OXFORD

Social Cognitive and Affective Neuroscience, 2019, 519–527

doi: 10.1093/scan/nsz028
Advance Access Publication Date: 17 April 2019
Original article

Contextual knowledge provided by a movie biases implicit perception of the protagonist

Mamdooh Afdile,[1,5] Iiro P. Jääskeläinen,[1,3] Enrico Glerean,[1,2,4,7], Dmitry Smirnov,[1] Jussi Alho,[1,6] Anna Äimälä,[1] and Mikko Sams[1,2,3]

[1]Department of Neuroscience and Biomedical Engineering, Aalto University, FI-00076, Espoo, Finland, [2]Department of Computer Science, School of Science, Aalto University, FI-00076, Espoo, Finland, [3]Advanced Magnetic Imaging Centre, Aalto NeuroImaging, Aalto University, FI-00076, Espoo, Finland, [4]Turku PET Centre, University of Turku, 20520 Turku, Finland, [5]Department of Media, School of Arts Design and Architecture, Aalto University, 11000, Espoo, Finland, [6]Department of Psychology and Logopedics, Medicum, Faculty of Medicine, Helsinki University, 00014, Helsinki, Finland, and [7]Helsinki Institute for Information Technology, Aalto University, FI-00076, Espoo, Finland

Correspondence should be addressed to Mamdooh Afdile, Department of Neuroscience and Biomedical Engineering, Aalto University, P.O. Box 12200, FI-00076, Espoo, Finland. E-mail: mamdooh.afadila@aalto.fi.

Abstract

We are constantly categorizing other people as belonging to our in-group ('one of us') or out-group ('one of them'). Such grouping occurs fast and automatically and can be based on others' visible characteristics such as skin color or clothing style. Here we studied neural underpinnings of implicit social grouping not often visible on the face, male sexual orientation. A total of 14 homosexuals and 15 heterosexual males were scanned in functional magnetic resonance imaging while watching a movie about a homosexual man, whose face was also presented subliminally before (subjects did not know about the character's sexual orientation) and after the movie. We discovered significantly stronger activation to the man's face after seeing the movie in homosexual but not heterosexual subjects in medial prefrontal cortex, frontal pole, anterior cingulate cortex, right temporal parietal junction and bilateral superior frontal gyrus. In previous research, these brain areas have been connected to social perception, self-referential thinking, empathy, theory of mind and in-group perception. In line with previous studies showing biased perception of in-/out-group faces to be context dependent, our novel approach further demonstrates how complex contextual knowledge gained under naturalistic viewing can bias implicit social perception.

Key words: implicit bias; face; movie character; in-group; out-group

Introduction

Picture yourself walking in the city center and accidentally seeing a male person who was just yesterday interviewed on TV about his daily experiences as a homosexual person. Would you go to him to express your appreciation of his openness, or would you rather look away to avoid any possibility of interaction? Would your behavior depend on your own sexual orientation?

We tend to join, trust and like group members with whom we share similarities (Yuki *et al.*, 2005). Although the similarity might be only illusory and temporary, it does bias our social behavior (Hornsey, 2008). Favoring people similar to us is probably rooted in biological survival mechanisms (Allport, 1954). We are better in predicting the behavior of in-group members (Bruner, 1957), which facilitates collaboration. However, in-group favoritism is often accompanied with negative attitude toward

Received: 11 April 2018; **Revised:** 4 April 2019; **Accepted:** 9 April 2019

out-group members. People are also less empathic to the pain and suffering of out-group individuals, which can support intergroup hostility (Xu *et al.*, 2009; Avenanti *et al.*, 2010).

Intergroup processes are based on a multitude of psychological mechanisms with complex neural basis (for a review, see Cikara and Van Bavel, 2014). Investigating why we favor one person over another faces many challenges, since critical neurocognitive mechanisms involved in this behavior are to a large extent automatic and implicit. Furthermore, the parallel implicit and explicit processing can lead to inconsistent findings when different paradigms and stimuli are used (for a review, see Amodio, 2014). For example, consequences of implicit prejudices can be inhibited by conscious suspension (Cunningham *et al.*, 2004; Payne, 2005) and thus tolerance at conscious explicit level does not necessarily mean that we do not have implicit prejudices (Olsson, 2005). Furthermore, low social status of oppressed individuals can cause favoring of high-status groups leading to out-group favoritism and bias against one's own group (Allport, 1954; Jost and Hunyady, 2003; Jost *et al.*, 2004).

Pictures of faces have been extensively used as stimuli in investigating implicit neural processes during social perception. We continuously evaluate our social environment based on what kind of faces we see (for a review, see Todorov *et al.*, 2015). It has been shown that a brief 32–100 ms exposure to a face is sufficient for formation of a social judgment (Bar *et al.*, 2006; Ballew and Todorov, 2007; Porter *et al.*, 2008; Borkenau *et al.*, 2009; Rule and Ambady, 2009; Todorov *et al.*, 2009, 2010). Face stimuli have been used to reveal intergroup bias among white as well as black participants (Hart *et al.*, 2000; Phelps *et al.*, 2000; Cunningham *et al.*, 2004; Avenanti *et al.*, 2010; Kubota *et al.*, 2012). However, most social groups, such as groups based on sexual orientation, religion or political views, are often characterized by factors other than salient facial features (Tajfel and Turner, 1979).

Recent studies have shown that face perception is dependent on contextual cues presented with the face. For example, verbal information given about a face (Schwarz *et al.*, 2013), visual background of the face (Righart and De Gelder, 2008) and seeing the face alone or with other faces (Mumenthaler and Sander, 2012) are all factors that can affect how we perceive a face. Such context dependency calls for stimuli that can better simulate real-life interactions involving faces (for a review, see Wieser and Brosch, 2012). Indeed, it has been shown that we are better at recognizing faces and individuals when presented with moving images approaching real-life perception in comparison to static images (O'Toole *et al.*, 2011).

A growing body of literature has demonstrated that movies can successfully be used as stimuli to study real-life-like social perception (Hasson, 2004; Malinen *et al.*, 2007; Jääskeläinen *et al.*, 2008; Wilson *et al.*, 2008; Lahnakoski *et al.*, 2014; Lähteenmäki *et al.*, 2015; Saarimäki *et al.*, 2016). Movie narratives can temporarily transport the viewers to the world of the protagonist (Hall and Bracken, 2011) and introduce them to complex emotional and contextual knowledge of the protagonist in a short time. Furthermore, watching the same movie synchronizes perception and underlying neural activity across viewers (Hasson, 2004; Hasson *et al.*, 2008; Jääskeläinen *et al.*, 2008; Bacha-Trams *et al.*, 2017) and makes them similar in interpretation of the protagonists and their behavior. Therefore, having subjects to view the same movie can be used to decrease variance in their perception, for example, in studying prejudices (Corrigan *et al.*, 2001).

Here we used a movie to reveal implicit brain activity, measured with functional magnetic resonance imaging (fMRI), triggered by the face of a movie character. Brains of homo- and heterosexual males were scanned with functional MRI while watching a movie about a homosexual man, whose face was also presented subliminally before and after the movie. Based on recent studies having shown negative attitude of heterosexuals toward homosexuals (Jellison *et al.*, 2004; Cullen and Barnes-Holmes, 2008) and in-group favoritism in both heterosexual and homosexual individuals (Banse *et al.*, 2001; Jellison *et al.*, 2004), we hypothesized that knowledge about the character's sexual orientation will bias the implicit neural response to his face in homosexual *vs* heterosexual subjects. More specifically, we expected homosexuals to show increased activity in regions associated with empathy and in-group perception such as temporal parietal junction (TPJ) (Saxe, 2006; Decety and Lamm, 2007; Decety *et al.*, 2012) and medial prefrontal cortex (mPFC) (Vogeley and Fink, 2003; Amodio and Frith, 2006; Mitchell *et al.*, 2006; Krueger *et al.*, 2009; Van Overwalle, 2009; Morrison *et al.*, 2012). On the other hand, based on behavioral study showing prejudice of heterosexual subjects toward homosexuals to be associated with feelings of disgust and pity (Cottrell and Neuberg, 2005), we expect the heterosexual subjects to have increased activity in the insula. Although insula activation has been reported in social pain and empathy and is related to various types of emotional processing, it is also sensitive to disgust (for a review, see Vicario *et al.*, 2017). More specifically, we expect an activation in the insular frontal operculum that is specifically associated with the processing of disgust (Calder *et al.*, 2000; Adolphs *et al.*, 2003; Carr *et al.*, 2003; Wicker *et al.*, 2003; Jabbi *et al.*, 2007, 2008). As another hypothesis, though it is not an indicator of implicit bias, we expected our homosexual subjects to self-report increased identification and feeling of closeness with the movie character after the movie reveals he is homosexual. Conversely, we expected heterosexual subjects to show decreased identification with the movie character after the movie reveals him as homosexual (Petta and Walker, 1992; Doosje *et al.*, 1995; Smith and Tyler, 1997; Spears *et al.*, 1997).

Material and methods

Participants

A total of 29 right-handed males (15 heterosexuals, mean age, 26; 14 homosexuals, mean age, 28; age range of participants, 20–46), volunteered for the experiment. All participants were Finnish speakers. The subjects reported neither history of neurological or psychiatric diseases nor medications affecting the central nervous system. The fMRI scanning was conducted between 10 am and 2 pm. The study protocol was approved by the Research Ethics Committee of the Aalto University, and each subject signed an informed consent form prior to participation.

Experimental design

During the fMRI scanning, the participants watched a 20 min movie consisting of scenes from the movie 'Priest' (Directed by Antonia Bird, 1994). The movie tells the story of a man struggling between his devotion to serve as a Roman Catholic Christian priest and his desire to be loved by another man, something that is forbidden for a Catholic priest. The movie was edited so that initially the priest appears to be heterosexual and after exactly 10 min the events in the movie reveal that he is actually homosexual. The structure of the movie was tested before the experiment on a pilot audience (9 male subjects, age 21–27) who unanimously indicated that the homosexuality of the character was identified only after he enters the gay bar (at 10 min).

After scanning, the subjects were asked to report how close they felt to the character in the first and in the second part of the film. Participants were not informed of the movie's content prior to the experiment. A professional film-maker (the first author of this paper) edited the original movie to assure that the stimulus has a story that flows naturally with an engaging plot line. The subjects reported that they had not seen the original film before and were not familiar in the type of experiment they participated in.

To test if knowledge of the sexual orientation of the priest modulates implicit perception of his face, we measured brain hemodynamic responses to 40 ms images of the character's face before and after viewing the movie. During the fMRI the subjects were asked to fixate at a mark in the center of the screen and watch a 4 min stimulus sequence that contained white noise and images of the character's face (see details of the stimuli below). The participants were not informed of the nature of the subliminal stimuli and were instructed with the following slide (written in Finnish): 'You will see a calibration clip. This clip is meant for calibrating the MRI scanner for your responses. The clip is only four minutes long and will look like white noise on a TV screen. Please keep your eyes fixated at the mark in the center of the screen until notified otherwise.'

The 4 min subliminal stimuli contained 16 blocks of 15 s each. Blocks were of two types: 'face' and 'objects'. Only blocks with the character's face were analyzed; the objects block served as a wash-out period. In pilot experiments, we used face blocks containing only faces. However, 4 out of 10 pilot participants could see that there were images of faces embedded in the stimulus. This could probably be due to higher sensitivity to face perception in some participants who therefore required further masking (Lähteenmäki *et al.*, 2015). We therefore decided to additionally mask the faces by inserting 40 ms images of objects between the faces that showed to be successful in our piloting. Object blocks consisted only of objects (Figure 1).

The face block consisted of 10 40 ms black-and-white images of faces presented at 1460 ms inter-stimulus interval filled with white noise. Each face image was followed by an object image. The object images were randomly selected from a set of 197 images at http://natural-scenes.cps.utexas.edu/db.shtml; (Geisler and Perry, 2011). A 15 s object block followed each face block. The two block types were presented alternatingly, starting with the object block.

Experimental procedure

The stimuli were delivered during the fMRI scanning using the Presentation software (Neurobehavioral Systems Inc., Albany, CA, USA). The stimuli were back-projected on a semitransparent screen using a Panasonic PT-DZ110X projector (Panasonic Corporation, Osaka, Japan), and from there via a mirror to the subject. Auditory stimulation was delivered through Sensimetrics S14 insert earphones (Sensimetrics Corporation, Malden, MA, USA) via plastic tubes through porous EAR-tip (Etymotic Research, ER3, IL, USA) earplugs. Sound intensity was adjusted individually to be comfortable but loud enough to be heard over the scanner noise.

After the fMRI scan, we asked the participants to indicate how close they felt to the character in the first part of the film (they did not know the character is homosexual) as well as in the second part of the film (they knew the character is homosexual). This measurement was done to assure that the change in identification with the character is a factor of knowing the character's sexual orientation (*vs* before) and not due to other possible factors that were present all along in the movie such as the character's age, nationality or occupation as a Catholic priest. The measurement of the identification with the character was done on a computer using an online form with a sliding continues measurement from 0 (no identification) to 1 (100% total identification). We subtracted the score of the second part from the first part of the movie to quantify the change in identification. For example, if a subject's identification level with the character in the first part of the movie is at 0.47 points and in the second part at 0.63 points, then the identification increased by 0.16 points. For this measurement two subjects were excluded from analysis for not providing their results for this section. Identification scores could have been obtained right after each part of the film. However, we wanted to exclude the possibility that such an inquiry would have influenced the viewing experience of the second part of the film. In our previous work, we have observed that this type of retrospective self-reporting can accurately capture at least experienced humorousness during watching of comedy clips, probably since re-seeing the clip serves as a memory cue (Jääskeläinen *et al.*, 2016). Furthermore, we conducted an additional behavioral measurement to assess the participants' unconscious bias to hetero- or homosexuality using the Implicit Association Task (IAT) using presentation software (details below). The IAT measured differences in reaction time (RT) in associating positive and negative words to images representing homosexual or heterosexual sexual orientation, as measurement of implicit bias (Greenwald *et al.*, 1998). In our experiment, IAT scores were collected to have an estimate of implicit bias of the subjects to homo- and heterosexuality.

fMRI data acquisition and preprocessing

MR imaging was performed at the Advanced Magnetic Imaging Centre at Aalto University. Images were acquired with a 3T Siemens MAGNETOM Skyra (Siemens Healthcare, Erlangen, Germany), using a standard 20-channel receiving head–neck coil. Anatomical images were acquired using a T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE) pulse sequence (TR 2530 ms, TE 3.3 ms, TI 1100 ms, flip angle 7°, 256×256 matrix, 176 sagittal slices, 1 mm^3 resolution). Whole-brain functional data were acquired with T2*-weighted echo-planar imaging (EPI) sequence sensitive to the BOLD contrast. Imaging parameters for the functional images were TR 1700 ms, TE 24 ms, flip angle 70°, FOV 217.6 mm, 64×64 matrix, 4.0 mm slice thickness with 1 mm gap between slices, 29 oblique slices acquired in interleaved ascending order covering the whole brain, resolution $3.4 \times 3.4 \times 4.0$ mm.

Standard fMRI preprocessing steps were applied using the FSL software (www.fmrib.ox.ac.uk, version 5.0.9) and custom MATLAB code (BRAMILA pipeline v2.0, available at https://version.aalto.fi/gitlab/BML/bramila). After slice timing correction, the functional images were realigned to the middle scan by rigid body transformations with MCFLIRT to correct for subject motion. Non-brain matter from functional and anatomical images was removed using Brain Extraction Tool (Smith, 2002). Functional images were registered to the MNI152 standard space template (Montreal Neurological Institute) with 2 mm resolution. The transformation parameters were acquired by first calculating transformations from structural to standard space (12 degrees of freedom) and from functional to structural space (9 degrees of freedom), and then concatenating these parameters. Next, these transformation parameters were used

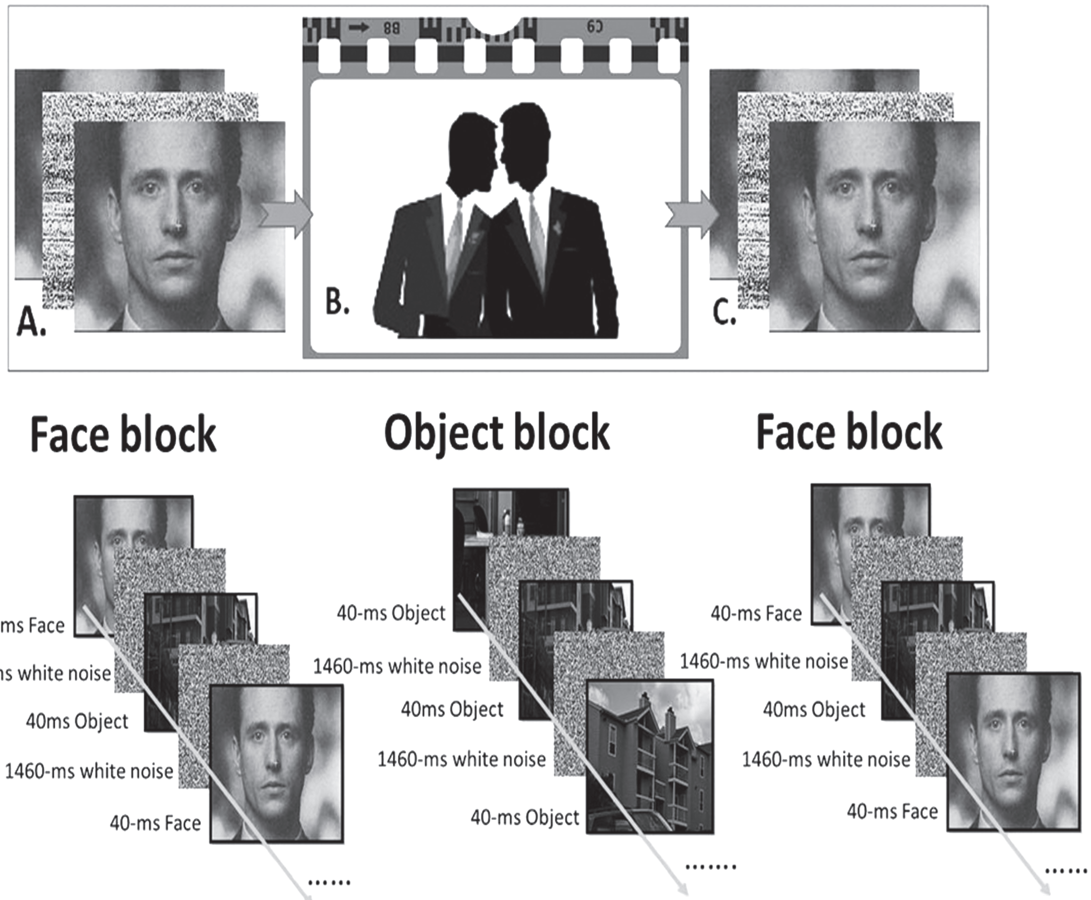

Fig. 1. Illustration of the experimental paradigm. 'Upper panel A and C'. Pictures of the face of the priest presented subliminally. 'Upper panel B'. A 20 min movie telling a story of a homosexual priest. 'Lower panel': first six pictures of face and object blocks.

to co-register functional datasets to the standard space. Both registration steps were performed using FLIRT (Jenkinson, 2002). To remove scanner drift, a 240 s long Savitzky–Golay filter (Çukur *et al.*, 2013) was applied. To control for motion and physiological artefacts, EPI time series were cleaned with linear regression using 24 motion-related signals, signal from deep white matter, ventricles and cerebral spinal fluid locations as described in Power *et al.* (2014). Additional spatial smoothing step with a Gaussian kernel of full width at half maximum (FWHM) 6 mm was also applied. The neural response to the face of the character was computed using a general linear model (GLM) implemented in FSL with film_gls with default parameters (i.e. prewhitening of time series).

GLM was run separately for the subliminal stimuli before and after the movie; the model contained two regressors for the face block (presentation of text and presentation of subliminal faces block and their corresponding temporal derivate) as well as two more regressors for the object block (presentation of text and presentation of subliminal object block and their corresponding temporal derivate). Regressors were convolved with the canonical hemodynamic response function. To analyze the differences between the homosexual and heterosexual group we run a two-way mixed-effect analysis of variance (ANOVA), which compares the differences in BOLD responses between the face after the film given the response before the film as a baseline. As a control analysis and to test if the results are unique for the face block we also re-ran the analysis using the object block. Finally, to correct for the multiple comparisons, we used a permutation-based non-parametric approach as implemented by FSL randomise with cluster-forming threshold at T-value = 3 and 5000 permutations. Visualization threshold were set at corrected $P < 0.05$. Significance threshold were set at cluster corrected $P < 0.05$.

The IAT

IAT was used to assess the subjects' implicit bias toward hetero- and homosexuals, measured via changes in RT to the different IAT conditions. During the IAT, the subjects were presented with images that represent heterosexual/homosexual orientation in conjunction with positive/negative words. The IAT included two random conditions (associate positive words with homosexuality and negative words with heterosexuality and for the other condition the other way around). Differences in RT caused by a certain association condition were interpreted as an implicit bias. Positive words in IAT were the following: joyful, beautiful, marvelous, wonderful, pleasure, glorious, lovely and superb. The negative words were the following: agony, terrible, horrible, humiliate, nasty, painful, awful and tragic. The heterosexual/homosexual orientation images were downloaded from http://www.millisecond.com/download/library/IAT, which has been previously used to address similar topics (Biele and Grabowska, 2006; Rowatt *et al.*, 2006). Within each condition the

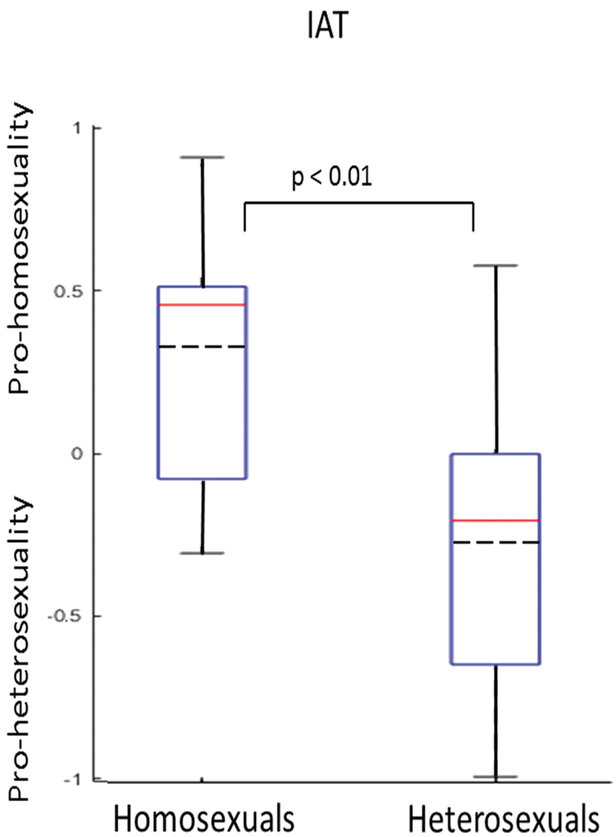

Fig. 2. Mean IAT scores (dotted lines, homosexuals = 0.3, heterosexuals = −0.26). Y-axis indicates the level of pro-homo/heterosexuality bias from 1 (very strong) to −1 (very weak). Red lines indicate medians (homosexuals = 0.45, heterosexuals = −0.20). Whiskers indicate the range of the scores.

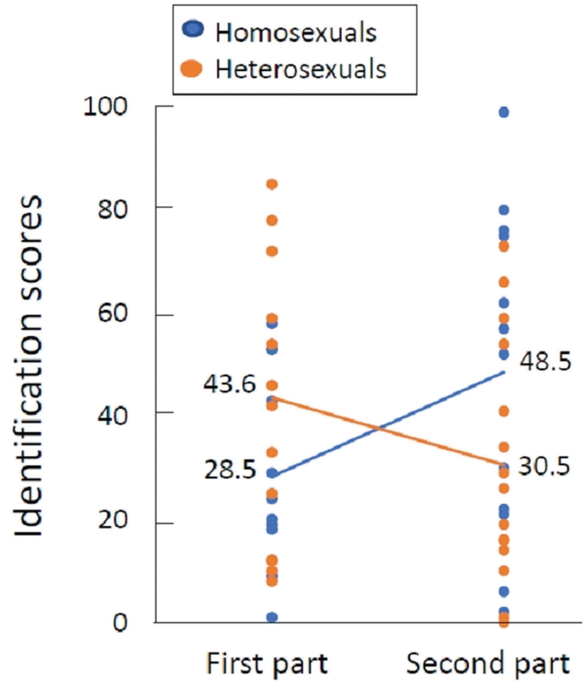

Fig. 3. Identification scores of all subjects in the first and second part of the film. Lines connect the mean values.

order of words and images were random. The subjects were instructed to respond as fast and as accurately as possible. The IAT was preceded by a 2 min training phase that assured that the subjects fully understood the instruction and the task.

Results

Figure 2 depicts the normalized IAT scores of homo- and heterosexual subjects. Mean scores were significantly different (homosexuals = 0.3, heterosexuals = −0.26, $t = 3.72$, $P < 0.01$). Furthermore, IAT scores of each group were significantly different from zero, (heterosexual $P = 0.043$ and homosexual $P = 0.0059$). Figure 3 depicts the identification scores of the hetero- and homosexual subjects from the first and second part of the movie, a line connects the mean scores. The mean score in the heterosexual subjects is lower in the second part of the movie but higher in the homosexual subjects. As depicted in Figure 3, the pair-wise comparisons of homosexual and heterosexual identification (IDN) scores (i) at first part of the film, and (ii) at second part of the film showed significant interaction between the groups and film parts ($F = 5.72$, $P = 0.02$). Changes in IDN and IAT scores showed a significant positive correlation ($r = 0.41$, $P = 0.033$).

A mixed-effect ANOVA analysis revealed significantly larger BOLD signals in the homosexual *vs* heterosexual subjects in response to the face after the movie viewing in mPFC, bilateral frontal pole (FP), anterior cingulate cortex (ACC), right temporal parietal junction and bilateral superior frontal gyrus (sFG), as depicted in Figure 4. Although the peak activation coordinates in mPFC is slightly dorsal, the activation in the present study did spread into the ventral medial regions. Furthermore, our control analysis using the object block (instead of the face block) failed to show any significant effects. We failed to see any significant activation in the reverse contrast heterosexuals *vs* homosexuals. The activation in insula did not survive cluster correction. Unthresholded result map can be found in NeuroVault at https://neurovault.org/collections/ELLZKWSV/images/61868. Other unthresholded maps of additional analysis that did not survive correction can be found at https://neurovault.org/collections/ELLZKWSV (i) a two-sample *t*-test (homosexual *vs* heterosexual) in response to the subliminal face presentation before watching the movie, (ii) a two-sample *t*-test (homosexual *vs* heterosexual) in response to the subliminal face presentation of the character after watching the movie and (iii) a two-way ANOVA using change in IDN scores as regressor on the subliminal face perception (iv) areas showing IAT-dependent differences on subliminal face perception (two-way ANOVA at different significant levels).

Discussion

Aiming to bring our experiment close to real social life, we used a movie to induce implicit social bias. As expected, by revealing the priest's homosexuality the film modulated how strongly the subjects identified with him. Furthermore, this knowledge modulated the implicit neural response to his face. Confirming our hypothesis, homosexual *vs* heterosexual subjects showed significantly stronger activation in response to the subliminally presented face of the homosexual character in mPFC and TPJ along with ACC, FP and sFG. This further demonstrates the important role of contextual knowledge on automatic and fast initial social evaluation of faces as was previously suggested (Freeman *et al.*, 2015; Stolier and Freeman, 2016).

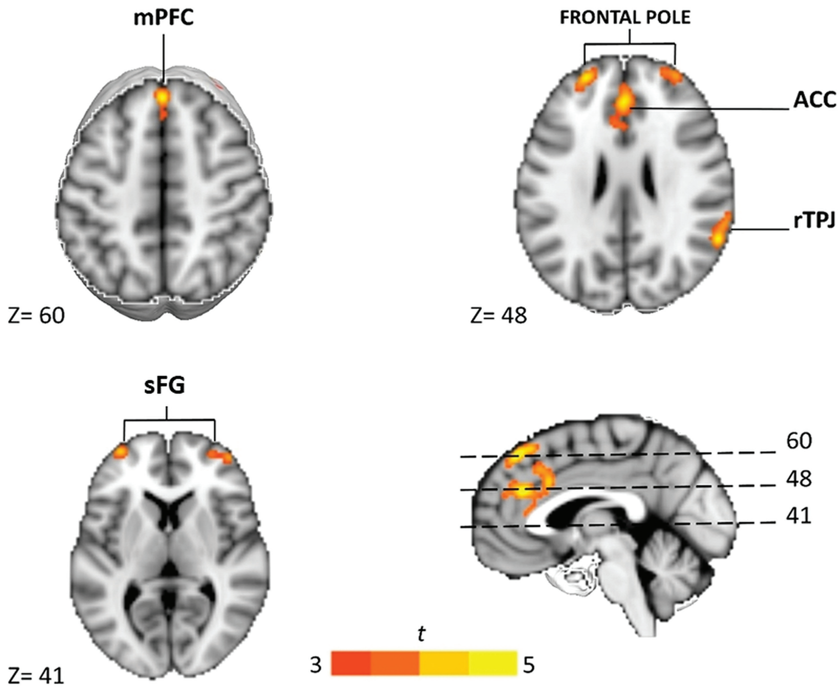

Fig. 4. Significantly stronger BOLD signals in homosexual *vs* heterosexual subjects to the subliminal presentation of the face of the character after viewing the movie ($P < 0.01$, cluster corrected).

Previous studies using subliminal presentation of faces show activity in brain areas thought to be related to social perception and empathy (for a review, see Brooks *et al.*, 2012). Our new results that accord with these findings suggest that neural mechanisms underlying intergroup empathy may occur already at the level of subliminal perception. The brain areas that were modulated in the present study have been previously reported to be involved in social perception including intergroup bias (for a review, see Amodio, 2014; Cikara and Bavel, 2014). For example, mPFC along with ACC has been reported as important regions for the processing of social information (Mitchell *et al.*, 2002; Amodio and Frith, 2006; Krueger *et al.*, 2009; Van Overwalle, 2009; Janowski *et al.*, 2013), as well as to the self-referential thinking (Vogeley and Fink, 2003; Mitchell *et al.*, 2006; Morrison *et al.*, 2012). Furthermore, mPFC has been reported to be involved in social information processing (see Grossmann, 2013 and in-group perception Volz *et al.*, 2009). In addition, TPJ has been consistently reported to support empathy, theory of mind and social perception (Saxe and Kanwisher, 2003; Saxe, 2006; Decety and Lamm, 2007; Decety *et al.*, 2012). Interestingly, the frontal gyrus and regions in the FP were also reported previously to be related to processing in-group faces as well as one's own face (Scheepers *et al.*, 2013). Thus, activations observed in the present study in the homosexual subjects in response to the face of the homosexual character might have been due to knowledge gained about him via the movie as an in-group member, which modulated implicit social perception and elicited empathy and in-group response.

Our IAT results indicate that heterosexual subjects had a bias toward homosexuality. In addition, identification scores show that heterosexual subjects experienced significant decrease of identification with the priest after knowing he is homosexual. As a consequence, we expected to find stronger neural activations in heterosexual (*vs* homosexual) subjects in the insular region after viewing the movie. A tendency for such activation was detected in insular region previously reported to be related to disgust (Stark *et al.*, 2007), yet the activation did not survive cluster correction.

An alternative explanation to the differences in brain activation between homosexual *vs* heterosexual subjects could be a difference in empathic sensitivity levels between the heterosexuals and homosexuals as suggested previously (Perry *et al.*, 2013). Since our film protagonist is a Catholic priest, which in the middle of the film is revealed to be homosexual, this might have caused an experience of violation of social norms in some of the subjects (Burgoon and Hale, 1988). Although it is unlikely to expect that one group experienced more violation of expectancy than the other, it is still possible that some of the results might be related to differences in experiencing a violation of expectancy (Le Poire and Burgoon, 1996). Another possibility might be an increase in perceived attractiveness. It has been shown that preferred-sex faces are more rewarding than non-preferred-sex faces (Hahn *et al.*, 2016). However, this preferred-sex bias was particularly pronounced when attractive faces were presented, while in our study the differences in activation were found to the same face and only after presenting contextual information via the movie. However, this does not exclude the possibility that the attractiveness of the priest's face increased in homosexual subjects after seeing the film and this could explain part of the observed activations such as the ones in the frontal gyrus (Turk *et al.*, 2004; Proverbio *et al.*, 2010).

Limitations of the study

The current study had some limitations. First, our participants were all Finnish men, so it is possible that choosing a Finnish movie would have evoked stronger responses among the subject groups, even though all subjects were sufficiently proficient in

English to follow the English movie with ease. Our choice of an English film from the 90s was to assure that none of our subjects had seen the film before, which would have damaged our experimental design where the main protagonist is revealed to be homosexual in the middle of the film. Secondly, our recruited subjects might have varied in their given implicit bias, i.e. some might have in-group bias but not out-group bias and vice versa, which we cannot discern based on the data collected in the present study. Finally, because the number of subjects in our experimental groups was relatively small, future experiments using larger sample sizes are needed to confirm our results and shed more light on possible additional brain regions that might be involved. To facilitate meta-analyses we have uploaded unthresholded statistical parametric maps to Neurovault.

Conclusion

We demonstrate here that subliminal presentation of the face of a film character with a specific social identity can be used to measure the brain's implicit neural responses to individuals with various social relationships to the viewer. Our approach can be a naturalistic alternative to investigate different types of intergroup biases.

Funding

This work was supported by the Academy of Finland (259952 and 276643).

Acknowledgements

We would like to thank Marita Kattelus and Toni Auranen from Advanced Magnetic Imaging Centre, Aalto NeuroImaging, Aalto University, Espoo, Finland for their help and support. The data reported in this paper are archived at the neurovault database: https://neurovault.org/collections/ELLZKWSV.

References

Adolphs, R., Tranel, D., Damasio, A.R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, **52**, 61–9.

Allport, G.W. (1954). Formation of in-groups. In: *The Nature of Prejudice*, Cambridge, MA: Addison-Wesley, 29–47.

Amodio, D.M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, **15**, 670–82.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, **7**, 268–77.

Avenanti, A., Sirigu, A., Aglioti, S.M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology*, **20**, 1018–22.

Bacha-Trams, M., Glerean, E., Dunbar, R., *et al.* (2017). Differential inter-subject correlation of brain activity when kinship is a variable in moral dilemma. *Scientific Reports*, **7**, 14244.

Ballew, C.C., Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 17948–53.

Banse, R., Seise, J., Zerbes, N. (2001). Implicit attitudes towards homosexuality: reliability, validity, and controllability of the IAT. *Experimental Psychology*, **48**, 145–60.

Bar, M., Neta, M., Linz, H. (2006). Very first impressions. *Emotion*, **6**, 269–78.

Biele, C., Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Experimental Brain Research*, **171**, 1–6.

Borkenau, P., Brecke, S., Möttig, C., *et al.* (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality*, **43**, 703–6.

Brooks, S.J., Savov, V., Allzén, E., *et al.* (2012). Exposure to subliminal arousing stimuli induces robust activation in the amygdala, hippocampus, anterior cingulate, insular cortex and primary visual cortex: a systematic meta-analysis of fMRI studies. *NeuroImage*, **59**, 2962–73.

Bruner, J.S. (1957). On perceptual readiness. *Psychological Review*, **64**, 123–52.

Burgoon, J.K., Hale, J.L. (1988). Nonverbal expectancy violations: model elaboration and application to immediacy behaviors. *Communication Monographs*, **55**, 58–79.

Calder, A.J., Keane, J., Manes, F., *et al.* (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, **3**, 1077–8.

Carr, L., Iacoboni, M., Dubeau, M.-C., *et al.* (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 5497–502.

Cikara, M., Van Bavel, J.J. (2014). The neuroscience of intergroup relations: an integrative review. *Perspectives on Psychological Science*, **9**, 245–74.

Corrigan, P.W., Edwards, A.B., Green, A., *et al.* (2001). Prejudice, social distance, and familiarity with mental illness. *Schizophrenia Bulletin*, **27**, 219–25.

Cottrell, C.A., Neuberg, S.L. (2005). Different emotional reactions to different groups: a sociofunctional threat-based approach to 'prejudice'. *Journal of Personality and Social Psychology*, **88**, 770–89.

Çukur, T., Nishimoto, S., Huth, A.G., *et al.* (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*, **16**, 763–70.

Cullen, C., Barnes-Holmes, D. (2008). Implicit pride and prejudice: a heterosexual phenomenon? In: Morrison, M.A., Morrison, T.G., editors. *The Psychology of Modern Prejudice*, New York: Nova Science, 195–223.

Cunningham, W.A., Johnson, M.K., Raye, C.L., *et al.* (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, **15**, 806–13.

Decety, J., Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *The Neuroscientist*, **13**, 580–93.

Decety, J., Norman, G.J., Berntson, G.G., *et al.* (2012). A neurobehavioral evolutionary perspective on the mechanisms underlying empathy. *Progress in Neurobiology*, **98**, 38–48.

Doosje, B., Ellemers, N., Spears, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of Experimental Social Psychology*, **31**, 410–36.

Freeman, J.B., Ma, Y., Barth, M., *et al.* (2015). The neural basis of contextual influences on face categorization. *Cerebral Cortex*, **25**, 415–22.

Geisler, W.S., Perry, J.S. (2011). Statistics for optimal point prediction in natural images. *Journal of Vision*, **11**, 14–4.

Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, **74**, 1464–80.

Grossmann, T. (2013). The role of medial prefrontal cortex in early social cognition. *Frontiers in Human Neuroscience*, **7**, 340.

Hahn, A.C., Fisher, C.I., DeBruine, L.M., *et al.* (2016). Sex-specificity in the reward value of facial attractiveness. *Archives of Sexual Behavior*, **45**, 871–5.

Hall, A.E., Bracken, C.C. (2011). I really liked that movie. *Journal of Media Psychology*, **23**, 90–9.

Hart, A.J., Whalen, P.J., Shin, L.M., *et al.* (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport*, **11**, 2351–5.

Hasson, U. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, **303**, 1634–40.

Hasson, U., Landesman, O., Knappmeyer, B., *et al.* (2008). Neurocinematics: the neuroscience of film. *Projections*, **2**, 1–26.

Hornsey, M.J. (2008). Social identity theory and self-categorization theory: a historical review. *Social and Personality Psychology Compass*, **2**, 204–22.

Jääskeläinen, I.P., Koskentalo, K., Balk, M.H., *et al.* (2008). Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *The Open Neuroimaging Journal*, **2**, 14–9.

Jääskeläinen, I.P., Halme, H.-L., Agam, Y., *et al.* (2016). Neural mechanisms supporting evaluation of others' errors in real-life like conditions. *Scientific Reports*, **6**, 18714.

Jabbi, M., Swart, M., Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage*, **34**, 1744–53.

Jabbi, M., Bastiaansen, J., Keysers, C. (2008). A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways J. *PLoS One*, **3**, ee2939.

Janowski, V., Camerer, C., Rangel, A. (2013). Empathic choice involves vmPFC value signals that are modulated by social processing implemented in IPL. *Social Cognitive and Affective Neuroscience*, **8**, 201–8.

Jellison, W.A., McConnell, A.R., Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: in group preferences and related behaviors and beliefs among gay and straight men. *Personality and Social Psychology Bulletin*, **30**, 629–42.

Jenkinson, M. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, **17**, 825–41.

Jost, J., Hunyady, O. (2003). The psychology of system justification and the palliative function of ideology. *European Review of Social Psychology*, **13**, 111–53.

Jost, J.T., Banaji, M.R., Nosek, B.A. (2004). A decade of system justification theory: accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, **25**, 881–919.

Krueger, F., Barbey, A.K., Grafman, J. (2009). The medial prefrontal cortex mediates social event knowledge. *Trends in Cognitive Sciences*, **13**, 103–9.

Kubota, J.T., Banaji, M.R., Phelps, E.A. (2012). The neuroscience of race. *Nature Neuroscience*, **15**, 940–8.

Lahnakoski, J.M., Glerean, E., Jääskeläinen, I.P., *et al.* (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage*, **100**, 316–24.

Lähteenmäki, M., Hyönä, J., Koivisto, M., *et al.* (2015). Affective processing requires awareness. *Journal of Experimental Psychology. General*, **144**, 339–65.

Le Poire, B.A., Burgoon, J.K. (1996). Usefulness of differentiating arousal responses within communication theories: orienting response or defensive arousal within nonverbal theories of expectancy violation. *Communication Monographs*, **63**, 208–30.

Malinen, S., Hlushchuk, Y., Hari, R. (2007). Towards natural stimulation in fMRI—issues of data analysis. *NeuroImage*, **35**, 131–9.

Mitchell, J.P., Heatherton, T.F., Macrae, C.N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 15238–43.

Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, **50**, 655–63.

Morrison, S., Decety, J., Molenberghs, P. (2012). The neuroscience of group membership. *Neuropsychologia*, **50**, 2114–20.

Mumenthaler, C., Sander, D. (2012). Social appraisal influences recognition of emotions. *Journal of Personality and Social Psychology*, **102**, 1118–35.

O'Toole, A.J., Jonathon Phillips, P., Weimer, S., *et al.* (2011). Recognizing people from dynamic and static faces and bodies: dissecting identity with a fusion approach. *Vision Research*, **51**, 74–83.

Olsson, A. (2005). The role of social groups in the persistence of learned fear. *Science*, **309**, 785–7.

Payne, B.K. (2005). Conceptualizing control in social cognition: how executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology*, **89**, 488–503.

Perry, D., Walder, K., Hendler, T., *et al.* (2013). The gender you are and the gender you like: sexual preference and empathic neural responses. *Brain Research*, **1534**, 66–75.

Petta, G., Walker, I. (1992). Relative deprivation and ethnic identity. *British Journal of Social Psychology*, **31**, 285–93.

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., *et al.* (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, **12**, 729–38.

Porter, S., England, L., Juodis, M., *et al.* (2008). Is the face a window to the soul? Investigation of the accuracy of intuitive judgments of the trustworthiness of human faces. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement*, **40**, 171–7.

Power, J.D., Mitra, A., Laumann, T.O., *et al.* (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, **84**, 320–41.

Proverbio, A.M., Riva, F., Martin, E., *et al.* (2010). Neural markers of opposite-sex bias in face processing. *Frontiers in Psychology*, **1**, 169.

Righart, R., De Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, **8**, 264–72.

Rowatt, W.C., Tsang, J.-A., Kelly, J., *et al.* (2006). Associations between religious personality dimensions and implicit homosexual prejudice. *Journal for the Scientific Study of Religion*, **45**, 397–406.

Rule, N.O., Ambady, N. (2009). She's got the look: inferences from female chief executive officers' faces predict their success. *Sex Roles*, **61**, 644–52.

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I.P., *et al.* (2016). Discrete neural signatures of basic emotions. *Cerebral Cortex*, **26**, 2563–73.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, **16**, 235–9.

Saxe, R., Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in 'theory of mind'. *NeuroImage*, **19**, 1835–42.

Scheepers, D., Derks, B., Nieuwenhuis, S., *et al.* (2013). The neural correlates of in-group and self-face perception: is there overlap for high identifiers? *Frontiers in Human Neuroscience*, **7**, 528.

Schwarz, K.A., Wieser, M.J., Gerdes, A.B.M., *et al.* (2013). Why are you looking like that? How the context influences evaluation and processing of human faces. *Social Cognitive and Affective Neuroscience*, **8**, 438–45.

Smith, S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, **17**, 143–55.

Smith, H.J., Tyler, T.R. (1997). Choosing the right pond: the impact of group membership on self-esteem and group-oriented behavior. *Journal of Experimental Social Psychology*, **33**, 146–70.

Spears, R., Doosje, B., Ellemers, N. (1997). Self-stereotyping in the face of threats to group status and distinctiveness: the role of group identification. *Personality and Social Psychology Bulletin*, **23**, 538–53.

Stark, R., Zimmermann, M., Kagerer, S., *et al.* (2007). Hemodynamic brain correlates of disgust and fear ratings. *NeuroImage*, **37**, 663–73.

Stolier, R.M., Freeman, J.B. (2016). Neural pattern similarity reveals the inherent intersection of social categories. *Nature Neuroscience*, **19**, 795–7.

Tajfel, H., Turner, J. (1979). An integrative theory of intergroup conflict. In: Austin, W.G., Worchel, S., editors. *The Social Psychology of Intergroup Relations*, Monterey, CA: Brooks/Cole, 33–47.

Todorov, A., Pakrashi, M., Oosterhof, N.N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, **27**, 813–33.

Todorov, A., Loehr, V., Oosterhof, N.N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, **39**, 514–32.

Todorov, A., Olivola, C.Y., Dotsch, R., *et al.* (2015). Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, **66**, 519–45.

Turk, D.J., Banfield, J.F., Walling, B.R., *et al.* (2004). From facial cue to dinner for two: the neural substrates of personal choice. *NeuroImage*, **22**, 1281–90.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, **30**, 829–58.

Vicario, C.M., Rafal, R.D., Martino, D., *et al.* (2017). Core, social and moral disgust are bounded: a review on behavioral and neural bases of repugnance in clinical disorders. *Neuroscience and Biobehavioral Reviews*, **80**, 185–200.

Vogeley, K., Fink, G.R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, **7**, 38–42.

Volz, K.G., Kessler, T., von Cramon, D.Y. (2009). In-group as part of the self: in-group favoritism is mediated by medial prefrontal cortex activation. *Social Neuroscience*, **4**, 244–60.

Wicker, B., Keysers, C., Plailly, J., *et al.* (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*, **40**, 655–64.

Wieser, M.J., Brosch, T. (2012). Faces in context: a review and systematization of contextual influences on affective face processing. *Frontiers in Psychology*, **3**, **3**, 1–1.

Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M. (2008). Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, **18**, 230–42.

Xu, X., Zuo, X., Wang, X., *et al.* (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, **29**, 8525–9.

Yuki, M., Maddux, W.W., Brewer, M.B., *et al.* (2005). Cross-cultural differences in relationship- and group-based trust. *Personality and Social Psychology Bulletin*, **31**, 48–62.