Social Cognitive and Affective Neuroscience Advance Access published May 9, 2016

Neural dissociations between meaningful and mere inconsistency in impression updating

Peter Mende-Siedlecki[1] and Alexander Todorov[2]

Affiliations: [1]Department of Psychology, New York University, New York, New York 10003, [2]Department of Psychology, Princeton University, Princeton, New Jersey 08542

Corresponding author: Peter Mende-Siedlecki, pms10@nyu.edu, [1]Department of Psychology, New York University, New York, New York 10003.

Abstract Word Count: 194
Main Text Word Count: 5947

© The Author (2016). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Abstract

Recent neuroimaging work has identified a network of regions that work in concert to update impressions of other people, particularly in response to inconsistent behavior. However, the specific functional contributions of these regions to the updating process remain unclear. Using fMRI, we tested whether increases in activity triggered by inconsistent behavior reflect changes in the stored representations of other people in response to behavioral inconsistency, or merely a response to the inconsistency itself. Participants encountered a series of individuals whose behavior either changed in an attributionally meaningful fashion or was merely inconsistent with the immediately preceding behavior. We observed that left ventrolateral prefrontal cortex and left inferior frontal gyrus were preferentially recruited in response to unexpected, immoral behavior, while a separate set of regions (including dorsal anterior cingulate cortex, posterior cingulate cortex, and temporoparietal junction/inferior parietal lobule) was preferentially recruited in response to more mundane inconsistencies in behavior. These results shed light on the distributed systems supporting impression updating. Specifically, while many regions supporting updating may primarily respond to moment-to-moment changes in behavior, a subset of regions (e.g., vlPFC and IFG) may contribute to updating person representations in response to trait-relevant changes in behavior.

Word count: 194/200
Keywords: impression updating, impression formation, social cognition, person perception

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Introduction

The ability to form and update impressions of others is a key social faculty, allowing us to predict how others will behave in the future and to tailor our own behavior towards those expectations. However, our understanding of the principles guiding the neural underpinnings of impression updating is in its infancy. Following foundational investigations into the neural bases of behavior-based impression formation (Cloutier, Kelley, & Heatherton, 2011a; Mitchell, Macrae, & Banaji, 2004; Mitchell, Macrae, & Banaji, 2005; Mitchell et al., 2006; Schiller et al., 2009), subsequent work has taken a more dynamic approach, examining how we change and update impressions over time based upon new information (Ames & Fiske, 2013; Baron et al., 2011; Bhanji & Beer, 2013; Cloutier et al., 2011b; Hackel, Doll, & Amodio, 2015; Harris & Fiske, 2010; Kim, Choi, & Jang, 2012; Ma et al., 2012; Mende-Siedlecki, Cai, & Todorov, 2013a; Mende-Siedlecki, Baron, & Todorov, 2013b; Stanley, 2015).

Updating is supported neurally by an extended network comprising regions involved in social cognition and impression formation, as well as regions associated with cognitive control and attention (Mende-Siedlecki et al., 2013a). Specifically, medial prefrontal cortex (mPFC, encompassing dorsomedial PFC and dorsal anterior cingulate [dACC]), lateral prefrontal cortex (lPFC, comprising both rostral and ventral aspects of lPFC), superior temporal sulcus (STS), temporoparietal junction/inferior parietal lobule (TPJ/IPL), and posterior cingulate cortex (PCC) show preferential increases in activity to behavior inconsistent with existing impressions.

Despite strong convergence observed across initial neuroimaging investigations (see Cloutier et al., 2011b; Hackel et al., 2015; Ma et al., 2012; Mende-Siedlecki et al.,

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

2013b), the specific neural contributions supporting impression updating remain unclear. Do neural responses triggered by inconsistent behaviors reflect an updated trait representation, or merely surprise at a moment-to-moment discrepancy in behavior? Several regions implicated in updating—dACC, TPJ/IPL, and PCC/precuneus—are associated with cognitive processes that might be more strongly linked to the latter account than the former. Meanwhile, other work suggests that a separate subset of regions involved in updating—including ventrolateral prefrontal cortex (vlPFC) and inferior frontal gyrus (IFG)—is recruited by inconsistencies that are diagnostic of an individual's character (Mende-Siedlecki et al., 2013b). We review evidence for these complementary hypotheses below, which, taken together, may reflect a dissociation in the neural bases of updating person impressions.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Neural responses to *mere* inconsistencies in behavior

Almost all recent neuroimaging investigations of impression updating have identified the dmPFC as playing a role in updating (Ames et al., 2013; Baron et al., 2011; Cloutier et al., 2011b; Ma et al., 2012; Mende-Siedlecki et al., 2013a; Mende-Siedlecki et al., 2013b). For example, Ma and colleagues observed activity in the dmPFC in response to inconsistent trait inferences, as well as similar activity in more posterior aspects of frontal cortex, including posterior medial frontal cortex (pmFC; e.g., Brodmann area 6) and dorsal anterior cingulate (dACC; e.g., Brodmann area 32), more typically associated with conflict monitoring and control (Ma et al., 2012). Likewise, our initial investigation of updating also observed a large cluster of updating-related activity encompassing the dmPFC, as well as dACC and pmFC (Mende-Siedlecki et al., 2013a). Additional analyses determined that dACC and pmFC were preferentially recruited immediately upon introduction of inconsistent behavioral information, potentially reflecting surprise in response to the inconsistency.

On one hand, the anterior portion of prefrontal cortex typically defined as dmPFC has been well-studied within the social neuroscience literature for its role in social cognition in general, (for review, see Amodio & Frith, 2006; Frith & Frith, 2006; Van Overwalle, 2009), and impression formation in particular (Cloutier et al., 2011a; Ferrari et al., in press; Gilron & Gutchess, 2012; Mitchell, et al., 2004; 2005; 2006; Schiller, et al., 2009), so its involvement in updating person representations is not surprising. On the other hand, more posterior regions of the mPFC like the dACC are linked with cognitive processes that are not explicitly social, such as conflict monitoring (e.g., Botvinick et al., 1999; Botvinick, Cohen, & Carter, 2004; Kerns et al., 2004), error monitoring (e.g., Carter et al., 1998; Kiehl, Liddle, & Hopfinger, 2000; van Veen et al., 2004), expectancy violation (e.g., Bolling et al., 2011; Hayden et al, 2011; Somerville, Heatherton, & Kelley, 2006), and uncertainty (e.g., Behrens et al., 2007; Cavanagh et al., 2012). Several integrative models have attempted to link these disparate interpretations (e.g., Alexander & Brown, 2011; Botvinick, 2007; Nee, Kastner, & Brown, 2011; Shenhav, Botvinick, & Cohen, 2013). One such account suggests that this region can be understood as an action-outcome predictor (Alexander & Brown, 2011), whose signals reflect “negative surprise” triggered by actions failing to generate an expected outcome. In this framework, activity in dACC during updating may not reflect an updated representation of a given individual’s character, but rather the fact that their behavior is inconsistent with what immediately preceded it.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

We note a similar tension arising with regards to parietal areas implicated in previous investigations of updating—IPL and TPJ (Bhanji & Beer, 2013; Cloutier et al., 2011b; Ma et al., 2012; Mende-Siedlecki et al., 2013a; Mende-Siedlecki et al., 2013b). These geographically neighboring but functionally distinct regions are traditionally associated with social cognitive processes that are likely instrumental to the task of impression updating. On the one hand, the IPL is observed to support the perception and understanding of intentional actions (Fogassi, Ferrari, Gesierich, Rozzi, Chersi, & Rizzolatti, 2005; Gallese, Keysers, & Rizzolatti, 2004; Iacoboni, Molnar-Szakacs, Gallese, Buccino, Mazziotta, & Rizzolatti, 2005; Montgomery & Haxby, 2008), in addition to working memory (Champod & Petrides, 2010; McNab et al., 2008; Van Hecke et al., 2010; Vergauwe, Hartstra, Barrouillet, & Brass, 2015). On the other hand, the TPJ is typically associated with theory-of-mind computations regarding the beliefs and motives of others (Decety & Grèzes, 2006; Samson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005; Spreng & Grady, 2010; Sommer, Döhnel, Sodian, Meinhardt, Thoermer, & Hajak, 2007; Völlm, Taylor, Richardson, Corcoran, Stirling, McKie, Deakin, & Elliott, 2006; Young, Dodell-Feder, & Saxe, 2010)—though it is not necessarily selective as such (Chang, Hsu, Tseng, Liang, Tzeng, Hung, & Juan, 2013; Corbetta & Shulman, 2002; Corbetta, Patel, & Shulman, 2008; Geng & Vossel, 2013; Mitchell, 2008a; Serences, Shomstein, Leber, Golay, Egeth, & Yantis, 2005; Rothmayr, Sodian, Hajak, Döhnel, Meinhardt, & Sommer, 2011).

In the context of social perception, a number of studies have tested direct comparisons that shed light on the functional distinctions between IPL and TPJ. For example, in the same set of participants, IPL was more active during an action-

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

understanding task, while TPJ was preferentially engaged by a false belief task (Gobbini, Koralek, Bryan, Montgomery, & Haxby, 2007). Moreover, in a task designed to dissociate between representations of the implementation of and motives subserving actions, Spunt and colleagues demonstrated that activity in IPL reflects “how” processing (i.e., action identification) while activity in TPJ reflects “why” processing (i.e., attribution of intention based on actions; Spunt, Falk, & Lieberman, 2010; Spunt, Satpute, & Lieberman, 2011, Spunt & Lieberman, 2012a; Spunt & Lieberman, 2012b). Finally, during a task where participants attempted to infer the emotional states of targets based on conflicting contextual and nonverbal (i.e., facial expression) cues, activity in TPJ tracked the influence of context, while IPL tracked the influence of nonverbal information (Zaki, Hennigan, Weber, & Ochsner, 2010).

Ultimately, while the roles of TPJ and IPL can be functionally dissociated based on the level of abstraction at which they contribute to action understanding and representation, both regions likely support processing the *temporary* (versus dispositional) states of other people (for meta-analysis, see Van Overwalle, 2009). Therefore, recruitment of TPJ/IPL during impression updating may reflect intentionality-related computations tied to specific, unexpected behaviors, as opposed to more global representations of how individuals are generally likely to behave.

Finally, we highlight the PCC/precuneus as another set of regions implicated by previous investigations of updating (Cloutier et al., 2011b; Ma et al., 2012; Mende-Siedlecki et al., 2013a; Mende-Siedlecki et al., 2013b) that may be responding to moment-to-moment discrepancies in behavior. While the PCC/precuneus is associated with social cognition, the precise nature of that role is somewhat underspecified (Amodio

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

& Frith, 2006; Mitchell, 2008b; Schilbach et al., 2008; Van Overwalle, 2009; Wolf, Dziobek, & Heekeren, 2010). The PCC/precuneus is also a key node in the default mode network (Gusnard & Raichle, 2001; Greicius et al., 2003; Raichle et al., 2001; Utevsky, Smith, & Huettel, 2014), and is implicated in various cognitive processes ranging from autobiographical memory to the representation of subjective value (for review, see Pearson et al., 2011). A recent reconceptualization of the PCC's function suggests that the PCC is responsible for detecting changes in the environment and motivating behavior in order to adapt accordingly (Pearson et al., 2011). In the context of impression updating, PCC/precuneus activity may reflect these changes in behavior, as opposed to the updates themselves.

The present study

We devised a task intended to dissociate neural responses driven by mere moment-to-moment inconsistencies in behavior and neural responses driven by attributionally meaningful inconsistencies. Participants updated their impressions of individuals whose behavior was either inconsistent with respect to a specific, yet quotidian action, or whose behavior was more globally inconsistent within the domain of morality.

We chose to focus on morality because behavior reflecting moral character dominates impression formation, compared to sociability or competence (Brambilla et al., 2012; Brambilla & Leach, 2014; Goodwin, Piazza, & Rozin, 2014; Wojciszke, Bazinska, & Jaworski, 1998). Moreover, there are clear behavioral predictions regarding what moral behavior is most meaningful: immoral behaviors are more diagnostic than behaviors indicating high moral character (Fiske, 1980; Reeder & Brewer, 1979;

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Skowronski & Carlston, 1987; Uhlmann, Pizarro, & Diermeier, 2015; Wojciszke, Brycz, & Borkenau, 1993), and trigger larger changes in impressions, both explicitly (e.g., Ybarra, 2001) and implicitly (e.g., Cone & Ferguson, 2014; Mann & Ferguson, 2015). Finally, our previous neuroimaging work suggests that the dominance of immoral behavior in impression updating is evident on a neural level. Activity in left vlPFC and IFG showed preferential increases when participants updated impressions based on immoral behavior (Mende-Siedlecki et al., 2013b; see **Supplemental Materials** for a replication of this result using data from our initial neuroimaging investigation of updating—Mende-Siedlecki et al., 2013a).

With regards to the present study, we predicted 1) that a distributed network of regions observed in previous studies would support updating impressions of both merely and meaningfully inconsistent individuals. Moreover, we predicted that activity in specific components of this network would dissociate as a function of the nature of the inconsistency. While 2) posterior aspects of PFC (i.e., dACC), TPJ/IPL, and PCC should respond preferentially when updating based on "mere" inconsistencies, 3) vlPFC and IFG should respond preferentially when updating based on "meaningful" inconsistencies. Specifically, this latter subset of regions should respond most strongly when updating positive impressions based upon new negative information, compared to updating negative impressions based upon new positive information.

Alternatively, it is possible that regions previously observed to support impression updating are simply responding to expectancy violations in general. The merely inconsistent individuals presented in the present study offer the possibility of testing this alternate hypothesis, because the inconsistencies established within these behaviors

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

concern specific actions—they are, in a definitional sense, clearer examples of pure inconsistency. In the absence of a trait concept, "merely inconsistent" cases might potentially instantiate stronger expectations about a person's future behavior, since identical behaviors should be more predictive of one another than behaviors that are abstractly related, but yet concretely quite different. For example, someone's bedtime on Monday night should predict their bedtime on Tuesday night, more so than "rescuing a kitten from a tree" should predict "donating to charity". That being said, trait concepts *do* typically guide social learning. As such, we predict that responses in brain regions previously identified to support impression updating in response to diagnostic trait information are unlikely to be explained by such low-level, statistical inconsistencies. Ultimately, this experiment gives a strong test of whether these regions are truly responsive to trait-relevant impression updates as conceptualized in social psychology, as opposed to prediction and expectancy violation as studied elsewhere in cognitive neuroscience.

Materials and Methods

Participants

Twenty-one (13 female) participants, ages 18 to 31 (*M* = 22.5, *SD* = 3.36), volunteered and received $30 for participation. (This sample size is consistent with our previous investigations of impression updating: N=24, Mende-Siedlecki et al., 2013a; N=23, Mende-Siedlecki et al., 2013b.) Participants were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological illnesses or abnormalities. We acquired informed consent for participation approved by the

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Institutional Review Board for Human Subjects at Princeton University, and debriefed participants upon completion.

Face and behavior stimuli

Each participant completed an in-scanner task (adapted from Mende-Siedlecki et al., 2013a, Mende-Siedlecki et al., 2013b) in which they learned about the behavior of a series of individuals. Each participant saw 50 male and female faces from the Karolinska Directed Emotional Faces set (Lundqvist, Flykt, & Ohman, 1998), paired with sets of behaviors either previously rated on kindness (Fuhrman, Bodenhausen, & Lichtenstein, 1989), or designed specifically for this paradigm to instantiate an inconsistency regarding a specific, everyday behavior.

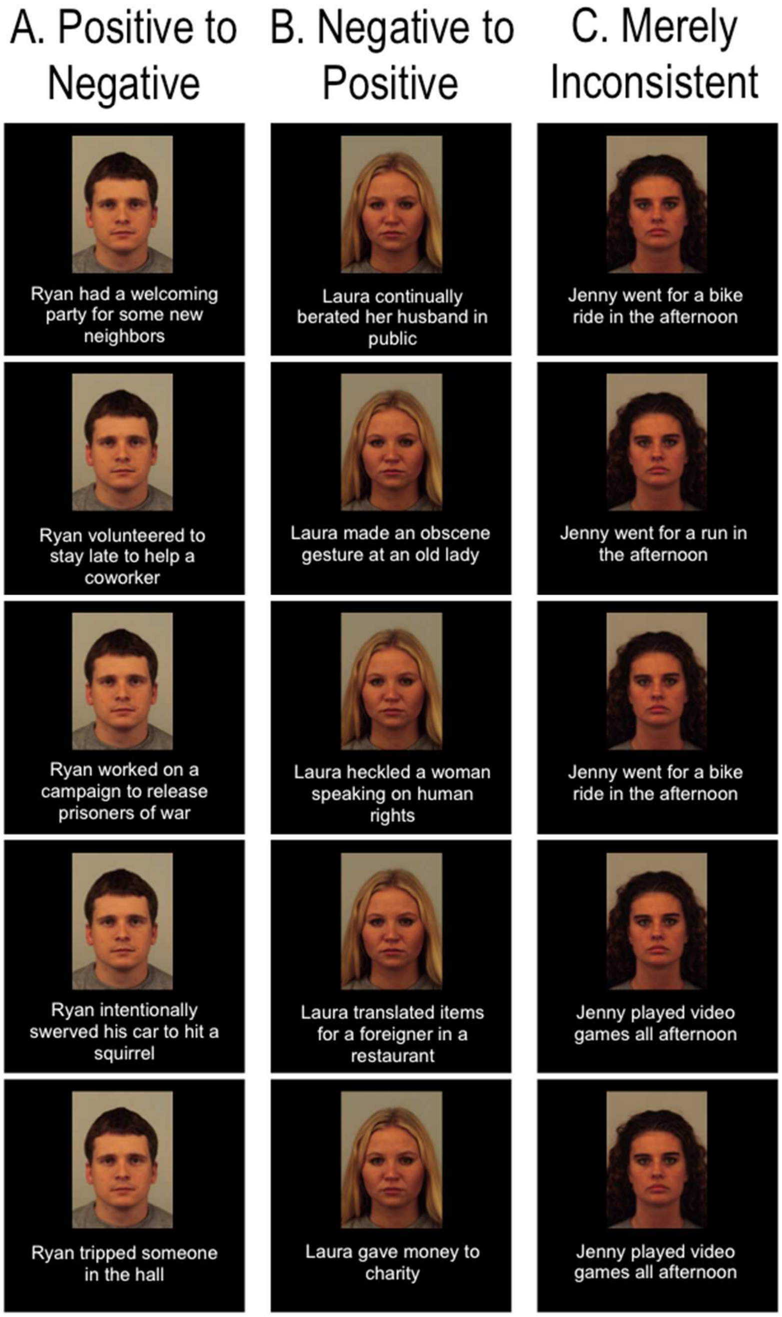
Each individual was represented by a male or female face paired with five consecutively-viewed behaviors. These sets of behaviors were internally inconsistent in a manner that elicits impression updating. “Meaningfully inconsistent” individuals were paired with behaviors that reflected moral character. These blocks of trials comprised a face paired with either three positive behaviors, followed by two negative behaviors (“Positive□to-Negative”, **Figure 1A**) or three negative behaviors, followed by two positive behaviors (“Negative-to□Positive”, **Figure 1B**). Positive-to-negative changes were expected to be more diagnostic of an individual’s moral character than negative-to-positive changes. For “merely inconsistent” individuals, information presented on the last two trials again conflicted with information seen during the first three trials, but in a way that was unlikely to trigger an updated representation of the individual’s moral character

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

(**Figure 1C**[1]). As in previous studies, participants also saw control individuals—faces paired with a sentence indicating the individual's name (i.e., "This man's name is Ron.")—to control for low-level stimulus attributes (e.g. faces paired with text on screen). In total, participants encountered 50 individuals—20 meaningfully inconsistent (10 Positive-to-Negative, 10 Negative-to-Positive), 20 merely inconsistent, and 10 control individuals.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

[1] While we originally intended for "merely inconsistent" sequences of behaviors to focus solely on one specific discrepancy in behavior (i.e., three instances of the same behavior, followed by two instances of a second, inconsistent behavior), behavioral pilot testing suggested that this design might be too monotonous, especially for a scanner task. As such, we introduced a degree of variability in these sequences, such that the second behavior was always a slight variation on the first and third. For example, someone described on Trials 1 and 3 as going to bed at 3 A.M., was described as having a 2:30 A.M. bedtime on Trial 2.



A. Positive to Negative

Ryan had a welcoming party for some new neighbors

Ryan volunteered to stay late to help a coworker

Ryan worked on a campaign to release prisoners of war

Ryan intentionally swerved his car to hit a squirrel

Ryan tripped someone in the hall

B. Negative to Positive

Laura continually berated her husband in public

Laura made an obscene gesture at an old lady

Laura heckled a woman speaking on human rights

Laura translated items for a foreigner in a restaurant

Laura gave money to charity

C. Merely Inconsistent

Jenny went for a bike ride in the afternoon

Jenny went for a run in the afternoon

Jenny went for a bike ride in the afternoon

Jenny played video games all afternoon

Jenny played video games all afternoon

Figure 1. Sample individuals from the positive-to-negative, negative-to-positive, and merely inconsistent conditions. Each individual face/behavior pair was presented on screen for 6 seconds, followed by a 6 second fixation cross. Once all five behaviors were presented, participants provided global ratings of each individual's trustworthiness (4 seconds) and surprisingness (4 seconds).

Procedures

We informed participants that they were participating in a study on impression formation, in which they would see a series of faces paired with behaviors. Participants

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

were asked to form an impression of each person, and were told that some information might run contrary to the impression they had formed so far. Finally, we told participants that picturing individuals performing behaviors might aid in forming impressions, and that they would make global ratings of trustworthiness and surprise, which should index their overall impression of each individual, taking into account everything they had learned about that person. Participants practiced one run of the task outside the scanner, where they encountered five individuals—comprising faces and behaviors not used in the scanner task.

The scanner task was fundamentally similar to those in our previous neuroimaging investigations of updating, with one critical change. While earlier studies included trial-by-trial ratings following *each* behavior, the present study replaced these trial-by-trial ratings with global ratings of each individual following the final behavior of each sequence. It is possible that trial-by-trial ratings employed in our previous studies (Mende-Siedlecki et al., 2013a, Mende-Siedlecki et al., 2013b) may impose an unrealistic demand to continually monitor one's impression of each individual and exaggerate neural activity associated with updating. Therefore, we felt it necessary to assess these evaluations more globally. While it is likely that no ratings need be required for updating to occur, these global ratings a) facilitated and indexed participants' task engagement, and b) instilled an explicit impression formation goal in our participants. However, given the global nature of these ratings, they cannot shed light on the magnitude of updates triggered by a given individual's inconsistent behavior. As such, we have chosen not to focus on these behavioral results, though they are reported in **Supplementary Materials.**

Each individual comprised five consecutive face/behavior pairs. Each single

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

face/behavior pair was presented on screen for 6 seconds, followed by a fixation cross (6 seconds). Following the fifth pair, participants rated each individual on trustworthiness (4 seconds, “How trustworthy is Sarah?”, 4-point scale, 1 = very untrustworthy, 4 = very trustworthy) and surprise (4 seconds, “How surprising is Sarah?”, 4-point scale, 1 = not at all surprising, 4 = very surprising). (Rating order was counterbalanced between subjects. Half always rated trustworthiness first and half always rated surprise first.) Participants were instructed that when rating surprise, they should take into account how consistent or predictable an individual was. Presentation order was pseudorandomized to ensure that one representative individual of each condition appeared per scanner run.

Imaging acquisition

Blood oxygenation level-dependent (BOLD) signal was used as a measure of neural activation. Echo planar images (EPI) were acquired using a Siemens 3.0 Tesla Allegra head-dedicated scanner (Siemens, Erlangen, Germany) with a standard ‘bird-cage’ head coil at a resolution of 3 × 3 × 4mm (TR=2000 ms, TE=30 ms, flip angle=80˚, matrix size=64x64). By using 32 interleaved 3-mm axial slices we achieved near whole-brain coverage. Prior to the primary data acquisition scans, a high-resolution anatomical image (T1-MPRAGE, TR=2500ms, TE=4.3 ms, flip angle=8˚, matrix size=256x256) was acquired for subsequent registration of functional activity to the participant’s anatomy and for spatially normalizing data across participants.

Imaging analyses

All fMRI data analysis was conducted using Analysis of Functional NeuroImages software (Cox, 1996). The first four EPI images of each run were discarded to allow signal to reach steady-state equilibrium. After slice scan-time correction, participants’

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

motion was corrected using a six-parameter 3D motion-correction algorithm. Transient spikes were removed from the signal using the AFNI program 3dDespike. Subsequently, data were low-pass filtered with a frequency cut-off of 0.1 Hz following spatial smoothing with a 6-mm full-width at half-maximum (FWHM) Gaussian kernel. Anatomical data was then aligned to unsmoothed functional data using the AFNI program align_epi_anat.py, and consequently transformed to Talairach space (Talairach and Tournoux, 1998) using the function @auto_tlrc. Finally, functional datasets were subjected to the same spatial transformation.

To generate parameter estimates, we performed voxel-wise multiple regression on each participant's preprocessed imaging data. Fifteen regressors of interest (5 6000-ms trials per individual × 4 types of individual: positive-to-negative, negative-to-positive, merely inconsistent, and control) were convolved with a canonical hemodynamic response function and entered into our general linear model (GLM). Additionally, we included several regressors of no interest, including head motion estimates and time points representing rating slide presentations.

Whole-brain analyses. First, we identified regions displaying increased activity during the last two (L2) trials, compared to the first three (F3) trials (L2>F3 contrast), collapsing across all individual conditions (positive-to-negative, negative-to-positive, merely inconsistent, and control) in a whole-brain contrast. This contrast identifies regions displaying a main effect of updating.

Moving forward, our primary goal was to identify functional regions of interest (fROIs) where updating activity differed as a function of condition. First, for each participant, we created maps comprising whole-brain activity within F3 and L2 trials,

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

separately, for each of the four individual conditions, resulting in eight maps per participant. These maps were submitted to a 2 (updating: L2 vs. F3) × 4 (condition: positive-to-negative, negative-to-positive, merely inconsistent, control) ANOVA (whole-brain) using the AFNI command line program 3dANOVA3 to identify brain regions displaying a) a main effect of updating (e.g., increased activity during L2 trials, compared to F3 trials—L2>F3), and b) more importantly, an interaction between updating and condition. In addition, 3dANOVA3 identified fROIs displaying c) a main effect of condition, reported in **Supplemental Materials**.

Unless otherwise noted, correction for multiple comparisons was performed using the program AlphaSim, which is part of the AFNI package. The Monte Carlo simulation indicated that a minimum cluster extent threshold of 15 voxels was needed to attain a corrected significance of p<.05, at a voxel-wise height threshold of p<.001.

Extracting parameter estimates from fROIs. 3dANOVA3 computes F-statistic maps indicating where activity differs significantly as a function of condition, or the interaction between conditions. To interpret these maps, parameter estimates must be extracted and plotted graphically. Since a) the main effect of updating is already tested by the L2>F3 contrast described above—which, as it is a directional test, does not require the extraction of parameter estimates for ease of interpretation—we report the results of the L2>F3 contrast for parsimony's sake in the results below. However, to interpret activity within fROIs showing b) an interaction between updating and condition (as well as, c) a main effect of condition), we extracted and analyzed parameter estimates from the F-statistic maps produced by 3dANOVA3, using the AFNI command line program 3dcalc.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Additional analyses. As a confirmatory step, we performed a series of targeted contrasts aimed at identifying clusters where activity increased from the F3 to L2 trials preferentially for one specific condition. Our primary interests were regions where activity changed preferentially in response to a) meaningful inconsistencies (e.g., inconsistent negative behavior, compared to inconsistent positive or merely inconsistent behavior) or b) mere inconsistencies (e.g., neutral but surprising behavioral information, compared to inconsistent negative or positive behavior). However, we also tested for regions where activity changed preferentially in response to c) inconsistent positive behavior, compared to inconsistent negative or merely inconsistent behavior or d) either inconsistent negative *or* positive behavior, compared to mere inconsistency. These data are thresholded as described above: corrected for multiple comparisons (*p*<.05) with a cluster-extent threshold of 15 voxels (determined by AlphaSim), at a voxel-wise threshold of *p*<.001.

Results

Neuroimaging results

Main effect of updating. We began by performing a whole-brain analysis testing the main effect of updating (L2>F3, *p*(corrected) < .05). This contrast allowed us to isolate regions recruited during updating, independent of condition, and to replicate the results of our previous neuroimaging investigations of updating. We observed a large set of regions that displayed an enhanced BOLD response during L2 trials, compared to F3 trials, including dmPFC, bilateral rostrolateral PFC (rlPFC), bilateral vlPFC extending through IFG, bilateral caudate nucleus, bilateral anterior temporal lobe (ATL), bilateral STS, bilateral IPL/TPJ, precuneus, and PCC (**Supplementary Table 1, Figure 2**).

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

(Regions showing an enhanced BOLD response during F3 trials are also detailed in **Supplementary Table 1**. See **Supplementary Materials** and **Supplementary Table 2** for information regarding regions displaying a main effect of condition.) Moreover, extensive clusters of activity in occipital and inferotemporal cortex (including bilateral fusiform gyrus) displayed an enhanced BOLD response during F3 trials, compared to L2 trials.

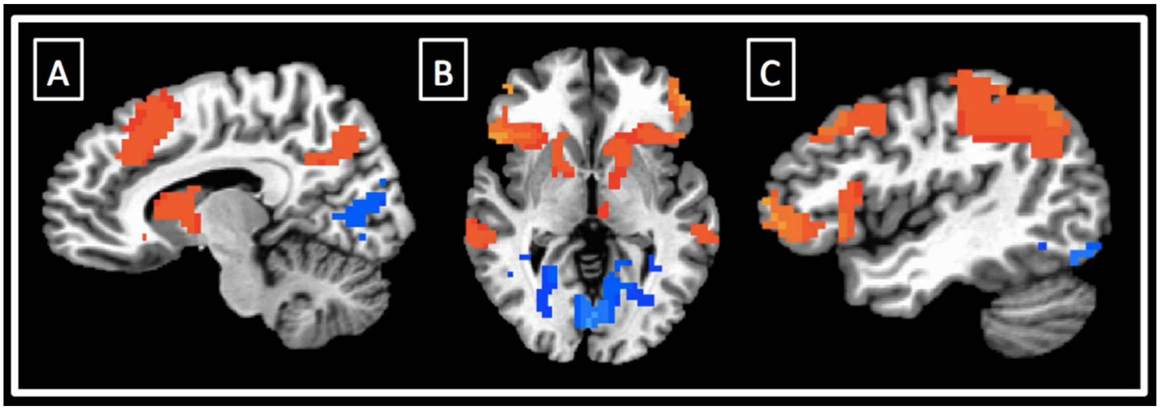


Figure 2. Main effect of updating (last two > first three trials, collapsed across individual type). We observed an extended network of regions that was recruited in response to inconsistent information presented during the last two trials (pictured in hot colors), including A) dmPFC, PCC/precuneus, B) bilateral rlPFC, bilateral vlPFC, bilateral caudate, bilateral STS, C) bilateral ATL, and bilateral TPJ/IPL. (Regions pictured in cool colors—including bilateral fusiform gyrus and cuneus—showed preferential responses to information presented on the first three trials, before inconsistent information was introduced.)

Interaction between updating and condition. Next, we tested for clusters where activity differed as a function of both updating and condition. This contrast allowed us to identify fROIs that might be preferentially recruited by updating impressions based on a specific type of inconsistency (i.e., meaningful vs. mere inconsistency). We identified 20 such regions (see **Table 1**), extracted parameter estimates from each, and categorized these regions based on the patterns we observed.

Left IFG, left vlPFC, & dmPFC (**Figure 3A, Table 1A**), in addition to three other fROIs, showed preferentially higher activity during L2 trials of positive-to-negative individuals, compared to merely inconsistent or control individuals (which, in turn,

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

showed higher activity compared to negative-to-positive individuals). Additionally, right mid-STS, left posterior STS (extending into TPJ), and pmFC (**Figure 3B, Table 1B**) showed preferentially higher activity during L2 trials of positive-to-negative, merely inconsistent, and control individuals, compared to negative-to-positive individuals.

Furthermore, dACC, mid-ACC, & PCC (**Figure 3C, Table 1C**), in addition to four other regions, showed preferentially higher activity during L2 trials of merely inconsistent and control individuals, compared to either negative-to-positive or positive-to-negative individuals. Finally, four regions (**Table 1D**) showed preferentially higher activity during L2 trials of control individuals, compared to merely inconsistent individuals (which, in turn, showed higher activity compared to both negative-to-positive or positive-to-negative individuals). See **Figure 3** for parameter estimates extracted from key regions identified by the updating-by-condition contrast. For full details on extracted parameter estimates, see the expanded version in **Supplementary Figure 1**.

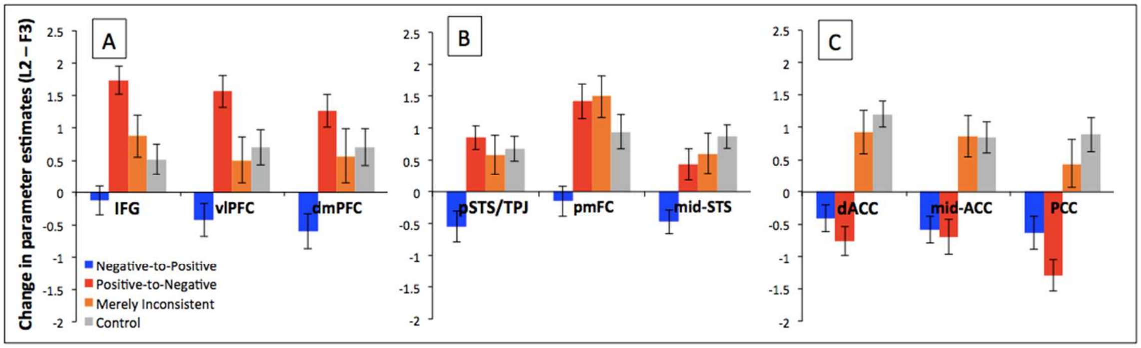


Figure 3. Parameter estimates extracted from regions displaying an interaction between updating and condition. Colored bars indicate the change in activity from the first three to the last two (i.e., update-provoking) trials. Here, we provide information on selected key regions emerging from the updating-by-condition interaction in which A) the updating-related change in activity was strongest for positive-to-negative individuals, B) this change in activity was stronger for positive-to-negative, merely inconsistent, and control individuals, compared to negative-to-positive individuals, and C) this change in activity was stronger for merely inconsistent and control individuals, compared to individuals from either meaningfully inconsistent condition. (Blue = negative-to-positive, red = positive-to-negative, orange = merely inconsistent, gray = control.) Error bars indicate +/-1 standard error. (For an expanded figure including all 20 regions identified by the updating-by-condition interaction contrast, see **Supplementary Figure 1.)**

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Preferential activity towards updates in response to meaningful inconsistencies. Next, we sought to confirm the results of the interaction contrast described above by performing targeted contrasts designed to isolate regions where updating-related activity was especially strong in response to one particular kind of inconsistency.

First, we tested for regions showing a preferential increase in activity when updating based on inconsistent negative behavioral information, compared to inconsistent positive, merely inconsistent, or control trials. A whole-brain contrast revealed that left IFG and left vlPFC, as well as left middle frontal gyrus (mFG) all displayed a bias towards updating based on immoral behaviors (**Table 2A**, **Figure 4A**).

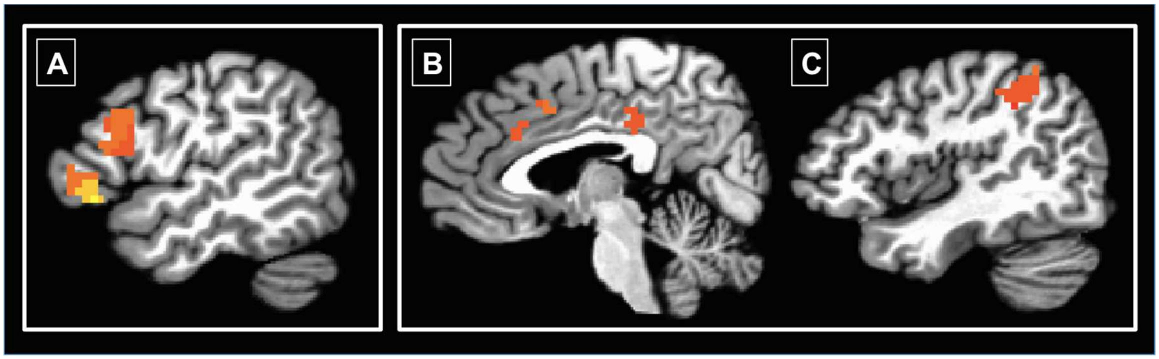


Figure 4. Preferential neural responses to meaningful and mere inconsistency. Left vlPFC and left IFG (Panel A), showed enhanced activity when updating based on immoral behaviors, compared to updates based on moral or merely inconsistent behaviors. dACC, mid-ACC, PCC (Panel B), and right TPJ/IPL (Panel C) all showed enhanced activity when updating based on merely inconsistent behaviors, compared to updates based on moral or immoral behaviors.

Preferential activity towards updates in response to mere inconsistencies. Next, we tested for regions showing a preferential increase in activity when updating based on mere inconsistencies, compared to inconsistent positive or negative behaviors, or control trials (**Table 2B**, **Figure 4B-C**). A whole-brain contrast revealed twelve regions displaying a pattern of mere inconsistency-specific activity, including dACC,

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

mid-ACC, and PCC/precuneus, as well as bilateral TPJ/IPL.

Additional contrasts. Finally, we tested whether any regions showed a preferential response to a) updating based on inconsistent positive behavioral information, compared to negative, merely inconsistent, and control trials, or b) updating based on *either* positive or negative behaviors, compared to merely inconsistent behaviors or control trials. Neither contrast revealed any significant clusters of activity.

Discussion

Our results suggest a dissociation between brain regions supporting behavior-based impression updating. While several regions (e.g., left vlPFC and IFG) displayed enhanced responses to meaningful, negative information particularly diagnostic of moral character, an additional subset of regions (e.g., dACC, TPJ/IPL, PCC) responded preferentially to inconsistencies of a more immediate, yet mundane nature.

Multiple analyses suggested that left vlPFC and left IFG displayed preferentially stronger activity when updating initially positive impressions of moral character with new negative information, relative to negative-to-positive updates, updates based on mere inconsistency, or control trials. This negativity bias during updating based on moral behavior is consistent with previous literature (Reeder & Coovert, 1986; Reeder & Spores, 1983; Skowronski & Carlston, 1987; Wojciszke, Brycz, & Borkenau, 1993; Wojciszke, 2005). These results also dovetail with our own neuroimaging investigations of updating. First, while our initial study did not test for interactions between valence and updating (Mende-Siedlecki et al., 2013a), re-analysis of these data confirms left vlPFC and left IFG showed preferential responses when updating based on immoral behaviors, compared to moral behaviors (see **Supplementary Materials, Supplementary Figure 2,**

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Supplementary Table 3). Moreover, a subsequent investigation comparing updating based on moral behaviors versus behaviors related to ability (Mende-Siedlecki et al., 2013b) determined that bilateral vlPFC and left IFG (as well as left STS) responded preferentially to diagnostic information from either domain (e.g., immoral behaviors and competent behaviors).

This previous study suggested that rather than displaying a pervasive negativity bias during updating, the brain is particularly responsive to behaviors that are perceived to be statistically infrequent in the environment—negative behaviors in the morality domain and positive behaviors in the ability domain. The diagnosticity of a given behavior, and thus its meaningfulness, is an emergent property of perceived frequency: less common behaviors have a stronger impact on our impressions of others. We interpret the present data through this lens, as negative, immoral behaviors are robustly perceived to occur with less frequency than their positive counterparts (Funder et al., 1987; Kanouse & Hanson, 1972; Mende-Siedlecki, Baron, & Todorov, 2013b; Rothbart & Park, 1986; Tausch et al., 2007). In a sense, echoing the work of Harold Kelley (Kelley, 1967; Kelley, 1973), one key determinant of what makes a behavior attributionally meaningful is its distinctiveness. Immoral behavior provides more informational value about the true character of an individual (Uhlmann, Pizarro, & Diermeier, 2015), and likely offers more predictive capability regarding their future behavior.

One interesting implication of the frequency account is that diagnosticity should be sensitive to context, varying from environment to environment. In an environment where deceptive, anti-social, or criminal behavior is common, learning that someone has behaved immorally may carry less informational value (e.g., Barclay, 2008;

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Peysakhovich & Rand, 2015). In such environments, positive, moral behaviors may be perceived as less frequent, and might therefore lead to stronger impression updates. By the same token, this logic extends to the sorts of behaviors employed by our "merely inconsistent" condition. In a context in which these behaviors are perceived to be particularly rare, or with a task framing that makes them more meaningful (e.g., learning about preferences), we might expect similar neural patterns of updating-related responses.

However, we do not mean to suggest that the contributions of the vlPFC and IFG to impression updating are "specifically social" in nature. Indeed, vlPFC and IFG have been extensively implicated in domain-general cognitive processes related to working memory and interference resolution (Jonides & Nee, 2006; Kan & Thompson-Schill, 2004; Thompson-Schill et al., 2002; Thompson-Schill, Bedny, & Goldberg, 2005). Additional work has clarified the contributions of these regions, distinguishing between the role of the anterior vlPFC (or IFG *pars orbitalis*, corresponding to the left vlPFC cluster observed in the present study) in the controlled, top-down retrieval of stored conceptual representations, and the role of the more posterior mid-vlPFC (IFG *pars triangularis*, corresponding to the left IFG cluster observed in the present study) in resolving competition between retrieved representations (Badre et al, 2005; Badre & Wagner, 2007; Souza, Donohue, & Bunge, 2009; Satpute, Badre, & Ochsner, 2014). Moreover, this region has been implicated in a process paralelling those isolated in the present task, yet without any social context—encoding changing representations of objects over time (Hindy, Altmann, Kalenik, & Thompson-Schill, 2012; Hindy, Solomon, Altmann, & Thompson-Schill, 2015). Ultimately, while we have observed evidence across three datasets that vlPFC and IFG support updating person impressions based upon

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

attributionally meaningful information, we would suggest that this represents an instance of a domain-general mechanism being recruited in service of a social process.

In contrast to the regions responding to attributionally meaningful behavioral information during updating, a separate subset of regions responded preferentially to mere inconsistencies in behavior. These included several clusters within the cingulate cortex (dACC, mid-ACC, PCC), as well as clusters in lateral parietal cortex comprising aspects of TPJ and IPL[2]. Moreover, we note that in general, these regions were less responsive to updates based on attributionally meaningful changes in behavior (though see Footnote 2). Taken together, these results suggest that in the context of updating person representations, these regions are particularly sensitive to changes in behavior that reflect more short-term states or intentions, as opposed to more dispositional, trait-level inconsistences, consistent with prevailing interpretations of the functions of these regions (e.g., outcome-specific surprise in dACC, Alexander & Brown, 2011; theory-of-mind computations in TPJ, Van Overwalle, 2009).

One might argue that a limitation of this investigation stems from potential differences in arousal across the different types of individuals participants learned about in the updating task. Indeed, while individuals in the merely inconsistent condition were rated as being significantly more surprising than the control condition, they were also rated as being less surprising than either the positive-to-negative or negative-to-positive conditions (see **Supplementary Materials**). If it were simply the case that no regions

[2] Though the targeted contrast yielded clusters in bilateral TPJ/IPL responding preferentially to merely inconsistent individuals, the updating-by-condition interaction did not produce activations in this vicinity. However, the interaction contrast identified a larger cluster comprising both left STS and TPJ, where updating-related activity did not differ between the positive-to-negative, negative-to-positive, and merely inconsistent conditions.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

responded to the mere inconsistencies presented in this task, an arousal explanation might be more compelling. However, we observed a number of such regions (e.g., dACC, TPJ/IPL, PCC), which we predicted would track specific, moment-to-moment changes in behavior based on previous accounts of their function (see **Introduction**). Moreover, for the difference in surprise ratings to explain the neural differences between responses to mere and meaningful inconsistency, there should also be a difference in surprise ratings between positive-to-negative and negative-to-positive individuals. Indeed, despite the differences in updating-related activity between positive-to-negative and negative-to-positive individuals, both types of individuals were rated as similarly surprising (see **Supplementary Materials**).

Regardless, we note that using global ratings did not produce divergent neuroimaging results compared to previous studies employing trial-by-trial ratings. This convergence suggests trial-by-trial ratings did not artificially inflate our previous imaging results by imposing an unnatural demand to update, and gives additional confidence in the results of our prior work (Mende-Siedlecki et al., 2013a, 2013b).

Conclusions

Ultimately, these data offer an answer to the question posed at the outset—whether activity in regions involved in impression updating reflects an updated trait representation of a given individual, or merely a response to an immediate discrepancy in behavior. The results demonstrate that when learning new information about other people, a subset of regions involved in updating impressions (left vlPFC and IFG) responds preferentially to attributionally meaningful information—particularly behaviors indicative of low moral character. However, a separate subset of regions (dACC,

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

IPL/TPJ, PCC) responds more strongly to more mundane, moment-to-moment inconsistencies in behavior. While this study marks a positive first step in understanding the different yet complementary roles of regions involved impression updating, further work should continue to pursue more fine-grained specifications of the computational contributions of these regions.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Acknowledgements

The authors thank Ashley Halkett for assistance in developing and piloting materials used in this study, Elizabeth Carnahan, Rui Costa, and Friederike Funk for assistance in data collection, and Virginia Falvello for assistance with subject recruitment. The authors also acknowledge a special debt of gratitude to Hillel Aviezer, whose helpful suggestions inspired this line of inquiry, and to Leor Hackel, for persistently illuminating comments and advice. This research was supported by National Science Foundation grant BCS-0823749 to A.T. and National Science Foundation grant DGE-1148900 to P.M-S.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Figure Legends

Figure 1. Sample individuals from the positive-to-negative, negative-to-positive, and merely inconsistent conditions. Faces and behaviors were presented on screen together for 6 seconds, followed by a 6 second fixation cross. Once all five behaviors were presented, participants provided global ratings of each individual's trustworthiness (4 seconds) and surprisingness (4 seconds).

Figure 2. Main effect of updating (last two > first three trials, collapsed across individual type). We observed an extended network of regions that was recruited in response to inconsistent information presented during the last two trials (pictured in hot colors), including A) dmPFC, PCC/precuneus, B) bilateral rlPFC, bilateral vlPFC, bilateral caudate, bilateral STS, C) bilateral ATL, and bilateral TPJ/IPL. (Regions pictured in cool colors—including bilateral fusiform gyrus and cuneus—showed preferential responses to information presented on the first three trials, before inconsistent information was introduced.)

Figure 3. Parameter estimates extracted from regions displaying an interaction between updating and condition. Colored bars indicate the change in activity from the first three to the last two (i.e., update-provoking) trials. Here, we provide information on selected key regions emerging from the updating-by-condition interaction in which A) the updating-related change in activity was strongest for positive-to-negative individuals, B) this change in activity was stronger for positive-to-negative, merely inconsistent, and control individuals, compared to negative-to-positive individuals, and C) this change in activity was stronger for merely inconsistent and control individuals, compared to individuals from either meaningfully inconsistent condition. (Blue = negative-to-positive,

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

red = positive-to-negative, orange = merely inconsistent, gray = control.) Error bars indicate +/-1 standard error. (For an expanded version of this figure including all 20 regions identified by the updating-by-condition interaction contrast, see **Supplementary Figure 1.)**

Figure 4. Preferential neural responses to meaningful and mere inconsistency. Left vlPFC and left IFG (Panel A), showed enhanced activity when updating based on immoral behaviors, compared to updates based on moral or merely inconsistent behaviors. dACC, mid-ACC, PCC (Panel B), and right TPJ/IPL (Panel C) all showed enhanced activity when updating based on merely inconsistent behaviors, compared to updates based on moral or immoral behaviors.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Tables

Table 1. Regions showing a significant interaction between updating and condition.

Region	Hemi	x	y	z	#Voxels
A. Positive-to-Negative > Merely Inconsistent & Control > Negative-to-Positive.					
IFG	L	-56	20	18	123
vlPFC	L	-47	23	-7	37
mFG	L	-41	-2	51	56
ATL	L	-41	8	-31	25
anterior temporal pole	L	-53	14	-10	21
dmPFC	-	-5	59	30	20
B. *Positive-to-Negative, Negative-to-Positive, Merely Inconsistent > Control*					
pSTS/TPJ	L	-59	-53	12	331
mid-STS	R	65	-41	3	64
posterior medial frontal cortex	-	-2	-2	66	58
C. *Merely Inconsistent & Control > Positive-to-Negative & Negative-to-Positive*					
dACC	-	2	23	30	85
mid-ACC	-	-2	2	42	80
superior frontal gyrus	L	-35	44	30	34
ATL	L	-56	5	-1	32
PCC	-	5	-38	48	29
mid-ACC	-	2	11	39	23
cerebellum	R	17	-35	-40	19
D. *Control > Merely Inconsistent > Positive-to-Negative & Negative-to-Positive*					
calcarine sulcus/cerebellum	-	-5	-56	9	1119
mPFC	-	2	56	12	39
parahippocampal gyus	R	29	-26	-22	26
mid-STS	R	65	-20	12	17

Group results (N = 21), corrected for multiple comparisons at a voxel-wise threshold of p<.001 and a cluster-extent threshold of 15 voxels (determined by AFNI's AlphaSim package). Coordinates refer to the peak voxel in Talairach space, and are rounded to the nearest integer. For each cluster, we report its hemisphere (Hemi) and size in voxels (#Voxels). Note: Rather than referring to specific contrasts run, Table 1 sub-headings are meant to be descriptive of the general patterns of parameter estimates extracted from each region displaying an interaction between updating and condition.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Table 2. Regions showing dissociation between updating based on meaningful information and a response to mere inconsistency

Region	Hemi	x	y	z	#Voxels
A. L2>F3$_{Positive\text{-}to\text{-}Negative}$ *> L2>F3*$_{Negative\text{-}to\text{-}Positive}$*, L2>F3*$_{Mere\ Inconsistency}$*, & L2>F3*$_{Control}$					
IFG	L	-56	20	18	98
vlPFC	L	-50	26	-4	47
mFG	L	-41	-2	51	19
B. L2>F3$_{Mere\ Inconsistency}$ *> L2>F3*$_{Positive\text{-}to\text{-}Negative}$*, L2>F3*$_{Negative\text{-}to\text{-}Positive}$*, & L2>F3*$_{Control}$					
TPJ/IPL	R	44	-50	48	66
TPJ/supramarginal gyrus	L	-44	-32	36	45
PCC	-	2	-29	30	31
dlPFC	L	-35	26	36	30
superior frontal gyrus	R	26	8	57	28
precuneus	R	14	-50	42	22
precuneus	R	11	-72	39	21
dACC	-	2	23	30	16
TPJ/IPL	L	-47	-53	39	16
mid-ACC	-	2	11	39	16

Group results (N = 21), corrected for multiple comparisons at a voxel-wise threshold of p<.001 and a cluster-extent threshold of 15 voxels (determined by AFNI's AlphaSim package). Coordinates refer to the peak voxel in Talairach space, and are rounded to the nearest integer. For each cluster, we report its hemisphere (Hemi) and size in voxels (#Voxels).

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

References

Alexander, W.H., Brown, J.W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, 14, 1338-1344.

Ames, D.L., Fiske, S.T. (2013). Outcome dependency alters the neural substrates of impression formation. *NeuroImage*, 83, 599-608.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268-277.

Badre, D., Wagner, A.D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*, 15, 2003–2012.

Badre, D., Wagner, A.D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, 45, 2883-2901.

Barclay, P. (2008). Enhanced recognition of defectors depends on their rarity. *Cognition*, *107*(3), 817-828.

Baron, S.G., Gobbini, M.I., Engell, A.D., Todorov, A. (2011). Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Social Cognitive and Affective Neuroscience*, 6, 572-581.

Behrens, T.E.J., Hunt, L.T., Woolrich, M.W., Rushworth, M.F.S. (2008). Associative learning of social value. *Nature*, 456, 245-249.

Bhanji, J.P., Beer, J.S. (2013). Dissociable neural modulation underlying lasting first impressions, changing your mind for the better, and changing it for the worse. *The Journal of Neuroscience*, *33*, 9337-9344.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Bolling, D.Z., Pitskel, N.B., Deen, B., Crowley, M.J., McPartland, J.C., Mayes, L.C., Pelphrey, K.A. (2011). Dissociable brain mechanisms for processing social exclusion and rule violation. *NeuroImage*, 54, 2462-2471.

Botvinick, M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective and Behavioral Neuroscience*, 7, 356-366.

Botvinick, M.M., Cohen, J.D., Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8, 539-546.

Botvinick, M.M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., (1999). Conflict monitoring versus selection for action in anterior cingulate cortex. *Nature*, 402, 179-181.

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology*, *51*(1), 149-166.

Brambilla, M., Leach, C. W. (2014). On the importance of being moral: the distinctive role of morality in social judgment. *Social Cognition*, *32*(4), 397-408.

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, 280, 747-749.

Carter, R.M., Huettel, S.A. (2013). A nexus model of the temporal–parietal junction. *Trends in Cognitive Sciences*, *17*(7), 328-336.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Cavanagh, J.F., Figueroa, C.M., Cohen, M.X., Frank, M.J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, 22, 2575-2586.

Champod, A. S., & Petrides, M. (2010). Dissociation within the frontoparietal network in verbal working memory: a parametric functional magnetic resonance imaging study. *The Journal of Neuroscience*, *30*(10), 3849-3856.

Chang, C. F., Hsu, T. Y., Tseng, P., Liang, W. K., Tzeng, O. J., Hung, D. L., & Juan, C. H. (2013). Right temporoparietal junction and attentional reorienting. *Human Brain Mapping*, *34*(4), 869-877.

Cloutier, J., Gabrieli, J.D.E., O'Young, D., Ambady, N. (2011b). An fMRI study of violations of social expectations: when people are not who we expect them to be. *NeuroImage*, 57, 583-588.

Cloutier, J., Kelley, W.M., & Heatherton T.F. (2011a). The influence of perceptual and knowledge-based familiarity on the neural substrates of face perception. *Social Neuroscience*, 6, 63-75.

Cone, J., Ferguson, M. J. (2014). He Did What? The Role of Diagnosticity in Revising Implicit Evaluations. *Journal of Personality and Social Psychology*, 108(1), 37-57.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201-215.

Corbetta, M., Patel, G., Shulman, G L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*, 306-324.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Cox, R. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29, 162–73.

Decety, J., & Grezes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain Research*, *1079*(1), 4-14.

Ferrari, C., Lega, C., Vernice, M., Tamietto, M., Mende-Siedlecki, P., Vecchi, T., Todorov, A., & Cattaneo, Z. (in press). The Dorsomedial Prefrontal Cortex Plays a Causal Role in Integrating Social Impressions from Faces and Verbal Descriptions. *Cerebral Cortex*, doi: 10.1093/cercor/bhu186.

Fiske, S.T. (1980). Attention and weight in person perception: the impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38, 889-906.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G. (2005) Parietal lobe: From action organization to intention understanding. *Science*, 308, 662-667.

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

Fuhrman, R.W., Bodenhausen, G.V., Lichtenstein, M. (1989). On the trait implications of social behaviors: Kindness, intelligence, goodness, and normality ratings for 400 behavior statements. *Behavior Research Methods, Instruments, & Computers*, 21, 587–597.

Funder, D. C., Dobroth, K. M. (1987). Differences between traits: properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, *52*(2), 409.

Gallese, V., Keysers, C., Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8, 396–403.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Geng, J.J., Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: Contextual updating?. *Neuroscience & Biobehavioral Reviews*, *37*, 2608-2620.

Gilron, R., Gutchess, A.H. (2012). Remembering first impressions: Effects of intentionality and diagnosticity on subsequent memory. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(1), 85-98.

Gobbini, M.I., Koralek, A.C., Bryan, R.E., Montgomery, K.J., Haxby, J.V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, 19, 1803-1814.

Goodwin, G. P., Piazza, J., Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148.

Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences, U.S.A*, 100, 253-258.

Gusnard, D.A., Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews Neuroscience*, *2*(10), 685-694.

Hackel, L.M., Doll, B.B., & Amodio, D.M. (2015) Instrumental learning of traits versus reward: dissociable neural correlates and effects on choice. *Nature Neuroscience,* 18, 1233–1235.

Harris, L.T., Fiske, S.T. (2010). Neural regions that underlie reinforcement learning are also active for social expectancy violations. *Social Neuroscience*, *5*, 76-91.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., Platt, M.L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *Journal of Neuroscience*, 31, 4178-4187.

Hindy, N.C., Altmann, G.T.M., Kalenik, E., Thompson-Schill, S.L. (2012). The effect of object state-changes on event processing: do objects compete with themselves?. *Journal of Neuroscience*, 32, 5795-5803.

Hindy, N.C., Solomon, S. H., Altmann, G.T., Thompson-Schill, S.L. (2015). A cortical network for the encoding of object change. *Cerebral Cortex*, *25*(4), 884-894.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., et al. (2005) Grasping the intentions of others with one's own mirror neuron system. PLoS Biology, *3*(3), e79.

Jonides, J., Nee, D.E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, 139, 181-193.

Judd, C. M., James-Hawkins, L., Yzerbyt, V., Kashima, Y. (2005). Fundamental dimensions of social judgment: understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899.

Kan, I.P., Thompson-Schill, S.L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 466-482.

Kanouse, D. E., Hanson, L. R, Jr. (1972). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attribution: Perceiving the causes of behavior (pp. 47-62). Morristown, NJ: General Learning Press.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Kelley, H. H. (1967). Attribution theory in social psychology. In *Nebraska symposium on motivation*. University of Nebraska Press.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S. (2004). Anterior cingulate conflict: monitoring and adjustments in control. *Science*, 303, 1023-1026.

Kiehl, K.A., Liddle, P.F., Hopfinger, J.B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. *Psychophysiology*, *37*, 216-223.

Kim, H., Choi, M.J., Jang, I.J. (2012). Lateral OFC activity predicts decision bias due to first impressions during ultimatum games. *Journal of Cognitive Neuroscience*, 24, 428-439.

Lundqvist, D., Flykt, A., Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Ma, N., Vandekerckhove, M., Baetens, K., Van Overwalle, F., Seurinck, R., Fias, W., (2012). Inconsistencies in spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, 7, 937-950.

Mann, T. C., Ferguson, M. J. (2015). Can We Undo Our First Impressions? The Role of Reinterpretation in Reversing Implicit Evaluations. Journal of Personality and Social Psychology, 108(6), 823-49.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

McNab, F., Leroux, G., Strand, F., Thorell, L., Bergman, S., & Klingberg, T. (2008). Common and unique components of inhibition and working memory: an fMRI, within-subjects investigation. *Neuropsychologia*, *46*(11), 2668-2682.

Mende-Siedlecki, P., Cai, Y., Todorov, A. (2013a). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience*, 8, 623-631.

Mende-Siedlecki, P., Baron, S.G., Todorov, A. (2013b). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *The Journal of Neuroscience*, *33*, 19406-19415.

Mitchell, J.P. (2008a) Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex*, 18, 262-271.

Mitchell, J.P. (2008b). Contributions of functional neuroimaging to the study of social cognition. *Current Directions in Psychological Science*, *17*, 142-146.

Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2004). Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience*, 24, 4912–4917.

Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2005). Forming impressions of people versus inanimate objects: Social-cognitive processing in the medial prefrontal cortex. *NeuroImage*, 26, 251–257.

Mitchell, J.P., Cloutier, J., Banaji, M.R., Macrae, C.N. (2006). Medial prefrontal dissociations during processing of trait diagnostic and nondiagnostic person information. *Social Cognitive and Affective Neuroscience*, 1, 49–55.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Montgomery, K. J., & Haxby, J. V. (2008). Mirror neuron system differentially activated by facial expressions and social hand gestures: a functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*(10), 1866-1877.

Nee, D.E., Kastner, S., Brown, J.W. (2011). Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex. *NeuroImage*, 54, 528-540.

Pearson, J.M., Heilbronner, S.R., Barack, D.L., Hayden, B.Y., Platt, M.L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends in Cognitive Sciences*, 15, 143-151.

Peysakhovich, A., Rand., D.G. (2016). Habits of virtue: creating norms of cooperation and defection in the laboratory. *Management Science,* 62(3), 631–647.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, USA*, *98*, 676-682.

Reeder, G.D., Brewer, M.B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, 86, 61–79.

Reeder, G.D., Spores, J.M. (1983). The attribution of morality. Journal of Personality and Social Psychology, 44, 736–745.

Reeder, G.D., Coovert, M.D. (1986). Revising an impression of morality. *Social Cognition*, 4, 1-17.

Rothbart, M., Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology*, *50*(1), 131.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Rothmayr, C., Sodian, B., Hajak, G., Döhnel, K., Meinhardt, J., Sommer, M. (2011). Common and distinct neural networks for false-belief reasoning and inhibitory control. *NeuroImage*, 56, 1705-1713.

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, *7*(5), 499-500.

Satpute, A.B., Badre, D., Ochsner, K.N. (2014). Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cerebral Cortex*, 24(5), 1269-77

Saxe, R., Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *NeuroImage*, 19, 1835–1842

Saxe, R., Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, 43, 1391-1399.

Schiller, D., Freeman, J.B., Mitchell, J.P., Uleman, J.S., Phelps, E.A. (2009). A neural mechanism of first impressions. *Nature Neuroscience*, 12, 508-514.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science*, *16*(2), 114-122.

Shenhav, A., Botvinick, M.M., Cohen, J.D., (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, 79, 217-240.

Skowronski, J.J., Carlston, D.E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52, 689-699.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Skowronski, J.J., Carlston, D. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105, 131–142.

Somerville, L.H., Heatherton, T.F., Kelley, W.M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, 9, 1007-1008.

Sommer, M., Döhnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. *NeuroImage*, *35*(3), 1378-1384.

Souza, M.J., Donohue, S.E., Bunge, S.A. (2009). Controlled retrieval and selection of action-relevant knowledge mediated by partially overlapping regions in left ventrolateral prefrontal cortex. *NeuroImage*, 46, 299-307.

Spreng, R. N., & Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *Journal of Cognitive Neuroscience*, *22*(6), 1112-1123.

Spunt, R.P., Falk, E.B., Lieberman, M.D. (2010). Dissociable neural systems support retrieval of how and why action knowledge. *Psychological Science*, 21, 1593-1598.

Spunt, R.P., Satpute, A.B., Lieberman, M.D., (2011). Identifying the what, why, and how of an observed action: an fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, 23, 63-74.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Spunt, R.P., Lieberman, M.D. (2012a) An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *NeuroImage*, 59, 3050–3059.

Spunt, R.P., Lieberman, M.D. (2012b). Dissociating modality-specific and supramodal neural systems for action understanding. *Journal of Neuroscience*, 32, 3575-3583.

Stanley, D. A. (2015). Getting to know you: general and specific neural computations for learning about people. *Social Cognitive and Affective Neuroscience*. Available online December 8, 2015. doi: 10.1093/scan/nsv145

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

Tausch, N., Kenworthy, J. B., Hewstone, M. (2007). The confirmability and disconfirmability of trait concepts revisited: Does content matter?. *Journal of Personality and Social Psychology*, *92*(3), 542.

Thompson-Schill, S.L., Jonides, J., Marshuetz, C., Smith, E.E., D'Esposito, M., Kan, I.P., Knight, R.T., Swick, D. (2002). Effects of frontal lobe damage on interference effects in working memory. *Journal of Cognitive, Affective & Behavioral Neuroscience*, 2, 109-120.

Thompson-Schill, S.L., Bedny, M., Goldberg, R.F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, 15, 219-224.

Uhlmann, E. L., Pizarro, D. A., Diermeier, D. (2015). A Person-Centered Approach to Moral Judgment. *Perspectives on Psychological Science*, *10*(1), 72-81.

Utevsky, A.V., Smith, D.V., Huettel, S.A. (2014). Precuneus is a functional core of the default-mode network. *The Journal of Neuroscience*, *34*, 932-940.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Van Overwalle F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, 30, 829-858.

Van Veen, V., Holroyd, C.B., Cohen, J.D., Stenger, V.A., Carter, C.S. (2004). Errors without conflict: Implications for performance. *Brain and Cognition*, 56, 267-276.

Van Hecke, J., Gladwin, T. E., Coremans, J., Destoop, M., Hulstijn, W., & Sabbe, B. (2010). Prefrontal, parietal and basal activation associated with the reordering of a two-element list held in working memory. *Biological Psychology*, *85*(1), 143-148.

Vergauwe, E., Hartstra, E., Barrouillet, P., & Brass, M. (2015). Domain-general involvement of the posterior frontolateral cortex in time-based resource-sharing in working memory: An fMRI study. *NeuroImage*, *115*, 104-116.

Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., ... & Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, *29*(1), 90-98.

Wojciszke, B., Brycz, H., Borkenau, P. (1993). Effects of information content and evaluative extremity on positivity and negativity biases. *Journal of Personality and Social Psychology*, 64, 327-335.

Wojciszke, B., Bazinska, R., Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251-1263.

Wojciszke, B. (2005). Morality and competence in person and self-perception. *European Review of Social Psychology*, 16, 155-188.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

Wolf, I., Dziobek, I., Heekeren, H.R. (2010). Neural correlates of social cognition in naturalistic settings: a model-free analysis approach. *NeuroImage*, *49*, 894-904.

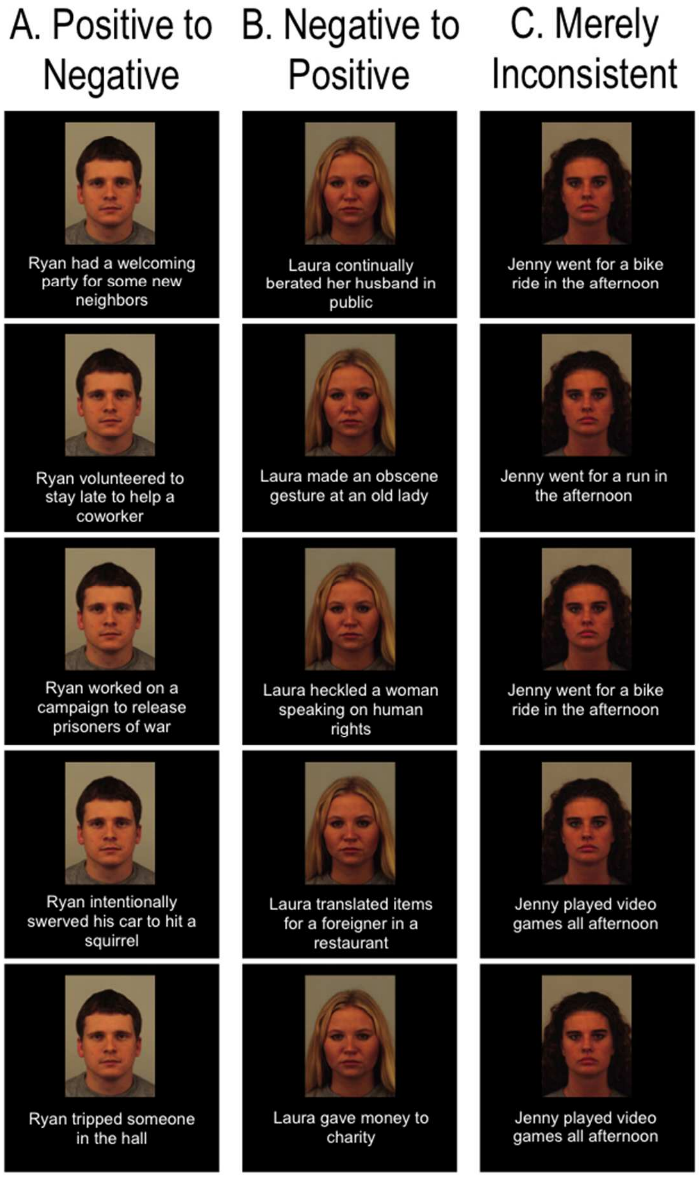
Ybarra, O. (2001). When first impressions don't last: The role of isolation and adaptation processes in the revision of evaluative impressions. *Social Cognition*, *19*(5), 491-520.

Young, L., Dodell-Feder, D., Saxe, R. (2010a). What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia*, 48, 2658-2664.

Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R. (2010b). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*(15), 6753-6758.

Zaki, J., Hennigan, K., Weber, J., Ochsner, K.N. (2010). Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. *Journal of Neuroscience*, 30, 8481-8488.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016



A. Positive to Negative

B. Negative to Positive

C. Merely Inconsistent

Ryan had a welcoming party for some new neighbors
Laura continually berated her husband in public
Jenny went for a bike ride in the afternoon
Ryan volunteered to stay late to help a coworker
Laura made an obscene gesture at an old lady
Jenny went for a run in the afternoon
Ryan worked on a campaign to release prisoners of war
Laura heckled a woman speaking on human rights
Jenny went for a bike ride in the afternoon
Ryan intentionally swerved his car to hit a squirrel
Laura translated items for a foreigner in a restaurant
Jenny played video games all afternoon
Ryan tripped someone in the hall
Laura gave money to charity
Jenny played video games all afternoon

Figure 1. Sample individuals from the positive-to-negative, negative-to-positive, and merely inconsistent conditions. Faces and behaviors were presented on screen together for 6 seconds, followed by a 6 second fixation cross. Once all five behaviors were presented, participants provided global ratings of each individual’s trustworthiness (4 seconds) and surprisingness (4 seconds).

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

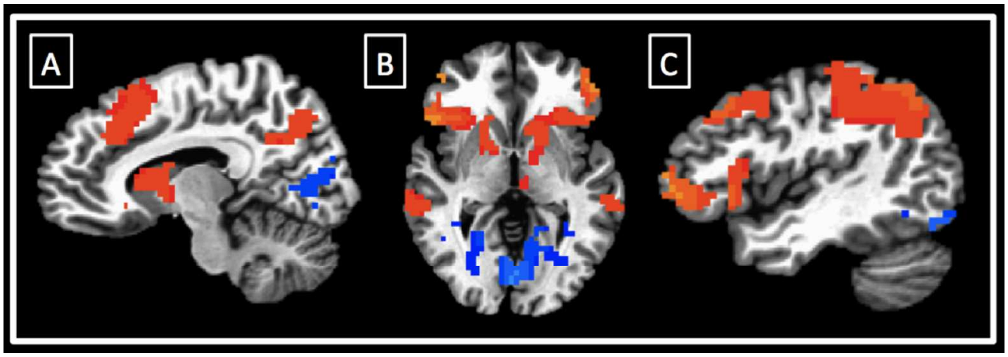


Figure 2. Main effect of updating (last two > first three trials, collapsed across individual type). We observed an extended network of regions that was recruited in response to inconsistent information presented during the last two trials (pictured in hot colors), including A) dmPFC, PCC/precuneus, B) bilateral rlPFC, bilateral vlPFC, bilateral caudate, bilateral STS, C) bilateral ATL, and bilateral TPJ/IPL. (Regions pictured in cool colors—including bilateral fusiform gyrus and cuneus—showed preferential responses to information presented on the first three trials, before inconsistent information was introduced.)

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

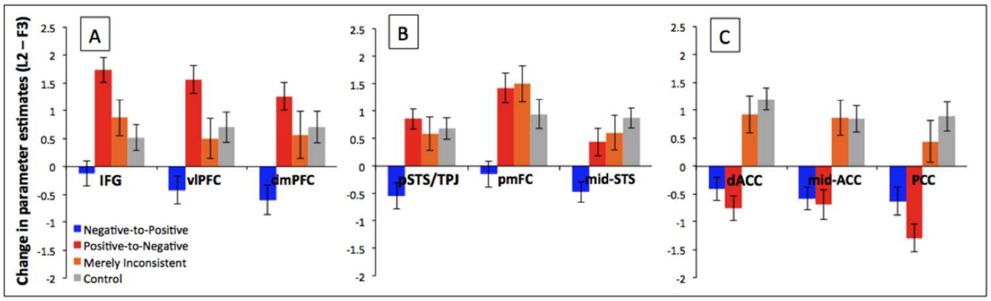


Figure 3. Parameter estimates extracted from regions displaying an interaction between updating and condition. Colored bars indicate the change in activity from the first three to the last two (i.e., update-provoking) trials. Here, we provide information on selected key regions emerging from the updating-by-condition interaction in which A) the updating-related change in activity was strongest for positive-to-negative individuals, B) this change in activity was stronger for positive-to-negative, merely inconsistent, and control individuals, compared to negative-to-positive individuals, and C) this change in activity was stronger for merely inconsistent and control individuals, compared to individuals from either meaningfully inconsistent condition. (Blue = negative-to-positive, red = positive-to-negative, orange = merely inconsistent, gray = control.) Error bars indicate +/-1 standard error. (For an expanded version of this figure including all 20 regions identified by the updating-by-condition interaction contrast, see Supplementary Figure 1.)

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016

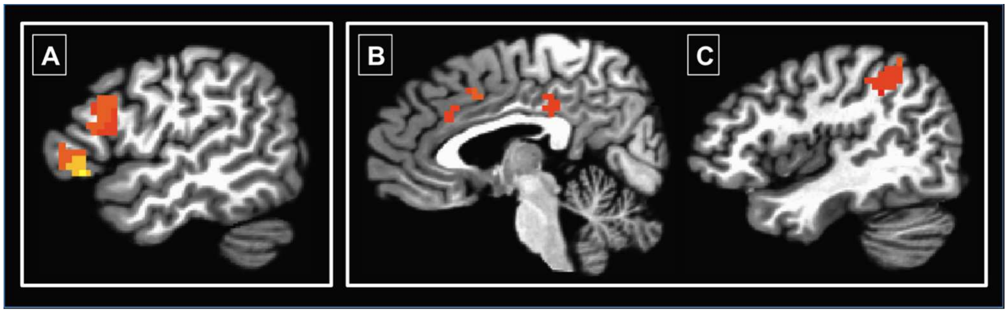


Figure 4. Preferential neural responses to meaningful and mere inconsistency. Left vlPFC and left IFG (Panel A), showed enhanced activity when updating based on immoral behaviors, compared to updates based on moral or merely inconsistent behaviors. dACC, mid-ACC, PCC (Panel B), and right TPJ/IPL (Panel C) all showed enhanced activity when updating based on merely inconsistent behaviors, compared to updates based on moral or immoral behaviors.

Downloaded from http://scan.oxfordjournals.org/ at University of Birmingham on May 17, 2016