFA; **(C)** Bilateral insular correlations of PNFA; **(E)** Right gyrus rectus for bvFTD. PNFA, progressive non-fluid aphasia, bvFTD, behavioral variant FTD.

clinical implications: first, it suggests that different assessment batteries should be designed and applied for targeting divergent cross-domains associations in each variant (Torralva et al., 2009); second, the possible impact of more basic cognitive processes on social domains should be considered for the neuropsychological assessment of PNFA and bvFTD patients. Finally, our results suggest that different cognitive and neuroanatomic pathways would affect ToM performance, in a relationship that is not only restricted to executive functions as reported in many papers, but also extended to basic face recognition and emotion processing (Ibanez et al., 2013a,b).

A limitation of this study is that although they are age, gender, and formal education matched, MRI control subjects were not the same as the ones in which task impairments were compared. Therefore, there is an indirect relationship between patients and controls' brain morphology on the one hand, and specific VBM associations with behavioral impairments, on the other hand. However, as the neural underpinnings of social cognition impairment in bvFTD are well-characterized and a common substrate of atrophy is present in both versions of FTD, we intended here to use bvFTD patients as a comparison sample to reveal the neural correlates of social cognition deficits in PNFA as done in previous reports (Seeley et al., 2008). Hence, we posit this is a more straightforward way of comparing both variants of this neurodegenerative disease. Another limitation is the small sample size which could be susceptible of spurious correlation and may have biased the results of regression analyses. However, several previous studies cited in this paper have used a similar sample size (Eslinger et al., 2007; Seeley et al., 2008; Zamboni et al., 2008).

## CONCLUSIONS

In brief, our results suggest that primary face and emotion recognition impairments would impact on ToM in PNFA, whereas in bvFTD the ToM deficits seem to be a sui generis impairment, with preservation of basic face recognition. Both results point to possible existence of alternative pathways to ToM impairment in these conditions. Hence, although not behaviorally dissociable, ToM seems to be dissociable neuroanatomically and this would suggest extended circuits that support this function.

Although relatively similar impaired performance in social cognition was observed in both FTD groups, those similar impairments can be related to different processes and atrophy patterns in PNFA and bvFTD. Particularly, the development of social cognition signatures in PNFA would not be solely related to the typical atrophy of frontal social nodes as in bvFTD. The basic recognition and emotion stages of face recognition impairments in PNFA patients (related to right temporal pole and insular cortex atrophy), could account for the processing of social skills such as ToM. In addition, different regions of the so called social context network model (SCNM; Ibáñez and Manes, 2012) are selectively affected in both FTD groups, suggesting their participation on the dynamic interplay between invariant-specific and context-dependent variables and stimuli.

## ACKNOWLEDGMENTS

We thank the patients, their families and to the control subjects, as well as the research group of the Memory and Cognition San Ignacio Hospital - Javeriana University Ageing Institute and all the National University researchers. This research was partially supported by grants CONICET, CONICYT/FONDECYT Regular (1130920), PICT 2012-0412; PICT 2012-1309, Javeriana University PUJ-2009; Colciencias: 545-2011 and 537-2011; and INECO Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of those grants.

## REFERENCES

Agosta, F., Scola, E., Canu, E., Marcone, A., Magnani, G., Sarro, L., et al. (2012). White matter damage in frontotemporal lobar degeneration spectrum. *Cereb. Cortex* 22, 2705–2714. doi: 10.1093/cercor/bhr288

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry–the methods. *Neuroimage* 11, 805–821. doi: 10.1006/nimg.2000.0582

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Bruce, V., and Young, A. (1986). Understanding face recognition. *Br. J. Psychol.* 77, 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Cloutman, L. L., Binney, R. J., Drakesmith, M., Parker, G. J., and Lambon Ralph, M. A. (2012). The variation of function across the human insula mirrors its patterns of structural connectivity: evidence from *in vivo* probabilistic tractography. *Neuroimage* 59, 3514–3521. doi: 10.1016/j.neuroimage.2011.11.016

Couto, B., Sedeno, L., Sposato, L. A., Sigman, M., Riccio, P. M., Salles, A., et al. (2013a). Insular networks for emotional processing and social cognition: comparison of two case reports with either cortical or subcortical involvement. *Cortex* 49, 1420–1434. doi: 10.1016/j.cortex.2012.08.006

Couto, B., Salles, A., Sedeño, L., Peradejordi, M., Barttfeld, P., Canales-Johnson, A., et al. (2013b). The man who feels two hearts: different pathways of interoception. *Soc. Cogn. Affec. Neurosci.* doi: 10.1093/scan/nst108. [Epub ahead of print].

Craig, A. D. (2009). How do you feel–now. The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Duval, C., Bejanin, A., Piolino, P., Laisney, M., De La Sayette, V., Belliard, S., et al. (2012). Theory of mind impairments in patients with semantic dementia. *Brain* 135, 228–241. doi: 10.1093/brain/awr309

Eslinger, P. J., Moore, P., Anderson, C., and Grossman, M. (2011). Social cognition, executive functioning, and neuroimaging correlates of empathic deficits in frontotemporal dementia. *J. Neuropsychiatry Clin. Neurosci.* 23, 74–82. doi: 10.1176/appi.neuropsych.23.1.74

Eslinger, P. J., Moore, P., Troiani, V., Antani, S., Cross, K., Kwok, S., et al. (2007). Oops! Resolving social dilemmas in frontotemporal dementia. *J. Neurol. Neurosurg. Psychiatry* 78, 457–460. doi: 10.1136/jnnp.2006.098228

Flombaum, J. I., and Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Curr. Biol.* 15, 447–452. doi: 10.1016/j.cub.2004.12.076

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647. doi: 10.1002/mrm.1910330508

Freiwald, W. A., and Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* 330, 845–851. doi: 10.1126/science.1194908

Frey, S., and Petrides, M. (2003). Greater orbitofrontal activity predicts better memory for faces. *Eur. J. Neurosci.* 17, 2755–2758. doi: 10.1046/j.1460-9568.2003.02714.x

Gainotti, G. (2007). Different patterns of famous people recognition disorders in patients with right and left anterior temporal lesions: a systematic review. [Review]. *Neuropsychologia* 45, 1591–1607. doi: 10.1016/j.neuropsychologia.2006.12.013

Garibotto, V., Borroni, B., Agosti, C., Premi, E., Alberici, A., Eickhoff, S. B., et al. (2011). Subcortical and deep cortical atrophy in Frontotemporal Lobar Degeneration. *Neurobiol. Aging* 32, 875–884. doi: 10.1016/j.neurobiolaging.2009.05.004

Gobbini, M. I., and Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia* 45, 32–41. doi: 10.1016/j.neuropsychologia.2006.04.015

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., and Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14, 21–36. doi: 10.1006/nimg.2001.0786

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., et al. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Ann. Neurol.* 55, 335–346. doi: 10.1002/ana.10825

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., et al. (2011). Classification of primary progressive aphasia and its variants. *Neurology* 76, 1006–1014. doi: 10.1212/WNL.0b013e31821103e6

Grossman, M., McMillan, C., Moore, P., Ding, L., Glosser, G., Work, M., et al. (2004). What's in a name: voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain* 127, 628–649. doi: 10.1093/brain/awh075

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., and Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cereb. Cortex* 13, 793–805. doi: 10.1093/cercor/13.7.793

Hornberger, M., Geng, J., and Hodges, J. R. (2011). Convergent grey and white matter evidence of orbitofrontal cortex changes related to disinhibition in behavioural variant frontotemporal dementia. *Brain* 134, 2502–2512. doi: 10.1093/brain/awr173

Ibanez, A., Aguado, J., Baez, S., Huepe, D., Lopez, V., Ortega, R., et al. (2013a). From neural signatures of emotional modulation to social cognition: individual differences in healthy volunteers and psychiatric participants. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst067. [Epub ahead of print].

Ibanez, A., Huepe, D., Gempp, R., Gutierrez, V., Rivera-Rei, A., and Toledo, I. (2013b). Empathy, sex and fluid intelligence as predictors of theory of mind. *Pers. Individ. Dif.* 54, 616–621. doi: 10.1016/j.paid.2012.11.022

Ibanez, A., Gleichgerrcht, E., and Manes, F. (2010). Clinical effects of insular damage in humans. *Brain Struct. Funct.* 214, 397–410. doi: 10.1007/s00429-010-0256-y

Ibáñez, A., and Manes, F. (2012). Contextual social cognition and the behavioral variant of frontotemporal dementia. *Neurology* 78, 1354–1362. doi: 10.1212/WNL.0b013e3182518375

Itier, R. J., and Batty, M. (2009). Neural bases of eye and gaze processing: the core of social cognition. *Neurosci. Biobehav. Rev.* 33, 843–863. doi: 10.1016/j.neubiorev.2009.02.004

Josephs, K. A., Whitwell, J. L., Vemuri, P., Senjem, M. L., Boeve, B. F., Knopman, D. S., et al., Jr. (2008). The anatomic correlate of prosopagnosia in semantic dementia. *Neurology* 71, 1628–1633. doi: 10.1212/01.wnl.0000334756.18558.92

Kanwisher, N., Mcdermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Kennedy, D. P., and Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends Cogn. Sci.* 16, 559–572. doi: 10.1016/j.tics.2012.09.006

Kumfor, F., Miller, L., Lah, S., Hsieh, S., Savage, S., Hodges, J. R., et al.

(2011). Are you really angry. The effect of intensity on facial emotion recognition in frontotemporal dementia. *Soc. Neurosci.* 6, 502–514. doi: 10.1080/17470919.2011.620779

Manes, F., Torralva, T., Ibanez, A., Roca, M., Bekinschtein, T., and Gleichgerrcht, E. (2011). Decision-making in frontotemporal dementia: clinical, theoretical and legal implications. *Dement. Geriatr. Cogn. Disord.* 32, 11–17. doi: 10.1159/000329912

Melloni, M., Lopez, V., and Ibanez, A. (2013). Empathy and contextual social cognition. *Cogn. Affect. Behav. Neurosci.* doi: 10.3758/s13415-013-0205-3

Mendez, M. F., and Shapira, J. S. (2013). Hypersexual behavior in frontotemporal dementia: a comparison with early-onset Alzheimer's disease. *Arch. Sex Behav.* 42, 501–509. doi: 10.1007/s10508-012-0042-4

Mesulam, M., Rogalski, E., Wieneke, C., Cobia, D., Rademaker, A., Thompson, C., et al. (2009). Neurology of anomia in the semantic variant of primary progressive aphasia. *Brain* 132, 2553–2565. doi: 10.1093/brain/awp138

Miller, L. A., Hsieh, S., Lah, S., Savage, S., Hodges, J. R., and Piguet, O. (2012). One size does not fit all: face emotion processing impairments in semantic dementia, behavioural-variant frontotemporal dementia and Alzheimer's disease are mediated by distinct cognitive deficits. *Behav. Neurol.* 25, 53–60.

Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., et al. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain* 133, 3256–3268. doi: 10.1093/brain/awq272

Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., et al. (2000). Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. *Brain* 123(Pt 9), 1903–1912. doi: 10.1093/brain/123.9.1903

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., et al. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* 51, 1546–1554. doi: 10.1212/WNL.51.6.1546

Nestor, P. G., Nakamura, M., Niznikiewicz, M., Thompson, E., Levitt, J. J., Choate, V., et al. (2012). In search of the functional neuroanatomy of sociality: MRI subdivisions of orbital frontal cortex and social cognition. *Soc. Cogn. Affect. Neurosci.* 8, 460–467. doi: 10.1093/scan/nss018

Nielson, K. A., Seidenberg, M., Woodard, J. L., Durgerian, S., Zhang, Q., Gross, W. L., et al. (2010). Common neural systems associated with the recognition of famous faces and names: an event-related fMRI study. *Brain Cogn.* 72, 491–498. doi: 10.1016/j.bandc.2010.01.006

O'Brien, L. M., Ziegler, D. A., Deutsch, C. K., Frazier, J. A., Herbert, M. R., and Locascio, J. J. (2011). Statistical adjustments for brain size in volumetric neuroimaging studies: some practical implications in methods. *Psychiatry Res.* 193, 113–122. doi: 10.1016/j.pscychresns.2011.01.007

Olson, I. R., Mccoy, D., Klobusicky, E., and Ross, L. A. (2013). Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc. Cogn. Affect. Neurosci.* 8, 123–133. doi: 10.1093/scan/nss119

Olson, I. R., Plotzker, A., and Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain* 130, 1718–1731. doi: 10.1093/brain/awm052

Omar, R., Henley, S. M., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., et al. (2011). The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *Neuroimage* 56, 1814–1821. doi: 10.1016/j.neuroimage.2011.03.002

Perry, R. J., and Miller, B. L. (2001). Behavior and treatment in frontotemporal dementia. *Neurology* 56, S46–S51. doi: 10.1212/WNL.56.suppl_4.S46

Piguet, O., Hornberger, M., Mioshi, E., and Hodges, J. R. (2011). Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. *Lancet Neurol.* 10, 162–172. doi: 10.1016/S1474-4422(10)70299-4

Rankin, K. P., Gorno-Tempini, M. L., Allison, S. C., Stanley, C. M., Glenn, S., Weiner, M. W., et al. (2006). Structural anatomy of empathy in neurodegenerative disease. *Brain* 129, 2945–2956. doi: 10.1093/brain/awl254

Rankin, K. P., Kramer, J. H., Mychack, P., and Miller, B. L. (2003). Double dissociation of social functioning in frontotemporal dementia. *Neurology* 60, 266–271. doi: 10.1212/01.WNL.0000041497.07694.D2

Rankin, K. P., Salazar, A., Gorno-Tempini, M. L., Sollberger, M., Wilson, S. M., Pavlic, D., et al. (2009). Detecting sarcasm from paralinguistic cues: anatomic and cognitive correlates in neurodegenerative disease. *Neuroimage* 47, 2005–2015. doi: 10.1016/j.neuroimage.2009.05.077

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., et al. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 134, 2456–2477. doi: 10.1093/brain/awr179

Ritchie, K., and Lovestone, S. (2002). The dementias. *Lancet* 360, 1759–1766. doi: 10.1016/S0140-6736(02)11667-9

Rohrer, J. D., Ridgway, G. R., Crutch, S. J., Hailstone, J., Goll, J. C., Clarkson, M. J., et al. (2010). Progressive logopenic/phonological aphasia: erosion of the language network. *Neuroimage* 49, 984–993. doi: 10.1016/j.neuroimage.2009.08.002

Rohrer, J. D., Sauter, D., Scott, S., Rossor, M. N., and Warren, J. D. (2012). Receptive prosody in nonfluent primary progressive aphasias. *Cortex* 48, 308–316. doi: 10.1016/j.cortex.2010.09.004

Rosen, H. J., Gorno-Tempini, M. L., Goldman, W. P., Perry, R. J., Schuff, N., Weiner, M., et al. (2002a). Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology* 58, 198–208. doi: 10.1212/WNL.58.2.198

Rosen, H. J., Kramer, J. H., Gorno-Tempini, M. L., Schuff, N., Weiner, M., and Miller, B. L. (2002b). Patterns of cerebral atrophy in primary progressive aphasia. *Am. J. Geriatr. Psychiatry* 10, 89–97.

Rosen, H. J., Pace-Savitsky, K., Perry, R. J., Kramer, J. H., Miller, B. L., and Levenson, R. W. (2004). Recognition of emotion in the frontal and temporal variants of frontotemporal dementia. *Dement. Geriatr. Cogn. Disord.* 17, 277–281. doi: 10.1159/000077154

Ross, L. A., and Olson, I. R. (2010). Social cognition and the anterior temporal lobes. *Neuroimage* 49, 3452–3462. doi: 10.1016/j.neuroimage.2009.11.012

Seeley, W. W. (2008). Selective functional, regional, and neuronal vulnerability in frontotemporal dementia. *Curr. Opin. Neurol.* 21, 701–707. doi: 10.1097/WCO.0b013e3283168e2d

Seeley, W. W., Crawford, R., Rascovsky, K., Kramer, J. H., Weiner, M., Miller, B. L., et al. (2008). Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. *Arch. Neurol.* 65, 249–255. doi: 10.1001/archneurol.2007.38

Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., and Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron* 62, 42–52. doi: 10.1016/j.neuron.2009.03.024

Shany-Ur, T., Poorzand, P., Grossman, S. N., Growdon, M. E., Jang, J. Y., Ketelle, R. S., et al. (2012). Comprehension of insincere communication in neurodegenerative disease: lies, sarcasm, and theory of mind. *Cortex* 48, 1329–1341. doi: 10.1016/j.cortex.2011.08.003

Shany-Ur, T., and Rankin, K. P. (2011). Personality and social cognition in neurodegenerative disease. *Curr. Opin. Neurol.* 24, 550–555. doi: 10.1097/WCO.0b013e32834cd42a

Simmons, W. K., Reddish, M., Bellgowan, P. S., and Martin, A. (2010). The selectivity and functional connectivity of the anterior temporal lobes. *Cereb. Cortex* 20, 813–825. doi: 10.1093/cercor/bhp149

Sollberger, M., Rankin, K. P., and Miller, B. L. (2010). Social cognition. *Continuum (Minneap Minn)* 16, 69–85.

Taylor, M. J., Arsalidou, M., Bayless, S. J., Morris, D., Evans, J. W., and Barbeau, E. J. (2009). Neural correlates of personally familiar faces: parents, partner and own faces. *Hum. Brain Mapp.* 30, 2008–2020. doi: 10.1002/hbm.20646

Torralva, T., Kipps, C. M., Hodges, J. R., Clark, L., Bekinschtein, T., Roca, M., et al. (2007). The relationship between affective decision-making and theory of mind in the frontal variant of fronto-temporal dementia. *Neuropsychologia* 45, 342–349. doi: 10.1016/j.neuropsychologia.2006.05.031

Torralva, T., Roca, M., Gleichgerrcht, E., Bekinschtein, T., and Manes, F. (2009). A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. *Brain* 132, 1299–1309. doi: 10.1093/brain/awp041

Valdes-Sosa, M., Bobes, M. A., Quinones, I., Garcia, L., Valdes-Hernandez, P. A., Iturria, Y., et al. (2011). Covert face recognition without the fusiform-temporal pathways. *Neuroimage* 57, 1162–1176. doi: 10.1016/j.neuroimage.2011.04.057

Viskontas, I. V., Possin, K. L., and Miller, B. L. (2007). Symptoms of frontotemporal dementia provide insights into orbitofrontal cortex function and social behavior. *Ann. N.Y. Acad. Sci.* 1121, 528–545. doi: 10.1196/annals.1401.025

Visser, M., Jefferies, E., Embleton, K. V., and Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *J. Cogn. Neurosci.* 24, 1766–1778. doi: 10.1162/jocn_a_00244

Von Der Heide, R. J., Skipper, L. M., and Olson, I. R. (2013). Anterior temporal face patches: a meta-analysis and empirical study. *Front. Hum. Neurosci.* 7:17. doi: 10.3389/fnhum.2013.00017

Whitwell, J. L., Przybelski, S. A., Weigand, S. D., Ivnik, R. J., Vemuri, P., Gunter, J. L., et al. (2009). Distinct anatomical subtypes of the behavioural variant of frontotemporal dementia: a cluster analysis study. *Brain* 132, 2932–2946. doi: 10.1093/brain/awp232

Whitwell, J. L., Sampson, E. L., Watt, H. C., Harvey, R. J., Rossor, M. N., and Fox, N. C. (2005). A volumetric magnetic resonance imaging study of the amygdala in frontotemporal lobar degeneration and Alzheimer's disease. *Dement. Geriatr. Cogn. Disord.* 20, 238–244. doi: 10.1159/000087343

Yang, J., Pan, P., Song, W., and Shang, H. F. (2012). Quantitative meta-analysis of gray matter abnormalities in semantic dementia. *J. Alzheimers Dis.* 31, 827–833.

Young, A. W., and Bruce, V. (2011). Understanding person perception. [Review] *Br. J. Psychol.* 102, 959–974. doi: 10.1111/j.2044-8295.2011.02045.x

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., and Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6430–6435. doi: 10.1073/pnas.0607061104

Zamboni, G., Huey, E. D., Krueger, F., Nichelli, P. F., and Grafman, J. (2008). Apathy and disinhibition in frontotemporal dementia: insights into their neural correlates. *Neurology* 71, 736–742. doi: 10.1212/01.wnl.0000324920.96835.95

Zhang, Y., Tartaglia, M. C., Schuff, N., Chiang, G. C., Ching, C., Rosen, H. J., et al. (2013). MRI signatures of brain macrostructural atrophy and microstructural degradation in frontotemporal lobar degeneration subtypes. *J. Alzheimers Dis.* 33, 431–444.

Zhou, J., Greicius, M. D., Gennatas, E. D., Growdon, M. E., Jang, J. Y., Rabinovici, G. D., et al. (2010). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain* 133, 1352–1367. doi: 10.1093/brain/awq075

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 02 June 2013; accepted: 26 July 2013; published online: 16 August 2013.*
*Citation: Couto B, Manes F, Montañés P, Matallana D, Reyes P, Velasquez M, Yoris A, Baez S and Ibáñez A (2013) Structural neuroimaging of social cognition in progressive non-fluent aphasia and behavioral variant of frontotemporal dementia. Front. Hum. Neurosci. **7**:467. doi: 10.3389/fnhum.2013.00467*
*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 11 October 2013
doi: 10.3389/fnhum.2013.00641

# Emotional modulation of the attentional blink and the relation to interpersonal reactivity

**Philipp Kanske[1]*†, Sandra Schönfelder[2]† and Michèle Wessa[2]**

[1] Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[2] Department of Clinical Psychology and Neuropsychology, Johannes Gutenberg University, Mainz, Germany

**Edited by:**
Maria Ruz, Universidad de Granada, Spain

**Reviewed by:**
Nicholas Thomas Van Dam, New York University School of Medicine, USA
Valentina Rossi, Ghent University, Belgium

***Correspondence:**
Philipp Kanske, Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, Leipzig 04103, Germany
e-mail: kanske@cbs.mpg.de

†These authors have contributed equally to this work.

The extent of the attentional blink effect on detection rates in rapid serial visual presentations is modulated by the emotionality of the stimuli. Emotionally salient stimuli are detected more often, even if presented in the attentional blink period, and elicit an enlarged P3 response, which has been interpreted as enhanced consolidation. This effect correlates with individual differences in trait affectivity such as anxiety or dysphoria. Here, we ask if it is also related to the capacity to detect emotions in others, i.e., to interpersonal social traits. We therefore presented emotional and neutral images depicting social scenes as targets in an attentional blink design and measured detection rates and event-related potentials. In addition, we recorded self-reports of empathy as measured by the Interpersonal Reactivity Index. The results show enhanced performance for emotional stimuli and increased P3 amplitudes, which correlated with individual differences in empathy. The data suggest that self-reported empathy goes along with enhanced processing of emotion in social stimuli, even under stimulus conditions that are suboptimal for conscious target detection.

**Keywords: P3 event-related potential, electroencephalography, event-related potentials, empathy, attentional blink, emotions**

## INTRODUCTION

The attentional blink phenomenon occurs when two stimuli are presented briefly one after the other, with the first stimulus impairing processing of the second stimulus because attentional resources cannot be sufficiently allocated (Raymond et al., 1992). The attentional blink effect is most widely observed in rapid serial visual presentation paradigms when the two target stimuli that need to be identified are embedded in a stream of distractor stimuli (Martens and Wyble, 2010). While identification of the first target (T1) is typically not affected, performance on the second target (T2) is impaired, if it is presented about 200 to 500 ms after the first target, but not in earlier or later time windows. The critical question regarding the stages at which processing of the second target is interrupted has been thoroughly addressed with the help of event-related potentials (ERPs) of the electroencephalogram. These data showed that the N1, P1, and N400 components of the ERP in response to T2 are not altered, which suggests that early sensory processing and semantic analysis of T2 are intact (Vogel et al., 1998). The P3 component however is typically found to be reduced for non-identified T2, which has been interpreted as impaired consolidation of the stimulus in working memory (Rolke et al., 2001; Vogel and Luck, 2002; Kranczioch et al., 2003). The suggestion that the P3 indexes working memory processes is already relatively old and based on data showing its sensitivity to the probability of task-defined stimulus categories (Donchin, 1981; Donchin and Coles, 1988). Nevertheless it is still consistent with the accumulated evidence, for example, through manipulations of memory load or subsequent recognition (for a review on P3 function, see Polich, 2007).

The size of the attentional blink effect is influenced by a number of different factors including personal relevance (Shapiro et al., 1997) and emotionality of the target stimuli (Anderson and Phelps, 2001). If the second target is emotionally negative or positive, the attentional blink is reduced, such that more of these stimuli are detected than neutral targets. In a series of experiments, Keil and Ihssen (2004) showed that this effect is related to emotional arousal rather than valence. In line with these behavioral effects, the amplitude of the P3 has been found to be enlarged in response to correctly identified emotional T2 stimuli (Trippe et al., 2007). These results have been interpreted as preferential selection of affective information, facilitating working memory consolidation (Keil and Ihssen, 2004).

Inter-individual variations in the extent to which emotions influence the attentional blink effect have also been repeatedly observed, particularly for individual differences in trait anxiety (Fox et al., 2005; Van Dam et al., 2012) or dysphoria (Koster et al., 2009). Further, the attentional blink effect is altered in mental disorders with clinically relevant changes in emotion processing, such as specific phobia (Trippe et al., 2007) or post-traumatic stress disorder (Amir et al., 2009). These correlations speak to the influence of a participant's own affect on the processing of emotional stimuli embedded in a rapidly presented visual stream. However, it remains an open question whether the emotional modulation of the attentional blink also relates to individual differences in the capacity to react to emotions in others (i.e., in interpersonal social traits). The term empathy has often been used in a broad sense, encompassing multiple facets of interpersonal reactivity, as in the frequently used Interpersonal Reactivity Index (IRI; Davis, 1983b), which asks for trait capabilities in emotional and more cognitive reactions towards others. Enhanced processing of emotional stimuli may be a basis for empathic reactions as the identification of the emotional content of a social scene necessarily precedes a reaction on the side of the observer. Interestingly, empathy has also been related to social reward sensitivity and social attention, which might mediate this relation. In a study of 8–12-year-old children, parent-reported empathic skills were related to behavioral benefits due to social reward, but not to monetary reward (Kohls et al., 2009). Data from individuals diagnosed with high functioning autism who showed deficits in empathic skills also indicates reduced sensitivity to social, but not monetary reward (Demurie et al., 2011; Delmonte et al., 2012). Probably related to this, autism is also characterized by drastically reduced attention to social stimuli and a preference to attend to non-social objects (for a review, see Dawson et al., 2012). Better differentiation of social stimuli could be a mechanism that enables them to be perceived as rewarding and to be attended more, which in turn allows for empathic reactions. There is already some indication that the reactivity to emotional stimuli correlates with empathy as measured by the IRI (Davis, 1983b). Silani et al. (2008) reported that in a group of individuals with high functioning autism an increase in activity of the amygdala in response to emotional images was related to empathy scores. Similarly, empathy scores correlated with amygdala responses to emotional faces in a developmental study of 10-year-old children (Pfeifer et al., 2008). The present study aims at extending these results to the attentional blink paradigm and tests healthy adults with a wider age range. The attentional blink task offers the advantage of testing emotion detection under stimulus conditions that are suboptimal for conscious target detection. Thereby it allows us to address the important question whether, under difficult stimulation conditions and high working memory load, individuals scoring high in empathic traits are more sensitive to emotional expressions and process these more deeply than those scoring low in empathy. We hypothesized that this difference is reflected in amplitude modulations of the P3, indicating enhanced consolidation of emotional T2 stimuli in high empathic individuals.

To this end, we used an attentional blink paradigm with emotional and neutral images as T2. They always followed neutral T1 stimuli and were embedded in streams of neutral distractors. The depicted scenes to be analyzed were always "social" in that humans were displayed. During the task, an electroencephalogram was recorded. In addition, we acquired the IRI (Davis, 1983b) from each participant and correlated the respective sum and subscale scores to the emotional modulation of the attentional blink effect. We expected (1) to find the behavioral attentional blink effect in lower T2 than T1 recognition rates, (2) a modulation of this effect by emotion such that emotional T2 are recognized more often than neutral T2, (3) a reflection of this effect in the ERP with larger P3 amplitudes to emotional than neutral T2 and (4) a positive correlation of this effect to inter-individual differences in empathy scores acquired with the IRI (Davis, 1983b). As only few studies have investigated the emotional modulation of the attentional blink with pictures (instead of written words, for example, Trippe et al., 2007) and mainly used objects rather than social scenes for the neutral condition, the results of the present study with social scenes in neutral and emotional conditions will allow broader generalizability, which is particularly relevant for testing its relation to interpersonal social traits.

## METHODS

### PARTICIPANTS

Twenty-seven healthy individuals (16 females) between 19 and 56 years (mean age = 31.07 years, SD = 11.13) took part in this study. All participants were native German speakers, right-handed according to the Edinburgh handedness inventory (Oldfield, 1971), and had normal or corrected-to-normal vision. Prior to their enrollment in this study, each participant was screened by telephone by an experienced clinical psychologist for exclusion criteria that included current or lifetime mental disorders, visual or hearing impairments, a lifetime history of head injury with loss of consciousness, brain damage or surgery, the presence of a cardiovascular disease, neurological illness, and regular use of medication (except for oral contraceptives). The presence of mental disorders, including alcohol or drug abuse, was evaluated by screening items that relied on the key diagnostic questions from the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID-I; German version: Wittchen et al., 1997). The study protocol was approved by the Ethics Committee of the Medical Faculty Mannheim, Heidelberg University, and written informed consent was obtained from each subject prior to the experimental session.

### INTERPERSONAL REACTIVITY INDEX

Subjects filled out the German translation of the IRI (original English version: Davis, 1980). This 28-item questionnaire assesses empathy in the form of statements that have to be agreed or disagreed on a Likert-type scale ranging from 1 ("does not describe me well") to 5 ("describes me very well"). The IRI can be subdivided into four subscales, each comprising seven items: "Empathic concern" and "Personal distress" refer to the two affective empathy dimensions that measure respondents' other-oriented feelings of compassion, warmth and concern for unfortunate others, and respondents' self-related discomfort and anxiety arising from observing other people's suffering, respectively. The subscales "Perspective taking" and "Fantasy" determine more about the cognitive empathy domain (Shamay-Tsoory et al., 2009) and inquire about the ability to mentally adopt the perspective of others as well as the tendency to identify with characters in fictional situations (e.g., movies and novels), respectively. Typically, an overall IRI score is calculated as index of the general capacity to empathize with others. The IRI subscales of the original American version (Davis, 1980) possess good psychometric properties with Cronbach's α coefficients of internal consistency ranging from 0.71 to 0.77 and test–retest reliability coefficients ranging from 0.61 to 0.81 (Davis, 1980). The basic psychometric quality of our German translation of the IRI was comparable in the present sample with satisfying Cronbach's α coefficients (i.e., 0.79 for "Fantasy," 0.82 for "Perspective taking," 0.78 for "Personal distress," and 0.68 for "Empathic concern").

### STIMULUS MATERIAL

The computerized attentional blink task comprised picture stimuli selected from the International Affective Picture System (IAPS;

Lang et al., 2005), a standardized in-house set of emotionally evocative pictures (Emotional Picture Set (EmoPicS); Wessa et al., 2010) and public internet photo libraries. The T2 stimuli were separated into three affective categories that all portrayed humans (see **Table 1** for mean valence and arousal ratings for the stimulus selection). Ten pictures depicted negative scenes of human violence, mutilation, loss, and illness, 10 neutral pictures showed human faces or people doing ordinary activities, and 10 pictures displayed positive scenes including happy families, erotic couples, and exciting sports. Based upon the normative data provided for the IAPS and EmoPicS databases, the emotional categories differed statistically with respect to valence [$F(2,87) = 1683.06$; $p < 0.001$] and arousal [$F(2,87) = 214.75$; $p < 0.001$]. To minimize sex differences (cf. Bradley et al., 2001), only pictures with relatively small gender differences in normative ratings (within 1.5 points on the 9-point scales for both affective valence and arousal) were included. To control for physical picture parameters, luminance, contrast, and color composition (red, blue, and green layer) values were extracted for each image using the histogram function of Adobe PhotoShop© software (version 9.0; Adobe Systems Inc., San Jose, CA, USA). A subsequent multivariate analyses of variance (ANOVA) performed on these measures did not reveal any significant differences between the categories [luminance: $F(2,87) = 1.16$; $p = 0.318$, contrast: $F(2,87) = 1.13$; $p = 0.327$, red layer: $F(2,87) = 0.79$; $p = 0.457$, green layer: $F(2,87) = 1.15$; $p = 0.320$, blue layer: $F(2,87) = 2.77$; $p = 0.068$]. The T2 stimuli were repeated such that there were 5 presentations of each neutral, positive, and negative image of a human as T2, leading to a total of 150 trials. Additionally, 150 trials with T2 stimuli depicting plants and animals were presented, each category appearing in 75 trials. These two semantic categories were added in order to reduce the chance level identification rate of T2 stimuli but were not part of the statistical data analysis. Overall, 40 neutral pictures of plants and 40 neutral pictures of animals served as stimuli for T1 and T2. Ten additional neutral pictures of humans were incorporated as T1. The attentional blink experiment comprised 300 trials with pictures of animals, plants, and humans occurring equally often as targets (T1 and T2). More specifically, each category was presented 200 times as target, such that human images appeared in 50 trials as T1 stimulus and in 150 trials as T2 (i.e., with 50 trials per emotion category), whereas plants and animals were shown in 125 of all trials as T1, and in 75 trials as T2, respectively. Finally, distractors encompassed 45 additional neutral pictures that depicted a wide range of plants, animals, and humans.

**Table 1 | Mean valence and arousal ratings and standard deviations (in parentheses) for T2 images.**

| | Normative IAPS and EmoPicS ratings | | Sample ratings | |
|---|---|---|---|---|
| | Valence | Arousal | Valence | Arousal |
| Negative | 2.02 (0.48) | 6.71 (0.82) | 2.23 (0.63) | 5.81 (1.31) |
| Neutral | 5.01 (0.36) | 3.16 (0.35) | 5.18 (0.28) | 1.98 (0.88) |
| Positive | 7.15 (0.32) | 6.30 (0.62) | 6.98 (0.86) | 4.76 (1.86) |

*Normative IAPS and EmoPicS ratings and the ratings of the present sample are displayed.*

### EXPERIMENTAL PARADIGM AND PROCEDURE

During the attentional blink paradigm, participants were asked to identify two target images (T1 and T2), which were present in a stream of distractor images. The second target (T2) was always presented shortly (310 ms) after the first (T1), so that it fell in the attentional blink period (Kranczioch et al., 2003). The sequence of events in a trial (see **Figure 1**) was the following: after the presentation of a fixation cross for 1500 ms, a stream of 2–10 distracter images indicated by a white frame was presented. Each of these images was presented for 155 ms. A first target image (T1), indicated by a red frame, was shown next, followed by a single distractor image, the second target image (T2), and another 10 distractor images. At the end of each trial participants saw two successive response screens asking whether T1 and T2 showed (1) an animal, (2) a human, (3) a plant, (4) or if they did not know the answer. They responded via button press with the right hand. The response screens were presented until the participant pressed a button. The total number of trials was 300. These were split up into five blocks with short breaks. The experiment lasted about 45 min.

### EEG DATA ACQUISITION AND ANALYSIS

Electroencephalography activity was recorded from 60 scalp sites (Fpz, Fz, Fcz, Cz, Cpz, Pz, POz, Oz, Fp1/2, AF3/4, AF7/8, F1/2, F3/4, F5/6, F7/8, FT7/8, Fc1/2, FC3/4, FC5/6, T7/8, C1/2, C3/4, C5/6, CP1/2, CP3/4, CP5/6, TP7/8, P1/2, P3/4, P5/6, P7/8, PO3/4, PO7/8, and O1/2), placed according to the extended International 10–20 system (Jasper, 1958; American Electroencephalographic Society, 1991), with Ag/AgCl-sintered electrodes and a ground positioned on AFz. Horizontal and vertical electroocular activity (EOG) was measured using a bipolar configuration lateral at the outer canthi of both eyes and from above and below the right eye. The right mastoid served as online reference. Electrode impedances were kept below 10 kΩ. Raw EEG signals were continuously registered with a sampling rate of 500 Hz (DC; 1000 Hz high-frequency cut-off) through two BrainAmp amplifiers (Brain Products GmbH, Munich, Germany).

Offline, data analysis was performed with Brain Vision Analyzer 2 software (Version 1.05; Brain Products GmbH, Munich, Germany). EEG data were initially re-referenced against the algebraic mean of the left and right mastoids and digitally filtered with a 0.1–25 Hz (48 dB/octave) bandpass filter. Eyeblink and horizontal ocular artifacts were subsequently corrected by an independent component analysis algorithm. Continuous EEG signals were segmented separately for positive, negative, and neutral T2 pictures into 2000 ms epochs (for the time period of 500 ms pre-stimulus to 1500 ms post-stimulus onset). All trials were semiautomatically screened for technical, muscle-related, or movement-related artifacts with amplitude deviations of ±80 μV and corrected relative to the 500 ms pre-stimulus baseline interval. In addition, trials were visually inspected and excluded if further artifacts were visible (e.g., extreme alpha activity). ERPs were obtained by averaging trials separately for each subject, electrode site and

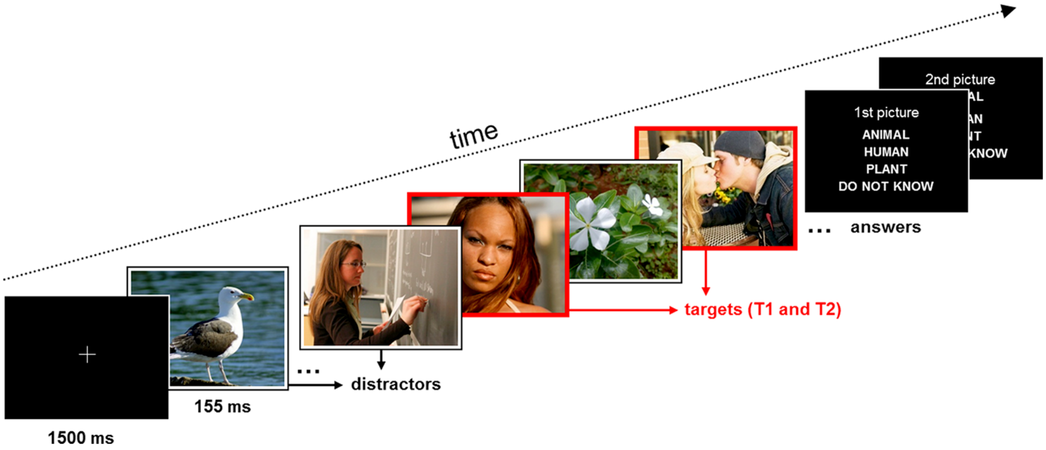


**FIGURE 1 | Schematic of the trial structure and time course of stimulus events in the experiment.** Participants were requested to identify the content of the two red-framed target pictures (T1 and T2; time interval between targets 310 ms) within a rapid series of white-framed distractor pictures.

T2 category (positive, neutral, negative) and for all trials where T2 was (a) correctly identified or (b) incorrectly identified or the answer choice "do not know," indicating that the target stimuli was not identified, was selected. ERPs to T2 presentation were only included in the analysis when the preceding T1 image had been correctly identified. Magnitudes of the P3 component were extracted from these averaged waveforms as mean activity in the pre-determined time interval of 300–800 ms after T2 onset from nine electrode sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4).

### STATISTICAL ANALYSES OF BEHAVIORAL DATA

Accuracy was analyzed with PASW (version 15.0, SPSS Inc., Chicago). A repeated-measures ANOVA with the factor emotion (negative, neutral, positive) was computed to elucidate the effects of emotion on picture recognition in the attentional blink. Only those trials in which T1 was correctly identified were included (see e.g., Trippe et al., 2007). Repeated pair-wise comparisons with Bonferroni correction were computed to test the differences between negative, neutral, and positive trials. All effects with a $p < 0.05$ were treated as statistically significant. Greenhouse-Geisser corrections were applied to significant $F$ ratios that did not meet Mauchly's sphericity assumption. Only interactions that yielded significant follow-up analyses are reported.

Further, we computed Pearson product-moment correlations between the IRI sum score and the attentional blink performance data as well as the P3 effect. Significant correlations with the composite IRI score were followed up by *post-hoc* correlations for the four IRI subscales. In order to quantify the P3 effect, we calculated difference scores by subtracting the ERP activity elicited by neutral images from the activity elicited by each emotion category (i.e., positive-neutral and negative-neutral) in the 300–800 ms time window and averaged across the analyzed nine electrode locations. The correlations were one-tailed because we had directional hypotheses (higher IRI scores are positively associated with the P3 increase for emotional T2). Correlations of the IRI subscales were treated as statistically significant when surviving a Bonferroni corrected $p$-value of $p < 0.0125$ ($p < 0.05/4$ IRI subscales).

## RESULTS

### BEHAVIORAL RESULTS

A total of 84.15% (SD = 6.26) of all T1 images were correctly identified, correct T2 identification rate (after correct T1 identification) was 51.88% (SD = 19.39) collapsed over all image categories (see **Figure 2**). Emotion significantly influenced T2 recognition rates [$F(2,50) = 14.251$, $p < 0.001$, partial $\eta^2 = 0.363$; positive$_{\text{mean(SD)}}$ = 64.24% (18.64), neutral$_{\text{mean(SD)}}$ = 55.13% (20.46), negative$_{\text{mean(SD)}}$ = 66.28% (23.44)], with higher recognition accuracy in negative and positive compared to neutral T2 images (both $p$'s = 0.001). T2 recognition rates for positive and negative pictures did not differ ($p > 0.15$).

### ERP RESULTS

Stimulus-locked ERP waveforms for correctly and incorrectly identified T2s following correctly identified T1 stimuli are displayed in **Figure 3**. Here, a repeated measures ANOVA with T2 identification accuracy (correct, incorrect) and two electrode site factors [lateralization (left, central, right); anteriority (anterior, central, posterior)] were calculated. All included participants had ≥10 trials for each condition. In accordance with previous research (e.g., Katayama and Polich, 1999), the scalp topography of the P3 showed a maximum voltage change over centro-parietal electrode sites [main effect of electrode site; anteriority $F(2,52) = 50{,}652$, $p < 0.001$, partial $\eta^2 = 0.661$; anterior$_{\text{mean}}$ = 0.556 μV, central$_{\text{mean}}$ = 3.750 μV, posterior$_{\text{mean}}$ = 4.521 μV; laterality $F(2,52) = 7{,}786$, $p < 0.001$, partial $\eta^2 = 0.230$; left$_{\text{mean}}$ = 4.418 μV, central$_{\text{mean}}$ = 4.434 μV, posterior$_{mean}$ = 3.680 μV]. The analysis also showed that correctly identified images elicited a larger P3 amplitude than incorrectly identified images [$F(1,26) = 15.560$, $p < .001$, partial $\eta^2 = 0.374$].

ERPs for trials with correctly identified positive, neutral, and negative T2s are depicted in **Figure 4**. To evaluate the effects of T2 emotion category on ERP magnitudes, a repeated-measures ANOVA with the within-subject factors emotion category (positive, negative, neutral) and the same electrode site factors as above was performed. Again, the topography of the P3 showed

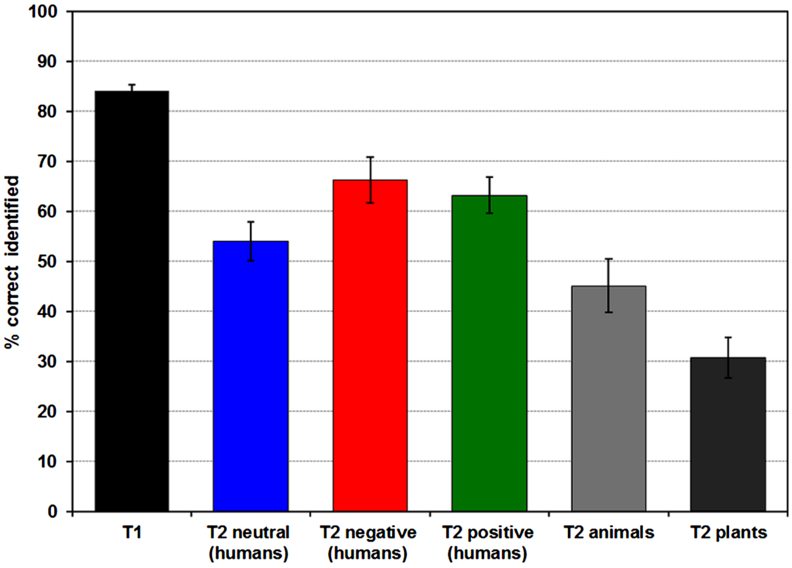


**FIGURE 2 | Diagram of mean percentage (and standard error) of correctly identified T1 and T2 pictures.** T2s were only included when the preceding T1 was detected correctly.

a maximum over centro-parietal electrode sites [main effect of electrode site; anteriority $F(2,52) = 18{,}138$, $p < 0.001$, partial $\eta^2 = 0.411$; $\text{anterior}_{\text{mean}} = 5.3626$ μV, $\text{central}_{\text{mean}} = 8.534$ μV, $\text{posterior}_{\text{mean}} = 7.755$ μV; laterality $F(2,52) = 14{,}387$, $p < 0.001$, partial $\eta^2 = 0.356$; $\text{left}_{\text{mean}} = 7.200$ μV, $\text{central}_{\text{mean}} = 7.947$ μV, $\text{posterior}_{mean} = 6.504$ μV]. Critically, the magnitude of the P3 amplitude differed as a function of emotion category [$F(2,52) = 16.783$, $p < 0.001$, partial $\eta^2 = 0.392$; Cz: $\text{positive}_{\text{mean(SD)}} = 10.55$ μV (6.05), $\text{neutral}_{\text{mean(SD)}} = 6.56$ μV (3.84), $\text{negative}_{\text{mean(SD)}} = 11.07$ μV (5.53)], with *post-hoc* pairwise comparisons indicating that the mean P3 amplitude to highly arousing positive and negative pictures differed from the P3 amplitudes to neutral pictures (negative vs. neutral: $p < 0.001$; positive vs. neutral: $p < 0.001$), but not from each other (negative vs. positive: $p = 1.00$). There were no interactions with the electrode site factors (all $p > 0.10$). There were also no effects of emotion category in the incorrectly identified T2 images ($p > 0.05$).

### CORRELATIONS OF EMPATHY QUESTIONNAIRE DATA WITH ATTENTIONAL BLINK PERFORMANCE AND ERP ACTIVITY

**Table 2** summarizes the empathy questionnaire scoring results of the present sample. For the behavioral attentional blink data we examined correlations between empathy and the emotional attentional blink effect by calculating difference scores of accuracy between the neutral and emotional picture categories (i.e., negative-neutral, positive-neutral). Neither positive nor negative images correlated significantly with the IRI total score (negative-neutral: $r = -0.07$, $p = 0.373$, positive-neutral: $r = -0.23$, $p = 0.147$). Hence, no further correlations were calculated. To correlate empathy to the P3 effects, we averaged across the analyzed electrodes and formed difference scores of negative-neutral and positive-neutral conditions. For the negative-neutral condition there was a significant correlation with the IRI total score ($r = 0.49$, $p = .009$). Similarly, there was a correlation for the positive-neutral condition ($r = 0.55$, $p = .003$), see **Figure 5**. Due to significant correlations with the IRI total score, we further examined the relation of P3 effects to the IRI subscales we found significant correlations for "fantasy" scores (with negative-neutral: $r = 0.52$, $p = 0.005$ and with positive-neutral: $r = 0.61$, $p = 0.001$) and for "perspective taking" scores (with negative-neutral: $r = 0.54$, $p = 0.004$ and with positive-neutral: $r = 0.77$, $p = 0.000009$). Analysis of "empathic concern" and "personal distress" scores did not reveal a significant correlation with P3 effect size (all $p$'s $> 0.25$).

## DISCUSSION

The present study showed a correlation of the P3 increase to emotional stimuli presented in the attentional blink with inter-individual differences in empathy as measured through self-reports. This suggests that individuals scoring high in interpersonal social traits show enhanced processing of emotional stimuli under stimulus conditions that are suboptimal for conscious target detection and high working memory load. The study also corroborates our knowledge about the attentional blink phenomenon by showing larger P3 amplitudes to correctly identified T2 stimuli and through demonstrating that emotional picture stimuli are detected more often, even if they are presented in the attentional blink period. Because of the very limited amount of attentional blink studies with pictorial stimuli and socio-emotional scenes in particular, the study also extends the phenomenon to more ecologically valid stimuli.

The influence of emotion on the attentional blink was apparent in the increased detection rates for emotional pictures and in the enlarged P3 amplitude to emotional T2 images (cf. Anderson and Phelps, 2001; Keil and Ihssen, 2004; Trippe et al., 2007). However, a correlation of the emotion effect with empathy

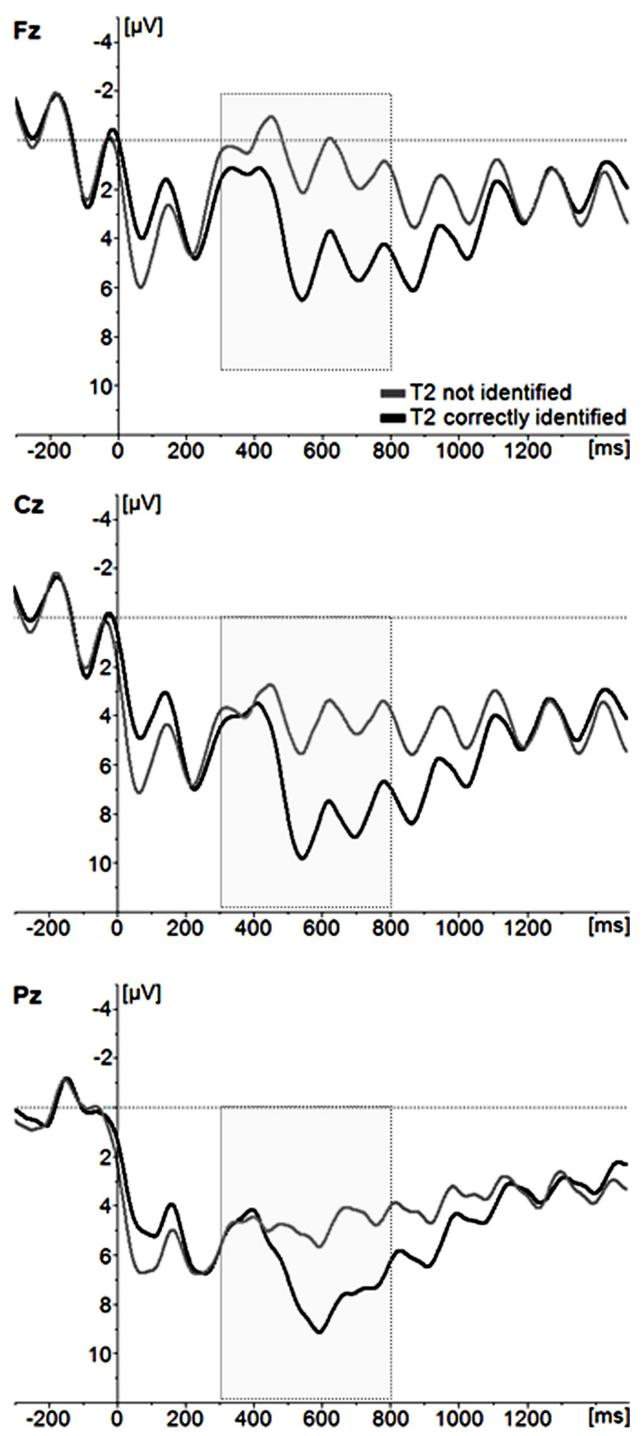


**FIGURE 3 | Grand average ERP waveforms for incorrectly identified and correctly identified T2 images.** For illustrative purposes, only the midline electrode sites Fz, Cz, and Pz are displayed.

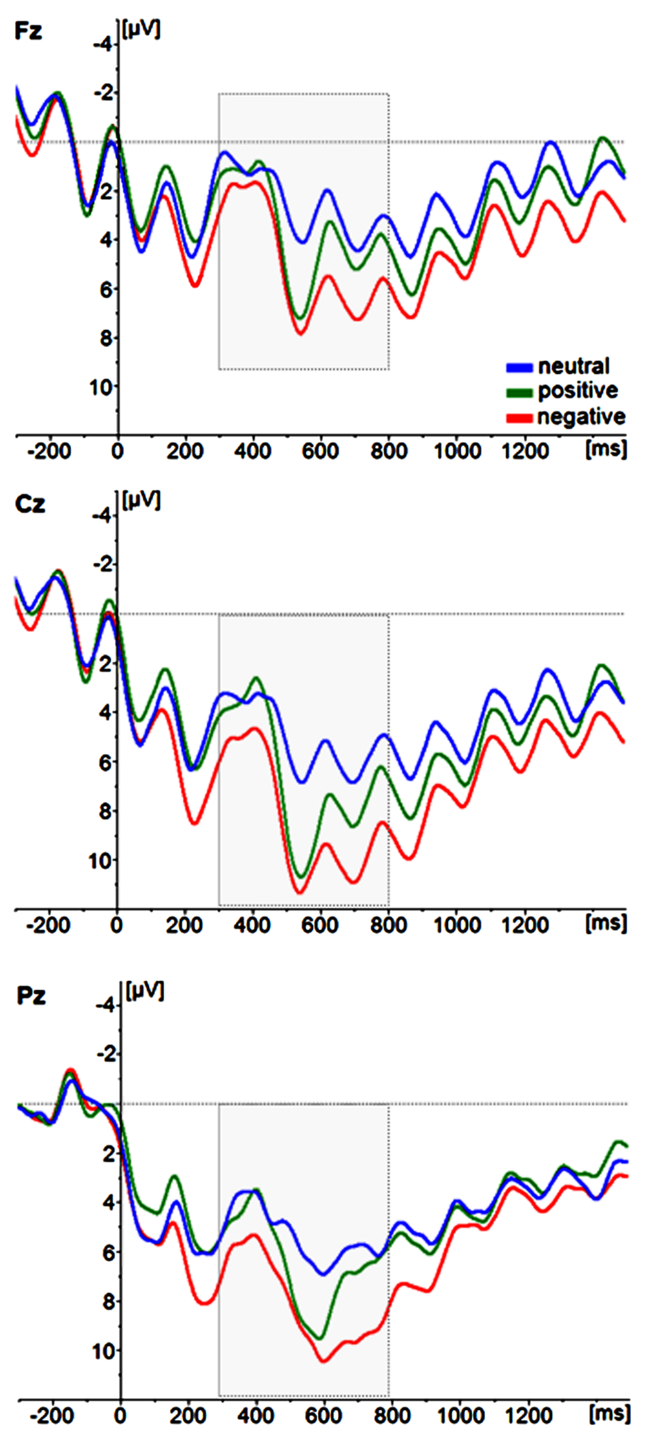


**FIGURE 4 | Grand average ERP waveforms of correctly identified T2 image categories (neutral, positive, negative).** For illustrative purposes, only the midline electrode sites Fz, Cz, and Pz are displayed.

scores (Davis, 1983b) was only observed in the P3 amplitude, not for the behavioral performance. Thus, individuals with high scores on empathy did not differ from participants scoring low on empathy in the percentage of correctly identified emotional targets relative to neural targets. P3 responses to those emotional images that were correctly identified, however, were enlarged. This pattern could suggest a dissociation of behavioral and ERP responses in the attentional blink, which has already been reported by Trippe et al.

**Table 2 | Scores on the empathy questionnaire.**

| Scale | M | SD | min | max |
|---|---|---|---|---|
| Perspective taking | 18.22 | 4.88 | 6 | 28 |
| Empathic concern | 18.83 | 4.07 | 11 | 25 |
| Fantasy | 15.56 | 5.57 | 2 | 24 |
| Personal distress | 12.17 | 5.17 | 4 | 25 |

(2007) who noted that they seem to be influenced through at least partially different neural processes. While the P3 is clearly linked to working memory consolidation, error rates represent a composite of multiple processing stages also including sensory processing and response preparation. This may yield the P3 potential a more sensitive measure and the present data suggest that particularly working memory consolidation is modulated by trait empathy (Vogel et al., 1998; Vogel and Luck, 2002).

The fact that the correlation pattern was observed for both positive and negative emotional images corroborates the findings for two different categories of stimuli. A similar influence of positive and negative emotional stimuli on the attentional blink (Keil and Ihssen, 2004; Ogawa and Suzuki, 2004; Trippe et al., 2007) and other attentional processes (Brosch et al., 2008; Kanske and Kotz, 2011a,b; Kanske, 2012) is in line with previous work and shows that putative reward and threat signals can possess similar salience. The present correlations indicate that the sensitivity of people scoring high in empathy applies to both positive and negative emotion in facial expressions and scenes. While it may be argued that such a relation is principally self-evident, the association of ERP with self-report data on how individuals habitually perceive themselves in reaction to others is noteworthy. Furthermore, the present data can specifically link this association to a particular processing stage, namely consolidation in working memory associated with the P3 component (Vogel and Luck, 2002; Martens and Wyble, 2010).

An interesting aspect of the present data is the specific correlation pattern between the IRI subscales and the emotional modulation of the attentional blink. In addition to the total score, the subscales fantasy and perspective taking were also individually correlated with the effect, while personal distress and empathic concern showed no significant correlation. Even though we had no a priori hypotheses regarding the subscale correlation, one might have expected to find a relation between all subscales, in particular personal distress and empathic concern, as they have been interpreted as representing the more affective aspect of the empathy construct (Davis, 1983a,b). However, it may be that the correlational pattern with the more cognitive, but not affective scales reflects the fact that the P3 indexes a rather late stage of stimulus processing, i.e., working memory consolidation. Possibly, earlier stages related to sensory processing and first classification of stimuli would show a different pattern, namely a correlation with the affective subscales of the IRI (Yamada and Decety, 2009; Decety et al., 2010). The present paradigm was not optimized to look at ERP components indexing these stages, as would for example an emotional dot probe paradigm (Pourtois et al., 2005). We did not observe such effects, but future research could address this point with more specialized designs. Nevertheless, it is also important to note that previous studies also reported correlation of the more cognitive subscale perspective taking to emotional reactivity as measured in amygdala activity (for example, Pfeifer et al., 2008; Silani et al., 2008). A more promising path to differentiating cognitive and affective social understanding may lie in assessing them experimentally, rather than with self-reports (Dziobek et al., 2006, 2008; Apperly et al., 2011). Performance on respective tasks could yield objective performance levels that are not subject to the biases inherent in self-report data.

There are some limitations to the present study. As already outlined, a more comprehensive insight into the relations of the presently discussed concepts could be gained by assessing emotional reactivity in more than one paradigm to allow better description of multiple stimulus processing stages (for example, Kanske et al., 2011). Additionally, applying multiple measures for empathy, in addition to self-report data, would potentially better separate cognitive and affective aspects (Shamay-Tsoory et al., 2009; Harari et al., 2010). A critical point regarding the presently used paradigm is that emotional stimuli are presented less frequently than neutral stimuli, in particular when considering the amount of distractor images. Thus, the P3 modulation and the better detection rates may just reflect a saliency through infrequency effect. While this is an inherent problem of all studies looking at an emotional modulation of the attentional blink, several studies reported differential effects for different discrete emotions, such as fear, anger, or happiness (Fox et al., 2005; Maratos et al., 2008; de Jong et al., 2010; Maratos, 2011). As these discrete emotion categories were presented equally often, frequency cannot explain the different effects, supporting the interpretation that it is the emotionality of the stimuli that gives rise to the differences. Furthermore, even though the present sample was relatively large, it could be argued that a larger spread in trait empathy could have helped in detecting more associations. In light of this, the lack of a correlation with the behavioral performance must, as with all null findings, be interpreted with care. Future studies could preselect participants to ensure greater variation in empathy scores, select extreme groups, or even test patient populations known to show impaired interpersonal reactivity, for example, individuals with autism spectrum disorder (Clark et al., 2008; Winkielman et al., 2009). Finally, the present study does not disentangle whether heightened empathy-related emotional responding is primarily related to the social content, the valence, or arousal aspect of the images. There has been some initial evidence showing that high empathic individuals, compared to low empathic individuals, are indeed more sensitive to cues of social versus monetary reward (Gossen et al., 2013) and that empathy-related responding is influenced by the phyologenetic similiarity of the stimuli to humans (Westbury and Neumann, 2008). A more systematic variation and experimentally testing of these variables (i.e., valence, arousal/salience aspect, and semantic content of stimuli) within one study design might be a worthwhile approach for future studies.

To conclude, the present results very convincingly affirm our knowledge about the influence of emotion on the attentional

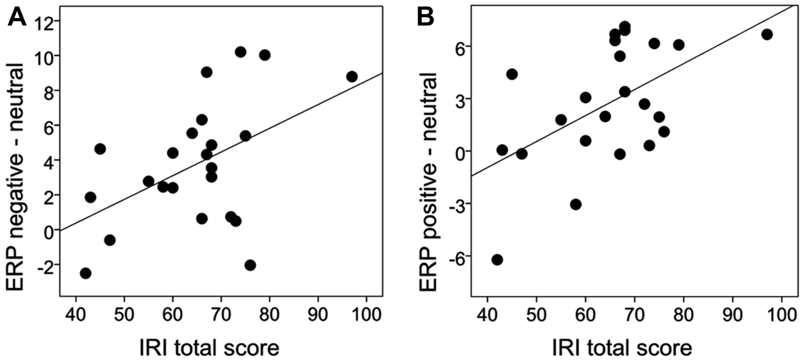


**FIGURE 5 | Scatterplot of the correlation between the P3 effect and IRI total score.** The P3 effect was quantified as the difference in the ERP waveform activity between emotional and neutral T2 pictures (i.e., negative-neutral, positive-neutral) in the time window of 300–800 ms.

blink with pictorial stimuli. Emotional stimuli are salient enough to reduce the attentional blink effect behaviorally and also elicit enlarged P3 amplitudes, indicative of enhanced stimulus consolidation. This effect is related to individual differences in self-reported empathy, suggesting that individuals scoring high in interpersonal social traits show enhanced processing of emotional stimuli. The correlation with empathy scores suggests that individuals with high empathy may also be more sensitive to emotions expressed by others in everyday situations. As the attentional blink paradigm tests for stimulus processing in stimulus conditions that are suboptimal for conscious target detection with very brief presentation time and embedding in streams of other stimuli, it may be that in difficult and potentially stressful situations high empathy enables better emotion detection and adequate responding.

## ACKNOWLEDGMENTS

The present work was funded by a grant from the Deutsche Forschungsgemeinschaft (We3638/3-1).

## REFERENCES

American Electroencephalographic Society. (1991). Guidelines for standard electrode position nomenclature. *J. Clin. Neurophysiol.* 8, 200–202. doi: 10.1097/00004691-199104000-00007

Amir, N., Taylor, C. T., Bomyea, J. A., and Badour, C. L. (2009). Temporal allocation of attention toward threat in individuals with posttraumatic stress symptoms. *J. Anxiety Disord.* 23, 1080–1085. doi: 10.1016/j.janxdis.2009.07.010

Anderson, A. K., and Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411, 305–309. doi: 10.1038/35077083

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., and Todd, S. (2011). Developmental continuity in theory of mind: speed and accuracy of belief-desire reasoning in children and adults. *Child Dev.* 82, 1691–1703. doi: 10.1111/j.1467-8624.2011.01635.x

Bradley, M. M., Codispoti, M., Sabatinelli, D., and Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion* 1, 300–319. doi: 10.1037/1528-3542.1.3.300

Brosch, T., Sander, D., Pourtois, G., and Scherer, K. R. (2008). Beyond fear: rapid spatial orienting toward positive emotional stimuli. *Psychol. Sci.* 19, 362–370. doi: 10.1111/j.1467-9280.2008.02094.x

Clark, T. F., Winkielman, P., and Mcintosh, D. N. (2008). Autism and the extraction of emotion from briefly presented facial expressions: stumbling at the first step of empathy. *Emotion* 8, 803–809. doi: 10.1037/a0014124

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog Sel. Doc. Psychol.* 10, 85.

Davis, M. H. (1983a). The effects of dispositional empathy on emotional-reactions and helping – a multidimensional approach. *J. Pers.* 51, 167–184. doi: 10.1111/j.1467-6494.1983.tb00860.x

Davis, M. H. (1983b). Measuring individual-differences in empathy – evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* 44, 113–126. doi: 10.1037/0022-3514.44.1.113

Dawson, G., Bernier, R., and Ring, R. H. (2012). Social attention: a possible early indicator of efficacy in autism clinical trials. *J. Neurodev. Disord.* 4, 11. doi: 10.1186/1866-1955-4-11

Decety, J., Yang, C. Y., and Cheng, Y. (2010). Physicians down-regulate their pain empathy response: an event-related brain potential study. *Neuroimage* 50, 1676–1682. doi: 10.1016/j.neuroimage.2010.01.025

de Jong, P. J., Koster, E. H., Van Wees, R., and Martens, S. (2010). Angry facial expressions hamper subsequent target identification. *Emotion* 10, 727–732. doi: 10.1037/a0019353

Delmonte, S., Balsters, J. H., Mcgrath, J., Fitzgerald, J., Brennan, S., Fagan, A. J., et al. (2012). Social and monetary reward processing in autism spectrum disorders. *Mol. Autism* 3, 7. doi: 10.1186/2040-2392-3-7

Demurie, E., Roeyers, H., Baeyens, D., and Sonuga-Barke, E. (2011). Common alterations in sensitivity to type but not amount of reward in ADHD and autism spectrum disorders. *J. Child Psychol. Psychiatry* 52, 1164–1173. doi: 10.1111/j.1469-7610.2010.02374.x

Donchin, E. (1981). Surprise!...Surprise? *Psychophysiology* 18, 493–513. doi: 10.1111/j.1469-8986.1981.tb01815.x

Donchin, E., and Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating. *Behav. Brain Sci.* 11, 357–374. doi: 10.1017/S0140525X00058027

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: a movie for the assessment of social cognition. *J. Autism Dev. Disord.* 36, 623–636. doi: 10.1007/s10803-006-0107-0

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., et al. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *J. Autism Dev. Disord.* 38, 464–473. doi: 10.1007/s10803-007-0486-x

Fox, E., Russo, R., and Georgiou, G. A. (2005). Anxiety modulates the degree of attentive resources required to process emotional faces. *Cogn. Affect. Behav. Neurosci.* 5, 396–404. doi: 10.3758/CABN.5.4.396

Gossen, A., Groppe, S. E., Winkler, L., Kohls, G., Herrington, J., Schultz, R. T., et al. (2013). Neural evidence for an association between social proficiency and sensitivity to social reward. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nst033 [Epub ahead of print].

Harari, H., Shamay-Tsoory, S. G., Ravid, M., and Levkovitz, Y. (2010). Double dissociation between cognitive and affective empathy in borderline personality disorder. *Psychiatry Res.* 175, 277–279. doi: 10.1016/j.psychres.2009.03.002

Jasper, H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.

Kanske, P. (2012). On the influence of emotion on conflict processing. *Front. Integr. Neurosci.* 6:42. doi: 10.3389/fnint.2012.00042

Kanske, P., and Kotz, S. A. (2011a). Conflict processing is modulated by positive emotion: ERP data from a flanker task. *Behav. Brain Res.* 219, 382–386. doi: 10.1016/j.bbr.2011.01.043

Kanske, P., and Kotz, S. A. (2011b). Positive emotion speeds up conflict processing: ERP responses in an auditory Simon task. *Biol. Psychol.* 87, 122–127. doi: 10.1016/j.biopsycho.2011.02.018

Kanske, P., Plitschka, J., and Kotz, S. A. (2011). Attentional orienting towards emotion: P2 and N400 ERP effects. *Neuropsychologia* 49, 3121–3129. doi: 10.1016/j.neuropsychologia.2011.07.022

Katayama, J., and Polich, J. (1999). Auditory and visual P300 topography from a 3 stimulus paradigm. *Clin. Neurophysiol.* 110, 463–468. doi: 10.1016/S1388-2457(98)00035-2

Keil, A., and Ihssen, N. (2004). Identification facilitation for emotionally arousing verbs during the attentional blink. *Emotion* 4, 23–35. doi: 10.1037/1528-3542.4.1.23

Kohls, G., Peltzer, J., Herpertz-Dahlmann, B., and Konrad, K. (2009). Differential effects of social and non-social reward on response inhibition in children and adolescents. *Dev. Sci.* 12, 614–625. doi: 10.1111/j.1467-7687.2009.00816.x

Koster, E. H., De Raedt, R., Verschuere, B., Tibboel, H., and De Jong, P. J. (2009). Negative information enhances the attentional blink in dysphoria. *Depress. Anxiety* 26, E16–E22. doi: 10.1002/da.20420

Kranczioch, C., Debener, S., and Engel, A. K. (2003). Event-related potential correlates of the attentional blink phenomenon. *Brain Res. Cogn. Brain Res.* 17, 177–187. doi: 10.1016/S0926-6410(03)00092-2

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). "*International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual.* Technical report A-6". Gainesville: Center for Research in Psychophysiology, University of Florida.

Maratos, F. A. (2011). Temporal processing of emotional stimuli: the capture and release of attention by angry faces. *Emotion* 11, 1242–1247. doi: 10.1037/a0024279

Maratos, F. A., Mogg, K., and Bradley, B. P. (2008). Identification of angry faces in the attentional blink. *Cogn. Emot.* 22, 1340–1352. doi: 10.1080/02699930701774218

Martens, S., and Wyble, B. (2010). The attentional blink: past, present, and future of a blind spot in perceptual awareness. *Neurosci. Biobehav. Rev.* 34, 947–957. doi: 10.1016/j.neubiorev.2009.12.005

Ogawa, T., and Suzuki, N. (2004). On the saliency of negative stimuli: evidence from attentional blink. *Jpn. Psychol. Res.* 46, 20–30. doi: 10.1111/j.1468-5884.2004.00233.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Pfeifer, J. H., Iacoboni, M., Mazziotta, J. C., and Dapretto, M. (2008). Mirroring others' emotions relates to empathy and interpersonal competence in children. *Neuroimage* 39, 2076–2085. doi: 10.1016/j.neuroimage.2007.10.032

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Pourtois, G., Thut, G., De Peralta, R. G., Michel, C., and Vuilleumier, P. (2005). Two electrophysiological stages of spatial orienting towards fearful faces: early temporo-parietal activation preceding gain control in extrastriate visual cortex. *Neuroimage* 26, 149–163. doi: 10.1016/j.neuroimage.2005.01.015

Raymond, J. E., Shapiro, K. L., and Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849–860. doi: 10.1037/0096-1523.18.3.849

Rolke, B., Heil, M., Streb, J., and Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology* 38, 165–174. doi: 10.1111/1469-8986.3820165

Shamay-Tsoory, S. G., Aharon-Peretz, J., and Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. doi: 10.1093/brain/awn279

Shapiro, K. L., Caldwell, J., and Sorensen, R. E. (1997). Personal names and the attentional blink: a visual "cocktail party" effect. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 504–514. doi: 10.1037/0096-1523.23.2.504

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., and Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Soc. Neurosci.* 3, 97–112. doi: 10.1080/17470910701577020

Trippe, R. H., Hewig, J., Heydel, C., Hecht, H., and Miltner, W. H. (2007). Attentional blink to emotional and threatening pictures in spider phobics: electrophysiology and behavior. *Brain Res.* 1148, 149–160. doi: 10.1016/j.brainres.2007.02.035

Van Dam, N. T., Earleywine, M., and Altarriba, J. (2012). Anxiety attenuates awareness of emotional faces during rapid serial visual presentation. *Emotion* 12, 796–806. doi: 10.1037/a0024648

Vogel, E. K., and Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychon. Bull. Rev.* 9, 739–743. doi: 10.3758/BF03196329

Vogel, E. K., Luck, S. J., and Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1656–1674. doi: 10.1037/0096-1523.24.6.1656

Wessa, M., Kanske, P., Neumeister, P., Bode, K., Heissler, J., and Schönfelder, S. (2010). EmoPicS: subjective and psychophysiological evaluation of new imagery for clinical biopsychological research. *Z. Klin. Psychol. Psychother. Suppl.* 1/11, 77 (available from Michele Wessa, e-mail: wessa@uni-mainz.de).

Westbury, H. R., and Neumann, D. L. (2008). Empathy-related responses to moving film stimuli depicting human and non-human animal targets in negative circumstances. *Biol. Psychol.* 78, 66–74. doi: 10.1016/j.biopsycho.2007.12.009

Winkielman, P., Mcintosh, D. N., and Oberman, L. (2009). Embodied and disembodied emotion processing: learning from and about typical and autistic individuals. *Emot. Rev.* 1, 178–190. doi: 10.1177/1754073908100442

Wittchen, H.-U., Zaudig, M., and Fydrich, T. (1997). *SKID-I. Strukturiertes Klinisches Interview für DSM-IV* [Structural clinical interview for DSM-IV axis I disorders]. Göttingen, Germany: Hogrefe.

Yamada, M., and Decety, J. (2009). Unconscious affective processing and empathy: an investigation of subliminal priming on the detection of painful facial expressions. *Pain* 143, 71–75. doi: 10.1016/j.pain.2009.01.028

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 24 June 2013; accepted: 16 September 2013; published online: 11 October 2013.*

*Citation: Kanske P, Schönfelder S and Wessa M (2013) Emotional modulation of the attentional blink and the relation to interpersonal reactivity. Front. Hum. Neurosci. 7:641. doi: 10.3389/fnhum.2013.00641*

*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 12 May 2014
doi: 10.3389/fnhum.2014.00257

# Monetary rewards modulate inhibitory control

**Paula M. Herrera**[1,2,3]*, **Mario Speranza**[1,4], **Adam Hampshire**[5,6] **and Tristán A. Bekinschtein**[6]

[1] Laboratoire ECIPSY - EA 4047, Université de Versailles Saint Quentin en Yvelines, Versailles, France
[2] Grupo de Investigación en Neurociencias (NeURos), Facultad de Ciencias de la Salud, Universidad del Rosario, Bogotá, Colombia
[3] Grupo de Investiga, Laboratorio de Psicología Experimental, Facultad de Psicología, Universidad El Bosque, Bogotá, Colombia
[4] Child and Adolescent Psychiatry Department, Centre Hospitalier de Versailles, Versailles, France
[5] Division of Brain Sciences, Department of Medicine, Imperial College London, London, UK
[6] Medical Research Council-Cognition and Brain Sciences Unit, Cambridge, UK

**Edited by:**
Agustin Ibanez, Institute of Cognitive Neurology, Argentina

**Reviewed by:**
Agustin Ibanez, Institute of Cognitive Neurology, Argentina
Sina Radke, University Hospital RWTH Aachen, Germany
Angelina Pilatti, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

***Correspondence:**
Paula M. Herrera, EA4047, Université de Versailles Saint Quentin en Yvelines, Service de Psychiatrie Infanto Juvenile, Hopital Mignot, 177 rue de Versailles, 78150 Le Chesnay, France
e-mail: polatija@gmail.com

The ability to override a dominant response, often referred to as behavioral inhibition, is considered a key element of executive cognition. Poor behavioral inhibition is a defining characteristic of several neurological and psychiatric populations. Recently, there has been increasing interest in the motivational dimension of behavioral inhibition, with some experiments incorporating emotional contingencies in classical inhibitory paradigms such as the Go/NoGo and Stop Signal Tasks (SSTs). Several studies have reported a positive modulatory effect of reward on performance in pathological conditions such as substance abuse, pathological gambling, and Attention Deficit Hyperactive Disorder (ADHD). However, experiments that directly investigate the modulatory effects of reward magnitudes on the performance of inhibitory tasks are scarce and little is known about the finer grained relationship between motivation and inhibitory control. Here we probed the effect of reward magnitude and context on behavioral inhibition with three modified versions of the widely used SST. The pilot study compared inhibition performance during six blocks alternating neutral feedback, low, medium, and high monetary rewards. Study One compared increasing vs. decreasing rewards, with low, high rewards, and neutral feedback; whilst Study Two compared low and high reward magnitudes alone also in an increasing and decreasing reward design. The reward magnitude effect was not demonstrated in the pilot study, probably due to a learning effect induced by practice in this lengthy task. The reward effect *per se* was weak but the context (order of reward) was clearly suggested in Study One, and was particularly strongly confirmed in study two. In addition, these findings revealed a "kick start effect" over global performance measures. Specifically, there was a long lasting improvement in performance throughout the task when participants received the highest reward magnitudes at the beginning of the protocol. These results demonstrate a dynamical behavioral inhibition capacity in humans, as illustrated by the reward magnitude modulation and initial reward history effects.

**Keywords: reward, inhibition (psychology), cognitive control, stop signal task, behavioral analysis**

## INTRODUCTION

Behavioral inhibition is an essential component of goal-oriented behavior, allowing the suppression of a pre-potent behavior in order to switch to a more suitable action when conditions change. Cancelling a planned action is also called "executive inhibition" (Nigg, 2001), as part of the inhibitory control network.

Measuring inhibition under experimental conditions has evolved progressively on the basis of core concepts as the negative priming (Tipper, 2001), interference control (Salo et al., 2001), mental withholding (Brass and Haggard, 2007), and allocation of attention (Hasher et al., 1999) among others. A non-exhaustive list of inhibition classical tests includes the Flankers (Wendt et al., 2014), the Go/NoGo (Bokura et al., 2001), the Continuous Performance Task (Ridderinkhof et al., 2004), and the Stop Signal Task (SST; Aron et al., 2003). Each of these tasks highlights a particular aspect of the inhibitory processes (for a review in the inhibition tasks, see MacLeod, 2007).

The SST presents a frequent Go stimulus (left or right), and a less frequent Stop stimulus between Go trials. It is widely considered that the main executive process of this task is the cancellation of the on-going action being triggered by a Go stimuli (Logan, 1994).

The SST is one of the most widespread measures of inhibitory control in the cognitive sciences (Li et al., 2006; Alderson et al., 2008; Chikazoe et al., 2009). It has long been used to investigate cognition in healthy individuals (Ramautar et al., 2004; Clark et al., 2005; Lansbergen et al., 2007; van Gaal et al., 2009), and is used as a diagnostic tool in several pathological conditions including Attention-Deficit/Hyperactivity Disorder (ADHD) (Stevens et al., 2002; Nichols and Waschbusch, 2004; Sonuga-Barke, 2005), Conduct Disorder (CD) (Oosterlaan et al., 1998), Oppositional Defiant Disorder (ODD) (Albrecht et al., 2005), substance abuse (Smith and Mattick, 2013), and personality disorders (Lipszyc and Schachar, 2010).

The SST is designed to enable the measurement of the inhibition process through its gold standard measure: the Stop Signal Reaction Time (SSRT). This score measures the time required for an individual to successfully stop their initiated action. Specifically successful behavior inhibition leaves no report (it is the lack of a response), and hence the inhibition measure has to be estimated by other behavioral markers closely related and dependent on the inhibition process. The SSRT is therefore calculated as the probability of inhibition (PI) (Liotti et al., 2005; Schmajuk et al., 2006) or the subtraction of the Mean reaction time (MRT) minus the Stop signal delay (SSD) (formulae SSRT = MRT − SSD) (Kok et al., 2004). The SSRT score is given in milliseconds, reflecting the time from the presentation of Go signal at which one starts to fail. In other words: how late can you receive the order to stop the ongoing action.

Both the MRT and the SSD are direct measures allowing an indirect calculation of the behavioral inhibition performance. Beyond their use on the SSRT calculation, these measures can also give crucial hints about the behavioral adjustments through the inhibition task (van Boxtel et al., 2001; Band et al., 2003). The relevance of these two measures will be discussed later in detail.

Besides these time related measures, the SST provides other useful information about the inhibition profile as the accuracy and number of performance errors (failed Go's, failed Stop's, left-right precision errors).

The SSRT—provides a sensitive behavioral marker that can be used to compare control groups with impulsivity disorders (Lijffijt et al., 2004). It has been shown that the ability to inhibit a response is present from early ages and that the SSRT improves through development (Carver and Scheier, 2001) until becoming a stable and individual measure in adult healthy participants (Cohen and Poldrack, 2008). SSRT values around 200 ms have been described as the normal range for adults (Logan and Cowan, 1984). SSRT beyond 400 ms have been reported in young children, elderly, and impulsive participants, as well as hyperactive children (Winstanley et al., 2006). SSRT can vary according to frontal lesions (Aron et al., 2003), and can be consistently altered in disorders such as pathological gambling (Lawrence et al., 2009), psychopathic personality (Masui and Nomura, 2011), or Attention Deficit/Hyperactive Disorder (ADHD) (Oosterlaan and Sergeant, 1998; Stevens et al., 2002; Banaschewski et al., 2003).

Previous works have stated that behavioral inhibition performance reaches mature development after childhood and it has been suggested that a similar behavioral trait should be evident across a range of contexts (Williams et al., 1999; Rubia et al., 2007). Nevertheless, some experimental results suggest that inhibition can change in response to emotional states due to feedback contingencies (Bechara et al., 1994), fear (Bush et al., 2000), or other motivational influences (Pessoa et al., 2012).

Motivation refers to the volitional engagement in a task and can come from either an internal and/or an external source (Panksepp, 2003). Character and temperament theories propose a distinction among individuals with an accentuated need for external rewards, while others would exhibit a more internal motivation driven behavior (Derryberry and Rothbart, 1997). Even if it is assumed that everyone has a distinctive behavioral pattern, specific situations can lead to unusual reactions, leading to popular phrases such as "money talks" (Living Colour, 1991) or "everybody has a price" (Jessie, 2011). The individual need of external incentives is part of the basic-stimulus response mechanism known as "reward dependency" (Cloninger, 1987), and can be overexpressed in pathological conditions such as gambling and compulsive buying (Avila and Parcet, 2001). Moreover, both clinical and experimental evidence support the view that ADHD children are particularly affected by immediate and salient rewards when engaging with a task (Michel et al., 2005; Sonuga-Barke, 2005; Potts et al., 2006; Groom et al., 2010; Luman et al., 2010).

The influence of motivation over behavior inhibition may be achieved through different strategies. One of the most prevailing is the "aversive/approach" system, allusive to the "Behavioral Inhibition System/Behavioral Activation System" (BIS/BAS) model (Gray, 1987; Quay, 1993). For instance, increased response times are observed when comparing neutral against punishment contingencies. This strategy can be convenient to heighten the likelihood of successful inhibition when avoiding punishment. In the case of reward contingencies, behavioral inhibition can be adjusted to increase the sum of fruitful inhibitions, thereby increasing the number of rewards (Boehler et al., 2012).

A range of prefrontal, sub cortical, and limbic structures have been implicated in behavioral inhibition in different task contexts. For example, premotor areas (Peterson et al., 1999), basal ganglia (Brown et al., 1999), and Anterior Cingulate Cortex (ACC) (Braver et al., 2001) are strong candidates as core anatomical structures enabling inhibitory motor control during the SST and its various analogs. By contrast, activation of the orbitofrontal cortex (OFC), caudate nucleus (Elliott et al., 2000), and limbic system structures (Etkin et al., 2006) has been reported when behavioral inhibition involves "hot" or emotional choices in response to punishment or reward. It has been suggested that the dorsolateral prefrontal (DLPF) cortex is involved in planned or "proactive" inhibition (Bechara et al., 1994; Dias et al., 1997). Executive and motivational inhibitory circuits are interconnected and share some anatomical pathways, but they also rely on independent structures (Nigg, 2000).

Despite being the focus of much research, behavioral inhibition and the mechanisms by which it is modulated remain poorly understood. Indeed, many authors have highlighted the inherent difficulty in taking a specific measure of inhibition due to other simultaneous processes that are tapped by classical inhibitory paradigms including perception, attention, and response planning (Rubia et al., 2003; Chen et al., 2013). Positive or negative emotions may interfere with the inhibition processes (Kalanthroff et al., 2013), as well as other elaborate cognitive processes related to education, culture, and environmental factors (Immordino-Yang and Damasio, 2007). Emotions may influence Inhibition by high order cognitive process such as reasoning, labeling, and voluntary modulation (van Reekum and Schaefer, 2011).

Even though the recent literature recognizes the role played by motivational aspects over inhibitory processes, few studies have explored the links between emotion and the inhibitory components of executive control and, more specifically, the effect of reward magnitude and context on inhibition capacity (Kalanthroff et al., 2013). The majority of those reports state simple contrasts, where inhibition is analyzed under rewarded vs. no-rewarded contingencies. Others have compared punishment against reward (Rubia et al., 2005), contrasting emotions such as erotic or painful stimuli (Yu et al., 2012), fear (Verbruggen and De Houwer, 2007; Sagaspe et al., 2011), or goal conflict (Neo et al., 2011).

The recent theoretical and experimental literature proposes the existence of two independent pathways for behavioral inhibition. The "cool" pathway, corresponding to deliberate executive control, and the "hot" pathway, related to affective or motivational modulations (Nigg, 2001; Zelazo et al., 2003). It has been suggested that most of the inhibitory tasks involve both executive and motivational pathways albeit to widely varying degrees (Geurts et al., 2006).

The overarching goal of this study was to obtain a clearer understanding of how reward and reward context modulate behavioral inhibition performance during the SST. The longer-term goal is to reproduce these studies with neuroimaging and EEG in order to explore the neural underpinnings of these effects. More specifically, we explored the motivational modulation of behavioral inhibition in normal adults using a SST with reward. A close temporal manipulation of reward size and contingencies was used to obtain a better understanding of the motivational dynamical adjustments of behavioral inhibition capacities. We highlight two specific aims: to clarify how important the magnitude of reward is (no reward, low or high reward) and what is the nature of that relationship. Is the value of the reward itself, strong enough to induce a similar level of behavioral inhibition performance, that is, a trait independent of context? The second aim was to determine whether the history of presentation of different reward magnitudes modulated behavior. What happens when opposing contrasted reward magnitudes at different times? Is there any difference when receiving a given reward at the beginning or at the end of the task?

On the basis of previous studies using reward contingencies in inhibition tasks in adults (Boksem et al., 2008; De Pascalis et al., 2010; Pessoa et al., 2012; Yu et al., 2012) one would hypothesize that the presence of a reward should improve inhibition performances compared to a neutral feedback. Hence, we expect to find that higher rewards would have a higher impact on inhibition independently from the order of presentation. Moreover, we predicted a reciprocal modulation effect relying on the history of presentation of rewards: a straight improvement in performances when presenting progressively increasing rewards, and a disengaging effect when moving from high to low reward.

## METHODS

### PARTICIPANTS

Young adult participants were recruited by informal community announcements among the staff and medicine school students attending at the Versailles General Hospital, the undergraduate students of the University of Nantes, and a mailing list of volunteers from the MRC-CBU in Cambridge.

One hundred and one participants were recruited (21 for pilot study, 41 for study one, and 39 for study two). The combined mean age for both men and women participating from the study was 24.7 (age range 20–33, $SD = 4.5$; sex ratio = 1.1), and had at least 2 years of Higher Education. They were screened for past and current psychiatric disorders ADHD, depression or bipolar disorders and schizophrenia, as these were part of the exclusion criteria. All participants gave written consent according to the procedures of the Ethical committee of the Versailles General Hospital (France) and the Cambridge Research and Ethics Committee (UK).

Before statistical analysis, all time responses (MRT, SSD, and SSRT) were screened for extreme values. A cutting point of $\pm 2$ standard deviations from the mean response value was considered as outlier. Three participants were excluded from study one following this criterion.

### PROCEDURE

Participants performed the experiment in a quiet room with a desk and a computer. After a short clinical interview to verify medical history, they were given a folder with questionnaires to fill up, written information about the study and a consent form in paper form.

All participants underwent a single 8 min acquisition block of a Go/NoGo task in order to take a base measure of the mean Go Reaction Time (Alderson et al., 2008; Cohen et al., 2010). Instructions for the Go/NoGo task were presented orally with a simple form. We used a standardized version of the Go/NoGo task, using green airplanes as go signals. Participants were told to hit the down arrow of the computer keyboard when the go signal appeared in the screen, and avoid responding when seeing a smiley face (the NoGo signal). NoGo's were randomly presented, but not consecutively 25% of the times a stimuli appeared.

Behavioral inhibition was examined using a SST paradigm that requires the cancellation of an already triggered go response. The experiment involves a routine motor reaction (hit a key) to a frequent go stimulus, with occasional cancellation of the routine response after an infrequent stop signal (Logan, 1994).

Instructions for the SST were presented in a standardized paper form. Participants were told that they were going to perform a video game-like task to determine how fast they were. They were told about the length of the task (6 acquisition blocks for the pilot study, 3 blocks for study one, and 4 blocks for study two) with a short pause between blocks (see **Figure 1**).

After reading the instructions, participants were asked to repeat the instructions to the evaluator and questions were answered. A brief training block of the SST without feedback was undertaken in order to ensure that the instructions had been fully understood.

In the present study we modified the SST developed by Rubia et al. (2003), which is, in turn, a faster visual variant of the Tracking SST (Logan et al., 1997).

As a modification to the SST, feedback is presented after each successful inhibition. Six types of feedback were exhibited: a smiley for the no-rewarded blocks, 1, 5, or 10 cents coin for the low incentive blocks, and a 20 or 50 cents coin for the high incentive blocks. The type of feedback was constant during each block. Participants were told about the order of presentation of rewards before the execution of the task.

These feedbacks were combined to build particular conditions, conducted along three experiments: pilot study, study one,

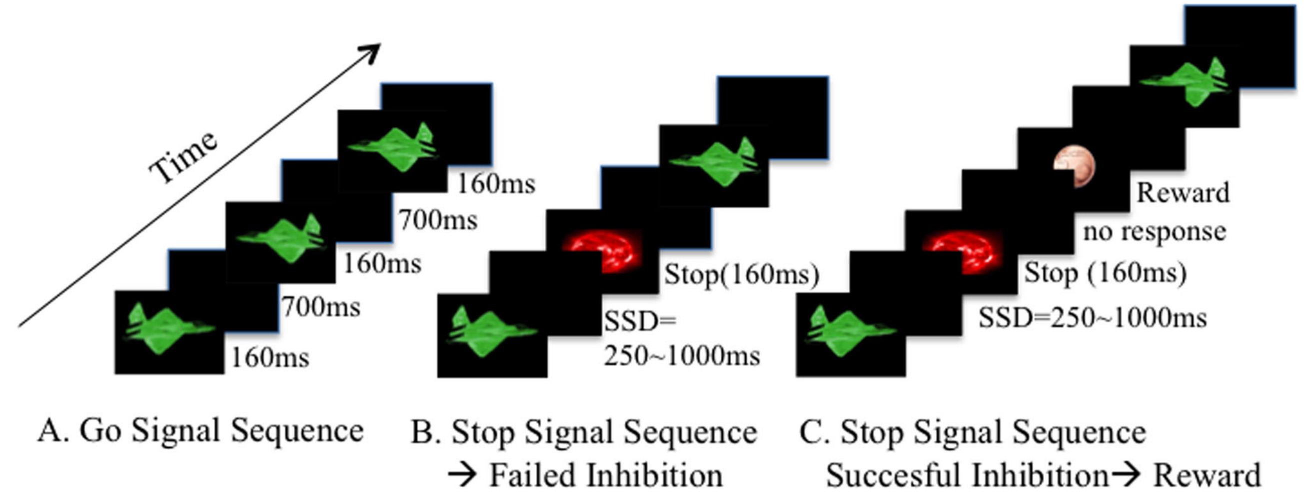


**FIGURE 1 | Experimental paradigm description.** The Stop Signal Task is composed by a Go stimulus, and a Stop signal after some Go stimuli, demanding the cancellation of an already triggered Go response. In our paradigm, each block has an estimated duration time of 8 min, with slight individual variations depending on the participant responses. Two hundred and ninety Go stimuli (green planes ~160 ms) and 59 Stop Signals (red planets ~160 ms) were presented in a counterbalanced order. **(A)** Go signal sequence: participants were asked to hit the right or left arrow of a keyboard, depending on the orientation of the plane. **(B)** Stop Signal appears after the Go Stimulus, in a delay between 250 and 1000 ms. This is called the Stop Signal Delay (SSD). If the participant fails to inhibit, no feedback is shown and the task continues with a new Go stimulus. **(C)** If the participant manages to withhold the motor response after the Stop Signal, a virtual reward is shown. Each block has a specific reward magnitude (smiley, 5 or 50 cts). Order of presentation of each block depends on the type of Study (see Pilot Study, Study One and Study Two for specific descriptions).

and study two (refer to each experiment descriptions after the Methods section).

Participants performed successive acquisition blocks, ~8 min each (6 blocks for the pilot study, 3 for the study one, and 4 for the study two). In each block two hundred and ninety green airplanes (Go-signals) were displayed on the middle of the screen for 300 ms. Participants were instructed to respond as quickly as they could by making left and right button presses (according to the direction of the plane). After the airplane, there was a blank screen for 700 ms except at ~20% of times. This accounted for the stop occurrences in the form of 59 red planets. They appeared after the airplane, at variable intervals, which corresponds to the SSD. The SSD changes in 50 ms steps, incrementing after successful inhibition, and decrementing after failed inhibitions. The SSD offset ranged from 250 up to 1000 ms.

Feedback was presented after each successful inhibition, at an offset of 250 ms after the stop stimulus disappearance. Predefined pseudo-randomized ISI occurred at 1600, 1700, 1800, 1900, or 2000 ms intervals and was not varied dynamically to balance for the frequency of successful vs. unsuccessful inhibition (Hampshire et al., 2010). For successful inhibition trials, where the feedback is presented, ISI expands dynamically depending on the predefined pseudo-randomized interval, and going up to 2400 ms till the next go stimulus trial presentation.

## DATA ANALYSES

Statistical analyses were performed with SPSS (IBM SPSS Software 19.0 Version, 2010). All data were checked for outliers, normal distribution, and homogeneity of variance. Critical alpha was set at 0.05 but frequently adjusted using Bonferroni corrections.

The dependent variables were three response time measures (MRT, SSD, and SSRT), and four task performance measures (number of failed inhibitions, missed go's, wrong keys, and number of rewards).

All variables were ready to be analyzed after recording, with the exception of the SSRT that was generated through a mathematical model proposed by Logan et al. (1997), following a subtraction of the MRT minus the SSD (formulae SSRT = MRT − SSD) (see Congdon et al., 2012, for a nice and detailed mathematical explanation).

A common model of a mixed ANOVA design was applied to the three studies. Each dependent variable was analyzed through the within-factors "order of blocks" (1, 2, 3 . . . given by the acquisition block order), "type of reward" (no reward, low reward, high reward) and the between-factor "condition" (increasing reward, decreasing reward).

Given the fact that each group condition was formed by different participants, we considered important to conduct a mixed ANOVA model in order to explore the behavior of the totality of participants, and then, separated One-Way ANOVA models, for each group condition (Increasing or Decreasing Reward) to better grasp the inner modulations of each independent condition, independently of the interaction effects (between groups) explored through the Two-Way ANOVA.

Since most task performance measures did not show a normal distribution, non-parametric tests were conducted for each independent group (Increasing and Decreasing Reward) through Wilcoxon paired-sample tests. Comparisons between each task measures and the condition groups for Increasing vs. Decreasing Reward groups were conducted through Kruskal–Wallis independent-sample tests (and corrected using Bonferroni).

## MATERIALS

Each participant performed the task on a DELL personal computer equipped with an Intell 2 processor. Individuals were seated 1 m from a 20″ screen, the nose aligned with the fixation cross so the reward stimuli would fall in the center of the visual field.

The stimuli presentation was programmed in Visual Basic 6.0. Each stimulus was presented against a black background at the center of a 15 inches standard screen.

The test was performed in a testing room, artificially lighted. At the beginning of the task, the participants underwent a short practice block, ensuring the correct visualization of every stimulus; luminosity was kept constant in the stimuli with no ambiguity. There was no need to measure the luminosity screen.

## GO/NOGO

Seventy-seven participants ($n = 38$ for study one, $n = 39$ for study two) underwent the Go/NoGo version in a single acquisition block of 8 min of duration. The Go/NoGo Task was not applied to the pilot study participants.

Mean and standard deviation to reaction time scores of the Go/NoGo Task per individual served as normalization parameter to the reaction time's obtained through the SST. Mean values were very consistent (study one $MRT = 282$ ms, $SD \pm 72$; study two $MRT = 223$ ms, $SD \pm 41$).

## THE PILOT STUDY

The aim of the pilot study was to test the effect of several rewards on inhibitory control as measured by different monetary rewards during the SST.

### *Procedure (pilot study)*

Twenty one participants [mean age 31 ($SD = 5.2$), gender ratio was 1.1] participated from the pilot study. Our in-house version of the SST was applied in 6 blocks. Four types of reward feedback were introduced: a no monetary reward with smiley, 1, 10, and 20 cents coins. Smiley feedback was always presented at the beginning of the protocol and for the odd blocks (blocks 1, 3, and 5). Monetary feedback was given on the even blocks (blocks 2, 4, and 6). The reward magnitudes for these even blocks were assigned in a random manner.

Given the distribution of reward through the task, participants were categorized in 3 conditions, relying on the progression trends of the reward magnitudes in time: "increasing condition" when the participant received low rewards at the beginning and then increasing reward magnitudes, "decreasing condition" when going from high rewards to low rewarded blocks, and "variable condition," with no specific reward progression pattern.

### *Analysis (pilot study)*

Two mixed ANOVA models were applied with the aim of explore diverse aspects of the reward effect over the inhibition profile. As described below, we examined the inhibition profile by response time and performance measures (MRT, SSD, SSRT, number of failed inhibitions, missed go's, wrong keys, and rewards, as described in the general methods section).

First, a One-Way ANOVA model was conducted to look at the influence of the reward magnitude *per se* on the time measures of the task, regardless of the order of reward type assignment or group type. Second, a 4(no reward, 1, 10, 20 cts) by 3(increasing condition, decreasing condition, variable condition) ANOVA was conducted to analyse the influence of the time history of reward assignment over the inhibition dependent variables.

### *Results (pilot study)*

The One-Way ANOVA revealed no modulatory effect of reward on the time measures of the SST although the descriptive results suggested differences that prompted further analyses.

The Two-Way mixed ANOVA did not show a significant effect of group for MRT or SSD. Even though there was a trend for group effect in SSRT [$F_{(5,\ 20)}1,\ 73$, $p = 0.052$]. There was no interaction effect between the group and the reward assignment conditions. However, this first experiment was not set to explore the group difference and for this was underpowered. We further test the hypotheses of group type and order of reward on the subsequent experiments.

Post paired tests for time measurements corrected for multiple comparisons, (MRT, SSD, SSRT) suggested differences between the first two blocks and the rest of the acquisition blocks set. The contrasts were significant for the SSD between first and second block [$df_{(1,\ 20)}$, $p = 0.002$] and the SSRT between the first and the last block [$df_{(1,\ 20)}$, $p = 0.001$]. Task performance variables showed a consistent and progressive improvement in gains, accuracy and less error.

### *Pilot study conclusions*

Pilot study analysis primarily showed a ceiling performance profile in all of the dependent variables, supposedly not influenced by the different reward magnitudes.

The progressive improvement trend most probably suggests a learning effect acquired through the lengthy task. This improvement seems to occur at a critical point where most of the performances had no more room for improvement. The SSRT appeared to improve till the last acquisition block. This training effect over the SSRT is not in agreement with previous studies theorizing that this value is a stable inhibition landmark (Cohen et al., 2010).

## STUDY ONE: EFFECT OF REWARD MAGNITUDE AND REWARD HISTORY

The aim of study one was to determine whether there is a modulatory effect over performance induced by different reward magnitudes, and the extent to which the order of presentation of rewards may modulate performance on subsequent blocks.

## PROCEDURE (STUDY ONE)

Thirty-eight participants [mean age 24 ($SD = 4$), gender ratio 1.1] were included. The study design was programmed after a close analysis of the pilot results. The fallout was a briefer protocol, a clearer reward assignment with only two types of monetary feedback sequences, instead of the four applied on the pilot study. The random reward assignment was replaced for a clear design where participant where allocated to an Increasing or Decreasing Reward Group.

The outcome was a three-block protocol with the modified version of the SST. Participants were distributed in two groups corresponding to either Increasing or Decreasing Reward.

Participants in the increasing reward condition began with a no monetary reward (smiley face feedback) block, second block, monetary feedback was 5 cents, and third was 50 cents. Participants in the Decreasing Reward condition undertook the same number of acquisition blocks, but rewards were presented in the reverse order (first block showing a 50 cents feedback, second block 5 cents, and finally the smiley). All participants were aware of the reward presentation order before beginning the task.

## RESULTS (STUDY ONE)

### *Time performance measures*

We first performed a One-Way ANOVA to evaluate the influence of the reward magnitude *per se* over the task measures regardless of the order of reward or group. The ANOVAs for SSD, MRT, and SSRT showed no significant differences between rewards, similarly to the initial results from the pilot study.

Two-Way mixed ANOVAs were conducted for each time measures (MRT, SSD, SSRT). A 2*3 design was applied, given the two condition groups (increasing and decreasing rewards), the three blocks of reward levels (first, second, and third block in one of the two orders defined by the design), or the three reward magnitudes (smiley, 5 cts, 50 cents) per Group. The first permitted the assessment of the effect of Order of reward and the second to test the reward effects *per se* while taking into account the Group factor. Means and SD are reported on **Table 1**.

Standard deviation scores seemed different for the time measures over the two condition groups, suggesting higher variance among participants in the Increasing reward condition (see **Table 1**). Despite similar slope changes for both condition groups, performance appeared slower and less variable for the Decreasing condition group (albeit not significant), we suspected from a different strategy for each condition group (**Figure 2A**).

The ANOVA model for the Reward magnitude analysis (independent of Order of reward), revealed a clear Interaction effect for SSD [$F_{(2,\ 72)} = 18.21$, $p < 0.001$] and SSRT [$F_{(2,\ 72)} = 7.52$, $p = 0.001$] between Reward magnitude and Group. There was no effect of reward *per se*.

The ANOVA testing Group, Order of reward and interaction showed a robust effect of Group for all three time measures (MRT, SSD, and SSRT) but no significant main effect of the Order of reward. Likewise, no interaction effect between the Group time measures and Order of reward was revealed (See **Table 2**).

*Post-hoc* comparisons (Bonferroni corrected) were performed between the three blocks independently from the factor "Group," assessing the effect of order *per se*, showing that paired comparisons between MRTs on blocks 1 and 2 significantly differ [$df_{(1,\ 37)}$, $p < 0.001$], but not MRT on block 1 and 3 [$df_{(1,\ 37)}$, $p = 0.36$] or block 2 and 3 [$df_{(1,\ 37)}$, $p = 0.19$]. Same paired-tests for SSD revealed significant differences between blocks 1 and 2 and block 1 and 3 [$df_{(1,\ 37)}$, $p < 0.001$] but not between blocks 2 and 3 [$df_{(1,\ 37)}$, $p = 0.031$]. Along with the ANOVA, these results suggest an effect specifically due to the chronological progression of the task, independently of the reward magnitude assigned on each block.

SSRT paired comparisons did not significantly differ between the first two blocks: block 1 vs. block 2 [$df_{(1,\ 37)}$, $p = 0.70$], but they significantly differed between the first and last blocks block 1 vs. block 3 [$df_{(1,\ 37)}$, $p = 0.003$], and showed a trend between the second and third blocks [$df_{(1,\ 37)}$, $p = 0.017$].

Additional *post-hoc* comparisons (Bonferroni corrected) were also performed between the three blocks from the factor "Reward," assessing the effect of reward magnitude *per se*, revealing a lack of significant effects for MRT, SSD, or SSRT measurements.

To further explore the effect of the reward magnitudes and order but independently for each Group, One-Way ANOVAs were conducted for Increasing and Decreasing rewards, separately. No significant differences were observed for time measures (MRT, SSD, and SSRT). So we cannot conclude that reward magnitude modulates inhibitory control when probing each group alone.

**Table 1 | Study One: Increasing and decreasing conditions.**

| **Order** | **Increasing condition (*n* = 18)** | | | **Decreasing condition (*n* = 20)** | | |
|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **1** | **2** | **3** |
| **Reward size** |  |  |  |  |  |  |
| | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** |
| MRT (ms) | 969 ± 224 | 1056 ± 261 | 1005 ± 263 | 1036 ± 216 | 1073 ± 228 | 1061 ± 208 |
| SSD (ms) | 700 ± 210 | 823 ± 197 | 835 ± 156 | 714 ± 229 | 778 ± 228 | 808 ± 221 |
| SSRT (ms) | 269 ± 106 | 232 ± 190 | 170 ± 210 | 321 ± 93 | 294 ± 133 | 253 ± 147 |
| Failed stops (*n*°) | 23.2 ± 6.2 | 21.9 ± 8.0 | 18.6 ± 8.5 | 24.6 ± 5.7 | 20.7 ± 7.7 | 18.8 ± 8.4 |
| Missed go (*n*°) | 7.4 ± 6.5 | 6.9 ± 5.4 | 8.8 ± 7.3 | 7.1 ± 8.3 | 7.2 ± 4.7 | 5.1 ± 3.3 |
| Wrong keys (*n*°) | 3.3 ± 3.9 | 1.9 ± 2 | 3.5 ± 3.4 | 1.5 ± 1.8 | 2.1 ± 1.9 | 1.3 ± 1.3 |
| Rewards (*n*°) | 30.7 ± 5.9 | 32.7 ± 7.2 | 36.3 ± 7.8 | 32 ± 5.4 | 33.5 ± 6.4 | 36.7 ± 8.4 |

*All Measures of performance in speed and accuracy.*

*MRT, Mean reaction time; SSD, Stop signal delay; SSRT, Stop signal reaction time.*

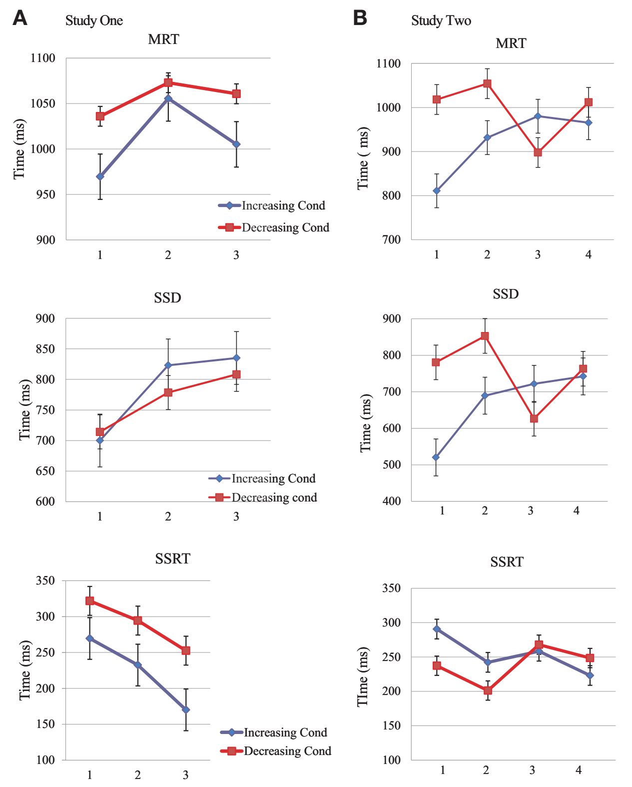


**FIGURE 2 | Order effects of Reward for MRT, SSD, and SSRT.** MRT (top), SSD (middle), and SSRT (bottom) means and standard errors for Study One and Study Two. **(A)** Study One blocks in the order of presentation (1, 2, and 3). Increasing condition participants (blue lines) received a smiley for successful inhibition on the first block, then 5 cents and finally 50 cents for the last block. Inverse reward for the decreasing condition group (red line). **(B)** Study Two blocks in the order of presentation (1, 2, 3, and 4). Increasing condition participants (blue lines) received two 5 cts block for successful inhibition and then two blocks of 50 cents. Inverse reward for the decreasing condition group (red line).

To disentangle the effects of the blocks we tested each pair of blocks between the two condition groups inside the ANOVA model with an univariate analysis for paired comparisons (Bonferroni corrected, block 1 from Increasing Reward group and Block 1 from Decreasing Reward group, and the same for blocks 2 and 3). No significant differences were revealed when comparing same presentation order blocks between the two condition groups. No firm conclusion can be drawn from the differences between the Decreasing and Increasing groups in pair comparisons either.

### *Task performance measures*

Errors were indexed by counting the number of inhibition errors (failed stops), missed go signals (missed go), and precision errors for the left-right decision (wrong keys). Gains correspond to the number of rewards.

Performance measures did not exhibit a normal distribution. Non-parametric tests (Kruskal–Wallis for independent samples) comparing each performance measure between the two condition groups did not show any significant differences.

**Table 2 | Study One: Two-Way ANOVA for time performance measures.**

| | Main effect of group | Main effect of order | Group*Order interaction |
|---|---|---|---|
| | **(*df* = 1.37)** | **(*df* = 1.37)** | **(*df* = 2.72)** |
| | ***F(p)*** | ***F(p)*** | ***F(p)*** |
| MRT (ms) | 6.62 (0.002) | 0.47 (0.49) | 1.19 (0.307) |
| SSD (ms) | 18.5 (<0.001) | 0.006 (0.94) | 1.49 (0.23) |
| SSRT (ms) | 7.87 (0.001) | 1.64 (0.207) | 0.177 (0.83) |

*MRT, Mean reaction time; SSD, Stop signal delay; SSRT, Stop signal reaction time. Order refers to the reward order assignment for each condition group.*

Wilcoxon test for related samples comparing performances between blocks for each group separately, showed significant differences for the inhibition errors (number of failed stops) between the first and last blocks $[df_{(1,37)}, p = 0.006]$ and between the second and last block $[df_{(1,\ 37)}, p = 0.004]$ for the increasing condition. This may correspond to a training effect as well as a motivational effect with the increasing reward magnitude. No significant differences were detected for the Missed Go or Wrong Key scores for the paired comparisons in the increasing condition.

Decreasing condition performance for the number of failed inhibitions were significantly different for paired blocks 1 and 2 $[df_{(1,\ 37)}, p = 0.003]$ and blocks 1 and 3 $[df_{(1,\ 37)}, p = 0.001]$, as well as for the number of Rewards between the first and last blocks $[df_{(1,\ 37)}, p = 0.007]$. A progressive improvement of inhibition errors and reward raw scores was observed.

Missed Go's paired comparisons were significantly different for the last two blocks inside the decreasing condition $[df_{(1,\ 37)}, p = 0.007]$. Missed Go's raw scores also show an improvement through the task, as for the number of failed inhibitions and rewards.

Again, against the hypothesis, participants from the decreasing condition group did not decrease their performance with diminishing rewards. We favor the explanation of the training/learning effect.

## COMMENTS ON STUDY ONE

The results suggest a modest modulatory effect of reward on performance for both condition groups, with a strong effect from the start when the highest reward is received in the first block of the task. A learning (or practice/training) effect may be the cause for the improvement of performance independently from rewards and its order.

Despite the clear main effect of group, the order of reward assignment (or an interaction effect between the groups) was not demonstrated. The main effect of group cannot necessarily be attributed to the fact that participants were exposed to one or the other reward assignment condition. Notwithstanding similar slope changes for both condition groups, averages seemed slower and less variable for the decreasing condition group (albeit not significant), leading us to suspect that each group may be applying a different strategy from the first block on.

Decreasing condition participants appeared to exhibit better global performances from the first block and consequently seemed more resistant to reward modulation afterwards. This early effect, which we refer to as the "kick start effect," may be caused by the high reward received in the first block. That is, the reward in the first block defines the primary performance, probably creating a mental set, and shaping a strategy that favors withholding of responses, and remains resistant to reward changes afterwards. The general performance seemed also modulated by a learning effect that was evident in the pilot study, where increasingly better scores were reported, regardless of reward or order of presentation of rewards. Participants may have learnt to proactively withhold their responses. The training effect trend was suggested by the SSD, number of failed inhibitions, and reward values for both condition groups, given by an apparently progressive improvement through the three blocks.

MRT for the increasing condition group, showed a steep slope between the first and the second block, corresponding to the transition between a no-rewarded condition to the first monetary reward condition. It could be hypothesized that the motivational boost effect of the first reward is strong. Participants in the increasing condition group also showed a clear slowing down of progression in the task (SSD), from block to block maybe signaling a better inhibitory capacity. However, it may be due to the training effect, which is also observed in performance measurements of the number of failed inhibitions. The suggested training effect could be masking the reward magnitude effect for the increasing condition group. SSRT scores also exhibited an improvement trend as suggested by the significant differences between first and second, and first and third blocks, also evidenced for both reward conditions, and the training effect could be at the origin of this trend or be the main contributor.

SSRT scores are the hallmark of the inhibition process measured through the SST, participants on the increasing condition group progressively obtained better inhibition scores (269 ms for the first trial, up to 170 ms for the last). The enhancement of the inhibition capacity has been described between reward cues and no-reward cues in the SST (Scheres et al., 2001; Lijffijt et al., 2004; Nigg, 2005). It may be difficult to disentangle whether the SSRT improvement on this experiment was induced primarily by the presence of increasing rewards or by the training effect.

Decreasing condition participants appeared to exhibit better global performances from the first block and consequently seemed to change to a lesser degree with subsequent blocks. This effect of high reward at the beginning of the task seemed to modulate performances from then on, reflecting our proposed "kick start effect."

Time measurements were not significantly different from block to block inside the same condition group. Even though MRTs had a slower starting point compared to increasing condition group (**Figure 2**). Likewise, SSRT were slower from the beginning, in comparison with SSRT from the increasing condition group (**Figure 2**).

In sum, the effect of the reward magnitude is suggested by an improvement in the global scores for the increasing reward group, with best performances for the higher rewarding blocks. Nevertheless, this magnitude reward effect was not backed up by an interaction effect, suggesting that there are other major factors influencing reward processing and inhibition. A training effect is a more likely explanation given a shared improvement pattern for Increasing as well as for Decreasing condition group. Moreover, a quick modulation of performances from the starting block, suggests a particular modulatory effect given by the impact of the history of reward assignment. In support of this "kick start" interpretation, it was observed that the Decreasing condition seemed to be influenced from the high-reward first block, thereby increasing the baseline performance.

## STUDY TWO: EFFECT OF REWARD MAGNITUDE

The aim of Study Two was to determine the effect of different levels of reward magnitudes (low vs. high rewards) and whether those contrasting magnitudes produced a modulation dependent on individual differences in SSRT performances.

### PROCEDURE (STUDY TWO)

Thirty-nine healthy participants underwent a two-rewards protocol [mean age 24.7($SD = 4.5$), gender ratio 1:1] presented in a four-block task: increasing reward ($n = 20$) and decreasing reward ($n = 19$) contingencies.

A four-block protocol was designed to contrast high and low rewards depending on order of presentation. Participants were allocated to one of two groups corresponding to either Increasing or Decreasing Rewards. In this experiment, the Increasing Group had four blocks with reward order as follows: 5, 5, 50, and 50 cts; while the Decreasing Group order was: 50, 50, 5, and 5 cts (see **Table 3**, for experimental design, means and STDs). There were no no-reward blocks in this second experiment and we also controlled for learning effects by including a repetition in each reward condition. All participants were aware of the reward presentation order before beginning the task.

### RESULTS (STUDY TWO)

#### *Time performance measures*

We first performed a One-Way ANOVA to evaluate the influence of the reward magnitude *per se* over the task measures regardless of the order of reward or group. The ANOVAs for SSD, MRT, and SSRT showed no significant differences between rewards, confirming the initial results from the pilot study and Study One.

Two-Way mixed ANOVAs were conducted for each time measures (MRT, SSD, SSRT), the first aiming and testing the reward effect *per se* by group and the second to evaluate the order effects. The Reward 2*2 mixed-model ANOVA was conducted between the 2 conditions (increasing and decreasing reward) and the 2 reward magnitudes (5 and 50 cents) and showed a main effect of group for MRT [$F_{(1,\ 37)} = 18.75$, $p < 0.001$] and SSD [$F_{(1,\ 37)} = 29.4$, $p > 0.001$], but no effect of Reward, and no interaction effects.

The 2*4 mixed-model ANOVA was applied to test the Order effects, with 2 conditions (increasing and decreasing Reward) and 4 blocks per time measure (MRT, SSD, SSRT). There were significant main effects of group in the all the time measures (MRT, SSD, SSRT), and significant interaction between group and order of reward assignment for MRT and SSD but not for SSRT (see **Table 4**).

**Table 3 | Study Two: Increasing and decreasing conditions.**

| Order | Increasing condition ($n = 20$) | | | | Decreasing condition ($n = 19$) | | | |
|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **1** | **2** | **3** | **4** |
| **Reward size** |  |  |  |  |  |  |  |  |
| | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** | ***M* ± *SD*** |
| MRT (ms) | 811 ± 263 | 932 ± 310 | 980 ± 315 | 965 ± 291 | 1018 ± 256 | 1054 ± 211 | 898 ± 241 | 1012 ± 255 |
| SSD (ms) | 520 ± 300 | 689 ± 299 | 722 ± 285 | 742 ± 265 | 781 ± 233 | 853 ± 166 | 626 ± 276 | 763 ± 238 |
| SSRT (ms) | 291 ± 156 | 242 ± 203 | 258 ± 162 | 223 ± 187 | 237 ± 168 | 201 ± 168 | 268 ± 146 | 248 ± 142 |
| Failed stops (*n*°) | 23.0 ± 7.8 | 17.4 ± 7.6 | 15.3 ± 6.5 | 15.1 ± 6.8 | 15.3 ± 7.2 | 13.9 ± 5.9 | 20.0 ± 7.1 | 15.6 ± 7.2 |
| Missed go (*n*°) | 5.6 ± 3.8 | 6.1 ± 4.9 | 5.3 ± 4.4 | 5.3 ± 3.6 | 6.1 ± 5.2 | 6.4 ± 7.0 | 4.0 ± 2.3 | 4.6 ± 3.2 |
| Wrong keys (*n*°) | 3.8 ± 4.5 | 2.8 ± 3.0 | 2.2 ± 2.8 | 2.8 ± 3.1 | 2.4 ± 2.5 | 2.4 ± 3.8 | 4.3 ± 5.6 | 2.7.0 ± 3.0 |
| Rewards (*n*°) | 28.0 ± 6.7 | 33.3 ± 8.2 | 36.4 ± 9.2 | 35.4 ± 8.5 | 35.9 ± 8.2 | 37.6 ± 8.8 | 32.0 ± 7.6 | 36.8 ± 8.3 |

*All Measures of performance in speed and accuracy.*
*MRT, Mean reaction time; SSD, Stop signal delay; SSRT, Stop signal reaction time.*

**Table 4 | Study Two: Two-Way ANOVA for time performance measures.**

| | Main effect of group | Main effect of order | Group* Order interaction |
|---|---|---|---|
| | **(*df* = 1.38)** | **(*df* = 1.38)** | **(*df* = 3.111)** |
| | ***F(p)*** | ***F(p)*** | ***F(p)*** |
| MRT (ms) | 5.81 (0.001) | 0.79 (0.38) | 14.94 (<0.001) |
| SSD (ms) | 4.45 (0.042) | 1.29 (0.26) | 24.17 (<0.001) |
| SSRT (ms) | 1.53 (0.020) | 0.11 (0.74) | 1.28 (0.284) |

*MRT, Mean reaction time; SSD, Stop signal delay; SSRT, Stop signal reaction time. Order refers to the reward order assignment for each condition group.*

Pairwise comparisons (Bonferroni corrected) *post-hoc* tests between the four blocks for the combined group (grouping the two conditions) tested the training effects *per se*, revealing significant differences for MRTs between blocks 1 and 2 [$df_{(1,\ 38)}$, $p = 0.003$], blocks 1 and 4 [$df_{(1,\ 38)}$, $p = 0.007$] and blocks 3 and 4 [$df_{(1,\ 38)}$, $p = 0.008$]. SSDs post-tests were also significant for the paired comparisons between blocks 1 and 2 [$df_{(1,\ 38)}$, $p < 0.001$], blocks 1 and 4 [$df_{(1,\ 38)}$, $p = 0.008$], and the last two blocks, block 3 vs. block 4 [$df_{(1,\ 38)}$, $p = 0.005$]. Pairwise comparison for the SSRT scores did not significantly differ. This exploratory analysis showed small but significant effects between blocks despite a negative main effect of order in the ANOVA. These findings also suggest dissimilar strategies between the two condition groups, supported by the interaction effect of group and order of presentation of rewards.

To further explore the effect of the reward magnitudes per condition group, One-Way ANOVAs were conducted for each independent condition group (Increasing or Decreasing Reward). No significant differences were observed for any time measures (MRT, SSD, and SSRT).

To better define the possible origin of the differences in the Order ANOVA, we conducted univariate comparisons in the ANOVA model (Tukey HSD) for first blocks, second, third and fourth blocks between the two conditions (block 1 from Increasing reward group vs. block 1 from Decreasing reward group, repeated for each of the 4 blocks). No significant differences were observed between each time-related block between the two conditions, for the MRT or the SSRT. However, univariate paired comparisons between time-related blocks on both condition groups, revealed that SSDs significantly differed between the first two blocks [$df_{(1,\ 38)}$, $p = 0.004$], suggesting a different strategy at the beginning of the task. This result may add some support to the kick start effect found in Study One.

While MRT and SSD showed significant interaction effects, SSRT failed to show differences. Furthermore, a lack of differences between blocks was probed independently through the One-Way ANOVA analyses.

An interpretation of these results will be presented in the discussion part, however, we think these findings suggest that the SSRT, as a compound measure, may lack sensitivity, and breakdown on MRT and SSD bears the potential to provide useful information on the behavioral adjustments, otherwise hidden by the SSRT scores.

### *Task performance measures*

Performance measures (Number of Failed Stops, Missing Go's, Wrong Keys, and Number of Rewards) did not exhibit a normal distribution. Non-parametric tests (Kruskal–Wallis for independent samples) showed significant differences between the number of failed stops for the first block between the two groups [$df_{(1,\ 38)}$, $p = 0.006$], as well as for rewards also for the first block comparison [$df_{(1,\ 38)}$, $p = 0.004$], consistent with the pairwise comparison between SSD time measurements. These results—again—support an early effect of the reward (kick start effect).

Wilcoxon test for related samples (Bonferroni corrected for multiple comparisons) were conducted to compare performances within each condition group. Failed Stops and Rewards were significantly different in several pair comparisons among blocks in the increasing condition reward group [failed stops between blocks 1–3 [$df_{(1,\,38)}$, $p = 0.001$], blocks 1–4 [$df_{(1,\,38)}$, $p = 0.003$]; number of rewards between blocks 1–2 [$df_{(1,\,38)}$, $p = 0.002$], blocks 1–3 [$df_{(1,\,38)}$, $p = 0.001$], Blocks 1–4 [$df_{(1,\,38)}$, $p = 0.001$], blocks 2–3 [$df_{(1,\,38)}$, $p = 0.004$]. Missed Go and Wrong Keys did not differ in any comparison. For the decreasing reward group, robust differences were found between the second and third blocks for Failed Stops [$df_{(1,\,38)}$, $p = 0.002$], Wrong Keys [$df_{(1,\,38)}$, $p = 0.006$], and Rewards [$df_{(1,\,38)}$, $p = 0.009$]. These findings support the hypothesis of the modulation of inhibitory control with reward magnitude since significant differences were found primarily among the strongest rewarded blocks (blocks 1 and 2) and the drop of reward magnitude in the third and fourth block.

## COMMENTS ON STUDY TWO

Analysis of Study Two revealed a robust main effect of group like in Study One, supporting the hypothesis of modulation of inhibitory control by history and context of reward, but also showed an interaction effect between group and order of presentation of rewards. In this experiment, the reward assignment seems to have induced differences in how performances change. However, the order effect *per se* was not significant.

SSD comparisons for the same block between groups showed a marked difference in Blocks 1 and 2; MRTs are clearly slower for the first blocks on the decreasing condition, suggesting an immediate behavior modulation for the highest rewarded blocks. When observing MRTs for the higher reward blocks on the increasing condition groups, there was also a withholding pattern, but the scores were not as slow compared to those obtained for the same rewarded blocks on the decreasing condition. These observations again support a kick-start effect induced by the highest reward at the beginning of the task (**Figure 2B**).

Performance measures suggest the modulatory effect of reward magnitude history highlighting two main phenomena: (1) participants in the Increasing condition improved their performances in a progressive manner throughout the task, and (2) participants in the Decreasing condition had a good performance from the beginning of the task (high reward) and a dramatic fall in all measures when they transitioned from high to low reward blocks. Furthermore, Decreasing group scores were higher, not only from the beginning, but also when compared to high reward blocks from the Increasing Group. This finding suggests a similar kick-start effect in both studies Two and One (**Figure 3**).

## DISCUSSION

The aim of this study was to investigate the behavioral effect of reward contingencies in the SST by manipulating the magnitude and order of reward. Little research has been done manipulating different reward magnitudes in an inhibition task (but see Shanahan et al., 2008) despite the common use of punishment and reward in learning to stop a particular behavior or inhibit an urge (Ridderinkhof et al., 2004). We argue that the experimental design from the two studies presented here provide deeper insight into the motivational mechanisms of the inhibitory processes, pushing experimental contingencies beyond the primary executive-motor dimension and shedding light into the mechanism underlying the modulation of cognitive control. Our study introduced multi-level reward magnitudes along with a dynamical presentation of those reward contingencies over two different experimental conditions. We intended to obtain additional clues to understand the motivational aspects of the manipulation of reward magnitudes in the same inhibition task. What is the impact of the reward magnitudes on cognitive inhibition? What is the overall motivational effect of giving rewards, independently of their magnitudes? Does reward size matters? We predicted two effects: a modulation of inhibition through reward *per se* and the modulation of the history (through order of presentation) of those rewards. The results observed for study one provided weak evidence in support of our hypothesis that there is a modulation on behavioral inhibition depending on the reward magnitude and order. Study two provided a more consistent confirmation of our reward effects hypotheses. Both studies demonstrated a strong modulation effect of the history of reward assignment.

In the pilot study, reward contingencies were masked by the randomization of reward magnitude blocks and furthermore, the

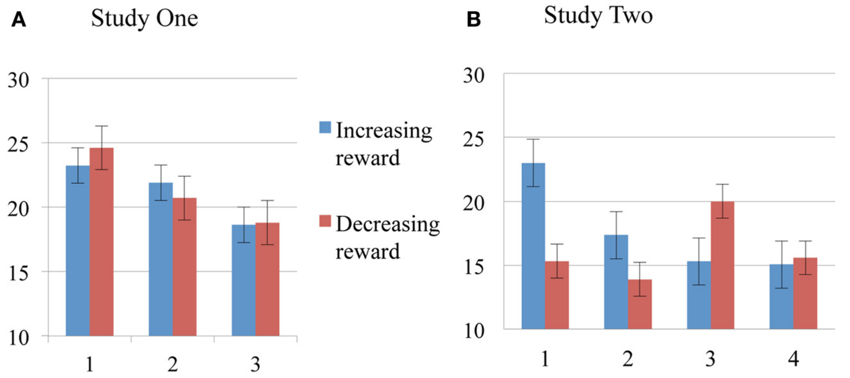


**FIGURE 3 | Order effects of Reward for Failed Inhibitions.** Means and standard errors for Study One **(A)** and Study Two **(B)**. Blue bars for Increasing Groups, red bars for Decreasing Groups.

experimental design was long (six blocks), allowing the development of a learning pattern that was evident across most of the performance measures, and particularly for the SSRT. Whilst the inhibitory measures seemed mostly independent of reward contingencies, the fact that they varied with practice, in conjunction with the results of other two experiments, demonstrate that the SSRT is not as stable a trait as originally claimed. Scheres et al. (2001) showed evidence on the improvement of SSRT scores due to reward contingencies on ADHD children. However, the SSRT seemed weak and unstable in the pilot study, leading us to postulate that by exploring the SSRT building blocks, the MRT and SSD, there is potential to better understand the inhibition modulation by reward.

The original hypotheses proposed a change, proportional to the reward magnitude and independent of the presentation order, nonetheless, it quickly became evident after the pilot study exploration that more specific hypotheses regarding the effects of reward on inhibitory control were needed to account for the results obtained and expected. This first pilot analysis permitted the design of specific experiments to test the effect of history of reward, and the reward magnitude itself.

Study One provided little evidence for the impact of reward magnitudes but strongly suggested a kick-start effect, a result that accords with the pilot study. A learning effect was also evident in Study One, with a progressive improvement on raw scores for both groups, that is, independent of the reward size assignment. Study Two offered clearer evidence of the effect of the reward modulation. We believe this discrepancy to be explained by the design of study being more prone to the masking effects of improvements due to task learning.

Studies One and Two both showed that participants on the decreasing conditions, that is, who start with higher rewards, presented better global scores on all measures. This result was unexpected and based on this we have proposed that there was an early modulatory effect induced by the reward. In study 1, the early modulatory effect was sustained across blocks even if rewards decreased or disappeared. Thus, we termed it the "kick start effect," as it seems to have a lasting influence over performances throughout the task. The results from study's one and two suggest that this kick-start effect works via a reward "boost," which impacts the behavioral markers of inhibition of the executive process in play during the SST. Thus, the presence of reward at the beginning of the task can lead to higher cognitive control over performance by moving the threshold of the capacity to withhold a response.

The results of study two demonstrated that the reward magnitude modulation seems to be independent of learning or practice effects but highly dependent on context. Participants were able to improve their scores when confronted to a higher reward on the Increasing condition, as expected, but there was a performance decrement from the second to the third block in the decreasing group, not caused by a lack of training but possibly due to a disengagement of motivation: participants that were exposed to 50 cents feedback for each successful inhibition, suffered a fall of 45 cts per trial after the third block. This change of reward seems to induce an override of the motivational effect that cannot be explained by practice.

The results in study one and two clearly suggest that behavioral adjustments may not only be related to the order of the reward magnitude, but also due to a "kick start effect" that modulates performance from the beginning and has consequences throughout the rest of the task. Other authors have described similar ideas in the literature of Stop Tasks with reward contingencies, using other terms such as the "Arousal Effect" (Shanahan et al., 2008) or the "Novelty effect" (Ronga et al., 2013). The concepts are not equivalent since the effects where not alike. Further theoretical efforts, based on wider meta-analyses and new experimental findings should help cement these concepts.

In Study One we were able to induce a modulatory effect of the order of reward magnitude that appeared stronger in Study Two. This effect, or rather its interaction, could have been diminished on Study One by several factors: (a) the presence of no-reward blocks in the same task, (b) a masking effect of the history of previous rewards, induced by the kick start effect observed on the decreasing condition, (c) the learning effect and its interaction. It is difficult to disentangle these probable causes but future experiments will necessarily do so when taking into account these factors in their experimental design.

The SSRT is the major index of inhibition pattern obtained for the SST. Many studies utilizing the SST have the tendency to report primarily the SSRT values as noted on the meta-analysis by Alderson et al. (2008). It is important to note that the SSRT is a combined measure obtained indirectly by the calculation of the optimal time up to which inhibition is still possible, according to a given SSD. Comparative analysis using only SSRT values for groups under different conditions may leave out the dynamical changes observed over MRT and SSD. The underlying idea of the SSRT is that it combines the reaction times and the recent history of the response withholding in one compound measure, however, in the present studies, we found no significant SSRT differences that created the impression of an absence of modulatory effects of reward contingencies. However, a more complete analysis revealed hidden patterns behind the MRT, SSD, and errors. We propose that a more detailed inspection of the measures obtained in the SST provides additional information of the differences in inhibitory performance between groups, otherwise hidden by the SSRT raw scores or by limited understanding of task measures.

Electrophysiological and neuroimaging studies should help to explore the underlying mechanisms of inhibitory control modulated by reward (Overbeek et al., 2005; Wiersema et al., 2005). The neurodynamics revealed by evoked potentials may be particularly helpful (Gehring et al., 1990) to determine whether there is a "novelty" or a "saliency" phenomenon with reward, and if there is an ERP magnitude correlation. These questions have been put forward in previous studies on ERN/Ne magnitudes using reward cues (Liotti et al., 2005; Holroyd et al., 2009), some studies even suggest that the ERN/Ne amplitude can also reflect the motivational value of a task, being elicited by individual sensitivities to reward magnitudes, punishments (Boksem et al., 2006; van Meel et al., 2011) or predicted rewards (Yasuda et al., 2004). Methodological adjustments should be taken into account in order to test the order of reward magnitude effect and the kick start effect.

Furthermore, a clinical application of the present protocols may be instrumental in exploring the cognitive (and neurophysiological) signatures in some psychiatric conditions, specifically when impulsivity is one of the core symptoms. There is a growing body of evidence suggesting that reward modulatory effects on inhibitory control could be stronger on Attention Deficit Hyperactive Disorder (ADHD) patients than in the normal population. Our experimental design will be further applied to address this question in ADHD patients in the near future.

## ACKNOWLEDGMENTS

This Research was supported by a Wellcome Trust Biomedical Research Fellowship WT093811MA (to Tristán A. Bekinschtein). We thank Corinne Bareham and Stanimira Georgieva for assisting with technical support and stimuli preparation and Charly Garcia for contributing valuable discussions and insights.

## REFERENCES

Albrecht, B., Banaschewski, T., Brandeis, D., Heinrich, H., and Rothenberger, A. (2005). Response inhibition deficits in externalizing child psychiatric disorders: an ERP-study with the Stop-task. *Behav. Brain Funct.* 1, 22. doi: 10.1186/1744-9081-1-22

Alderson, M. R., Rapport, M. D., Sarver, D. E., and Kofler, M. J. (2008). Attention deficit hyperactive disorder (ADHD) and behavioral inhibition: a re-examination of the stop-signal task. *J. Abnorm. Child Psychol.* 36, 989–998. doi: 10.1007/s10802-008-9230-z

Aron, A. R., Fletcher, P. C., Bullmore, T., Sahakian, B. J., and Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci.* 6, 115–116. doi: 10.1038/nn1003

Avila, C., and Parcet, M. A. (2001). Personality and inhibitory deficits in the stop-signal task: the mediating role of Gray's anxiety and impulsivity. *Pers. Individ. Dif.* 31, 975–986. doi: 10.1016/S0191-8869(00)00199-9

Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., and Rothenberger, A. (2003). Association of Attention Deficit Hyperactive Disorder (ADHD) and conduct disorder–brain electrical evidence for the existence of a distinct subtype. *J. Child Psychol. Psychiatry* 44, 356–376. doi: 10.1111/1469-7610.00127

Band, G. P., Van Der Molen, M. W., and Logan, G. D. (2003). Horse-race model simulations of the stop-signal procedure. *Acta Psychol. (Amst).* 112, 105–142. doi: 10.1016/S0001-6918(02)00079-3

Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–15. doi: 10.1016/0010-0277(94)90018-3

Boehler, C., Hopf, J., Stoppel, C. M., and Krebs, R. M. (2012). Motivating inhibition—reward prospect speeds up response cancellation. *Cognition* 125, 498–503. doi: 10.1016/j.cognition.2012.07.018

Boksem, M., Tops, M., Wester, A., Meijman, T., and Lorist, M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Res.* 1101, 92–101. doi: 10.1016/j.brainres.2006.05.004

Boksem, M. A., Tops, M., Kostermans, E., and De Cremer, D. (2008). Sensitivity to punishment and reward omission: evidence from error-related ERP components. *Biol. Psychol.* 79, 185–192. doi: 10.1016/j.biopsycho.2008.04.010

Bokura, H., Yamaguchi, S., and Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clin. Neurophysiol.* 112, 2224–2232. doi: 10.1016/S1388-2457(01)00691-5

Brass, M., and Haggard, P. (2007). To do or not to do: the neural signature of self-control. *J. Neurosci.* 27, 9141–9145. doi: 10.1523/JNEUROSCI.0924-07.2007

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., and Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb. Cortex* 11, 825–836. doi: 10.1093/cercor/11.9.825

Brown, J., Bullock, D., and Grossberg, S. (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *J. Neurosci.* 19, 10502–10511.

Bush, G., Luu, P., and Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222. doi: 10.1016/S1364-6613(00)01483-2

Carver, C. S., and Scheier, M. F. (2001). *On the Self-Regulation of Behaviour*. Cambridge, UK: Cambridge University Press.

Chen, A., Tang, D., and Chen, X. (2013). Training reveals the sources of stroop and flanker interference effects. *PLoS ONE*, 8:e76580. doi: 10.1371/journal.pone.0076580

Chikazoe, J., Jimura, K., Hirose, S., Yamashita, K. I., Miyashita, Y., and Konishi, S. (2009). Preparation to inhibit a response complements response inhibition during performance of a stop-signal task. *J. Neurosci.* 29, 15870–15877. doi: 10.1523/JNEUROSCI.3645-09.2009

Clark, L., Roiser, J. P., Cools, R., Rubinsztein, D. C., Sahakian, B. J., and Robbins, T. W. (2005). Stop signal response inhibition is not modulated by tryptophan depletion or the serotonin transporter polymorphism in healthy volunteers: implications for the 5-HT theory of impulsivity. *Psychopharmacology (Berl).* 182, 570–578. doi: 10.1007/s00213-005-0104-6

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: a proposal. *Arch. Gen. Psychiatry* 44, 573–588. doi: 10.1001/archpsyc.1987.01800180093014

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., et al. (2010). Decoding developmental differences and individual variability in response inhibition through predictive analyses across individuals. *Front. Hum. Neurosci.* 4:47. doi: 10.3389/fnhum.2010.00047

Cohen, J. R., and Poldrack, R. A. (2008). Automaticity in motor sequence learning does not impair response inhibition. *Psychon. Bull. Rev.* 15, 108–115. doi: 10.3758/PBR.15.1.108

Congdon, E., Mumford, J., Cohen, J., Galvan, A., Canli, T., and Poldrack, R. A. (2012). Measurement and reliability of response inhibition. *Front. Psychol.* 3:37. doi: 10.3389/fpsyg.2012.00037

De Pascalis, V., Varriale, V., and D'Antuono, L. (2010). Event-related components of the punishment and reward sensitivity. *Clin. Neurophysiol.* 121, 60–76. doi: 10.1016/j.clinph.2009.10.004

Derryberry, D., and Rothbart, M. K. (1997). Reactive and effortful processes in the organization of temperament. *Dev. Psychopathol.* 9, 633–652. doi: 10.1017/S0954579497001375

Dias, R., Robbins, T. W., and Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *J. Neurosci.* 17, 9285–9297.

Elliott, R., Dolan, R. J., and Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb. Cortex* 10, 308–317. doi: 10.1093/cercor/10.3.308

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., and Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51, 871–882. doi: 10.1016/j.neuron.2006.07.029

Gehring, W. J., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1990). The error-related negativity: an event-related brain potential accompanying errors. *Psychophysiology* 27, S34.

Geurts, H., van der Oord, S., and Crone, E.(2006). Hot and cool aspects of cognitive control in children with Attention Deficit Hyperactive Disorder (ADHD): decision-making and inhibition. *J. Abnorm. Child Psychol.* 34, 811–822. doi: 10.1007/s10802-006-9059-2

Gray, J. A. (1987). "The neuropsychology of emotion and personality," in *Cognitive Neurochemistry*, eds S. M. Stahl, S. D. Iversen, and E. C. Goodman (New York, NY: Oxford University Press), 171–190

Groom, M. J., Scerif, G., Liddle, P. F., Batty, M. J., Liddle, E. B., Roberts, K. L., et al. (2010). Effects of motivation and medication on electrophysiological markers of response inhibition in children with attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 67, 624–631. doi: 10.1016/j.biopsych.2009.09.029

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., and Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage* 50, 1313–1319. doi: 10.1016/j.neuroimage.2009.12.109

Hasher, L., Lustig, C., and Zacks, R. (1999). "Inhibitory mechanism and the control of attention," in *Variation in Working Memory*, ed C. Jarrold (New York, NY: Oxford Press), 293–308.

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., and Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cogn. Affect. Behav. Neurosci.* 9, 59–70. doi: 10.3758/CABN.9.1.59

Immordino-Yang, M. H., and Damasio, A. (2007). We feel, therefore we learn: the relevance of affective and social neuroscience to education. *Mind Brain Educ.* 1, 3–10. doi: 10.1111/j.1751-228X.2007.00004.x

Jessie, J. (2011). *From Who Are You. "Price Tag"*. Los Angeles, CA: Lava Island.

Kalanthroff, E., Cohen, N., and Henik, A.(2013) Stop feeling: inhibition of emotional interference following stop-signal trials. *Front. Hum. Neurosci.* 7:78. doi: 10.3389/fnhum.2013.00078

Kok, A., Ramautar, J. R., De Ruiter, M. B., Band, G. P., and Ridderinkhof, K. R. (2004). ERP components associated with successful and unsuccessful stopping in a stop—signal task. *Psychophysiology* 41, 9–20. doi: 10.1046/j.1469-8986.2003.00127.x

Lansbergen, M. M., Böcker, K. B., Bekker, E. M., and Kenemans, J. L. (2007). Neural correlates of stopping and self-reported impulsivity. *Clin. Neurophysiol.* 118, 2089–2103. doi: 10.1016/j.clinph.2007.06.011

Lawrence, A. J., Luty, J., Bogdan, N. A., Sahakian, B. J., and Clark, L. (2009). Impulsivity and response inhibition in alcohol dependence and problem gambling. *Psychopharmacology (Berl).* 207, 163–172. doi: 10.1007/s00213-009-1645-x

Li, C. S. R., Huang, C., Constable, R. T., and Sinha, R. (2006). Gender differences in the neural correlates of response inhibition during a stop signal task. *Neuroimage* 32, 1918–1929. doi: 10.1016/j.neuroimage.2006.05.017

Lijffijt, M., Bekker, E. M., Quik, E. H., Bakker, J., Kenemans, J. L., and Verbaten, M. N. (2004). Differences between low and high trait impulsivity are not associated with differences in inhibitory motor control. *J. Atten. Disord.* 8, 25–32. doi: 10.1177/108705470400800104

Liotti, M., Pliszka, S. R., Perez, R., Kothmann, D., and Woldorff, M. G. (2005). Abnormal brain activity related to performance monitoring and error detection in children with Attention Deficit Hyperactive Disorder (ADHD). *Cortex* 41, 377–388. doi: 10.1016/S0010-9452(08)70274-0

Lipszyc, J., and Schachar, R. (2010). Inhibitory control and psychopathology:a meta-analysis of studies using the stop signal task. *J. Int. Neuropsychol. Soc.*16, 1064–1076. doi: 10.1017/S1355617710000895

Living Colour. (1991). *EP Biscuit, "Money Talks."* Los Angeles, CA: A&M Recording.

Logan, G. D., and Cowan, W. B. (1984). On the abilitiy to inhibit thought and action: a theory of an act of control. *Psychol. Rev.* 91, 295–327. doi: 10.1037/0033-295x.91.3.295

Logan, G., Schachar, R., and Tannok, R. (1997). Impulsivity and inhibitory control. *Psychol. Sci.* 8, 60–64. doi: 10.1111/j.1467-9280.1997.tb00545.x

Logan, G. D. (1994). "On the ability to inhibit thought and action: a user's guide to the stop signal paradigm," in *Inhibitory Processes in Attention, Memory and Language*, eds D. Dagenbach and T. H. Carr (San Diego, CA: Academic Press), 189–239.

Luman, M., Tripp, G., and Scheres, A. (2010). Identifying the neurobiology of altered reinforcement sensitivity in ADHD: a review and research agenda. *Neurosci. Biobehav. Rev.* 34, 744–754. doi: 10.1016/j.neubiorev.2009.11.021

MacLeod, C. M. (2007). "The concept of inhibition in cognition," in *Inhibition in Cognition*, eds D. S. Gorfein and C. M. MacLeod (Washington, DC: American Psychological Association), 3–23.

Masui, K., and Nomura, M. (2011). The effects of reward and punishment on response inhibition in non-clinical psychopathy. *Pers. Individ. Dif.* 50, 69–73. doi: 10.1016/j.paid.2010.08.024

Michel, J. A., Kerns, K. E., and Mateer, C. A. (2005). The effect of reinforcement variables on inhibition in children with Attention Deficit Hyperactive Disorder (ADHD). *Child Neuropsychol.* 11, 295–302. doi: 10.1080/092970490911270

Neo, P., Thurlow, J., and McNaughton, N. (2011). Stopping, goal-conflict, trait anxiety and frontal rhythmic power in the stop-signal task. *Cogn. Affect. Behav. Neurosci.* 11, 485–493. doi: 10.3758/s13415-011-0046-x

Nichols, S. L., and Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of Attention Deficit Hyperactive Disorder (ADHD). *Child Psychiatry Hum. Dev.* 34, 297–315. doi: 10.1023/B:CHUD.0000020681.06865.97

Nigg, J. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 125, 220–249. doi: 10.1037/0033-2909.126.2.220

Nigg, J. (2001). Is Attention Deficit Hyperactive Disorder (ADHD) a inhibitory disorder? *Psychol. Bull.* 127, 571–598. doi: 10.1037/0033-2909.127.5.571

Nigg, J. T. (2005). Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: the state of the field and salient challenges for the coming decade. *Biol. Psychiatry* 57, 1424–1435. doi: 10.1016/j.biopsych.2004.11.011

Oosterlaan, J., Logan, G. D., and Sergeant, J. A. (1998). Response inhibition in AD/HD, CD, comorbid AD/HD+ CD, anxious, and control children: a meta-analysis of studies with the stop task. *J. Child Psychol. Psychiatry* 39, 411–425. doi: 10.1017/S0021963097002072

Oosterlaan, J., and Sergeant, J. (1998). Effects of reward and response cost on response inhibition in Attention Deficit Hyperactive Disorder (ADHD), disruptive, anxious, and normal children. *J. Abnorm. Child Psychol.* 26, 161–174. doi: 10.1023/A:1022650216978

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, R. K. (2005). Dissocialbe components of error processing. *J. Psychophysiol.* 19, 319–329. doi: 10.1027/0269-8803.19.4.319

Panksepp, J. (2003). At the interface of the affective, behavioural, and cognitive neurosciences: decoding the emotional feelings of the brain. *Brain Cogn.* 52, 4–14. doi: 10.1016/S0278-2626(03)00003-4

Pessoa, L., Padmala, S., Kenzer, A., and Bauer, A. (2012). Interactions between cognition and emotion during response inhibition. *Emotion* 12, 192–197. doi: 10.1037/a0024109

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., and Gore, J. C. (1999). An fMRI study of Stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. *Biol. Psychiatry* 45, 1237–1258. doi: 10.1016/S0006-3223(99)00056-6

Potts, G. F., George, M. R. M., Martin, L. E., and Barratt, E. S. (2006). Reduced punishment sensitivity in neural systems of behaviour monitoring in impulsive individuals. *Neurosci. Lett.*. 397, 130–134. doi: 10.1016/j.neulet.2005.12.003

Quay, H. C. (1993). The psychobiology of undersocialized aggressive conduct disorder: a theoretical perspective. *Dev. Psychopathol.* 5, 165. doi: 10.1017/S0954579400004326

Ramautar, J. R., Kok, A., and Ridderinkhof, K. R. (2004). Effects of stop-signal probability in the stop-signal paradigm: the N2/P3 complex further validated. *Brain Cogn.* 56, 234–252. doi: 10.1016/j.bandc.2004.07.002

Ridderinkhof, R. K., van den Wildenberga, W. P. M., Segalowitzd, S. J., and Carter, C. S.(2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016

Ronga, I., Valentini, E., Mouraux, A., and Iannetti, G. D. (2013). Novelty is not enough: laser-evoqued potentials are determined by stimulus saliency, not absolute novelty. *J. Neurophysiol.* 109, 692–701. doi: 10.1152/jn.00464.2012

Rubia, K., Smith, A. B., Brammer, M., and Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage* 20, 351–358. doi: 10.1016/S1053-8119(03)00275-1

Rubia, K., Smith, A. B., Brammer, M., and Taylor, E. (2007). Temporal Lobe dysfunctions in medication-naïve boys with Attention Deficit Hyperactive Disorder (ADHD) during attention allocation and its relation to response variability. *Biol. Psychiatry* 62, 999–1006. doi: 10.1016/j.biopsych.2007.02.024

Rubia, K., Smith, A. B., Brammer, M., Toone, B., and Taylor, E. (2005). Abnormal brain activation during inhibition and error detection in medication-naive adolescents with ADH. *Am. J. Psychiatry* 162, 1067–1075. doi: 10.1176/appi.ajp.162.6.1067

Sagaspe, P., Schwartz, S., and Vuilleumier, P. (2011). Fear and stop: a role for the amygdala in motor inhibition by emotional signals. *Neuroimage* 55, 1825–1835. doi: 10.1016/j.neuroimage.2011.01.027

Salo, R., Henik, A., and Robertson, L. C. (2001). Interpreting stroop interference: an analysis of differences between task versions. *Neuropsychology* 15:462. doi: 10.1037/0894-4105.15.4.462

Scheres, A., Oosterlaan, J., and Sergeant, J. A. (2001). Response inhibition in children with DSM-IV subtypes of AD/HD and related disruptive disorders: the role of reward. *Neuropsychol. Dev. Cogn. Sect. C Child. Neuropsychol.* 7, 172–191. doi: 10.1076/chin.7.3.172.8746

Schmajuk, M., Liotti, M., Busse, L., and Woldorff, M. G. (2006). Electrophysiological activity underlying inhibitory control processes in normal adults. *Neuropsychologia* 44, 384–395. doi: 10.1016/j.neuropsychologia.2005.06.005

Shanahan, M. A., Pennington, B. F., and Willcutt, E. W. (2008). Do motivational incentives reduce the inhibition deficit in Attention Deficit Hyperactive Disorder (ADHD)? *Dev. Neuropsychol.* 33, 137–159. doi: 10.1080/87565640701884238

Smith, J. L., and Mattick, R. P. (2013). Evidence of deficits in behavioral inhibition and performance monitoring in young female heavy drinkers. *Drug Alcohol Depend*. 133, 398–404. doi: 10.1016/j.drugalcdep.2013.06.020

Sonuga-Barke, E. (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol. Psychiatry* 57, 1231–1238. doi: 10.1016/j.biopsych.2004.09.008

Stevens, J., Quittner, A. L., Zuckerman, J. B., and Moore, S. (2002). Behavioral inhibition, self-regulation of motivation, and working memory in children with attention deficit hyperactivity disorder. *Dev. Neuropsychol.* 21, 117–139. doi: 10.1207/S15326942DN2102_1

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *Q. J. Exp. Psychol. A* 54, 321–343. doi: 10.1080/713755969

van Boxtel, G. J., van der Molen, M. W., Jennings, J. R., and Brunia, C. H. (2001). A psychophysiological analysis of inhibitory motor control in the stop-signal paradigm. *Biol. Psychol.* 58, 229–262. doi: 10.1016/S0301-0511(01)00117-X

van Gaal, S., Ridderinkhof, K. R., van den Wildenberg, W. P., and Lamme, V. A. (2009). Dissociating consciousness from inhibitory control: evidence for unconsciously triggered response inhibition in the stop-signal task. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1129. doi: 10.1037/a0013551

van Meel, C. S., Heslenfeld, D. J., Oosterlaan, J., Luman, M., and Sergeant, J. A. (2011). ERPs associated with monitoring and evaluation of monetary reward and punishment in children with Attention Deficit Hyperactive Disorder (ADHD). *J. Child Psychol. Psychiatry* 52, 942–953. doi: 10.1111/j.1469-7610.2010.02352.x

van Reekum, C. M., and Schaefer, S. M. (2011). Aging is associated with positive responding to neutral information but reduced recovery from negative information. *Soc. Cogn. Affect. Neurosci.* 6, 177–185. doi: 10.1093/scan/nsq031

Verbruggen, F., and De Houwer, J. (2007). Do emotional stimuli interfere with response inhibition? Evidence from the stop signal paradigm. *Cogn. Emot.* 21, 391–403. doi: 10.1080/02699930600625081

Wendt, M., Kiesel, A., Geringswald, F., Purmann, S., and Fischer, R. (2014). Attentional adjustment to conflict strength. *Exp. Psychol.* 61, 1–13. doi: 10.1027/1618-3169/a000227

Wiersema, J. R., Van Der Meere, J. J., and Roeyers, H. (2005). ERP correlates of impaired error monitoring in children with Attention Deficit Hyperactive Disorder (ADHD). *J. Neural Transm.* 112, 1417–1430. doi: 10.1007/s00702-005-0276-6

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., and Tannock, R. (1999). Development of inhibitory control across the life span. *Dev. Psychol.* 35, 205–213.

Winstanley, C. A., Eagle, D. M., and Robbins, T. W. (2006). Behavioral models of impulsivity in relation to Attention Deficit Hyperactive Disorder (ADHD): translation between clinical and preclinical studies. *Clin. Psychol. Rev.* 26, 379–395. doi: 10.1016/j.cpr.2006.01.001

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., and Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *Neuroreport* 15, 2561–2565. doi: 10.1097/00001756-200411150-00027

Yu, J., Hung, D. L., Tseng, P., Tzeng, O. J., Muggleton, N. G., and Juan, C. H. (2012). Sex differences in how erotic and painful stimuli impair inhibitory control. *Cognition* 124, 251–255. doi: 10.1016/j.cognition.2012.04.007

Zelazo, P. D., Muller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function in early childhood. *Monogr. Soc. Res. Child Dev.* 68,vii–137. doi: 10.1111/j.0037-976X.2003.00269.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 08 October 2013; accepted: 07 April 2014; published online: 12 May 2014.*
*Citation: Herrera PM, Speranza M, Hampshire A and Bekinschtein TA (2014) Monetary rewards modulate inhibitory control. Front. Hum. Neurosci.* ***8****:257. doi: 10.3389/fnhum.2014.00257*
*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 06 February 2014
doi: 10.3389/fnhum.2014.00031

# Social information and personal interests modulate neural activity during economic decision-making

***Anna Moser, Celia Gaertig and María Ruz****

*Department of Experimental Psychology, Mind, Brain and Behavior Research Center, University of Granada, Granada, Spain*

***Edited by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

***Reviewed by:***
*Steve W. C. Chang, Duke University, USA*
*Claudio Lavin, Universidad Diego Portales, Chile*

****Correspondence:***
*María Ruz, Department of Experimental Psychology, Mind, Brain and Behavior Research Center, University of Granada, Campus Universitario Cartuja s/n, 18071 Granada, Spain*
*e-mail: mruz@ugr.es*

In the present study we employed electrophysiological recordings to investigate the levels of processing at which positive and negative descriptions of other people bias social decision-making in a game in which participants accepted or rejected economic offers. Besides social information, we manipulated the fairness of the assets distribution, whether offers were advantageous or not for the participant and the uncertainty of the game context. Results show that a negative description of the interaction partner enhanced the medial frontal negativity (MFN) in an additive manner with fairness evaluations. The description of the partner interacted with personal benefit considerations, showing that this positive or negative information only biased the evaluation of offers when they did not favor the participant. P300 amplitudes were enhanced by advantageous offers, suggesting their heightened motivational significance at later stages of processing. Throughout all stages, neural activity was enhanced with certainty about the personal assignments of the split. These results provide new evidence on the importance of interpersonal information and considerations of self-interests relative to others in decision-making situations.

**Keywords: positive and negative interpersonal information, economic decision-making, fairness, uncertainty, MFN, P300**

## INTRODUCTION

It is widely acknowledged that human decisions are not purely rational and outcome maximizing (Camerer, 2003) but that other factors, such as emotions or information about other people, are sources of bias in social decision-making (Harlé and Sanfey, 2007; Ruz et al., 2011). The aim of the current study is to shed light on the neural basis of the bias that evaluative information about other people exerts in interpersonal choices and its inter-dependence with fairness and personal benefit considerations.

Economic games are often used to study strategic human decisions in interpersonal contexts. One of them, the Ultimatum Game (UG; Güth et al., 1982), allows for investigation of how people react to unfair situations. In this game, two players share a certain amount of money. One player offers a split, and the other can then decide whether to accept or reject it. If she/he accepts, both players get the assigned part, whereas if she/he rejects it neither one gets any pay-off. Although the rational decision would be to accept any offer, since a small part of the split is still better than nothing, empirical results show that people reject about 50% of the unfair offers (Camerer, 2003). A prominent explanation of those results refers to the influence of emotions on decision-making (e.g., Pillutla and Murnighan, 1996). Unfair offers elicit negative emotions towards the proposer and trigger punishment of such antisocial behavior; they also activate brain regions linked to emotion processing (Sanfey et al., 2003) and social norm violation (Güroğlu et al., 2010).

Results from Ruz and Tudela (2011) and Ruz et al. (2013) suggest that the congruency of emotions and behavior of interaction partners exerts influence on the decisions we make in a social context. Further, Harlé and Sanfey (2007) found that even incidentally felt emotional states can influence decisions in economic bargaining situations. Social information about the person we interact with can also influence choices made in the UG. In a series of experiments, Ruz et al. (2011; see also Gaertig et al., 2012) showed that verbal descriptions of the personality of interaction partners influence the decisions in a modified UG. More precisely, positive descriptions of alleged partners were shown to increase the acceptance rates of both fair and unfair offers, and this effect was mainly present in an uncertain context in which people did not know the distribution of the offer between the two players.

With electroencephalography (EEG) it is possible to track changes in cortical activation with very high temporal resolution. Therefore, this method gives the opportunity to investigate the stages of processing at which biasing factors exert their influence. The medial frontal negativity (MFN), a specific event related potential (ERP), is of particular interest. This potential summarizes a family of ERPs that are thought to reflect reward prediction errors (the Error Related Negativity) and performance feedback (Feedback Related Negativity; Van Noordt and Segalowitz, 2012). The MFN is thought to originate in the anterior cingulate cortex, a region associated with cognitive control (Carter et al., 1998), negative emotional states (Sanfey et al., 2003) and the processing of both physical and social pain (Rainville et al., 1997). It has been related to performance monitoring and the outcome of decisions (Gehring and Willoughby, 2002), especially the emotional or motivational evaluation of negative outcomes, such as monetary losses (Hajcak et al., 2005). One interpretation of the MFN therefore is that it reflects the evaluation or appraisal of an outcome on a good-bad dimension (Yeung and Sanfey, 2004).

In electrophysiological studies of fairness perception in the UG, Boksem and De Cremer (2010; see also Polezzi et al., 2008; Hewig et al., 2011; Van der Veen and Sahibdin, 2011) found that the presentation of unfair offers elicited a more negative-going MFN compared to the presentation of fair offers. These results suggest that the MFN reflects a fast initial distinction as to whether outcomes adhere to an equity norm. The MFN has been found to also encode other social factors, such as the relationship between the interacting individuals (Ma et al., 2011) or inter-individual differences in social traits (Pfabigan et al., 2011). Campanhã et al. (2011) found that the MFN is strongly responsive to social distance. Crucially, when the economic offer is believed to come from a friend, the polarity of the MFN is reversed, and acceptance rates for unfair offers increase. The authors follow that unfair offers coming from a friend are perceived as less unfair and generate less "social pain" than offers from strangers. This suggests a strong association between the MFN and the ultimate appraisal of an outcome. The authors note, however, that friendship involves much more than social closeness, but also elements like similarity, sympathy and trustfulness. Further, the impact that negative social information might have on the perception of offer fairness remains unclear.

Therefore, a main interest of the present study was investigating whether positive and negative information about otherwise unknown partners modulated the evaluation of the fairness of their offers as reflected on the MFN potential or whether these effects exerted an additive influence. In addition, we manipulated other factors that are also known to affect choices in interpersonal situations. The introduction of advantageous and disadvantageous offers in which the participant either gets the higher or the lower part of the split allows for differentiation between effects elicited by offer *fairness* and those elicited by offer *advantageousness*. When disentangled, the rejection of unfair offers independent of advantageousness reflects an aversion to a violation of an impersonal equity rule (Fehr and Schmidt, 1999). Effects of advantageousness refer to the satisfaction of personal interests in a social comparative context instead. Finally, given that in real life we seldom know all the consequences of our decisions, we also manipulated the uncertainty of the context. Previous results suggest that in an uncertain context, in which participants lack full knowledge about the assignments of the split, social information has a higher impact on choices (Platt and Huettel, 2008; Ruz et al., 2011; Gaertig et al., 2012). The use of electrophysiological recordings allowed us to explore whether social information received enhanced processing in the uncertain context, as previous behavioral data would predict.

Although the MFN was the central potential of interest in our study, there are other ERP deflections that may provide valuable information regarding the levels of information processing at which social information modulates interpersonal decision-making. The P300 peaks around 300–600 ms on centro-parietal sites. It is understood as representing higher-order cognitive operations like decision-making (Nieuwenhuis et al., 2005), or attentional resource allocation (Donchin and Coles, 1988). In gambling studies it has been related to reward magnitude (Yeung and Sanfey, 2004) and reward valence (Hajcak et al., 2005). Employing the UG, Wu et al. (2012) associated the P300 with increased attention depending on the emotional/motivational significance of an outcome in asset distribution. In our study, analysis of the P300 might shed light on the encoding of personal benefit considerations in the human brain.

The present study was designed to investigate if the perception of fair and unfair offers in a modified UG could be altered by previous knowledge about interaction partners in contexts of varying certainty, extending behavioral results about the influence of social information on choices in the UG (Marchetti et al., 2011; Ruz et al., 2011; Gaertig et al., 2012). The study therefore manipulates the *offer fairness* (fair vs. unfair), the *social information* about the interaction partner (positive vs. negative), the *context certainty* (certain vs. uncertain) and the *advantageousness of the offer* (advantageous vs. disadvantageous). At the behavioral level we predict a replication of previous findings of our group concerning the influence of social information on choices in classic (Gaertig et al., 2012) and modified versions (Ruz et al., 2011) of the UG, showing higher acceptance rates for fair and unfair offers following a positive partner description. At the neural level, we hypothesize that the MFN will be modulated by the social information about the interaction partner. We further hypothesize that the P300 will be enhanced by the advantageous offers, given their enhanced motivational significance.

## METHODS

### PARTICIPANTS

Twenty-four students from the University of Granada (14 female, mean age: 22.9, age range: 18–34) participated in the study. All subjects had normal or corrected to normal vision. They signed a consent form approved by the Ethics Committee of the University of Granada and received course credits and a chocolate token in exchange for their participation.

### TASK

Participants played a modified game used previously by the authors (Ruz et al., 2011) in which they had to either accept or reject economic offers made by a partner. Participants were told that in each trial their partner, the proposer, received an initial amount of fictional money and split it into two parts, one for each of them. The participant then had to either accept or reject the offer. If she/he accepted it they would both earn their share, whereas if she/he decided to reject the proposer's offer, none would add money for that trial. To enhance closeness to reality, participants were told that offers used in the experiment were made by participants in previous experiments. In addition, to stress that participants' decision could not influence the offer on the next trial, they were told that they would play with a different proposer on each trial. To introduce the variable of social information, each proposer was described with a positive or negative adjective before the offer was presented. Additionally, and with the goal of getting participants to pay attention to personal benefits, they were told to try to accumulate more fictional money than all their partners together. Finally, we manipulated the certainty of the context in which choices were made. Participants had either full (certain context) or incomplete (uncertain context) information about the outcome of their decisions (see Ruz et al., 2011).

### STIMULI AND PROCEDURE

Offers were displayed in the center of the screen as two single-digit numbers (from 1 to 9), one for the proposer and one for the responder, separated by a slash symbol. The two numbers were never the same, and their difference was either 1 (fair offers) or 4 (unfair offers). Half of the offers were advantageous, which means that the participant received the higher part of the split, and the other half disadvantageous, assigning the smaller amount of the split to the participant. Participants responded pressing a button on a keypad with the index and middle fingers of their dominant hand (button assignment was counterbalanced across participants). They were instructed to respond as fast as they could, and that the higher part of the split would be added to the amount of the partner if they did not respond within 1500 ms. This information was given to maximize the effect of the verbal descriptions of the partners and to replicate the paradigm employed in previous behavioral studies (Ruz et al., 2011).

For the trait-valenced descriptions, the same 48 words used in a previous study by the authors (Ruz et al., 2011) were selected from the Spanish translation of the ANEW database (Redondo et al., 2007). Half of the words had a positive valence (7.26 in average) and the other half a negative valence (2.19 in average). Words were matched in number of letters (6.5 in average), arousal ratings (5.67 in average) and frequency of usage (26.4 in average; Kucera and Francis, 1967).

To manipulate the certainty of the context, the task was divided into a certain and an uncertain block. Numbers in one block were displayed in different colors (green vs. blue) and in the other block in different font styles (bold vs. underlined). The assignment of color vs. font style to the certain or uncertain conditions was counterbalanced across participants. In the certain block participants were informed of their color/font style and therefore knew which part of the split corresponded to them, whereas this information was not provided in the uncertain block. Even though the different colors/font styles did not reveal any information to the participant in the uncertain block, they were still used to hold visual input constant across blocks. The order of the certain and uncertain blocks (with 384 trials each, and breaks every 96 trials) was counterbalanced across participants. In total, participants received 768 offers. Each participant saw the same word 16 times, each time associated with a different offer. Participants took approximately 70 min to complete the whole task.

Each trial comprised a fixation cross (with a variable duration between 2000 and 3000 ms; +; 0.5°), then the positive or negative adjective for 200 ms (average 1.15°), another fixation cross (variable duration between 600 and 1000 ms) and finally the offer for 1500 ms (0.6°; see **Figure 1**).

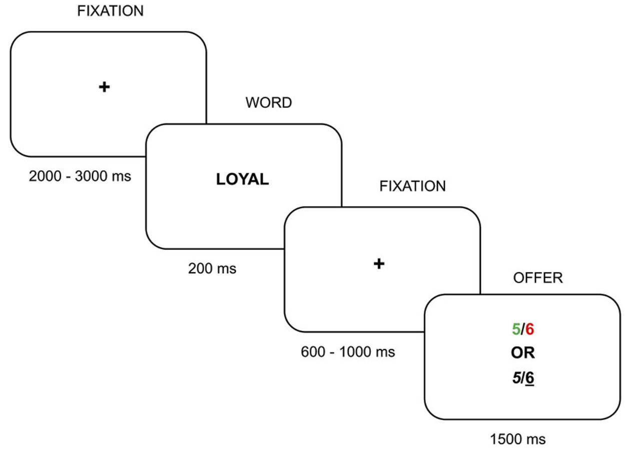


**FIGURE 1 | Schematic display of a trial sequence.**

### EVENT RELATED POTENTIAL (ERP) RECORDING AND ANALYSIS

An EEG net with 128 electrodes (Geodesics Sensor Net from *Electrical Geodesics, Inc.*, EGI) referenced to the vertex channel was used for the electrophysiological recordings. Participants sat in a dimly illuminated and electrically shielded room in front of a 50 cm distant computer screen. After receiving the instructions verbally and in written form, participants first performed a training block to familiarize themselves with the task. To secure good recording quality, participants were instructed to avoid eye-blinking and eye-movements during stimulus presentation. The channels above, beneath and beside the eyes were used as electrooculograph leads to detect eye blinks and movements. Signals were passed through an AC-coupled, high-input impedance amplifier (200 MΩ), and impedance was kept below 50 kΩ for all electrodes. The signal was amplified (0.1–100 Hz band pass), digitized at a sampling rate of 250 Hz (16 bits/D-converter) and stored for off-line analyses.

A 40 Hz lowpass filter was applied to the EEG to remove electrical noise, and afterwards data were segmented into epochs beginning 200 ms before and ending 800 ms after offer onset. Artifact detection was made for eye blinks and saccades (±70 μV threshold) and bad channels (±80 μV threshold). Data were further inspected manually to eliminate remaining bad segments not detected by the software. Channels were replaced with a spherical interpolation algorithm (Pernier et al., 1989) when more than 20% of the trials were bad for a specific channel. Trials with no behavioral response were excluded from the analysis. In order to maintain an acceptable signal-to-noise ratio a criterion of at least 25 trials per condition and subject was established. Data was re-referenced to the average (Tucker et al., 1994), and a single averaged segment was calculated for every condition and subject.

We focused our analyses on 19 anterior-frontal electrodes (number 3, 4, 5, 6, 7, 10, 11, 12, 13, 16, 19, 20, 21, 24, 25, 107, 113, 119, 124) for the MFN and on 16 central-posterior electrodes (number 52, 53, 54, 55, 60, 61, 62, 67, 68, 73, 78, 79, 80, 86, 87, 93) for the P300, where these potentials were maximally distributed. Electrode selection was also aided by localizations reported on previous studies (Boksem and De Cremer, 2010; Wu et al., 2011b). Average amplitudes over these electrodes were calculated with reference to a 200 ms pre-stimulus baseline, with time windows according to those reported in previous literature (Luck, 2005; Boksem and De Cremer, 2010), and also to visual inspection of the timing of waveforms in the present experiment.

For the behavioral results, the choices made by the participants (% of acceptance) were analyzed in a 2 (Context: certain vs. uncertain) × 2 (Fairness: fair vs. unfair) × 2 (Valence: positive vs. negative description of the partner) repeated measures ANOVA. To analyze effects of advantageousness, which refers to personal benefits of an offer compared to the outcome of the partner, data of only the certain block were analyzed in a 2 (Fairness: fair vs. unfair) × 2 (Valence: positive vs. negative description of the partner) × 2 (Advantageousness of the offer: advantageous vs. disadvantageous) repeated measures ANOVA. ERP analyses were performed analogously, submitting the mean amplitudes averaged across channels and temporal windows to the ANOVAs. The Greenhouse-Geisser correction for violations of the assumption of sphericity was used where appropriate and Bonferroni corrections were applied for multiple comparisons.

## RESULTS

### BEHAVIORAL RESULTS

Participants responded on time in 97.1% of the trials. The average acceptance rate of the offers was 52.7%. There was a main effect of fairness. Participants accepted more fair ($M$ = 65.9%, SE = 3.8%) than unfair ($M$ = 39.5%, SE = 2.6%) offers ($F_{1,23} = 21.20$, $p < 0.001$). Valence of the word also had a significant effect on the choice. Participants accepted offers preceded by a positive adjective ($M$ = 54.4%, SE = 1.6%) more often than those following a negative adjective ($M$ = 51.0%, SE = 1.9%; $F_{1,23} = 4.30$, $p < 0.05$). There was an interaction between the context and the fairness of the offer ($F_{1,23} = 20.73$, $p < 0.001$). The effect of fairness (i.e., acceptance rates of fair minus acceptance rates of unfair offers) was larger in the uncertain (44.7%, $F_{1,23} = 23.83$, $p < 0.001$) than in the certain condition (8.23%, $F_{1,23} = 4.72$, $p < 0.05$). In addition, there was an interaction between the context and the valence of the words ($F_{1,23} = 6.30$, $p < 0.05$). The effect of valence was significant only in the uncertain context (7.5%, $F_{1,23} = 5.33$, $p < 0.05$ vs. $F_{1,23} = 1.08$, $p = 0.31$ in the certain context). There was also a three-way interaction between context, fairness and valence ($F_{1,23} = 6.47$, $p < 0.05$). In both contexts the interaction between fairness and valence was significant (certain: $F_{1,23} = 5.76$, $p < 0.05$; uncertain: $F_{1,23} = 3.35$, $p < 0.05$). In the certain context, acceptance rates of fair offers were marginally higher when preceded by a negative ($M$ = 59.35%, SE = 3.82%) than by a positive ($M$ = 57.75%, SE = 3.89%) partner description ($F_{1,23} = 3.94$, $p = 0.06$). There was no difference for unfair offers ($F < 1$). In the uncertain condition, acceptance rates of fair offers were higher when preceded by positive ($M$ = 78.21%, SE = 4.41%) than by negative ($M$ = 68.21%, SE = 5.68%) words ($F_{1,23} = 5.51$, $p < 0.05$). Acceptance rates of unfair offers were marginally higher when preceded by positive ($M$ = 30.99%, SE = 5.68%) than by negative ($M$ = 26.08%, SE = 5.02%) words ($F_{1,23} = 3.89$, $p = 0.06$; see **Figure 2**).

The additional analysis yielded a main effect of advantageousness ($F_{1,23} = 639.6$, $p < 0.001$) with higher acceptance rates for advantageous ($M$ = 94.6%, SE = 1.5%) than for disadvantageous offers ($M$ = 14.0%, SE = 3.2%). An interaction between the advantageousness and the fairness of the offer ($F_{1,23} = 28.3$, $p < 0.001$) showed that when offers were advantageous, unfair offers were accepted more often (97.9%) than fair offers (91.4%; $F = 5.31$, $p < 0.05$). When offers were disadvantageous, fair offers were accepted more often (25.7%) than unfair offers (2.4%; $F = 15.17$, $p < 0.001$). Finally, the effects found in the main analysis were confirmed, showing an effect of fairness ($F_{1,23} = 4.93$, $p < 0.05$) and an interaction between fairness and valence ($F_{1,23} = 5.76$, $p < 0.05$).

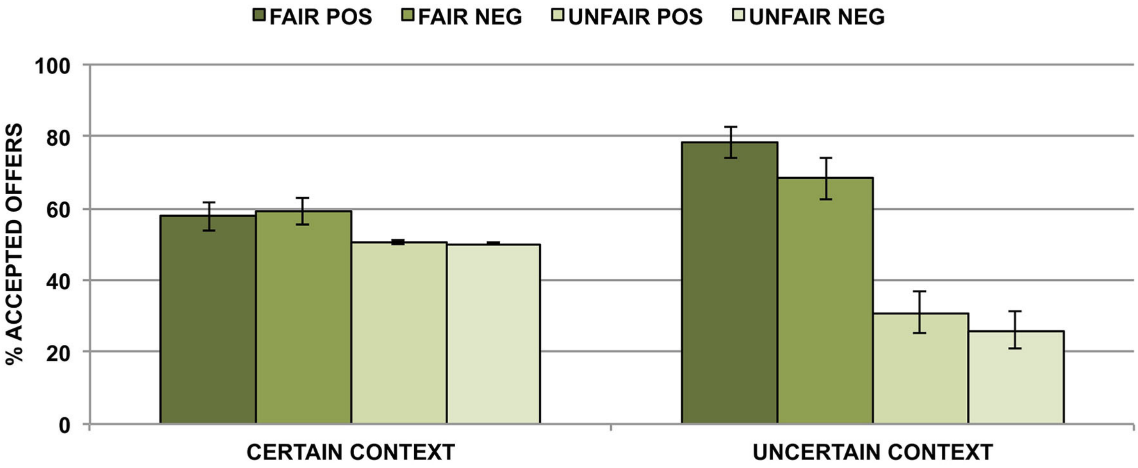


**FIGURE 2 | Acceptance rates for fair and unfair offers following positive and negative descriptions of the interaction partners in certain and uncertain contexts.** Error bars represent standard error of the mean.

## ELECTROPHYSIOLOGICAL RESULTS

### *Medial frontal negativity (MFN)*

The MFN peaked at 385 ms in fronto-central electrodes and was analyzed in a 370–400 ms temporal window. The analysis revealed a main effect of context, with a more pronounced MFN in the certain ($-0.61$ μV) as compared to the uncertain context ($-0.32$ μV; $F_{1,23} = 5.10$, $p < 0.05$; see **Figure 3**). Further, there was a main effect of fairness, as unfair offers elicited a more negative MFN ($-0.57$ μV) than fair offers ($-0.36$ μV; $F_{1,23} = 8.92$, $p < 0.01$). There was also a main effect of valence, because a negative description of the proposer elicited a more negative MFN ($-0.56$ μV) than a positive description ($-0.38$ μV; $F_{1,23} = 15.92$, $p = 0.001$; see **Figure 4**; also **Figure 5**). The interaction between fairness and valence was not significant ($F < 1$).

The additional analysis confirmed the results of the main analysis. A main effect of fairness ($F_{1,23} = 10.94$, $p < 0.01$) and a main effect of valence ($F_{1,23} = 9.83$, $p < 0.01$) were found. Furthermore, there was a significant interaction between the variables valence and advantageousness ($F_{1,23} = 4.36$, $p < 0.05$). Planned contrasts revealed that the effect of valence was only present when preceding disadvantageous offers (negative $-0.86$ μV vs. positive $-0.44$ μV; $F_{1,23} = 12.76$, $p < 0.01$), but not when preceding advantageous offers ($-0.64$ μV vs. $-0.58$ μV; $F < 1$).

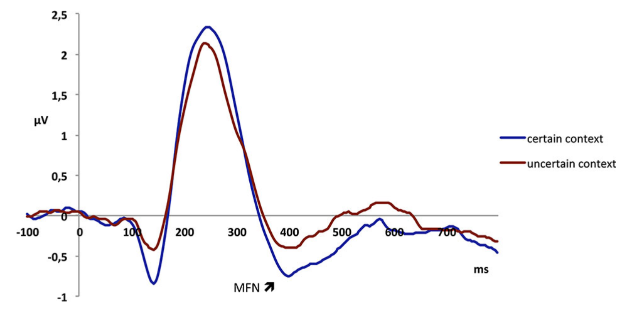


**FIGURE 3 | Electrophysiological data shows that offers presented in the certain context elicit a more negative MFN than those presented in the uncertain context.**

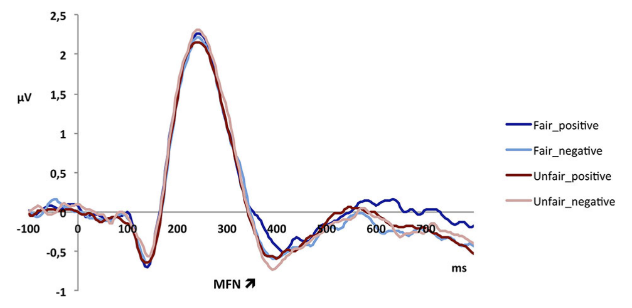


**FIGURE 4 | Electrophysiological data shows that unfair offers elicit a more negative MFN than fair offers, and offers preceded by a negative description of the interaction partner elicit a more negative MFN than those preceded by a positive description.** The effects of fairness and valence of the partner description are additive but do not interact.

### *P300*

The P300 was analyzed in centro-parietal electrodes in a 370–650 ms time window. The analysis revealed a significant main effect of context, with a higher amplitude in the certain (2.63 μV) than in the uncertain context (2.25 μV; $F_{1,23} = 4.79$, $p < 0.05$; see **Figure 6**). Further, there was a significant interaction between context and fairness ($F_{1,23} = 11.43$, $p < 0.01$). Planned contrasts showed that in the certain condition, unfair offers elicited a significantly larger P300 than fair offers (2.75 vs. 2.53 μV; $F_{1,23} = 4.37$, $p < 0.05$ ). Conversely, in the uncertain condition, fair offers elicited a marginally significant higher P300 than unfair offers (2.37 vs. 2.15 μV; $F_{1,23} = 4.00$, $p = 0.06$).

The additional analysis yielded a main effect of fairness ($F_{1,23} = 4.31$, $p < 0.05$), with unfair offers eliciting a higher P300 (2.73 μV) than fair offers (2.51 μV), confirming the effect found in the main analysis. The analysis also revealed a main effect of advantageousness ($F_{1,23} = 7.73$, $p < 0.05$), indicating that advantageous offers elicited a higher P300 (2.83 μV) than disadvantageous offers (2.41 μV; see **Figure 7**; also **Figure 5**). There were no other main effects or interactions (all $ps > 0.05$).

Visual inspection of the waveform suggested that there might be separate processes reflected in the early and the late phase of the P300. We therefore performed an additional analysis of an early (370–450 ms) and a late time windows (450–650 ms) of the component. The analyses confirmed the main effect of context (early time window: $F_{1,23} = 4.41$, $p < 0.05$; late time window: $F_{1,23} = 4.00$, $p = 0.06$) and the interaction between context and fairness (early time window: $F_{1,23} = 4.93$, $p < 0.05$; late time window: $F_{1,23} = 9.24$, $p < 0.01$) of the main analysis. However, in the early time window, the effect of advantageousness did not reach significance ($p > 0.05$), whereas this effect was significant in the analysis of the late time window ($F_{1,23} = 11.70$, $p < 0.01$).

## DISCUSSION

The present study was designed to investigate whether social information about other people modulates neural activity at the same neural stages as fairness and personal benefit considerations during interpersonal choices. While EEG was recorded, participants received fair and unfair offers from people previously described either positively or negatively. Both the fairness of the offer and, crucially, the social information about the partner were found to modulate electrophysiological responses, but without interacting between them. Furthermore, the advantageousness of an offer accounted for differential processing of social information about the interaction partner, underscoring the role of personal interest evaluations. Throughout all the stages under study, processing of the offer was increased in the certain as compared to the uncertain context.

At the behavioral level, results were similar to classic findings of the UG (Camerer, 2003), showing that people rejected more than half of the unfair offers. Furthermore, the results confirmed previous findings on the influence of social information on interpersonal choices (Ruz et al., 2011; Gaertig et al., 2012), showing that people accept more offers when these are believed to come from a positively as compared to negatively described person. This shows that non-predictive social information about interaction partners can bias decision-making in interpersonal situations.

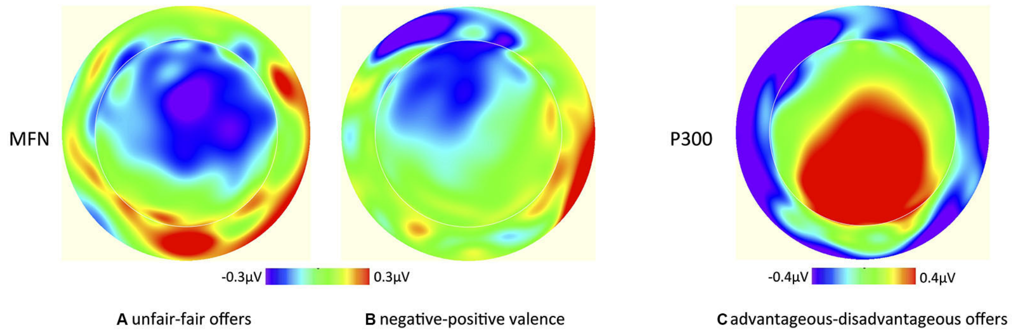


**FIGURE 5 |** Scalp potential topographies of the average voltage differences between **(A)** unfair and fair offers and **(B)** negative and positive partner description for the MFN, and between **(C)** advantageous and disadvantageous offers for the P300.

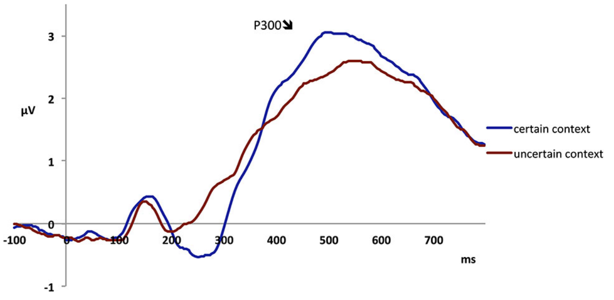


**FIGURE 6 | Electrophysiological data shows that offers presented in the certain context elicit higher P300 amplitudes than those presented in the uncertain context.**

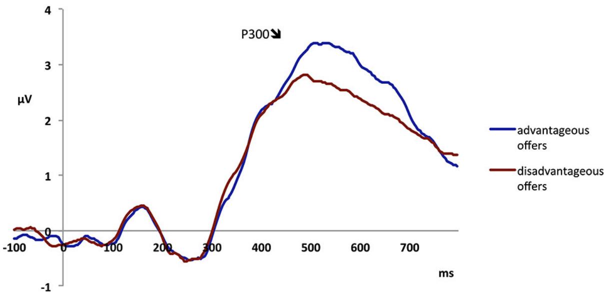


**FIGURE 7 | Electrophysiological data shows that advantageous offers, in which the participant is offered the higher part of the split, elicit higher P300 amplitudes than disadvantageous offers.**

This behavioral effect was only present in an uncertain context, in which participants lacked complete information about the outcome of their choices. In such uncertainty about the consequences of the decisions, participants seem to make use of each piece of information, independent of its actual validity as a predictor for optimal choice. We also found that the advantageousness of the offer influenced choices and participants accepted more offers when they were assigned the higher amount of the split. This effect interacted with the fairness of the offer, and participants preferred unfair offers when they were assigned the greater part of the split, and fair offers when they were assigned the smaller part of the split. This shows that participants tended to opt for choices which brought them more fictional money than their interaction partners, and, if that was not possible, they preferred offers in which the difference in gains was only small, which is conform to the instructions and also to natural self-interest. A three-way interaction between context, offer fairness and valence indicates that the influence of both fairness and partner description is much more pronounced in the uncertain context. This suggests that when the consequences of an action are less predictable, sources of additional information, such as characteristics of the offer and the interaction partner, have more influence on the decision at hand.

The MFN has been related to the affective appraisal of negative outcomes, such as unfair offers in an UG (Boksem and De Cremer, 2010). Our results replicated the finding of a more negative MFN for unfair offers than for fair ones. Most importantly, the valence of the social information about the interaction partner also had a significant effect on this potential. Negative as compared to positive partner descriptions enhanced the amplitude of the MFN. This effect indicates that offers are evaluated differentially depending on the character of the person that makes the offer. It suggests that as soon as the economic offer is evaluated, it is appraised as a more negative outcome when coming from an unlikable person.

Crucially, we found no interaction between the valence and the fairness of an offer. This indicates that having positive or negative information about the interaction partner does not change the evaluation of the fairness of the offer *per se*. Rather, our results suggest that both fairness and social information add up to generate an overall evaluation of the offer in a positive-negative continuum. This result is in line with the appraisal hypothesis of the MFN (Yeung and Sanfey, 2004), which suggests that the MFN reflects the ultimate appraisal of an outcome.

An alternative theory about the MFN is the Reinforcement Learning approach (Holroyd and Coles, 2002). It refers to expectancy violations and predicts a more pronounced MFN for situations in which previously generated expectations e.g., of fairness are not met. However, our neural results do not support this theory, because in our modified game we find no interaction between the valence of the partner description (which would reasonably inform fairness expectations) and the fairness of the offer. The data therefore rather suggest that the valence of the partner description and the offer fairness independently bias the evaluations of the offer as reflected in the MFN.

Furthermore, our additional analysis including advantageousness in the certain context allowed us to study whether the social information interacted with personal benefit considerations at this stage of processing. In this case, negative social information only enhanced the MFN for offers in which the participant received the lower amount of the split (disadvantageous offers). This effect indicates that social information did not bias the perception of an offer when the sum was split up in a way that privileged the participant. In the condition in which personal interests were satisfied, the personal character of the interaction partner did not seem to have an effect on the affective appraisal of the offer. In contrast, disadvantageous offers from partners described in a negative manner generated a MFN of more negative amplitude than those coming from partners preceded by positive information, which suggests that the offer is appraised more negatively in the former than in the latter case. This result demonstrates the priorities given to the different components of an interpersonal interaction, highlighting in first place personal benefit considerations. It suggests that the character of the interaction partner is considered only when those are not satisfied. When an offer is beneficial, people take less account of the character of the interaction partner.

Another interesting result is that we do not find an interaction between the fairness and the advantageousness of the offer. This suggests that the fairness of the offer modulates the MFN independently of its advantageousness. This is especially interesting because it provides insight into the role of the MFN as a reflection of fairness considerations that are not limited to self-interest. In other studies (e.g., Boksem and De Cremer, 2010) the fairness of the offer was always linked to an advantageous split. Crucially, our design enabled us to distinguish between impersonal offer fairness and personal advantageousness, showing a cleaner effect in the MFN. Our results suggest that the MFN actually reflects an evaluation of fairness, which at this stage of processing is independent of self-benefit considerations. In this line of thought, results of other studies are interesting, showing that witnessing negative outcomes for other people can also elicit MFN responses, which suggest a possible relationship between the MFN and empathy (Thoma and Bellebaum, 2012). It remains a subject for future research to examine the differential MFN depending on outcomes concerning oneself vs. others.

Compared to other reports of this ERP (e.g., Gehring and Willoughby, 2002), the MFN effects found in our study occur relatively late. This might be due to the more complex task design in the current study involving interpersonal information, and/or to the small size of the stimuli employed and lighting conditions (e.g., Wijers et al., 1997). Other studies employing similar tasks (e.g., Wu et al., 2012) tend to also report rather late MFN effects, though not as late as our findings suggest. Yet, it is not fully clear why we find such late MFN effects and future replications are needed to better understand the timing of the component.

At a later stage of processing, the P300 revealed a significant interaction between the context and the fairness of the offer. In the certain condition, where the allocation of the split was disclosed, unfair offers, which are characterized by a greater difference in outcome between both interaction partners, seemed to receive particular attention and elicited a larger P300 than fair offers. The effect was only present in the certain context, which allowed for an outcome comparison with the interaction partner. This suggests that an earlier focus on an impersonal equity rule as reflected in the MFN shifted to social comparative considerations concerning personal interests in the P300. Knowledge about the personal allocation of the split was crucial for enhanced processing of unfair offers, which indicates a role for the P300 in evaluating stimuli relevant to personal interests in a socially comparative setting. In the uncertain condition the effect of offer fairness seemed to be reversed, showing marginally significant higher P300 amplitudes for fair offers. Here, fair offers might have enhanced motivational significance (Yeung and Sanfey, 2004) for the proposer, because even without revealing allocations, these offers do not hold the risk of inequitable treatment. P300 amplitudes were also higher for offers in which the proposer received the higher part of the split (advantageous offers). Evidence showing that the P300 encodes the valence of a stimulus, i.e., win or loss (e.g., Hajcak et al., 2007) suggests that advantageousness in this study could be understood as a social comparative account of stimuli valence. Here, advantageous offers represent an economical benefit in comparison to the gain of the interaction partner. Our P300 results therefore suggest an involvement of the P300 in higher order social cognitive processes (Wu et al., 2011b), in particular social comparison.

It is striking, however, that the effects of fairness and advantageousness in the certain condition were both opposite to results from Wu et al. (2011a, 2012). They found higher P300 amplitudes for equal as compared to unequal splits, as well as for disadvantageous as compared to advantageous unequal splits. The authors interpreted their results by suggesting that participants devoted more attention to disadvantageous offers, because participants might have had to reflect more upon whether to accept or reject such an offer. However, their and our results consistently showed no influence of social information or social distance on the P300 amplitude. This suggests that the early impact of sympathy towards interaction partners was later replaced by strategic or social comparative considerations about outcomes. The separate analysis of early and late time windows of the P300 indicated that the advantageousness of the offer is processed rather late even within the component. This might again reflect the complexity of social comparative considerations in decision-making situations, which leads to a relatively late timing.

Both the MFN and P300 potentials showed a main effect of context, suggesting that throughout all stages of processing, events in the certain context engaged more cognitive resources. Increased amplitudes in the certain condition suggest that attention was drawn to the stimuli that could provide information and that social information was processed in both certain and uncertain conditions, but informed behavioral choices mainly in the uncertain context. Thought to reflect feedback evaluation, the MFN in our experiment was less pronounced in the uncertain condition, where less feedback about outcomes was provided.

There are, however, some limitations in the present study that warrant further investigations. In the first place, the modifications that were made to the UG limit the extension of our findings to the classic game. At the behavioral level, however, our study replicates common findings in the UG. In addition, social information about the partners has also been found to modulate choices in a classic UG setting (Gaertig et al., 2012). Therefore, it would be desirable to replicate our main ERP findings employing a classic UG. Future studies should also improve the ecological validity involving more naturalistic settings and providing a more inherently social environment for the study of interpersonal decision-making, since the current experimental setting has the drawback of artificiality. Previous studies (e.g., Pillutla and Murnighan, 1996; Sanfey et al., 2003) emphasized the role of emotions in decision-making. Regarding the present experiment, it is possible that personal information about interaction partner elicited positive and negative emotions, which might have influenced offer perception and behavioral choices. However, this is only speculative, and further neuroimaging research is needed to determine the active brain regions and associated cognitive processes to integrate the present findings into a bigger picture. Future research could be aimed at exploring how personal interest considerations interact with social information to bias outcome evaluations. Also here, it would be of particular interest to identify the brain regions reflecting such bias. Further, including a neutral condition could help to better understand the impact of positive and negative social information on interpersonal choices.

In summary, our findings underscore the role of social information in interpersonal decision-making situations and show that it affects information processing at several neural stages. As shown, positive and negative character traits of the interaction partner change the appraisal of the offer (MFN potential), in an additive fashion to fairness considerations. Interestingly, whereas such influence is no longer present when economic offers favor personal interests of the responder, personal descriptions modulate valence evaluations of the offers not beneficial to the participant. At a later stage of processing (P300 potential), social comparison mechanisms and personal benefits considerations seem to outweigh influences of the personal characteristics of the interaction partners. Our findings provide new evidence on the importance of social information on the appraisal of outcomes in interpersonal decision situations and its conjoint effects with fairness and personal benefit considerations.

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiment: María Ruz and Anna Moser. Performed the experiment: Anna Moser and Celia Gaertig. Analyzed the data: Anna Moser. Wrote the paper: Anna Moser and María Ruz.

## ACKNOWLEDGMENTS

We thank Alexander Moscicki for helpful comments on earlier versions of this manuscript. Financial support to this research came from the Spanish Ministry of Science and Innovation through a "Ramón y Cajal" research fellowship (RYC-2008-03008) and grant PSI2010-16421 to María Ruz, and also from the European Commission through a "Leonardo da Vinci" fellowship (DE/10/LLP-LdV/PLM/282611) to Anna Moser.

## REFERENCES

Boksem, M., and De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Soc. Neurosci.* 5, 118–128. doi: 10.1080/17470910903202666

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.

Campanhã, C., Minati, L., Fregni, F., and Boggio, P. S. (2011). Responding to unfair offers made by a friend: neuroelectrical activity changes in the anterior medial prefrontal cortex. *J. Neurosci.* 31, 15569–15574. doi: 10.1523/jneurosci.1253-11.2011

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. (1998). Anterior cingulate cortex, error detection and the online monitoring of performance. *Science* 280, 747–749. doi: 10.1126/science.280.5364.747

Donchin, E., and Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357–374. doi: 10.1017/s0140525x00058027

Fehr, E., and Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Q. J. Econ.* 114, 817–868. doi: 10.1162/003355399556151

Gaertig, C., Moser, A., Alguacil, S., and Ruz, M. (2012). Social information and economic decision-making in the ultimatum game. *Front. Neurosci.* 6:103. doi: 10.3389/fnins.2012.00103

Gehring, W. J., and Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295, 2279–2282. doi: 10.1126/science.1066893

Güroğlu, B., van den Bos, W., Rombouts, S. A., and Crone, E. A. (2010). Unfair? It depends: neural correlates of fairness in social context. *Soc. Cogn. Affect. Neurosci.* 5, 414–423. doi: 10.1093/scan/nsq013

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388. doi: 10.1016/0167-2681(82)90011-7

Hajcak, G., Moser, J. S., Holroyd, C. B., and Simons, R. F. (2007). It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology* 44, 905–912. doi: 10.1111/j.1469-8986.2007.00567.x

Hajcak, G., Moser, J. S., Yeung, N., and Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology* 42, 151–160. doi: 10.1111/j.1469-8986.2005.00270.x

Harlé, K. M., and Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion* 7, 876–881. doi: 10.1037/1528-3542.7.4.876

Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M., Holroyd, C. B., et al. (2011). Why humans deviate from rational choice. *Psychophysiology* 48, 507–514. doi: 10.1111/j.1469-8986.2010.01081.x

Holroyd, C. B., and Coles, M. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709. doi: 10.1037/0033-295X.109.4.679

Kucera, H., and Francis, W. N. (1967). *Computational Analysis of Present-Day American English*. Providence: Brown University Press.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Ma, Q., Shen, Q., Xu, Q., Li, D., Shu, L., and Weber, B. (2011). Empathic responses to others' gains and losses: an electrophysiological investigation. *Neuroimage* 54, 2472–2480. doi: 10.1016/j.neuroimage.2010.10.045

Marchetti, A., Castelli, I., Harlé, K. M., and Sanfey, A. G. (2011). Expectations and outcome: the role of proposer features in the Ultimatum Game. *J. Econ. Psychol.* 32, 446–449. doi: 10.1016/j.joep.2011.03.009

Nieuwenhuis, S., Aston-Jones, G., and Cohen, J. D. (2005). Decision making, the P3 and the locus coeruleus-norepinephrine system. *Psychol. Bull.* 131, 510–532. doi: 10.1037/0033-2909.131.4.510

Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187. doi: 10.1016/0013-4694(89)90180-6

Pfabigan, D. M., Alexopoulos, J., Bauer, H., Lamm, C., and Sailer, U. (2011). All about the money? External performance monitoring is affected by monetary, but not by socially conveyed feedback cues in more antisocial individuals. *Front. Hum. Neurosci.* 5:100. doi: 10.3389/fnhum.2011.00100

Pillutla, M. M., and Murnighan, J. K. (1996). Unfairness, anger and spite: emotional rejections of ultimatum offers. *Organ. Behav. Hum. Decis. Process.* 68, 208–224. doi: 10.1006/obhd.1996.0100

Platt, M. L., and Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nat. Neurosci.* 11, 398–403. doi: 10.1038/nn2062

Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., et al. (2008). Mentalizing in economic decision-making. *Behav. Brain Res.* 190, 218–223. doi: 10.1016/j.bbr.2008.03.003

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., and Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* 277, 968–971. doi: 10.1126/science.277.5328.968

Redondo, J., Fraga, I., Padrón, I., and Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behav. Res. Methods* 39, 600–605. doi: 10.3758/bf03193031

Ruz, M., Madrid, E., and Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Soc. Cogn. Affect. Neurosci.* 8, 838–844. doi: 10.1093/scan/nss080

Ruz, M., Moser, A., and Webster, K. (2011). Social expectations bias decision-making in uncertain interpersonal situations. *PLoS One* 6:e15762. doi: 10.1371/journal.pone.0015762

Ruz, M., and Tudela, P. (2011). Emotional conflict in interpersonal interactions. *Neuroimage* 54, 1685–1691. doi: 10.1016/j.neuroimage.2010.08.039

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758. doi: 10.1126/science.1082976

Thoma, P., and Bellebaum, C. (2012). Your error's got me feeling - how empathy relates to the electrophysiological correlates of performance monitoring. *Front. Hum. Neurosci.* 6:135. doi: 10.3389/fnhum.2012.00135

Tucker, D. M., Liotti, M., Potts, G. F., Russell, G. S., and Posner, M. I. (1994). Spatiotemporal analysis of brain electrical fields. *Hum. Brain Mapp.* 1, 134–152. doi: 10.1002/hbm.460010206

Van der Veen, F. M., and Sahibdin, P. P. (2011). Dissociation between medial frontal negativity and cardiac responses in the ultimatum game: effects of offer size and fairness. *Cogn. Affect. Behav. Neurosci.* 11, 516–525. doi: 10.3758/s13415-011-0050-1

Van Noordt, S. J. R., and Segalowitz, S. J. (2012). Performance monitoring and the medial prefrontal cortex: a review of individual differences and context effects as a window on self-regulation. *Front. Hum. Neurosci.* 6:197. doi: 10.3389/fnhum.2012.00197

Wijers, A. A., Lange, J. J., Mulder, G., and Mulder, L. J. (1997). An ERP study of visual spatial attention and letter target detection for isoluminant and nonisoluminant stimuli. *Psychophysiology* 34, 553–565. doi: 10.1111/j.1469-8986.1997.tb01742.x

Wu, Y., Hu, J., van Dijk, E., Leliveld, M. C., and Zhou, X. (2012). Brain activity in fairness consideration during asset distribution: does the initial ownership play a role? *PLoS One* 7:e39627. doi: 10.1371/journal.pone.0039627

Wu, Y., Leliveld, M. C., and Zhou, X. (2011a). Social distance modulates recipient's fairness consideration in the dictator game: an ERP study. *Biol. Psychol.* 88, 253–262. doi: 10.1016/j.biopsycho.2011.08.009

Wu, Y., Zhou, Y., van Dijk, E., Leliveld, M. C., and Zhou, X. (2011b). Social comparison affects brain responses to fairness in asset division: an ERP study with the Ultimatum game. *Front. Hum. Neurosci.* 5:131. doi: 10.3389/fnhum.2011.00131

Yeung, N., and Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *J. Neurosci.* 24, 6258–6264. doi: 10.1523/jneurosci.4537-03.2004

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 06 November 2013; accepted 14 January 2014; published online: 06 February 2014.*

*Citation: Moser A, Gaertig C and Ruz M (2014) Social information and personal interests modulate neural activity during economic decision-making. Front. Hum. Neurosci. **8**:31. doi: 10.3389/fnhum.2014.00031*

*This article was submitted to the journal Frontiers in Human Neuroscience.*

frontiers in
HUMAN NEUROSCIENCE

GENERAL COMMENTARY
published: 13 February 2014
doi: 10.3389/fnhum.2014.00050

# Erratum: Valence of emotions and moral decision-making: increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults

***Martina Carmona-Perera*[1]*, Celia Martí-García*[2]*, Miguel Pérez-García*[1,3,4] *and Antonio Verdejo-García*[1,5,6,7]****

*[1] Department of Personality, Assessment and Psychological Treatment, University of Granada, Granada, Spain*
*[2] School of Health Sciences, University of Granada, Granada, Spain*
*[3] Centro de Investigación Mente, Cerebro y Comportamiento, University of Granada, Granada, Spain*
*[4] Centro de Investigación Biomédica en Red de Salud Mental, University of Granada, Granada, Spain*
*[5] Institute of Neuroscience F. Oloriz, University of Granada, Armilla, Spain*
*[6] Red de Trastornos Adictivos, Instituto Carlos III, University of Granada, Spain*
*[7] School of Psychology and Psychiatry, Monash University, Victoria, Australia*
**Correspondence: averdejo@ugr.es*

***Edited by:***
*Maria Ruz, Universidad de Granada, Spain*

***Reviewed by:***
*Agustin Ibanez, Institute of Cognitive Neurology, Argentina*

**Keywords: arousal, moral emotions, moral-decision making, utilitarian choices, valence**

**An erratum on**

**Valence of emotions and moral decision-making: increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults**
*by Carmona-Perera, M., Martí-García, C., Pérez-García, M., and Verdejo-García, A. (2013). Front. Hum. Neurosci. 7:626. doi: 10.3389/fnhum.2013.00626*

We have noticed that we made an incorrect interpretation of the correlation between valence ratings to unpleasant pictures and proportion of utilitarian choices (Figure 1, Panel A, left side). The Figure is correct but our interpretation was wrong. We originally interpreted that decreased unpleasantness was associated with higher utilitarism, but the data points to the opposite direction: increased unpleasantness is associated with higher utilitarism. This interpretation error impacts the third line of the published title ("*decreased* unpleasantness to unpleasant images" should be "*increased* unpleasantness to unpleasant images"), line 11 of the abstract ("*less* unpleasantness to negative stimuli" should be "*more* unpleasantness to negative stimuli"), and one line in the Results subsection "ASSOCIATION BETWEEN SUBJECTIVE REACTIVITY TO EMOTIONAL STIMULI AND UTILITARIAN CHOICES AND DIFFICULTY RATINGS TO DILEMMAS": the sentence "experiencing *less* unpleasantness in response to unpleasant images (both moral and non-moral), and more pleasantness in response to pleasant images were associated with more utilitarian choices (see Figure 1A)" should be "experiencing *more* unpleasantness in response to unpleasant images (both moral and non-moral), and more pleasantness in response to pleasant images were associated with more utilitarian choices (see Figure 1A)." Moreover, in the first paragraph of the Discussion, finding (2) "*lower* experience of unpleasantness" should be "*higher* experience of unpleasantness." Therefore, while we originally argued that findings "support the notion that *diminished* experience of unpleasantness favors utilitarian choice patterns," the correct interpretation is that "*increased* experience of unpleasantness favors utilitarian choice patterns."

During replication of the whole set of statistical analyses in SPSS v. 20, we have noticed additional errors in the Results section. The $F$-value of the within-measures comparison of valence ratings should be 1143.97 instead of 143.97. Several $r$-values from correlation analyses were incorrect and the Bonferroni correction to $p$-values was incorrectly applied. Therefore, here, we present the correct $r$-values and exact $p$-values of these analyses, and we clarify the Bonferroni-corrected alpha level to interpret results. The exact $p$-value of the correlation between moral choices and valence ratings to unpleasant moral images should be 0.004 instead of the published value (0.016). The exact $p$-value of the correlation between moral choices and valence ratings to unpleasant non-moral images should be 0.011 instead of the published value (0.043). The $r$ and exact $p$-values of the correlation between moral choices and valence ratings to pleasant images should be $r = 0.21$ and $p = 0.040$ instead of the published values ($r = 0.26$, $p = 0.047$). The $r$ and exact $p$-values of the correlation between moral choices and arousal ratings to unpleasant moral images should be $r = 0.31$ and $p = 0.002$ instead of the published values ($r = 0.34$, $p = 0.004$). The $r$ and exact $p$-values of the correlation between difficulty ratings and dominance ratings to unpleasant non-moral images should be $r = -0.24$ and $p = 0.020$ instead of the published values ($r = -0.26$, $p = 0.043$). The $r$ and exact $p$-values of the correlation between difficulty ratings and dominance ratings to unpleasant moral images should be $r = -0.31$ and $p = 0.002$ instead of the published values ($r = -0.29$, $p = 0.016$). When applying a Bonferroni correction taking into account 2 types of dilemmas and 4 image conditions, the alpha value for all the correlations reported must be established at 0.006. Three previously unreported correlations concerning impersonal moral dilemmas were detected utilizing these parameters: a negative correlation between unpleasantness ratings to unpleasant pictures and utilitarian choices ($r = -0.33$, $p = 0.001$), a negative correlation between unpleasantness ratings to moral-laden pictures and utilitarian choices ($r = -0.37$, $p < 0.001$) and a positive correlation between arousal ratings to moral-laden images and utilitarian choices ($r = 0.32$, $p = 0.002$).

To recap, in the published article we reported four main findings, namely: (1) individual differences in self-reported emotional experience correlate with decision-making in moral scenarios, but not in non-moral scenarios; (2) *lower* experience of unpleasantness to both moral and non-moral unpleasant images and higher experience of pleasantness to pleasant images are associated with utilitarian choice patterns; (3) higher experience of arousal (specifically in response to moral laden images) are associated with more utilitarian choices, and (4) lower dominance over emotions is significantly associated with higher perceived difficulty to make decisions in moral scenarios. In this Erratum, we note that finding (2) should be reappraised as follows: *higher* experience of unpleasantness, *mainly to moral unpleasant images*, is associated with utilitarian choice patterns. This association was observed for the combined measure of utilitarian choices to moral dilemmas, and for utilitarian choices to impersonal dilemmas. Moreover, the correlation between pleasantness ratings and utilitarian choices would not survive a Bonferroni correction. Our revamped conclusion about the direction of correlations is that increased negative valence, increased arousal and lower dominance to emotional stimuli correlate with utilitarian choices. This interpretation cannot be easily framed in the dual-process theory (paragraph 3 of the Discussion) which has been useful to explain utilitarian choices in clinical populations, but actually fits better with the notion that certain types of negative emotions are linked to utilitarian choices in healthy populations.

## ACKNOWLEDGMENTS

We would like to thank Dr. Indrajeet Patil for making us aware of the interpretation error concerning Figure 1.

*Received: 04 December 2013; accepted: 22 January 2014;*

*Citation: Carmona-Perera M, Martí-García C, Pérez-García M and Verdejo-García A (2014) Erratum: Valence of emotions and moral decision-making: increased pleasantness to pleasant images and decreased unpleasantness to unpleasant images are associated with utilitarian choices in healthy adults. Front. Hum. Neurosci.* ***8****:50. doi: 10.3389/fnhum.2014.00050*
*This article was submitted to the journal Frontiers in Human Neuroscience.*