

MIT Open Access Articles

*An fMRI study of violations of social expectations:
When people are not who we expect them to be*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Cloutier, J., J.D.E. Gabrieli, D. O'Young, and N. Ambady. "An fMRI Study of Violations of Social Expectations: When People Are Not Who We Expect Them to Be." *NeuroImage* 57, no. 2 (July 2011): 583–588.

As Published: <http://dx.doi.org/10.1016/j.neuroimage.2011.04.051>

Publisher: Elsevier

Persistent URL: <http://hdl.handle.net/1721.1/102462>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-NonCommercial-NoDerivs License



**An fMRI study of violations of social expectations: When people are not
who we expect them to be**

J. Cloutier¹, J. D. E. Gabrieli², D. O'Young², & N. Ambady³

¹Dept. of Psychology, University of Chicago

²Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology;

³Dept. of Psychology, Tufts University

Address Correspondence To:

Jasmin Cloutier

Department of Psychology

University of Chicago

Chicago, IL

Email: jcloutier@uchicago.edu

Tel. : (773) 834-1363

Abstract

The current study examines the effect of violations of social expectancies on the neural substrates of person perception. In an event-related fMRI experiment, participants were presented with the photographs of either Republican or Democrat politicians paired with either typical Republican or Democrat political views (e.g., “wants a smaller government” or “wants liberal supreme court judges”). Subjects were asked to form an impression of the targets using information about both their political affiliation and their political views. Of interest was the contrast between stereotypically congruent trials and stereotypically incongruent trials. The results reveal that brain regions previously involved in mentalizing (i.e., temporoparietal junction and medial prefrontal cortex) are preferentially recruited when viewing incongruent social targets.

Key words: person perception; expectancy violation; mentalizing; individuation, fMRI

The ubiquitous use of social expectations when perceiving others is well established. Social cognitive investigations have repeatedly demonstrated how impression formation based on categories and stereotypes (i.e., information that is expected to describe social targets belonging to a specific social group) often take precedence over construal based on individuating information (i.e., information that is specific to a social target) (Devine, 1989; Macrae & Bodenhausen, 2001). Nevertheless, social expectations are often violated during impression formation (Hamilton, Driscoll, & Worth, 1989; Hastie & Kumar, 1979; Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999; Sherman, Bessenoff, & Frost, 1998). As a consequence, we routinely are required to override our social expectations and instead create individuated impressions of others.

The implementation of such individuation processes following the violations of social expectations has been extensively documented (Brewer, 1988; Fiske & Neuberg, 1990; Macrae et al., 1999; Hastie & Kumar, 1979; Srull & Wyer, 1989). When individuated, social targets are construed as complex social agents with their personal constellation of beliefs, personality characteristics and intentions, as opposed to stereotypical members of a particular social group. Individuation, therefore, requires the attribution of unique characteristics, such as intentions and mental states, to social targets.

From a social cognitive perspective, studying violations of social expectations during impression formation has revealed many of the requirements and consequences of flexibly construing others (Macrae & Bodenhausen, 2000;

Smith, 1998). However, although fMRI has been utilized to explore the brain regions supporting categorical or stereotype-based responses (Mitchell, Ames, Jenkins, & Banaji, 2009; Wheeler & Fiske, 2005; Quadflieg, Turk, Waiter, Mitchell, Jenkins, & Macrae, 2009; Richeson et al., 2003), few studies have investigated the perception of violations of social expectations using the same method.

Previous fMRI studies have examined the congruency of affective associations towards social targets (Harris & Fiske, 2010; Knutson, Wood, Spampinato, & Grafman, 2006; Westen, Blagov, Harenski, Kilts, & Hamann, 2006) and, using electroencephalography (EEG), both regulation of racial bias and perceived violations of social expectations have been investigated (Amodio, Harmon-Jones, Devine, Curtin, Hartley, & Covert, 2004; Amodio, Kubota, Harmon-Jones, & Devine, 2006). Furthermore, with the help of EEG, the neural operations underlying the processing of words or sentences that are either congruent or incongruent in terms of gender stereotypes have been studied (Osterhout, Bersick, & McLaughlin, 1997; White, Crites, Taylor, & Corral, 2009). Nevertheless, the neural correlates of social cognitive processes recruited when pre-existing social expectations are violated during impression formation have yet to be investigated using fMRI (see Amodio & Lieberman, 2009; for a recent review of the literature). Accordingly, the current study aims to identify brain regions recruited by fundamental social cognitive processes during the perception of targets violating social expectation.

The medial prefrontal cortex (MPFC) and temporoparietal junction (TPJ) appear to be central components of a constellation of brain regions supporting social cognition (Adolph, 2009; Amodio & Frith, 2006; Decety & Lamm, 2007; Spreng, Mar, & Kim, 2008). In particular, multiple lines of investigations suggest that these regions support processes enabling perceivers to perform, in one way or another, mental inferences about encountered individuals (Adolph, 2009; Frith & Frith, 2006; Mitchell, Macrae, & Banaji, 2006; Saxe & Wexler, 2005; Spreng, Mar, & Kim, 2008). Of particular relevance to the current investigation, tasks requiring the attribution of specific mental states or access to person-knowledge about social targets have been shown to recruit these brain regions (Cloutier, Kelley, & Heatherton, 2011; Frith & Frith, 2006; Spreng, Mar, & Kim, 2008; Todorov, Gobbini, Evans, & Haxby, 2007). Following social cognitive theorization, such processes should be extremely useful when forming an impression tailored to an individual for which pre-existing expectations are not applicable (Brewer, 1988; Fiske & Neuberg, 1990; Macrae et al., 1999).

Accordingly, processes supported by the TPJ and MPFC are often mentioned as prime candidates to support the individuation of social targets (Amodio & Lieberman, 2009; Freeman, Schiller, Rule, & Ambady, 2010; Harris & Fiske, 2007). It is therefore surprising that little research has been done to test this possibility. Motivating the current study is the hypothesis that these brain regions will be preferentially engaged during the perception of incongruent social targets. When perceiving violations of social expectations, both the MPFC and

TPJ are expected to support mental inferences necessary to form individuated impressions of the social targets.

To explore this possibility, the current study used an event-related fMRI design to identify brain regions underlying the processing of socially incongruent social targets. To this end, participants were presented with photographs of unknown politicians, assigned to either the Democrat or Republican parties, who endorsed either typically Democrat or Republican views (Figure 1). Of particular interest were the brain regions preferentially engaged when perceivers were presented with incongruent trials (i.e., Democrats endorsing typical Republican views and Republicans endorsing typical Democrat views). Crucially, because the congruent Republican views were also incongruent Democrat views and vice-versa, all the information conferred by the faces and sentences (i.e., the political views) contributed equally to congruent and incongruent trials across participants.

METHODS

Participants

Twenty participants were recruited from the local MIT community. Of these twenty participants, two were excluded from subsequent analyses (the first subject excluded reported discomfort during the scan and difficulty performing the task, the second subject was the only one to report identification with the

Republican party). The remaining eighteen were between the ages of 19 and 30 years (9 male, mean age = 20.7 years), reported no significant abnormal neurological history and had normal or corrected-to-normal visual acuity. Sixteen participants were right-handed as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). Participants were paid for their participation and gave informed consent in accordance with the guidelines set by the Committee on the Use of Humans as Experimental Subjects at MIT.

Material and Pre-rating Task

In a pilot study, participants (N=24) rated a list of sentences created to represent either typical Democrat (i.e., *prioritizes environmental policies*) or Republican views (i.e., *wants to privatize Social Security*). As a group, the participants considered themselves affiliated with the Democrat party (M = 2.7; s.d. = 1: on a 7 point scale with 1= “Extremely Democrat” to 7 “Extremely Republican”) and as having liberal views (M = 2.8; s.d. = 1: on a 7 point scale with 1=“Extremely liberal” to 7 =“Extremely conservative”). Their task was to rate how “stereotypically Democrat or Republican” they believed the views described by sentences were (on 7 point scale, 1=“Very stereotypically Democrat” to 7=“Very stereotypically Republican”). From this pilot study, we identified two lists of sentences (typical Democrat views: mean (s.d.) = 2.25 (0.39); typical Republican views: mean (s.d.) = 4.33 (0.30)) that were subsequently used in the functional imaging task. Because the sentences were created and rated by individuals identifying themselves mostly as Democrats, the resulting sentences

can be construed to represent typical Democrat or Republican views from the perspective of Democrat individuals.

Functional Imaging Task and Procedure

During the fMRI experiment, participants formed impressions of either Democrat or Republican politicians (80 unique targets were created using photographs of unknown politicians paired with background colors ascribed to each political party) paired with either typical Democrat or Republican political views (40 sentences of each type of view were paired with politicians of each political affiliation) (Figure 1). Each face was presented twice with two unique sentences of the same condition. This resulted in 40 unique congruent-Democrat trials, 40 unique incongruent-Democrat trials, 40 unique congruent-Republican trials and 40 unique incongruent-Republican trials. Each trial consisted of a photograph of an unknown politician with a colored background (indicating a political affiliation) paired with a sentence describing a political view and was presented for 3500ms. Following each stimulus presentation, a fixation cross was presented for 500ms. Null events consisting of a fixation cross for 2000ms were pseudorandomly interspersed to introduce jitter into the fMRI time-series to create ITIs of either 500ms, 2500ms, 4500ms or 6500ms. Participants were instructed to form impressions of the politicians based on the information available to them (i.e., the portrait, the party affiliation associated with the background color and the political views represented by the sentence). The pictures were gray-scaled photographs of unfamiliar politicians used in a previous

study. These pictures were presented in the center of the screen at a size of 100-128 pixels wide by 150 pixels tall. The photographs and the color backgrounds were counterbalanced across participants to ensure that they would be equally represented in each trial type. Participants took part in practice trials prior to the fMRI session to ensure that they would efficiently associate the background color with the appropriate political affiliations.

Following previous fMRI investigations using impression formation instructions, participants were simply asked to press response buttons held in both hands once they felt they completed the task. The behavioral response was requested mainly to ensure that participants were paying attention to the task and participants were therefore not asked to perform their response as quickly as possible.

Importantly, the information communicated by the faces and the sentences was counterbalanced across participants to ensure they would not create confounds when comparing trials based either on congruency or party affiliation. Indeed, across participants, the sentences contributed equally to the congruent and incongruent conditions as well as to the Democrat and Republican conditions. This ensured that factors such as sentence content, difficulty or length would not affect the results. Similarly, counterbalancing across participants ensured that the perceptual information afforded by the faces did not impact comparison across conditions.

Following the functional imaging session, participants were presented with the photographs once again and were asked if they recognized any of the politicians. This was done to verify that the individuals in photographs were indeed unfamiliar to the participants.

Functional Imaging Acquisition

Anatomical and functional whole-brain imaging was performed on a Siemens 3T Tim Trio Scanner using a phase-array 32-channel head coil (Siemens Medical, Erlangen, Germany). An Apple Macbook Pro running the Psychophysics Toolbox extensions in Matlab (The Mathworks, Natick, MA) was used to present stimuli to the participants. Anatomical images were acquired using a high-resolution MPRAGE sequence (128 sagittal slices, TE = 3ms, TR = 2500ms, flip angle = 7°, 1x1x1mm voxels). Functional images were collected in 4 functional runs of 146 time points each, using a gradient echo, echo planar sequence sensitive to BOLD contrast (T2*) (32 axial slices per whole-brain volume, 2mm in-plane resolution, 4mm thickness, 0.8mm skip, TR = 2000ms, TE = 30ms, flip angle = 90°).

Data Analysis

Functional MRI data was analyzed using SPM5 (Wellcome Department of Cognitive Neurology, London, UK). Prior to statistical analysis, images were

preprocessed to remove sources of noise and artifacts. Functional data were realigned within and across runs to correct for head movement and transformed into a standard anatomical space (3mm isotropic voxels) based on the ICBM 152 brain template (Montreal Neurological Institute). Normalized data were then spatially smoothed (8mm full width at half maximum) using a Gaussian kernel. Finally, using in-house artifact detection software, individual runs were analyzed (on a subject-by-subject basis) to find outlier timepoints as measured by two criteria: we excluded from further analysis volumes during which subject head motion exceeded 1mm or .75 degree, and volumes in which the overall signal for that timepoint fell more than three standard deviations outside the mean global signal for the entire run. Outlier time-points were excluded from the GLM analysis via the use of subject-specific regressors of no interest. Each subject's data were high-pass filtered at 128 sec. Analyses took place at two levels: formation of statistical images and regional analysis of hemodynamic responses. In the first analysis, a GLM incorporating task effects for the 4 trial types of interest (congruent-Democrat, incongruent-Democrat, congruent-Republican, incongruent-Republican) and covariates of no interest (a session mean, six movement parameters derived from realignment corrections, and regressors to deweight individual outlier volumes) was used to compute parameter estimates (β) and t-contrasts images (containing weighted parameter estimates) for each comparison at each voxel and for each subject.

To determine which brain regions showed preferential activation to incongruent trials, we conducted a random effects analysis in which individuals' first-level contrast images for incongruent vs. congruent conditions were submitted separately to second-level, one-sample t-tests. This analysis produced a group-level t-contrast with minimum clusters of 10 voxels and only brain regions surviving FDR corrections for multiple comparisons are reported. Bilateral TPJ (Left: -54, -54, 18); (Right: 54, -49, 19) and MPFC (-3, 52, 20) spherical regions of interest (ROI) of 6 mm were defined based on Talairach coordinates taken from a recent meta-analysis of Theory of Mind studies (Spreng, Mar, & Kim, 2008). Parameter estimates from contrast images comparing each of the 4 trial types (congruent-Democrat, incongruent-Democrat, congruent-Republican, incongruent-Republican) to the baseline control (fixation) were extracted from the ROIs, submitted to statistical analysis and plotted to further characterize the activations for all trial types in these brain regions.

RESULTS

Behavioral Results

There were no significant differences in response time for the four trial types [congruent-Democrat, mean (s.d.) = 2134 msec (434 msec); incongruent-Democrat, mean (s.d.) = 2104 msec (422 msec); congruent-Republican, mean (s.d.) = 2130 msec (438 msec); incongruent-Republican, mean (s.d.) = 2089

msecs (405 msecs)]. This lack of behavioral difference is not surprising considering that participants were not instructed to answer as quickly as possible and simply pressed buttons held in both hands after forming impressions of the targets. A behavioral test following the scan confirmed that participants were not familiar with any of the politicians presented during the fMRI task.

fMRI Results

The first analysis identified brain regions preferentially activated when forming impressions of incongruent compared to congruent targets. Greater activation was found for brain regions associated with social cognition (e.g., TPJ bilaterally, MPFC and precuneus) and for a number of lateral prefrontal regions ostensibly involved in some forms of cognitive control (Aron, Robbins, & Poldrack, 2004; Braver, Reynolds, & Donaldson, 2004; Cools, Clark, Owen, & Robbins, 2002; Ochsner & Gross, 2005) (Table 1 and Figure 2). There were no activations greater for congruent compared to incongruent trials and no differences between Democrat and Republican targets.

Based on coordinates obtained from a recent meta-analysis of Theory of Mind studies (Spreng, Mar, & Kim, 2008), ROI analyses were employed to characterize activations for all trial types in brain regions hypothesized to be preferentially involved in the perception of socially incongruent targets (i.e., MPFC and bilateral TPJ). For each subject, signal intensities for the ROIs were calculated separately for the four trial types and examined statistically to directly

compare activation to the congruent and incongruent targets affiliated with each political party (Figure 2).

These analyses confirmed that areas of the MPFC and TPJ previously shown to be involved in mentalizing about others (Spreng, Mar, & Kim, 2008) were preferentially recruited when participants formed impressions of socially incongruent targets, irrespective of their political party of affiliation. Indeed, the results revealed a main effect of social congruence [MPFC: $F(17) = 8.95, p=.008$; Right TPJ: $F(17) = 11.44, p=.004$; and Left TPJ: $F(17) = 11.87, p=.003$] but no difference in activation based on party affiliation [MPFC: $F < 1$; Right TPJ: $F < 1$; and Left TPJ: $F < 1$] and no interaction between social congruence and party affiliation [MPFC: $F < 1$; Right TPJ: $F < 1$; and Left TPJ: $F(17) = 1.55, p=.23$].

DISCUSSION

Brain regions previously shown to support mentalizing about others (i.e., TPJ and MPFC) were preferentially recruited when participants perceived individuals violating social expectations. These findings further specify the social cognitive processes underlying the individuation of social targets during person perception. In agreement with models of person perception, additional mental inferences were ostensibly required when forming impressions of social targets violating expectations (Brewer, 1988; Fiske & Neuberg, 1990; Macrae et al., 1999). These results suggest that processes supporting the individuation of

social targets, as postulated by social cognitive researchers, overlap with processes supported by brain regions involved in mentalizing about others.

Individuation and mental state inferences

In the current experiment, the bilateral TPJ and the MPFC, brain regions involved in mentalizing about others, were preferentially recruited when forming impressions of social targets violating social expectations. Numerous studies now point towards these brain regions playing central roles in social cognition (Adolph, 2009; Amodio & Frith, 2006; Decety & Lamm, 2007; Frith & Frith, 2006; Mitchell, Macrae, & Banaji, 2006; Saxe & Wexler, 2005; Spreng, Mar, & Kim, 2008). The TPJ, frequently bilaterally, is consistently involved in tasks requiring the attribution of mental states to social targets (Saxe & Wexler, 2005; Saxe, 2006; Spreng, Mar, & Kim, 2008) and damage to this brain area has been shown to impair performance on Theory of Mind (ToM) tasks (Samson, Apperly, & Humphreys, 2004). The MPFC is also central to many social cognitive processes (Amodio & Frith, 2006), including those supporting impression formation (Mitchell, Macrae, & Banaji, 2004), ToM (Frith & Frith, 2006) and the perception of faces for which person-knowledge is available (Cloutier, Kelley, & Heatherton, 2011; Todorov, Gobbini, Evans, & Haxby, 2007).

Accordingly, preferential involvement of the TPJ and MPFC for incongruent social targets provides additional evidence in support of person perception models positing the frequent necessity to individuate targets violating social

expectations (Brewer, 1988; Fiske & Neuberg, 1990; Macrae et al., 1999). Indeed, in light of the social cognitive functions ascribed to the TPJ and MPFC, the current findings suggest that additional attributions of mental states was required to individuate socially incongruent targets. These conclusions are in agreement with recent studies suggesting the involvement of these regions in different contexts requiring the individuation of social targets (Freeman, Schiller, Rule, & Ambady, 2010; Harris & Fiske, 2007).

In addition to the TPJ and MPFC, the right superior temporal sulcus and the precuneus, two regions also believed to support social cognitive tasks (Adolphs, 2009), were preferentially recruited during socially incongruent trials. Amongst other functions, the precuneus is believed to play a role in ToM operations (Saxe, Moran, Scholz, & Gabrieli, 2006) and the right STS is believed to play a role in social perception (Allison, Puce, & McCarthy, 2000; Pelphrey & Morris, 2006). The STS was also shown to be preferentially recruited when perceivers form impressions of social targets paired with meaningful person-knowledge (Mitchell, Cloutier, Banaji, & Macrae, 2006). It is therefore possible that increased activation in STS to socially incongruent trials was a consequence of the particular relevance of the information provided for the purpose of individuating the social targets.

Additional requirements when perceiving incongruent targets

Although social expectations often guide person perception, flexibility is required when considering the multiple levels at which others can be construed and the great variability of personal characteristics they possess. This cognitive flexibility is believed to involve more effortful and controlled processes (Devine, 1989; Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2001). In contrast to the previously described mentalizing operations, these processes have typically been ascribed to areas of the prefrontal cortex not believed to be specific to social cognition (Cunningham et al., 2004; Macrae et al. 1999). It was therefore not surprising to uncover various lateral prefrontal brain regions preferentially recruited by the socially incongruent targets. These brain regions may likely support various cognitive control operations (Aron, Robbins, & Poldrack, 2004; Braver, Paxton, Locke & Barch, 2009; Cools et al., 2002; D'Esposito, 2007; Dosenbach et al., 2007; Kerns, et al., 2004; Koechlin, Ody, & Kouneilher, 2003; Kringelbach & Rolls, 2003; Ochsner & Gross, 2005) necessary to create distinct impressions of social targets violating social expectations. For example, such cognitive processes are required to override inconsistencies between existing expectations about a social target's group (i.e., Republican or Democrat party) and the unexpected information available about the same individual (i.e., beliefs that go against the positions typically adopted by the party in question) (Hastie & Kumar, 1979; Srull & Wyer, 1989).

It is noteworthy that the ACC, a region believed to play an important role in

conflict detection and cognitive control (Barch, Braver, Akbudak, Conturo, Ollinger, Snyder, 2001; Botvinick, Cohen, & Carter, 2004; Kerns, Cohen, MacDonald, Cho, Stenger, & Carter, 2004), was not found to be preferentially responsive to stereotypically incongruent trials.

The ACC has previously been shown to support error-monitoring operations (Carter et al., 1998), with increased activity in the region being at times reported irrespective of the commission of an error from the participants. The tasks used in these studies typically involve the possibility of an incorrect answer and/or require the selection of one amongst multiple actions (Barch et al., 2001). The impression formation task of the current study did not involve the selection of one amongst many presented responses. Therefore, absence of preferential ACC activation to socially incongruent trials may suggest that this region is indeed involved in conflict detection at the level of response selection (Botvinick, Cohen, & Carter, 2004). Accordingly, in the context of person perception operations, involvement of the ACC might be indicative of efforts from perceivers to regulate prejudicial responses (Amodio et al., 2004, Amodio, Kubota, Harmon-Jones, & Devine, 2006), rather than suggestive of the implementation of social cognitive processes required to individuate social targets.

As the stimuli in the congruent and incongruent conditions are the same across participants (i.e., the congruent Democrat statements were also presented as incongruent Republican statements and vice-versa), the observed differences in brain activity cannot be explained by the material presented in each

experimental condition. This strongly suggests that the obtained results are truly a consequence of perceived violations of social expectations and are not driven by differences in the visual (i.e., faces) or semantic (i.e., sentences) information provided to the perceivers.

Nevertheless, there are inherent limits the design used in the current study. Although the behavioral responses required of the participants did not reveal any differences in reaction time across conditions, social cognitive investigations have repeatedly found evidence of increased processing demands required by socially incongruent targets (Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2001; Macrae, et al., 1999; Sherman, Bessenoff, & Frost, 1998). As such, investigations incorporating experimental manipulations and/or behavioral responses sensitive to the distinct effortful processes recruited by the perception of socially incongruent targets will be required. These investigations will not only help to further specify the function of various brain regions preferentially recruited by socially incongruent targets, but could also distinguish between the required operations at different stage of processing (Cunningham et al., 2004).

Additionally, the fact that participants in the current study held liberal views may limit the generalizability of our findings. Because conservatism is typically associated with less tolerance of ambiguity, less openness to experience and an increase need for structure and order (Jost, Glaser, Kruglanski, & Sulloway, 2003), differential brain activations during person perception may also be

expected from more conservative perceivers (Amodio, Jost, Master, & Yee, 2007; Brosch, Coppin, Scherer, Schwartz, & Sander, In Press; Knutson et al., 2006).

Conclusion

As hypothesized, the perception of socially incongruent targets recruited brain regions involved in mentalizing about others. The observed preferential activation of the MPFC and TPJ suggests the occurrence of further mental inferences when forming impressions of socially incongruent targets. These findings once again underscore the importance of the MPFC and TPJ for social cognition. Additionally, they suggest an overlap between the processes underlying the individuation of social targets (Brewer, 1988; Fiske & Neuberg, 1990; Hastie & Kumar, 1979; Macrae & Bodenhausen, 2000; Srull & Wyer, 1989) and the processes involved in mentalizing about others (Adolph, 2009; Amodio & Frith, 2008; Frith & Frith, 2006; Mitchell, Macrae, & Banaji, 2006; Saxe & Wexler, 2005). More speculatively, the recruitment of additional prefrontal brain regions when perceiving incongruent social targets might index the involvement of cognitive control operations in response to violations of social expectations.

As the number of brain-imaging studies investigating different facets of person perception increase, we should gain a better understanding of the so-called social brain. The context in which others are encountered often dictates how we construe them. For this reason, studying complex social cognitive

phenomenon, such as the modulations of social expectations during impression formation, demonstrates the flexibility with which specific brain regions are recruited to make sense of our social environment. Importantly, much of these investigations can benefit from the insights gained by behavioral studies of social cognition. Conversely, through the integration of different research perspectives, brain-imaging studies have the potential not only to increase our understanding of the brain but also to provide new insights into the social-cognitive processes involved in person perception.

Acknowledgments

We thank Tom Meagher, Carlos Cardenas, Rebecca Martin for their help with data collection as well as William Cunningham and anonymous reviewers for their helpful comments.

References

- Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology, 60*, 693-716.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267–278.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*, 268 – 277.
- Amodio, D. M. & Lieberman, M. D. (2009). Pictures in our heads: Contributions of fMRI to the study of prejudice and stereotyping. In Todd D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination*. New York: Psychology Press.
- Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Sciences, 15*, 88-93.
- Amodio, D. M., Jost, J. T., Master, S. L., & Yee, C. M. (2007). Neurocognitive correlates of liberalism and conservatism. *Nature Neuroscience, 10*, 1246-1247.
- Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience, 1*, 26-36.

- Aron, A. R., Robbins, T. W. & Poldrack, R. A. (2004). Inhibition and the right inferior frontal. *Trends in Cognitive Sciences*, 8, 170-177.
- Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex*, 11, 837-848.
- 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.
- Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences*, 106 (18), 7351-7356.
- Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2004). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, 39, 713–726.
- Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum.
- Brosch, T., Coppin, G., Scherer, K. R., Schwartz, S. & Sander, D. (In Press). Generating value(s): Psychological value hierarchies reflect context-dependent sensitivity of the reward system. *Social Neuroscience*.

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*: 5364, 747-749.

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

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience, 22*, 4563-4567.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science, 15*, 806-813.

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *The Neuroscientist, 13*, 580-593.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18.

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society, 362*, 761-772.

- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23*, 1 – 74.
- Freeman, J. B., Schiller, D., Rule, N. O., & Ambady, N. (2010). The neural origins of superficial and individuated judgments about ingroup and outgroup members. *Human Brain Mapping, 31*, 150-159.
- Frith, C.D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50*, 531 – 534.
- Hamilton, D. L., Driscoll, D. M., & Worth, L. T. (1989). Cognitive organization of impressions: Effects of incongruity in complex representations. *Journal of Personality and Social Psychology, 57*, 925-939.
- Harris, L. T. & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in the mPFC. *Social Cognitive Affective Neuroscience, 2*, 45-51.
- Harris, L. T. & Fiske, S. T. (2009). Neural regions that underlie reinforcement learning are also active for social expectancy violations. *Social Neuroscience, 5*, 76-91.
- Hastie, R., & Kumar, P. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37*, 25-38.

- Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129*, 339-375.
- 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*: 5660, 1023-1026.
- Knutson, K. M., Wood, J. N., Spampinato, M. V., & Grafman, J. (2006). Politics on the brain: An fMRI investigation. *Social Neuroscience, 1*, 25-40.
- Kringelbach, M. L., & Rolls, E. T. (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *Neuroimage, 20*, 1371-1383.
- Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology, 76*, 200-213.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93-120.
- Mitchell, J. P., Ames, D. L., Jenkins, A. C., & Banaji, M. R. (2009). Neural correlates of stereotype application. *Journal of Cognitive Neuroscience, 21*, 594-604.
- 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. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron, 50*, 655-663.
- 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.
- Ochsner, K. N. & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences, 9*, 242-249.
- Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25*, 273-285.
- Pelphrey, K. A. & Morris, J. P. (2006). Brain mechanisms for interpreting the actions of others from biological-motion cues. *Current Direction in Psychological Sciences, 15*, 136-140.
- Quadflieg, S., Turk, D. J., Waiter, G. D., Mitchell, J. P., Jenkins, A. C., & Macrae, C. N. (2009). Exploring the neural correlates of social stereotyping. *Journal of Cognitive Neuroscience, 21*, 1560-1570.
- Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., & Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience, 6*, 1323-1328.

- Samson, D., Apperly, I., & Humphreys, G. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, 7, 499-500.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16, 235–239
- Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, 43, 1391 – 1399.
- Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. D. E. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, 1, 229-234.
- Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology*, 75, 589-606.
- Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 391-445). Boston: McGraw-Hill.
- Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2008). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21, 489-510.

- Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review*, 96, 58-83.
- Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, 45, 163-173.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185: 4157, 1124-1131.
- Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice. *Psychological Science*, 16, 56-62.
- White, K. R., Crites, S. L., Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience*, 4, 191-198.

Table 1. Identification of BOLD signal differences between Congruent and Incongruent conditions.

| Brain Region | <i>Incongruent > Congruent</i> | P FDR-corr | T | X | y | z |
|---------------------|--|-------------------|----------|----------|----------|----------|
| BA 9 | R Superior Frontal Gyrus | 0.038 | 6.51 | 6 | 52 | 44 |
| BA 6 | L Middle Frontal Gyrus | 0.098 | 5.75 | -48 | 2 | 54 |
| BA 9 | Medial Prefrontal Cortex | 0.038 | 5.71 | 2 | 54 | 30 |
| BA 39 | L Temporoparietal Junction | 0.038 | 5.55 | -44 | -58 | 32 |
| BA 45 | L Inferior Frontal Gyrus | 0.038 | 5.52 | -52 | 22 | 8 |
| BA 40 | R Temporoparietal Junction | 0.038 | 5.50 | 50 | -60 | 34 |
| BA 7 | Precuneus | 0.038 | 5.45 | 0 | -60 | 36 |
| BA 10 | L Ventromedial Prefrontal Cortex | 0.038 | 5.40 | 2 | 60 | -2 |
| BA 9 | R Middle Frontal Gyrus | 0.038 | 5.35 | 36 | 10 | 32 |
| BA 41 | R Superior Temporal Sulcus | 0.038 | 4.89 | 44 | -36 | 8 |
| BA 47 | L Ventrolateral Prefrontal Cortex | 0.038 | 4.87 | -44 | 30 | -16 |
| BA 19 | R Lingual Gyrus | 0.038 | 4.85 | 30 | -62 | 2 |
| BA 21 | L Superior Temporal Sulcus | 0.038 | 4.35 | -52 | -12 | -16 |
| BA 21 | R Middle Temporal Gyrus | 0.039 | 4.22 | 54 | 2 | -32 |
| BA 11 | R Ventrolateral Prefrontal Cortex | 0.039 | 4.18 | 38 | 36 | -16 |
| BA 38 | L Temporal Pole | 0.039 | 4.16 | -28 | 20 | -26 |
| BA 8 | L Middle Frontal Gyrus | 0.039 | 4.15 | -32 | 16 | 40 |
| BA 21 | R Superior Temporal Sulcus | 0.043 | 3.97 | 54 | -10 | -20 |
| BA 9 | R Middle Frontal Gyrus | 0.045 | 3.87 | 46 | 26 | 30 |

Activations determined to be significant ($p < 0.001$, uncorrected; clusters ≥ 10 voxels) following FDR-correction are listed along with the best estimate of their location. BA = approximate Brodmann's area location. X,Y,Z values represent MNI coordinates. Locations of the activations are determined based on the functional responses superimposed on averaged anatomical MRI images.

Figure 1. Figures displaying an example of a stereotypically congruent trial (left) and an example of a stereotypically incongruent trial (right). In these trials, orange was predetermined to signify that a target was a Republican and green that the target was a Democrat.

Figure 2. Sagittal section (top left) and coronal section (top right) illustrating regions believed to be involved in mentalizing displaying increased activation to socially incongruent trials [bilateral TPJ and MPFC]. Graphs at the bottom of the image display signal change (parameter estimates extracted from spherical ROIs identified from a meta-analysis by Spreng, Mar, & Kim, 2008) for each trial type (congruent-Democrat, incongruent-Democrat, congruent-Republican, incongruent-Republican) for each of these brain regions. Inspection of these figures confirms that preferential activation was obtained for both socially incongruent conditions.