Patrolling the boundaries of social domains: neural activations to violations of expectations for romantic and work relationships

Bland, A. R.[1], Zahn, R.[2], Elliott, R.[3,4], Taylor, J. R. [4] and Hill, J.[5*]

[1]Department of Psychology, Manchester Metropolitan University, UK
[2]Centre for affective disorders, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, UK
[3]Neuroscience and Psychiatry Unit, University of Manchester, UK
[4]Division of Neuroscience and Experimental Psychology, School of Biological Sciences, University of Manchester, UK
[5]School of Psychology and Clinical Language Sciences, University of Reading, UK.

*Corresponding author: j.hill@reading.ac.uk

Keywords –fMRI, fronto-insular, salience, social-inappropriateness, social-interaction

Word count: 4781

Abstract

According to the social domains hypothesis, we reduce the information-processing demands of complex social cues by classifying them into a limited number of domains, each with distinct sets of expectations. This requires rapid identification of violations of the boundaries between domains. We hypothesised that these violations are likely to be associated with neural activation of the salience system. Using fMRI we compared responses of 20 adults to expected and unexpected everyday social scenarios in personal and work interactions. The vignettes exemplified different kinds of scenarios presented in the work setting, i.e., task-focused scenarios which are expected at work and scenarios with a personal focus which are unexpected at work. The key contrast between task and personal focussed scenarios presented in the work setting was associated with fronto-insular activation. Perceived inappropriateness of the unexpected scenarios, and shorter response time to judgement of inappropriateness, were also associated with fronto-insular activation, after controlling for unpleasantness. This study indicates specific neural responses to violations of expectations in different social situations. Our findings suggest that the fronto-insular region is implicated in rapid detection of behaviours that cross the boundaries of social domains which are hypothesised to be necessary for efficient social information processing.

Introduction

Managing and respecting boundaries in social situations is a crucial skill for forming and maintaining different kinds of human relationships. For instance, a work-related question asked by a colleague falls within the expected boundaries of the social situation, while a deeply personal question is likely to feel inappropriate and uncomfortable. This, we have proposed, is a particular instance of the general way in which humans classify social experiences into social domains, such as work and romantic settings, in order to guide the interpretation of social cues and the range of possible actions in each domain (Hill, Pilkonis and Bear, 2010). It reflects the broader human capability to select from, and summarise, complex sensory inputs in order to underpin appropriate and timely action (Bolton and Hill 2004, Gigerenzer, 2008, Kahneman et al 1982). Social domains provide a rapid and reasonably accurate basis for social interactions, provided that others' social behaviours are monitored to detect whether they conform to the expectations of the domain, or whether they cross the boundary. Behaviours that cross the boundary such as a colleague interacting in a very personal way, will give rise to uncertainty in how to respond and will require greater attentional and information processing resources while behaviours are evaluated and the identity of the domain reappraised. We hypothesised that this ability will be underpinned by neural structures implicated in evaluating salience, (Uddin, 2015) and by neurones found in species with more complex social organizations (Allman et al 2010). In this study we tested, using fMRI, the prediction that social stimuli that cross social domains boundaries will be associated with amygdala, dorsal anterior cingulate and fronto-insular activation.

It seems likely that there is overlap between brain regions implicated in responding to salient social and non-social events. Indeed, some studies have found that that responses to social stimuli involve the same salience network structures as non-social stimuli (Luo et al, 2014). The potential importance of the salience network to accurate social responding was illustrated in a study of adolescents, where unexpected mismatches between facial emotional expressions were associated with right AIC and dACC activations specifically in early adolescence, a period of increasing social sensitivity (Rosen et al 2017). Furthermore, better social functioning was associated with right AIC and right dACC activation. However, social salience may not be unitary, with different aspects of social stimuli activating different brain regions. For example, in a study of responses to emotionally arousing pictures, amygdala activation was associated with participants' ratings of their emotional intensity, but not their ratings of how much they related to the pictures (Phan et al 2004). Personal relatedness, however, was associated with dorsal medial frontal and insula activation. Furthermore, social stimuli used in experiments rarely refer to day-to-day social experiences and so we may need to look to more indirect evidence in generating hypotheses for social salience. One likely informative route is to identify regions and locations of neurons that are confined to animals with complex social organisations or even distinct to humans. In relation to the insula, humans and apes can be distinguished from monkeys and small mammals by the presence of Von Economo Neurones (VEN) in the ventral anterior (agranular) insula, with a particular abundance of these neurons in humans (Rose, 1928). Comparative studies have demonstrated VEN phylogenetic restriction to humans and other large-brained mammals with complex sociality, including great apes, cetaceans, and elephants (Nimchinsky et al., 1999; Hof and Van der Gucht, 2007; Butti et al., 2009; Hakeem et al., 2009; Allman et al., 2010). Contemporary functional imaging, lesion-deficit, and anatomical studies link the VEN rich regions of the anterior cingulate and fronto-insular cortices to a host of social-emotional functions (Craig 2010; Critchley 2005; Heimer and Van Hoesen 2006; Seeley et al. 2007; Williamson and Allman 2011). In the light of this evidence the fronto-insular region as well as the anterior insula may be a candidate for processing social salience.

The salience of many social stimuli depends not only on their physical characteristics but also on their 'priority' in relation to high-level representations such as goals or expectations (Ptak, 2012). These expectations may be formulated as more or less explicit rules for interactions, with violations of the rules likely to activate the salience system. Neural correlates of responses to violations of social norms have been investigated using monetary exchange paradigms where norms are established between 'investors' and 'trustees' based on offers that are likely to benefit both participants. Offers by investors that make mutual benefits less likely therefore violate the expected norm. Detection of this violation by trustees is essential to their ability to find ways of restoring cooperation. Several studies have demonstrated AIC activation in trustees in response to these violations by investors, suggesting its key role in detecting violations of social norms (Spitzer et al 2007, Rilling & Sanfey 2011). The question then arises as to whether the same mechanism is employed in more naturalistic social exchanges.

We have argued that information processing in day-to-day social exchanges is made manageable in humans by our ability to partition social interactions into different social domains. These "algorithms of social life" provide implicit rules that guide the interpretation of social events and social responding (Bugental 2000, Hill et al 2008, Hill et al 2010). For example, the implicit rules of an educational or work environment include that each person has a task focused role and that expressions of emotions can generally be interpreted in relation to that role. By contrast, according to the implicit rules of a romantic relationship emotions need to be interpreted in relation to a much broader set of possibilities. The domains are efficient vehicles for social interaction, ensuring for example that, in work, people can cooperate efficiently immediately after meeting, while romantic relationships can provide emotional resources not available in work but rely on time consuming building of relationships. Social domain functioning will only be efficient if domains are implemented jointly by social participants. This is ensured to some degree by social structures such as employment contracts and marriage, but these are unlikely to guarantee shared moment-to-moment implementation of domains. This, we suggest, is maintained by constant monitoring of whether social behaviours conform to the expectations of the domain, or whether they cross the boundary. Behaviours that cross the boundary will require greater attentional and information processing resources while they are evaluated and the identity of the domain reappraised. The implication is that unexpected, domain incongruent behaviours will be highly salient. In this study we set out to examine whether brain regions of the salience network are activated by social domain incongruent stimuli.

The key contrast for the study of the neuroscience of domains functioning is between social cues that are congruent to a setting where a narrow range of emotions and behaviours can be accommodated, and those that are domain incongruent, because of their focus or intensity. For the present study, this is a contrast between different kinds of social cues presented in the work setting, on the one hand task-focused cues that are domain congruent, and on the other hand personal-focused cues that are domain incongruent. Equally, it is predicted that this contrast of task versus personal focus will not have the same implications within a marital or partner relationship where a wide range of emotions and behaviours is expected and congruent. Thus, we include contrasts not only between different kinds of social cue within work, but also with the same set of scenarios depicted within a partner relationship.

It is plausible that mechanisms in detecting and responding to domain incongruent social cues are both unconscious, in order to underpin rapid responses, and conscious, in order to facilitate reappraisal of the assumed rules for interaction and action. According to some recent hypotheses the AIC is involved in initial unconscious processing that determines priorities for higher level attentional resources in networks associated with consciousness (Michel, 2017). We therefore sought to evaluate whether salience network activation is associated with reaction time (a likely index of initial unconscious processes) to making a judgement about a stimulus and with conscious evaluation of its inappropriateness.

In this fMRI study we first tested the hypothesis that social cues which are congruent in marital or partner relationships will be associated with ratings of domain incongruence when displayed within a work setting and will therefore be associated with greater activation in the amygdala, dorsal anterior cingulate and fronto-insular regions. Second, we predicted that shorter reaction times from the appearance of domain incongruent stimuli to making the judgement of their inappropriateness will be associated with greater salience network activation in the domain incongruent, compared to domain congruent conditions. Third, the extent to which domain incongruent stimuli are judged as inappropriate will be associated with greater activation in the amygdala, dorsal anterior cingulate and fronto-insular regions to incongruent stimuli. As inappropriateness may be associated with unpleasantness, likelihood and uncertainty, these were examined as potential confounds.

We also examined frontopolar and anterior temporal activations which have been shown to represent different aspects of knowledge relevant for understanding the rules underpinning social behaviour. There is evidence to support the role of the frontopolar cortex as representing sequential aspects (Zahn et al., 2017) and the anterior temporal cortex as representing abstract conceptual aspects (Zahn et al., 2009). We predicted that there would be no selective activations in these regions for our domain-incongruence contrasts, because activations of representations of social rules are entailed in both domain-congruent and domain-incongruent conditions.

Materials and Methods

Participants

Twenty healthy volunteers (11 female) took part in this study. Participants were included if they met the following criteria: 18-50 years old, no self-reported previous or current psychiatric disorders including depression, anxiety, eating disorders and drug/alcohol addiction, no neurological disorders, no significant head injury resulting in unconsciousness, no current use of medication known to affect mood or cognition, no first degree relatives suffering from any psychiatric disturbances, smoking less than 5 cigarettes per day, drinking less than the UK government guidelines for weekly alcohol intake (males: 3-4 units per day; females: 2-3 units per day) and fluent in English. Participants were further interviewed using the M.I.N.I. International Neuropsychiatric Interview (Sheehan et al., 1998) and were excluded if they met the criteria for a current psychiatric diagnosis. All participants included in the study had at least 3-months work experience and had at least 3-months romantic experience. At the time of scanning, 8 participants were single and 12 in a relationship varying between 6 months – 7.5 years. At the time of scanning, 9 participants were in full time employment, 2 were in part-time employment and 9 were in full time University education. A structural MRI scan was taken before fMRI scanning to aid analysis and to exclude data from participants who showed any evidence of structural abnormality. No participants were excluded from the subsequent analysis on this basis.

Procedure

Prior to their visit for scanning, participants read the information sheet, gave informed consent and completed questionnaires online. On arrival they were briefed about the scanning process, and then completed a series of practice trials in order to familiarise themselves with the task. The task was then presented in the scanner, and again out of the scanner together with self-report questions about their responses to the task items. The study was approved by the University of Manchester Research Ethics Committees and all participants provided written informed consent after the study procedures were explained. Participants were reimbursed for their time and travel expenses.

Measures

Social domains assessment

The task was displayed on a Dell XT3 laptop using PsychoPy software (Peirce, 2007). Participants were presented in the scanner with short vignettes describing social interactions that differed by their congruence with the social domain, task focus being congruent with both work and partner relationships, and personal focus only to partner relationships. The vignettes were designed to exemplify four conditions: work congruent behaviours seen in work (WORK-TASK FOCUS) and in a partner relationship (PARTNER-TASK FOCUS), and behaviours that are congruent in a partner relationship (PARTNER-PERSONAL FOCUS) but not in work (WORK-PERSONAL FOCUS). An example of WORK-TASK FOCUS was "A work colleague gets angry with you for leaving the light on" and of PARTNER-TASK FOCUS "Your partner gets angry with you for leaving the light on". An example of PARTNER-PERSONAL FOCUS was "Your partner gets angry with you for not noticing that they have changed their hairstyle" and of WORK-PERSONAL FOCUS "A work colleague gets angry with you for not noticing that they have changed their hair style". The WORK-PERSONAL FOCUS vs WORK-TASK FOCUS contrasts were used to test the hypotheses, with the PARTNER–TASK FOCUS vs PARTNER-PERSONAL FOCUS contrasts controlling for differences of focus that did not introduce domain incongruence, and the PARTNER–PERSONAL FOCUS controlling for simple effects of personal focus.

Participants were asked to report how unpleasant they would feel ("mildly" or "very") by pressing a button within 6 seconds. If they responded before the end of the 6 seconds their response remained on the screen and the alternative disappeared. This was followed by a jittered intertrial interval with a mean duration of 4 seconds. This was jittered in 9 steps of 500ms around the mean interval with equal distribution of intertrial intervals across different stimulus types. There were 26 vignettes for each of the four conditions that had been selected from a battery of 60 vignettes (total 240) to provide stimuli that were matched as closely as possible on degree of unpleasantness as rated by a panel of 20 volunteers.

Self-report responses to the task

After the imaging session, participants provided self-report ratings for each trial. Perceived inappropriateness was rated for each social scenario on a Likert scale (0 = not at all inappropriate, 100 = highly inappropriate), and reaction times from display of the social scenarios to judgement of inappropriateness were recorded. Likert scales included as potential confounds were of: unpleasantness (0 = not at all unpleasant, 100 = extremely unpleasant), likelihood of it happening (0 = very unlikely, 100 = very likely), and how certain they would be of how to respond (0 = very uncertain, 100 = very uncertain). Reaction times for making these judgements were recorded.

MRI scanning

MRI data were acquired on a 3 T Philips Achieva scanner (Philips Medical Systems, the Netherlands) with an eight-channel coil. A series of echo-planar images (EPI) was acquired for each participant using a sequence optimized for detecting ventral frontal signals (240 volumes; 40 axial slices; 3 mm slice thickness; ascending sequential acquisition; repetition time: 2000 ms; echo time: 22 ms; field of view: 240 × 240 × 120 mm; acquisition matrix: 80 × 80 voxels; reconstructed voxel size: 3 mm^3; flip angle: 90°). A three-dimensional T1-weighted magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) structural image was also acquired for each participant (160 axial slices; 0.9 mm slice thickness; repetition time: 8.4 ms; echo time: 3.9 ms; field of view: 240 × 191 × 144 mm; acquisition matrix: 256 × 163 voxels; reconstructed voxel size: 0.94 × 0.94 × 0.9 mm; flip angle: 8°). In order to rule out clinically significant neurological abnormalities, T2-weighted structural images were also acquired. Statistical Parametric Mapping (SPM12; https://www.fil.ion.ucl.ac.uk/spm/software/spm12/) and the general linear model were used for image analysis (Friston, Frith, Turner, & Frackowiak, 1995; Worsley & Friston, 1995).

Statistical analyses

Functional images were realigned, unwarped, coregistered to the participant's T1-weighted images, and normalized to the SPM12 (2014) template using the transformation parameters for the T1-weighted image. A smoothing kernel of 8-mm full-width at half maximum was then applied. At the first level, the four conditions; PARTNER-PERSONAL, PARTNER–TASK FOCUS, WORK-PERSONAL FOCUS, WORK-TASK FOCUS were modelled as t-contrasts for each subject. For the second level we then set up an ANOVA model with the four conditions as the within-subject factor. In order to test the first hypothesis, within-group differences on WORK-PERSONAL FOCUS vs WORK-TASK FOCUS were examined. A single bilateral combined *a priori* region of interest (ROI) mask was defined for small volume multiple comparison correction at a voxel-based FWE-corrected level of $p<0.05$ which was composed of the Automated Anatomical Labelling (AAL) atlas-defined anterior cingulate cortex (ACC), bilateral amygdala, bilateral insula, frontal pole (BA10, Talairach Daemon atlas as implemented into the WFU Pick-Atlas version 2.4) and temporal poles. We report all activations which either survived a voxel-based familywise error corrected $p<0.05$ over the whole brain or the ROI mask. All data underwent registration and 12-parameter affine normalization and were transformed into standard MNI space using a 2 × 2 × 2 isotropic resolution. Data were smoothed using an 8 mm full width at half maximum Gaussian kernel. In a further analysis, we confirmed that the resulting clusters were not driven by the PARTNER context by exclusively masking with PARTNER-PERSONAL FOCUS and PARTNER-TASK FOCUS at $p=0.001$ uncorrected.

In further confirmatory analyses, we extracted the cluster averages for regions surviving multiple comparison correction (voxel-based few $p<0.05$ as described above) for analyses in SPSS. This allowed us to test the second hypothesis by regressing the difference between WORK–TASK FOCUS and WORK–PERSONAL FOCUS activations in each cluster on the differences in reaction times (recorded outside of the scanner) to judgement of inappropriateness across the conditions. For the third hypothesis the difference in activations was regressed on to perceived inappropriateness, and for the fourth it was regressed on to the contrast in expressions of anger or distress across the conditions. All analyses controlled for perceived unpleasantness, uncertainty or likelihood. In view of the sample size separate models were examined for each of the three confounds.

Results

Self-report measures

Table 1 shows the mean scores for perceived inappropriateness of the four conditions, and reaction times to self-report of domain inappropriateness, together with potentially confounding self-report variables. Respondents reported that scenarios depicting disclosure of personal information or strong emotions in the work setting (WORK-PERSONAL FOCUS) were more inappropriate, less likely, and would create less certainty as to how to respond, than WORK-TASK FOCUS and PARTNER–PERSONAL FOCUS scenarios ($p<0.01$). WORK–PERSONAL FOCUS scenarios were reported to be more unpleasant than WORK–TASK FOCUS scenarios ($p<0.001$). Reaction times to the judgement of appropriateness of WORK-PERSONAL did not differ from reaction times to the other types of scenario.

fMRI results consistent across participants

The contrast between WORK-PERSONAL FOCUS vs WORK-TASK FOCUS was associated with a selective activation in the bilateral fronto-insular cortex (Figure 1) with the highest level of activation on the right hemisphere (Table 2, Figure 2). Further analyses showed that this cluster remained significant, when exclusively masking with activations associated with PARTNER-PERSONAL FOCUS and PARTNER-TASK FOCUS, showing that this region was not activated by personal focus per se, or by the differences in focus in general.

In confirmatory data analyses using the extracted cluster averages, there was a significant main effect of condition [$F(3, 57)=2.96$, $p<0.04$, $\eta_p^2=0.14$] on right fronto-insular activation illustrated in Figure 2, and in post hoc analyses, only WORK-TASK FOCUS and WORK-PERSONAL FOCUS differed significantly [$t(1,19) = 5.41$, $p<0.001$]. In similar analyses, hypothesised effects on dACC activations were not found, and nor were there effects on the amygdala, frontopolar and anterior temporal regions.

Individual difference effects on fMRI results and the role of potential confounders

Table 3 shows associations between scores for the differences between WORK-PERSONAL FOCUS and WORK-TASK FOCUS activations in the right fronto-insular and differences between the same conditions in self-report and reaction times for judgement of inappropriateness. Right front-insular activation was associated with greater perceived inappropriateness of WORK-PERSONAL FOCUS, and lower reaction times for judgement of its inappropriateness. Perceived inappropriateness of the WORK–PERSONAL FOCUS over WORK–TASK FOCUS was associated with lower likelihood, while greater unpleasantness of WORK–PERSONAL FOCUS was associated with shorter reaction time to judgement of inappropriateness.

In multiple linear regression, greater perceived inappropriateness of the WORK-PERSONAL FOCUS compared to the WORK–TASK FOCUS scenarios, predicted right fronto-insular activation for the same comparison, controlling for unpleasantness (Table 4), and in separate models controlling for likelihood (inappropriateness β=0.55, p=0.04) and uncertainty (inappropriateness β =0.56, p=0.03).

Shorter reaction times also predicted right fronto-insular controlling for perceived unpleasantness (Table 5), and controlling for unlikelihood (β =-0.67, p <.01) and uncertainty (β =-0.60, p<0.01). When examined jointly, without confounds, both perceived inappropriateness and shorter reaction times to indication of inappropriateness made independent contributions to right fronto-insular activation in the WORK-PERSONAL FOCUS vs WORK-TASK FOCUS comparison (Table 6).

Discussion

Social scenarios depicting interactions with work colleagues with a personal focus, predicted to be inappropriate to the domain, and hence in need of prioritising for reappraisal and action, were associated with increased fronto-insular activation, consistent with activation of the salience network. This selectivity was not seen in the ldACC region. Reported inappropriateness of the WORK-PERSONAL FOCUS compared to the WORK-TASK FOCUS stimuli, and shorter reaction time to indication of inappropriateness across these conditions, were independently associated with elevated fronto-insular activation.

As far as we are aware, this is the first study to examine whether violations of expectations involving complex social stimuli typical of those encountered in everyday life are associated with activation of the salience system. The findings are consistent with the hypothesis that everyday interactions are underpinned by shared rules regarding expected behaviours and that departures from these expectations are rapidly detected. Salience depends not only on behaviour but also the domain in which it occurs. Importantly, scenarios with personal focus did not activate the right fronto-insular when portrayed in the context of a partner relationship, only in interaction with a work colleague. Furthermore, the finding that both shorter reaction time to indicating inappropriateness and the judgement of inappropriateness were independently associated with fronto-insular activation is consistent with its role in unconscious and conscious mapping of social domains incongruence.

Our findings are also relevant to understanding psychopathology. We have previously reported that limitations in the demarcation of social domains, likely to be associated with impaired 'patrolling' of domains boundaries, 'domain disorganization', is elevated in adults with Borderline Personality Disorder (BPD) compared to psychiatric patients with other personality disorders and no personality disorder (Hill et al 2008, Morse et al 2009). In a study comparing adults with Borderline Personality Disorder with healthy controls while playing an economic exchange game, violations of expectations of fair exchange were associated with anterior insula activation in the controls but not in those with BPD (King-Casas et al. 2008). BPD patients therefore may fail to activate neural mechanisms associated with detection of norm violations, including norms associated with domain congruence.

The study had a number of strengths that enabled us to examine the specificity of the findings. While the key comparison was between behaviours with personal focus seen in partner relationships where they are to be expected, and in work relationships where they are not, the further comparison of task focussed behaviours between the domains allowed us to control for general effects of domain and focus. Furthermore, we were able to account for the extent to which any differences arose from the perceived unpleasantness of the behaviours. By obtaining self-report and reaction time data we were able to examine whether the observed activation was associated with both conscious awareness of appropriateness and performance. The main limitation of the study was the modest sample size, which while adequate for examination of limited numbers of possible confounds, was not sufficiently large to include all those of interest in the same model. In addition, although unpleasantness (i.e. valence) is associated with arousal, we did not specifically gather participant's ratings of arousal.

The theoretical framework on which this study was based has features in common with other current theories regarding social decision making (Rilling & Sanfey 2011). These hypothesise that interpretations of others' behaviours are based on expectations derived from social norms, which as outlined earlier speeds up social information processing. Neural mechanism that track violations alert the person to reassess the expectations. Our findings are consistent with previous studies that have identified fronto-insular activation in response to violations but different from others that have for other regions, for example the anterior cingulate cortex (Chang & Sanfey 2013). However, they have to be evaluated in the light of the differences, in theoretical framework and experimental tasks, compared to previous studies of social decision making. While other theoretical approaches assume that qualities such as fairness and trust apply generally across social interactions, the domains framework proposes that there are classes of social interaction each with their distinctive and different norms and values. Consequently, the task that we employed tested, not for violations of rules of exchange, but for violations of theoretically predicted demarcations of social domains. In spite of the differences from previous findings, those described here are consistent with the indication that there may be general mechanisms for responding to norm violations that are employed to patrol the boundaries of social domains.

Acknowledgements

The authors thank Emily Pye and staff at the NIHR/Wellcome Trust Clinical Research Facility, Manchester, UK.

References

Allman, J. M., Tetreault, N. A., Hakeem, A. Y., Manaye, K. F., Semendeferi, K., Erwin, J. M., ... & Hof, P. R. (2010). The von Economo neurons in frontoinsular and anterior cingulate cortex in great apes and humans. *Brain Structure and Function*, *214*(5-6), 495-517.

Bugental, D. B. (2000). Acquisition of the algorithms of social life: A domain-based approach. *Psychological bulletin*, *126*(2), 187.

Butti, C., Sherwood, C. C., Hakeem, A. Y., Allman, J. M., & Hof, P. R. (2009). Total number and volume of Von Economo neurons in the cerebral cortex of cetaceans. *Journal of Comparative Neurology*, *515*(2), 243-259.

Bolton D. and Hill J. (2004) *Mind, Meaning, and Mental Disorder: The Nature of Causal Explanation in Psychology and Psychiatry* (Oxford: Oxford University Press).

Chang, L. J., & Sanfey, A. G. (2013). Great expectations: neural computations underlying the use of social norms in decision-making. *Social cognitive and affective neuroscience*, *8*(3), 277-284.

Chen, T., Cai, W., Ryali, S., Supekar, K., & Menon, V. (2016). Distinct global brain dynamics and spatiotemporal organization of the salience network. *PLoS biology*, *14*(6), e1002469.

Craig, A. D. (2010). The sentient self. *Brain structure and function*, *214*, 563-577.

Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2009). At the heart of the ventral attention system: the right anterior insula. *Human brain mapping*, *30*(8), 2530-2541.

Eysenck HJ, Eysenck SBG. 1991. Manual of Eysenck personality scales (EPS adult): comprising the EPQ-Revised (EPQ-R), EPQ-R short scale, Impulsiveness (IVE) questionnaire. London; Sydney: Hodder and Stoughton.

Gigerenzer, G. (2008) 'Why Heuristics Work', *Perspectives on Psychological Science* 3: 20-29.

Hakeem, A. Y., Sherwood, C. C., Bonar, C. J., Butti, C., Hof, P. R., & Allman, J. M. (2009). Von Economo neurons in the elephant brain. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology*, *292*(2), 242-248.

Heimer, L., & Van Hoesen, G. W. (2006). The limbic lobe and its output channels: implications for emotional functions and adaptive behavior. *Neuroscience & Biobehavioral Reviews*, *30*(2), 126-147.

Hill, J., Pilkonis, P., Morse, J., Feske, U., Reynolds, S., Hope, H., ... & Broyden, N. (2008). Social domain dysfunction and disorganization in borderline personality disorder. *Psychological medicine*, *38*(01), 135-146.

Hill, I., Pilkonis, P. A., & Bear, I. (2010). Social Domains, Personality, and Interpersonal Functioning. Horowitz, L. M., & Strack, S. (Eds.). (2010). *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions*. John Wiley & Sons.*Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions*, 281.

Hof, P. R., & Van der Gucht, E. (2007). Structure of the cerebral cortex of the humpback whale, Megaptera novaeangliae (Cetacea, Mysticeti, Balaenopteridae). *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology*, *290*(1), 1-31.

John OP, Donahue EM, Kentle RL. 1991. The big five inventory—versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

Kahneman, D., Slovic, P., & Tversky, A. (1982) *Judgment Under Uncertainty: Heuristics and Biases* (New York: Cambridge University Press).

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *science*, *321*(5890), 806-810.

Luo, Y., Qin, S., Fernandez, G., Zhang, Y., Klumpers, F., & Li, H. (2014). Emotion perception and executive control interact in the salience network during emotionally charged working memory processing. *Human brain mapping*, *35*(11), 5606-5616.

Michel, M. (2017). A role for the anterior insular cortex in the global neuronal workspace model of consciousness. *Consciousness and Cognition*, *49*, 333-346.

Morse, J. Q., Hill, J., Pilkonis, P. A., Yaggi, K., Broyden, N., Stepp, S., ... & Feske, U. (2009). Anger, preoccupied attachment, and domain disorganization in borderline personality disorder. *Journal of Personality Disorders*, *23*(3), 240-257.

Nimchinsky, E. A., Gilissen, E., Allman, J. M., Perl, D. P., Erwin, J. M., & Hof, P. R. (1999). A neuronal morphologic type unique to humans and great apes. *Proceedings of the National Academy of Sciences*, *96*(9), 5268-5273.

Patton JH, Stanford MS. 1995. Factor structure of the Barratt Impulsiveness Scale. J Clin Psychol. 51(6); 768-774.

Phan, K. L., Taylor, S. F., Welsh, R. C., Ho, S. H., Britton, J. C., & Liberzon, I. (2004). Neural correlates of individual ratings of emotional salience: a trial-related fMRI study. *Neuroimage*, *21*(2), 768-780.

Pierce JW. 2007. PsychoPy—psychophysics software in Python. J Neurosci Methods. 162(1); 8-13.

Ptak, R. (2012). The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. *The Neuroscientist*, *18*(5), 502-515.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual review of psychology*, *62*, 23-48.

Rose M. 1928. Die Inselrinde des Menschen und der Tiere. J Psychol Neurol. 37:467--624.

Rosen, M. L., Sheridan, M. A., Sambrook, K. A., Dennison, M. J., Jenness, J. L., Askren, M. K., ... & McLaughlin, K. A. (2017). Salience network response to changes in emotional expressions of others is heightened during early adolescence: relevance for social functioning. *Developmental Science*.

Seeley WW, Carlin DA, Allman JM, Macedo MN, Bush C, Miller BL, DeArmond SJ. 2006. Early frontotemporal dementia targets neurons unique to apes and humans. Ann Neurol. 60(6); 660-667.

Seeley WW, Allman JM, Carlin DA, Crawford RK, Macedo MN, Greicius MD, DeArmond SJ, Miller BL. 2007. Divergent social functioning in behavioral variant frontotemporal dementia and Alzheimer disease: reciprocal networks and neuronal evolution. Alzheimer Dis Assoc Disord. 21(4); S50-S57.

Seeley WW. 2008. Selective functional, regional, and neuronal vulnerability in frontotemporal dementia. Curr Opin Neurol. 21(6); 701.

Shacham S. 1983. A shortened version of the Profile of Mood States. J Pers Assess. 47(3); 305-306.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. 1998. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*. 59(suppl 20); 22-33.

Spielberger CD, Gorsuch RL, Lushene RE. 1970. State-trait anxiety inventory: STAI. Consulting Psychologists Press.

Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007). The neural signature of social norm compliance. *Neuron*, *56*(1), 185-196.

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, *16*(1), 55-61.

[SPM12] Wellcome Trust Centre for Neuroimaging [Internet]. 2014. Available from: http://www.fil.ion.ucl.ac.uk/spm/software/spm12/

Wechsler D. 2008. Wechsler adult intelligence scale-fourth. San Antonio, TX: Pearson.

WFU Pick-Atlas version 2.4; Wake Forest UniversityWinston-Salem, NC, USA. [Internet]. Available from:
http://www.fmri.wfubmc.edu/downloads/WFU_PickAtlas_User_Manual.pdf

Whiteside SP, Lynam DR. 2003. Understanding the role of impulsivity and externalizing psychopathology in alcohol abuse: application of the UPPS impulsive behavior scale. Exp Clin Psychopharmacol. 11(3); 210.

Williamson, P., & Allman, J. M. (2011). *The human illnesses: neuropsychiatric disorders and the nature of the human brain*. Oxford University Press, USA.

Zahn, R., Moll, J., Iyengar, V., Huey, E. D., Tierney, M., Krueger, F., & Grafman, J. (2009). Social conceptual impairments in frontotemporal lobar degeneration with right anterior temporal hypometabolism. *Brain*, *132*(3), 604-616.

Zahn, R., Green, S., Beaumont, H., Burns, A., Moll, J., Caine, D., ... & Ralph, M. A. L. (2017). Frontotemporal lobar degeneration and social behaviour: Dissociation between the knowledge of its consequences and its conceptual meaning. *Cortex*, *93*, 107-118.

Tables

Table 1: Mean self-report ratings of the social scenarios and reaction times to report of domain inappropriateness

	Work personal focus	Work task focus	Partner personal focus	Partner task focus
Inappropriateness	70.35 (10.19)	43.80 ^{c} (10.08)	24.80 (11.58) ***	28.20 (11.32) ***
Reaction time	2.76 (1.18)	3.01 (1.32)	2.66 (1.07)	2.59 (1.20)
Unpleasantness	62.45 (8.84)	50.90 (11.43) ***	54.75 (12.58)	44.30 (15.80) ***
Likelihood	34.80 (18.63)	51.25 (15.92) ***	47.35 (12.55) ***	54.35 (19.20) ***
Certainty	40.25 (12.50)	52.35 (15.8) ***	55.55 (13.35) ***	61.55 (17.00) ***

Reaction times were times to make the judgement of inappropriateness when the stimuli were viewed outside of the scanner.
Means and standard deviations, in parenthesis are shown. Means were compared using repeated measures ANOVA. Models for unpleasantness, likelihood, certainty, and inappropriateness were significant at $p < 0.001$, and for reaction time at $p = 0.010$. Superscripts indicate significance of comparisons with the domain incongruent WORK – PERSONAL FOCUS as the reference condition.
* = $p<0.05$ ** = $p<0.01$ *** = $p<0.001$

Table 2: fMRI comparison of Work Context: Personal Focus vs. Work Context: Task Focus

Hemisphere	Region	Peak MNI Coordinates X	Y	Z	Peak *z* score	Cluster Size	FWE-Corrected *p* Value
Right	Fronto-insular cortex	45	26	-3	4.69	88	0.003*
	Dorsal anterior cingulate cortex	15	23	28	4.24	37	0.019
Left	Fronto-insular cortex	-27	23	-10	4.38	62	0.011*

Results were thresholded at p=0.001 uncorrected, only regions surviving additional Familywise error (FWE) Small Volume Correction at the voxel-level <.05 over our combined apriori ROI mask were reported. No significant voxels remained when correcting over the volume of the whole brain.
* Remained significant at p=.05 voxel-wise FWE correction after additional exclusive masking by Work Context: Personal Focus vs. Partner Context: Work Focus and Work Context: Personal Focus vs. Work Context: Work Focus.

Table 3 Bivariate associations between self-report and reaction time differences and right fronto-insular activation differences, contrasting WORK-PERSONAL FOCUS and WORK-TASK FOCUS conditions

	Right fronto-insular activation	Inappropriate-ness	Reaction time (RT)	Unpleasant-ness	Unlikeli-hood
Inappropriateness	0.52* (p = 0.02)				
RT	0.62** (p = 0.007)	-0.21			
Unpleasantness	0.33	0.41 (p = 0.07)	-0.43 (p = 0.06)		
Unlikelihood	0.19	0.47* (p = 0.04)	0.22	-0.22	
Uncertainty	0.09	0.39	0.05	0.18	0.54* (p = 0.012)

Values shown are Pearson's *r*.
Values of $p < 0.1$ are shown in parentheses. * = $p<0.05$ ** = $p<0.01$
Reaction times (RT) were times to make the judgement of inappropriateness when the stimuli were viewed outside of the scanner.

Table 4 Summary of multiple linear regression models predicting right fronto-insular activation from self-report domain inappropriateness

	ΔR^2	*p*		β	*p*
Model 1	0.10	0.19	Unpleasantness	0.31	0.19
Model 2	0.19	0.05	Unpleasantness	0.14	0.54
			Inappropriateness	0.47	0.05

In separate regression analyses the effect of inappropriateness controlling for unlikelihood was β = 0.55, p = 0.04, and controlling for uncertainty β = 0.56, p = 0.03
For all models the scores were the differences between self-report or right fronto-insular activations during WORK – PERSONAL FOCUS and WORK – TASK FOCUS scenarios.

Table 5 Summary of multiple linear regression models predicting right fronto-insular activation from reaction time to the judgement of domain inappropriateness

	ΔR^2	*p*		β	*p*
Model 1	0.10	0.19	Unpleasantness	0.31	0.19
Model 2	0.26	0.02	Unpleasantness	0.05	0.84
			Reaction time	-0.58	0.02

In separate regression analyses the effect of inappropriateness controlling for unlikelihood was $\beta = -0.67$, $p = 0.003$, and for uncertainty $\beta = -0.60$, $p = 0.008$.
For all models the scores were the differences between reaction time, self-report, or right fronto-insular activations during WORK – PERSONAL FOCUS and WORK – TASK FOCUS scenarios

Table 6 Summary of multiple linear regression models predicting right fronto-insular activation from rating of domain inappropriateness and reaction time to the indication of inappropriateness

	ΔR^2	*p*		β	*p*
Model 1	0.27	0.02	Inappropriateness	0.52	0.02
Model 2	0.24	0.01	Inappropriateness	0.41	0.04
			Reaction time	-0.50	0.01

For all models the scores were the differences between reaction time, self-report, or right fronto-insular activations during WORK – PERSONAL FOCUS and WORK – TASK FOCUS scenarios.

Figures

Figure 1 shows an axial slice through the bilateral fronto-insular clusters of activation for the contrast of Work: Personal Focus vs. Work: Task Focus.

Maps were thresholded at p=0.001 uncorrected with no voxel threshold and inclusively masked with the a priori ROI. MRIcron (Rorden et al., 2007) was used for visual display.

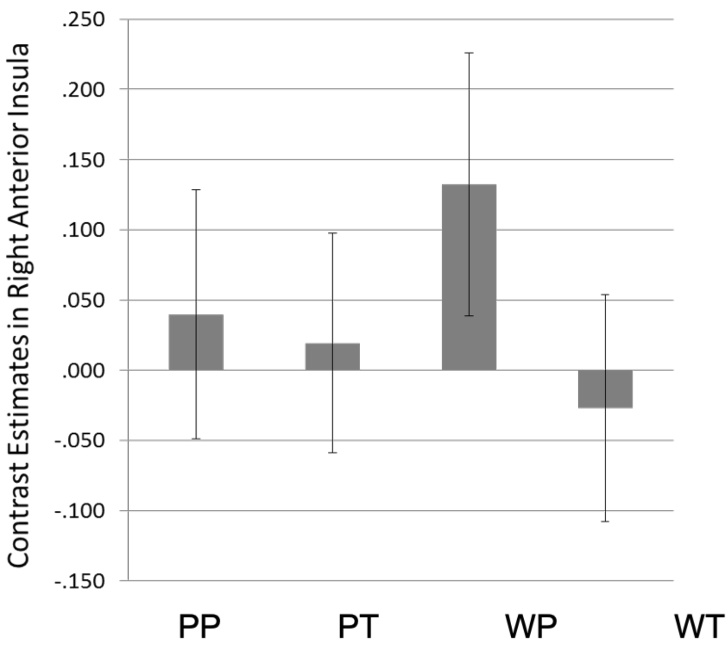

Figure 2: Plot to show right fronto-insular cortex activation across each of the conditions. Error bars represent one standard error of the mean