



HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2015 June 24.

Published in final edited form as:

Neuroimage. 2013 December ; 83: 599–608. doi:10.1016/j.neuroimage.2013.07.001.

Outcome dependency alters the neural substrates of impression formation

Daniel L. Ames and Susan T. Fiske

Princeton University, USA

Abstract

How do people maintain consistent impressions of other people when other people are often inconsistent? The present research addresses this question by combining recent neuroscientific insights with ecologically meaningful behavioral methods. Participants formed impressions of real people whom they met in a personally involving situation. fMRI and supporting behavioral data revealed that outcome dependency (i.e., depending on another person for a desired outcome) alters previously identified neural dynamics of impression formation. Consistent with past research, a functional localizer identified a region of dorsomedial PFC previously linked to social impression formation. In the main task, this ROI revealed the predicted patterns of activity across outcome dependency conditions: greater BOLD response when information confirmed (vs. violated) social expectations if participants were outcome-independent and the reverse pattern if participants were outcome-dependent. We suggest that, although social perceivers often discount expectancy-disconfirming information as noise, being dependent on another person for a desired outcome focuses impression-formation processing on the most diagnostic information, rather than on the most tractable information.

Keywords

dmPFC; Impression formation Social cognition; fMRI; Inconsistency Outcome dependency

Introduction

Consistent impressions of inconsistent people

The ability to distill the vast amount of interpersonal information that people encounter each day into compact impressions is critical for making sense of the social world. As such, a central goal of cognitive neuroscientists studying social processes has been not only to define the mental operations and neural processes that give rise to social impressions, but also to characterize the nature of these impressions themselves. One consistent observation from behavioral research has been that not all social information counts equally – rather, certain pieces of information come to compose central expectations about people, and these expectations exert a strong pull over how subsequent information is interpreted. Historically, social psychologists have expressed this observation in terms of trait centrality (Asch, 1946),

trait primacy (Asch, 1946), implicit personality theory (Rosenberg and Sedlak, 1972), social schemata (Delia and Crockett, 1973), prototypes (Cantor and Mischel, 1979), and various theories of stereotyping (Hamilton and Sherman, 1996).

In parallel, cognitive neuroscience has long viewed this drive toward coherent representations as a general property of cognition and perception (e.g., Sporns et al., 2004; Tononi et al., 1998). Likewise, the importance of perceivers' expectations in guiding these integrative processes has been expressed in numerous theoretical contexts (and numerous brain regions), including visual perception (e.g., feature integration theory; Schoenfeld et al., 2003; Treisman and Gelade, 1980), language acquisition (e.g., native language neural commitment; Kuhl, 2004; Saffran et al., 1996), discourse comprehension (Martín-Loeches et al., 2008) and memory formation (e.g., hippocampal/neocortical interactions theory; Wang and Morris, 2010).

These various perspectives all predict (correctly) that people will tend to form coherent, stable impressions of other people, objects, and scenes. This is adaptive, because representing the world as coherent and stable makes the world more comprehensible and easier to act on. However, the brain's proclivity to extract structure and patterns from noisy inputs leads to more-than-occasional cognitive missteps. People see coherent objects where none exist, confidently invest money to capitalize on illusory stock market patterns, and construct memories that comport well with expected event structures, but poorly with actual events (e.g., Bartlett, 1932; Whitson and Galinsky, 2008).

Given the brain's general (over)zealousness for building coherence, it is unsurprising that (at least according to the dominant models in social psychology) people typically construe other agents as consistent entities whose actions are guided primarily by stable dispositions (Gilbert and Malone, 1995; Jones and Harris, 1967; c.f. Malle, 2006). This personality-driven construal (notably, a primarily Western phenomenon, Choi et al., 1999) is, in many ways, unrealistic. People are, in fact, remarkably variable in their behavior across time and situations (Ross and Nisbett, 1991). Yet “knowing” that people are variable does not necessarily diminish the drive toward stable social impressions – just as “knowing” that the stairs in M.C. Escher's *Ascending and Descending* (1960) are logically irreconcilable does not diminish the drive to construct a visually coherent staircase. An important question then, is how people maintain consistent impressions of other people when other people are so inconsistent.

The tools of cognitive neuroscience can be usefully applied to this question, having already delineated the biological underpinnings of several “coherence problems” (see the *Achieving consistency by discounting the inconsistent* and *Achieving consistency by integrating the inconsistent* sections), as well as many of the structures that contribute to social impression formation. By far the most consistent area to emerge in studies of impression formation is a dorsal region of medial prefrontal cortex (dmPFC; for meta-analyses and reviews, see Denny et al., 2012; Mitchell, 2009; van Overwalle, 2009; Wagner et al., 2012). Several other regions, including the temporo-parietal junction (TPJ), amygdala, posterior cingulate cortex (PCC), inferior frontal gyrus (IFG), and superior temporal sulcus (STS), have also been implicated in impression formation processes (Cloutier et al., 2011; Freeman et al., 2010;

Ma et al., 2011; Mende-Siedlecki et al., in press; Mitchell et al., 2005; Schiller et al., 2009). Yet this research reveals very little about what specific processing strategies might be deployed to resolve what is arguably the fundamental problem of impression formation (Hamilton and Sherman, 1996): creating highly coherent representations from highly divergent information. Moreover, neuroimaging studies that attempt to examine impression formation under relatively naturalistic conditions are all but absent from the literature. This is perhaps puzzling, since the functional value of impression formation (at least as described by some cognitive neuroscientists) lies largely in being able to understand and predict other people. By understanding others and predicting their behavior, one can improve one's social interactions, and better achieve desired outcomes – both social and material. Yet these studies rely on forming impressions of “people” (usually face databases and/or invented names) with whom participants will never interact, who cannot help perceivers to desired outcomes, and who may not even be regarded as “real.” Thus, the question of how dMPFC (or other regions) might respond under more involving conditions remains unanswered.

Achieving consistency by discounting the inconsistent

A review of relevant research points toward two (conflicting) approaches that people might use to create and sustain coherent social impressions. The first is to discount or explain away information that does not conform to preconceived expectations. Well-established theories from neuroscience (Kersten et al., 2004), cognitive psychology (Anderson, 1998) and social psychology (Fiske and Linville, 1980; Snyder and Swann, 1978) converge on the notion that selectively discounting expectancy-disconfirming information is an efficient learning strategy, relieving people of the burden of interpreting information that is difficult to process and that, given what is already “known” seems more likely to represent noise than signal (consistent with a Bayesian learning approach; Anderson, 1998).

Achieving consistency by integrating the inconsistent

However, not all expectancies are accurate; therefore, not all expectancy-disconfirming information is noise. Inaccurate impressions arise partially from the fact that people often form these impressions based on minimal evidence. For example, people can provide a judgment of others' trustworthiness after seeing their face for as little as 33 ms (Todorov et al., 2009). The amygdala, orbitofrontal cortex, and anterior insula have been frequently implicated in these rapid, intuitive impressions. These judgments, though not necessarily accurate, nonetheless predict important outcomes, including political elections and criminal sentences (for an overview, see Ames et al., 2011). This and other research (e.g., Ambady and Rosenthal, 1993; Devine, 1989) highlights the fact that social expectancies, while strongly felt and demonstrably influential, are often based on scant evidence. Thus, under-informed expectancies routinely become the lenses through which other people are viewed. In principle, information that violates these expectancies provides a means of correcting the prescription of these lenses, delivering valuable cues as to when impressions may be erroneous, while simultaneously provisioning the perceiver with the raw materials for building a more nuanced understanding. Revising impressions takes effort, however, and often the core goal of maintaining cognitive consistency trumps the objective of perceiving the world accurately (Hamilton and Sherman, 1996), people being cognitive misers (Fiske and Taylor, 2013).

Still, people do sometimes attend more to unexpected information than to expected information, with inferior frontal and temporoparietal cortices often playing a key role in reorienting visual attention toward expectancy violations (e.g. Corbetta and Shulman, 2002; Mitchell, 2008; Schank and Abelson, 1977), and posterior STS frequently observed in conjunction with unexpected changes in social gaze or movement (Frith and Frith, 2010; Pelphrey et al., 2003; Saxe et al., 2004). Moreover, these violations, when attended to, can inform social impressions (Srull and Wyer, 1989). Some of these findings appear to conflict with the literature reviewed in the previous section, suggesting that people may sometimes employ a second impression formation strategy, one that maintains coherent impressions, not by explaining away incongruous information, but by adjusting the impression to accommodate that information.

Which one when?

In sum, there are at least two competing approaches by which people might maintain consistent impressions of other people – explaining away inconsistency to preserve the impression, and altering the impression to fit the inconsistency. Given that previous literature provides examples of both approaches to maintaining coherent impressions (both social and nonsocial), it seems likely that the appropriate question is not *which* approach people use, but rather *which one when*.

Neural predictions and rationale

The present study investigates one answer to this question (recognizing that there may be more than one). It begins with the following premise: people pay attention to information that helps them get what they want. This suggests that when people depend on someone else for a desired outcome (when they are *outcome-dependent* on that person), perceivers may attend to information about the other person that they would ordinarily ignore (including, perhaps, expectancy-disconfirming information). This idea is supported by behavioral research showing that people selectively allocate limited cognitive resources toward people who are most apt to have functional implications (Ackerman et al., 2006; Rodin, 1987; Sporer, 2001). Also consistent with this hypothesis, several behavioral studies from the social attention literature reveal that being outcome-dependent focuses interpersonal attention (as measured in looking time) on inconsistencies (Erber and Fiske, 1984; Neuberg and Fiske, 1987; Ruscher and Fiske, 1990).

But while the thesis that outcome dependency increases attention to otherwise ignored information is well supported, the hypothesis that outcome dependency alters impression formation processes has (perhaps surprisingly) received little support from this literature (see the *Discussion* section). However, this may be largely explained by methodological limitations. Prior investigations into the effect of outcome dependency on impression formation have lacked a dependent variable that measures, in real time, the extent to which any given piece of information engages the cognitive processes subserving impression-formation. Cognitive neuroscience provides such a measure, as well as a large corpus of data indicating what areas of the brain most reliably index these processes. As noted in the *Consistent impressions of inconsistent people* section, the most consistently observed region in impression formation tasks has been dmPFC. While the responsiveness of this region to

social impressions has been long established, new studies have continued to articulate further the region's critical role in processing others' traits and personalities. For example, a recent experiment employing multivoxel pattern analysis (MVPA) required participants to learn the personalities of four individuals. A whole-brain searchlight procedure revealed that the specific identity of the person being thought about could be reliably decoded from dmPFC, but not from any other region (Hassabis et al., in press).

More generally, dmPFC is well-suited to perform the various cognitive functions required by impression formation, having been implicated in the online integration of information across time (Hasson et al., 2007), understanding the motivations behind other agents' actions (Spunt et al., 2011), encoding memories related to people (Mitchell et al., 2004), and goal-directed retrieval of semantic information (Binder et al., 2009). Though it makes good theoretical sense that the computations subserving these diverse processes would factor into impression formation, the fact that dmPFC is a large area of cortex with a large set of functions presents interpretational challenges. For this reason, the present study employs a well-established impression-formation task as a localizer to identify voxels specifically sensitive to impression-formation. Convergent and divergent evidence for this region's functional profile is then sought through independent behavioral measures; analysis of the main task is then restricted to these a priori-defined voxels.

This approach provides an opportunity to test the hypothesis that depending on another person for a desired outcome leads to changes in the neural dynamics of impression formation – with expectancy -confirming or -disconfirming information differentially driving dmPFC depending on whether the perceiver's ability to achieve desired outcomes depends on the target. Specifically, we predict that voxels in dmPFC that subserve impression formation should be more recruited in processing expectancy-confirming over -disconfirming information under *outcome-independent* conditions, but should show the reverse pattern under *outcome-dependent* conditions. This is the first study to examine the role of asymmetrical outcome dependency in neural systems of impression formation.

Ecological validity

Several methodological features of the present experiment differentiate it from previous studies. The most important of these concerns ecological validity. First, participants formed impressions of real people whom they had met in person several minutes prior to scanning. Second, participants expected to work cooperatively with these individuals immediately following scanning. Third, a desired reward was contingent on this work; thus, the paradigm presented participants with a genuine motivation to form impressions (rather than simply instructions to do so), as their understanding of their partners would likely affect their ability to work with them and to obtain the desired outcome. Fourth, the information that participants saw during the impression formation task was functionally relevant to the participants themselves, since it pertained to the interaction they expected to have a few minutes later. Finally, participants performed no explicit task during scanning, allowing us to study impression formation under somewhat more naturalistic conditions.

Methods

Behavioral procedure Cover story

Participants ($N = 19$, 6 male; $M = 19.1$ years, $SD = 1.2$) were told that the purpose of the study was to investigate “how experts and non-experts work together to solve problems creatively.” Prior to scanning, the experimenter introduced the participant to two confederates, explaining that the confederates were studying education at a neighboring university. It was further explained that they had been hired as “expert consultants” because the present study involved educational themes. To support the ostensible university affiliation of the confederates, they made subtle use of authentic props that had been purchased from the neighboring university's bookstore (e.g. a lanyard featuring that university's name pulled incidentally from a confederate's pocket). The task (adapted from Erber and Fiske, 1984), which would supposedly be completed after the fMRI scan, would require using a set of engaging wind-up toys (which participants saw on the experimenter's desk) to invent educational games for first-graders – for example, the concept of addition might be explored by having two of the toys hop toward a larger set. The experimenter noted that the participant and confederates would first generate ideas on their own, and then each expert-novice pair would work together to produce a joint product from their individual ideas (i.e., the participant would work with each confederate in turn).

Outcome dependency manipulation

To manipulate participants' outcome dependency, the experimenter explained that a \$50.00 prize would be awarded to the non-expert participant who performed best in the task involving the toys (confederates, being experts, would be paid a flat consulting fee). This was justified by explaining that one of the study's objectives was to explore the role of incentive structures in expert/non-expert collaboration. In order to allow the study to compare two different incentive structures, the participant would work with each of the two “experts” under slightly different reward conditions. When the participant worked with Confederate I (*outcome-independent* condition), the participant's own eligibility for the \$50.00 prize would depend only on what the participant did in the first phase of the task (generating ideas individually). In contrast, when the participant worked with Confederate D (*outcome-dependent* condition), the participant's own eligibility for the prize would depend on what the participant and confederate did together during the second phase of the task. Thus, participants were *outcome-dependent* on one confederate (because participants depended on this confederate for a desired outcome – gaining \$50.00), but *outcome-independent* with respect to the other confederate (because the second confederate's contributions would not be weighed in determining prize eligibility).

Four points should be noted with respect to the outcome dependency manipulation. First, participants expected to do exactly the same activities with both confederates (first generating ideas and then working on those ideas collaboratively); the only difference would be the criteria for prize eligibility (that is, outcome dependency). Second, in neither condition was the collaboration competitive between partners because the confederates were not eligible for the prize – again, only the basis of prize eligibility for the participant differed across conditions. Third, because the confederates were not eligible for the prize, the

participants' outcome dependency status (the variable of interest) was not confounded with factors of non-interest, such as the desire to help the expert partner maximize their outcomes. Finally, confederate identity and outcome dependency condition were counterbalanced across participants (i.e., the actors playing the roles of the education experts traded off which part they played).

Establishing expectancies

Following the outcome dependency manipulation, participants and confederates completed a *Participant Information Form*. Reading the confederates' forms allowed participants to naturalistically generate expectations about each of them (so that expectancy -confirming and -disconfirming information could be shown during subsequent scanning). The form included several demographic items, and a blank space in which participants and confederates were to predict how well they would do on the toy task, and also to discuss their background (if applicable) in this sort of work. Confederates wrote prepared statements in this space—one describing herself as “lov[ing] teaching K and first-grade” and as being “a camp instructor for 6 – 8-year-olds,” and the other describing herself as “prefer[ring] to work with middle-schoolers” and “now specializing [with that age group]” (see supplemental materials for complete self-descriptions). For convenience, we refer to these as expectancies as the “grade-school expectancy” and the “middle-school expectancy,” respectively. The expectancies themselves were of no scientific interest (fully counterbalanced); expectancies were simply necessary so that the effects of expectancy -confirming and -disconfirming information could be examined. Analyses revealed no difference in liking for the two expectancy-paired confederates (assessed by having participants make a hashmark on a line indicating how much they liked each confederate; *grade-school expectancy*, $M = 9.4$, $SD = 1.5$; *middle-school expectancy*, $M = 8.1$, $SD = 2.5$; *ns*).

Confederates then left (“to work on their teaching questionnaires”) while the participant prepared for the fMRI portion of the experiment. Prior to scanning, participants reviewed each confederate's personal statement paired with her snapshot, a manipulation that had reliably produced the predicted interpersonal expectations in pretesting.

Attention measure: glancing behavior

After the outcome-expectancy manipulation and before dismissing the confederates, an assistant surreptitiously recorded (using paper and pen) how often the participant looked at each of the two confederates, who were seated on either side of the participant (fully counterbalanced with all other variables; see supplementary materials for additional procedural details). Post-experimental interviews confirmed that both participants and confederates were unaware of this procedure during data collection.

Main task

Immediately prior to fMRI scanning, participants were given the following instructions:

One of the things that might influence how people work with expert partners is what they know about those partners and the background from which they derive

their expertise. To that end, we're going to show you some "teaching notes." We're interested in comparing your brain responses as you're learning about these people to what you do when you work with these people.

During fMRI scanning, participants viewed sentences about the two confederates without any explicit task. Participants believed that these sentences were excerpts of the confederates' "teaching evaluation notes" from the past semester (participants watched the confederates hand these notes on flash drives to the experimenter, and watched the experimenter upload them to the computer that presented the scanning task). Ostensibly, these notes had been taken by other student teachers while the confederates worked with classrooms of first graders. In truth, these statements had been invented for the study. Half of these ($n = 60$) had been pretested as confirming the *grade-school expectancy* (e.g., "[She] read in different voices for the different characters;" "If anything, she spent too much time on making [the lesson] fun"). These same 60 items were also pretested on separate subjects as *disconfirming the middle-school expectancy*. The other half of the stimuli ($n = 60$) were pretested as confirming the *middle-school expectancy*, and (again, separately) as also *disconfirming the grade-school expectancy* (for example, by referencing teaching methods well-suited to middle-school students, but poorly suited for grade-school students; "Her turtle example was interesting, but probably too complicated;" "She maybe forgot that not all of the kids knew how to tell time"). Conditions were matched for statement length (number of words).

During scanning, participants viewed each of the 120 statements once (approx. 3450 ms per statement, jittered). The intertrial interval separating statements was variable from 1250 ms to 9250 ms in steps of 2000 ms to obtain temporal jitter required to deconvolve event-related fMRI response. For each participant, a random half of these statements were paired with the *grade-school expectancy* confederate, and the other half were paired with the *middle-school expectancy* confederate. The snapshot that appeared with a given statement indicated to whom that statement pertained. Thus, half ($n = 30$) of the statements that were paired with each confederate *confirmed* the participant's expectancy about that confederate, and the other half ($n = 30$) *disconfirmed* that expectancy. Across participants, each of the 120 stimuli was equally likely to be assigned to any of the four cells in the 2 (dependent/independent) \times 2 (consistent/inconsistent) design (i.e., all stimuli served as their own controls); and, only *outcome dependency* and *expectancy-confirmation/disconfirmation* (the two variables of interest) varied across conditions.

Functional localizer

Following the main task, participants completed a second, well validated experiment used to identify regions of the brain preferentially recruited for forming social impressions. In line with prior neuroimaging (Mitchell et al., 2004, 2005) and behavioral studies (e.g., Hamilton et al., 1980; Klein and Loftus, 1990), participants saw a photograph of an unknown person and a piece of information pertaining to that person on each trial. On "form impression" trials ($n = 20$), participants were asked to form an impression of the person pictured. On "remember order" trials ($n = 20$), participants were asked to remember the order in which the information appeared. Following previous experiments, participants believed that they would be tested on both conditions after scanning. Although both conditions require

participants to relate each piece of information to a specific individual, only the form-impression condition requires impression formation. Recent data from other researchers (Ma et al., 2011, 2012) show that these deliberate “form impression” instructions produce the same patterns of activation in dmPFC as spontaneous impression formation. That is, regardless of whether participants deliberately use information to form impressions or encounter impression-relevant information with no explicit task demands, dmPFC apparently contributes the same functions (while major differences emerge in other areas, including a more posterior region of medial frontal cortex [pmFC] and right prefrontal cortex [rPFC]).

As with the main task, assignment of stimuli to condition was randomized across participants such that each stimulus served as its own control. This secondary experiment provided an independently defined impression-formation region of interest (ROI) identified from completely separate scans. Limiting analyses of the main task to this region reduces the need for subjective judgments of functional localization, providing a more principled test of the main hypothesis. Previous experiments employing this task have consistently identified a region of dmPFC (Mitchell et al., 2004, 2005). The present investigation employed identical procedures and stimuli to those used in these previous studies. This design does not include a baseline condition of faces without information.

fMRI data acquisition

BOLD signal was used to assess regional brain activation. EPI were acquired with a Siemens 3.0-T Allegra scanner (Siemens, Erlangen, Germany) with a standard head coil (TR = 2000 ms, TE = 30 ms, flip angle = 90°, matrix size = 64 × 64). Whole brain coverage was achieved with 34 interleaved 3.6-mm axial slices with an interslice gap of 0.36 mm. At the end of each scan session, a high-resolution anatomical image was acquired (T1-MPRAGE, TR = 2000 ms, TE = 30 ms, flip angle = 8°, matrix size = 256 × 256) (Fig. 1).

Preprocessing

fMRI data were preprocessed using BrainVoyager software (Brain Innovation, version 1.8). Preprocessing included slice time correction, linear trend removal and high-pass filtering (frequencies below 3 cycles per functional run removed). To correct for head motion, we used a 3D algorithm that adjusts for small head movements via rigid body transformations of all slices to the first reference volume. All functional images were transformed to standard Talairach space so that corresponding brain regions were roughly spatially aligned. To further overcome misregistration across subjects, the data were spatially smoothed with a Gaussian filter of 6-mm full width at half-maximum value. To remove transient nonspecific signal elevation effects at the beginning of the experiment, we excluded the first 5 timepoints of each functional run.

In order to correct for false positives, activations were cluster-corrected using standard software: AFNI's *alphaSim*. Specifically, we applied a Monte Carlo simulation that takes into account the matrix size and smoothness of the specific dataset under consideration in order to determine the appropriate cluster size threshold (Forman et al., 1995). For the present dataset, a minimum cluster size of 22 voxels combined with a voxelwise threshold of

$p < .005$ was found to be required to control for false positives at a corrected alpha of $p < .05$. Importantly, the estimate of the spatial correlation across voxels was modeled using the program 3dFWHMx, not the amount of Gaussian smoothing applied during preprocessing (see Bennett et al., 2009).

Individual difference measures

Following scanning, participants completed the personal need for structure scale (PNS; Neuberg and Newsom, 1993), a measure of individuals' preference for structure and internal consistency. A straight-forward interpretation of this scale predicts that people exhibiting greater tolerance for inconsistency should be more likely to use (rather than ignore or explain away) expectancy-inconsistent information. We therefore capitalized on the existence of this scale to provide an independent means of validating our region of interest. Specifically, we predicted that PNS scores would correlate positively with BOLD response to expectancy-confirming vs. -disconfirming information (that is, greater difference scores of average betas). In order to show discriminant validity, participants also completed the implicit person theory scale (IPT Levy et al., 1998). This frequently-used scale shares some psychometric characteristics with the PNS, including a six-point response range (1–6) and identically-worded anchors (“strongly disagree”–“strongly agree”). Critically, if both scales index acceptance of discrepant information—however the PNS focuses on the belief that people can change over time, a theme not especially related to the present study, since all “teaching evaluation notes” were ostensibly from the same semester. We predicted that the PNS, but not the IPT would correlate with activity in the ROI. Two participants did not complete these questionnaires because their scanning sessions ran over time.

Functional characterization of region of interest

Thus, prior to testing our main hypotheses, we establish three separate lines of evidence to suggest that the region of cortex we investigated in this study supports the processes of interest (though it certainly supports other processes as well). First, dozens (if not hundreds) of imaging studies have now implicated dmPFC in impression formation (Denny et al., 2012; Mitchell, 2009; van Overwalle, 2009; Wagner et al., 2012). Second, we identified an ROI a priori using a functional localizer. This task has been validated in both behavioral and neuroimaging research and consistently identifies a region of dmPFC. Moreover, the region identified by this task correlates with successful encoding (i.e., memory) for person information when participants are forming an impression of the target (but not when performing a sequencing task; Mitchell et al., 2004). Third, we added an additional layer of independence by validating the *function* of this region (in addition to its location) by testing whether that function correlated with an individual differences measure of personal need for structure (but not with a superficially similar scale)

Results

fMRI results

Functional localizer—We began our analysis by identifying regions of cortex subserving impression formation in the functional localizer task using a standard random effects GLM contrast (*form impression* N *sequencing*). Consistent with previous studies, the functional

localizer identified a region of dorsomedial prefrontal cortex ($[-12, 25, 54]$, 22 voxels, $3 \times 3 \times 3$ mm, $p < .05$ cluster-corrected) as preferentially responsive to the *form impression* condition vs. the *sequencing* condition. A more ventral region of mPFC and right anterior lobe also emerged from this contrast (see Table 1). As noted in Functional characterization of region of interest section, activity in this same region of dmPFC, defined by the same task and using the same stimuli and instructions, correlates with correct recall of person information during impression formation trials, but not during sequencing trials.

Science is cumulative, and rather than replicate this recall effect, we sought to build upon it, first by showing a new behavioral correlate for this now-standard impression formation region (see Personal need for structure scale section), and then by using the functional localizer to test our primary hypotheses concerning outcome dependency and social expectations.

Main analysis—In the main analysis, we again employed a standard random effects GLM approach, modeling the data as a 2 (outcome-dependent, outcome-independent) \times 2 (expectancy-confirming, expectancy-disconfirming) ANOVA. The predicted crossover interaction was observed in the dmPFC ROI, $F(1, 18) = 8.40$, $p < .01$, partial $\eta^2 = .33$. When participants did not depend on the target, this region was more active in processing *expectancy-confirming* information than *expectancy-disconfirming* information, $t(18) = 2.34$, $p < .05$, uncorrected. This suggests that, when outcome-independent, participants selectively marshaled impression information resources to process information that seemed more likely to represent signal than noise, given what they already knew. However absent outcome dependency, participants discounted the unexpected as noise, thereby fulfilling the goal of maintaining coherence without engaging in the effortful and complex task of changing the impression (Table 2).

The pattern reversed, however, when participants depended upon the target for something desirable (money), with expectancy-*disconfirming* information driving the impression formation ROI marginally more than expectancy-confirming information, $t(18) = 1.9$, $p = .07$, uncorrected. One likely interpretation of this reversal is that when participants had something at stake (i.e., when they depended on the confederate for their financial outcomes), they may have had other goals beyond maintaining coherence—such as forming an accurate impression to help them better predict (and perhaps influence) the confederate and thereby obtain the desired outcome. They may have therefore taken seriously the unexpected information and used it to update their impressions (Mende-Siedlecki et al., in press), while mobilizing limited impression formation resources relatively less for expectancy-confirming information (which is, by definition, redundant, and therefore less diagnostic). This interpretation is supported by previous research demonstrating that cognitive and perceptual resources are typically allocated toward individuals with the most significant implications for the perceiver (e.g., Ackerman et al., 2006; Rodin, 1987; Sporer, 2001). Post-hoc analysis of the more ventral mPFC region revealed a similar pattern of results; however the interaction was nonsignificant, $p = .12$.

It is worth reiterating that, across subjects, all four conditions contained exactly the same words and pictures pertaining to exactly the same people presented in exactly the same

format. The only differences were those that the participants themselves brought into the scanner in the form of (a) their preconceived expectations about the two targets and (b) their differential dependency on those targets for a desired outcome.

No main effect of dependency was observed ($F < 1$), suggesting that outcome dependency does not simply lead to more impression formation, but rather to a qualitatively different style of impression formation—one that (we suggest) shifts neural resources from a focus on information that is redundant but easy to assimilate, to information that is more difficult to make sense of, but perhaps also more diagnostic. Likewise, no main effect of consistency was observed ($F < 1$), supporting our claim that the question of how people maintain coherent impressions should not be treated as a question of which kind of information is privileged, but rather which kind of information best serves the perceiver's goals across different social contexts.

Behavioral results

Glancing behavior—People attend to those on whom they depend (see Integration with existing behavioral research section). Thus, if the outcome dependency manipulation in this study was effective, participants should have attended more to the confederate on whom they were outcome-dependent than to the confederate from whom they were outcome-independent. One ecologically valid means of assessing a person's attentional focus is to observe where the person is looking (Rayner, 1998). As predicted, participants looked more often at the confederate on whom they were outcome-dependent ($M = 2.5$ glances, $SD = 2.3$) than at the participant from whom they were outcome-independent ($M = 1.6$, $SD = 2.0$), $t(16) = 2.4$, $p < .05$, $d = .42$. The durations of individual glances were not recorded.

Personal need for structure scale—It was predicted that people scoring higher in personal need for structure would show a preferential neural response for information that cohered with the structure of their expectations about the targets. Specifically, we expected this differentiation between expectancy-confirming/-disconfirming information to emerge in our a priori-defined subregion of dmPFC, with higher PNS scores corresponding to greater relative responses to expectancy-confirming vs. -disconfirming. This prediction was confirmed, $r(17) = .55$, $p < .05$ (independent condition: $r = .52$, $p < .05$; dependent condition: $r = .23$, $p = .37$, ns). Also as predicted, no such correlation was observed when the same analysis was conducted using the implicit person theory scale. To verify that PNS but not IPT significantly predicted dmPFC response, the data were reanalyzed using multiple regressions, with PNS and IPT entered simultaneously as predictors of dmPFC response. The overall model was significant, $R^2_{adj} = .26$, $F(2, 14) = 3.76$, $p < .05$. As predicted, PNS emerged as a significant predictor of dmPFC response, $\beta = .506$, $SE = .02$, $t(14) = 2.29$, $p < .05$, while IPT did not, $\beta = .214$, $SE = .03$, $t(14) = 0.97$, $p = .35$, ns. PNS, IPT and glancing behavior were not significantly correlated with one another ($ps > .3$).

This suggests that, at least in the present investigation, dmPFC BOLD response during social perception relates predictably to differences in how individuals react to expected and unexpected information (PNS); this result cannot be accounted for by more general lay theories about the fixedness of others' personalities and behaviors (IPT). These findings add

to a large set of evidence that the region of dmPFC defined by the impression-formation localizer indexes greater recruitment of impression formation resources (as reviewed in Achieving consistency by discounting the inconsistent section). More importantly for present purposes, these data corroborate the proposal that the region of dmPFC under investigation is functionally sensitive to the specific kinds of information (expectancy-confirming vs. -disconfirming) that are most useful to individual perceivers. We hasten to add that correlations with relatively few observations are notoriously unreliable, and that this novel effect requires replication. However, while this individual differences analysis is correlational, the main analysis tests this same relationship experimentally.

Discussion

The consistency problem

In an influential analysis of the (then) 50-year history of research in social impression formation, Hamilton and Sherman (1996) identified the tendency toward unity and coherence as “the fundamental postulate” of impression formation: “*The perceiver assumes unity in the personalities of others, and persons are seen as coherent entities; therefore, one’s impression of another person should reflect that unity and coherence*” (p. 337; italics in original). This view still describes the dominant perspective within social psychology. Nevertheless, a substantial body of research suggests that people are in fact neither unified nor coherent—that human beings are most aptly characterized as a collection of multiple selves, and that their cognitions and behaviors vary greatly with the situation (Roberts and Donahue, 1994; Ross and Nisbett, 1991). This returns us to the question posed in the introduction: how do people maintain unified and consistent impressions of other people when other people are often inconsistent?

To help answer this question, this study drew upon methods and perspectives from neuroscience, cognitive psychology and social psychology. All three converge on the notion that one approach for solving the consistency problem is to discount or disengage from information that contradicts what is already “known” (Anderson, 1998; Kersten et al., 2004; Snyder and Swann, 1978), and which may therefore seem more likely to represent noise rather than signal.¹ A second approach exists however: instead of viewing inconsistency as “noise,” one may see it as a clue that one’s current impression is inaccurate or incomplete. In this case, one may choose to focus one’s limited processing resources on the expectancy-*disconfirming* information—which provides the raw materials for constructing more complete impressions—while treating expectancy-*confirming* information as redundant (Schank and Abelson, 1977; Srull and Wyer, 1989). This approach, though potentially more effortful (Fiske and Neuberg, 1990), may be more useful when people depend on another person to reach a desired outcome. Specifically, a more nuanced understanding of another

¹This approach suggests an essentially Bayesian method of integration. A growing body of evidence suggests that Bayesian integration aptly characterizes neural processes in many domains such as object perception, motor learning, and sensory integration in spatial encoding (Battaglia et al., 2003; Kersten et al., 2004; Knill and Pouget, 2004; Körding and Wolpert, 2004). We suggest that Bayesian theoretical models currently being employed to understand lower-level aspects of brain function might be profitably extended for understanding social cognition (see also Baker et al., 2006). Of course, the term “Bayesian” encompasses a broad range of theoretical models. I. J. Good (1971) calculated (at least half-jokingly) that there were at least 46,656 different kinds of “Bayesians”. Formal tests of any specific Bayesian model of social processing in the brain await future investigation.

person that takes into account conflicting information might help one predict or influence that person. This ability to predict and influence could, in turn improve one's chances of gaining the desired outcome from the person in question.

What information counts when?

Integration with existing neuroimaging research—The results reported in this paper suggest that people in fact use both of these approaches, and that they flexibly (though not necessarily consciously) switch between them as a function of social dependency. Impression formation researchers have long recognized that not all information contributes equally to a person's understanding of others' traits. The question of how this information is weighted and, specifically, how this weighting is instantiated in the brain, is a topic of considerable scientific interest (e.g., Cloutier et al., 2011; Harris et al., 2005; Ma et al., 2011, 2012; Mende-Siedlecki et al., in press; Mitchell et al., 2006; Schiller et al., 2009; Van Duynslaeger et al., 2007). The present findings offer a framework for systematically predicting and organizing this weighting (while recognizing that there may be other useful frameworks as well)—specifically, as a function of a person's (a) expectations about the target and (b) dependency on the target.

Past investigations of the weight given to different types of social information have focused on two relatively stable factors: (a) the content of the information itself (“he kicked a dog” is more diagnostic than “he kicked a soccer ball”) and (b) individual differences in observers (people may vary in their moral beliefs about harming dogs) (Schiller et al., 2009). Without denying the importance of these factors, the present research emphasizes more dynamic aspects of this weighting process, showing that the diagnostic value of any given piece of social information can change depending on (a) social dependency and (b) social expectations.

At first blush, this dynamic shifting of dmPFC activity may appear to conflict with the results of Schiller et al. (2009). Though Schiller and colleagues provided further evidence that dmPFC is recruited for processing social information, their investigation did not show modulations of dmPFC as a function of which information seemed to influence individuals' positive/negative valence judgments of targets. A second study by Mende-Siedlecki et al. (in press) also found no effect of valence in dmPFC. Valence, however, is only one component of social impressions; and it is unclear why one should predict that dmPFC (rather than other regions already known to subservise valuation) should be expected to encode the evaluative component of social impressions. In fact, these are the conclusions reached by Schiller and colleagues, who suggest that “the dmPFC is not essential for the evaluative component of impression formation,” and that this evaluative component rather “recruits brain regions that are not socially specialized but are more generally involved in valuation and emotional processes” (p. 512). Thus, previous research points out that informational relevance to valence-specific appraisals does not produce functional modulations in dmPFC, while the present work shows that social dependency and social expectations do.

Meanwhile, the role of interpersonal expectations in impression formation has been the focus of three recent fMRI studies employing more traditional paradigms. Cloutier et al.

(2011) observed greater dmPFC BOLD response to pictures of politicians who were paired with statements that contradicted their (ostensible) political group's norms (e.g., a picture of an ostensible Democrat paired with "He wants smaller government" vs. a picture of an ostensible Republican paired with "He is morally conservative," p. 584). Similarly, Mende-Siedlecki et al. (in press) and Ma et al. (2012) observed increased dmPFC BOLD amplitude in response to abrupt reversals in the valence of information paired with faces or names (e.g., showing participants the statement "Tolvan gave her sister a hug" followed by the statement "Tolvan gave her mother a slap;" Ma et al., 2012, p. 939). The present investigation aligns well with these studies in suggesting that the same information can be processed differently in dmPFC as a function of one's social expectations. However, Cloutier and colleagues observed a main effect such that unusual (and presumably unexpected) political beliefs elicited greater dmPFC activation. Likewise Mende-Siedlecki and colleagues and Ma and colleagues observed greater dmPFC response when information about a target switched from positive to negative or vice versa (as compared with when no such valence shift occurred). Each of these findings would seem to predict that the present experiment should find a main effect of expectancy-consistency, rather than the socially modulated reversal that was, in fact, observed.

One possible explanation is that when people view large sets of relatively anonymous faces or names, those faces that "stick out" as representing unusual individuals prompt significantly more concerted effort at impression formation (see Cloutier et al. for a useful discussion of this individuation processes and Mende-Siedlecki et al. for a useful discussion of impression updating). This interpretation is consistent with research on social attention (Erber and Fiske, 1984) and person memory (Ackerman et al., 2006) and aligns with the intuition that contradictory people are more interesting.

Of course, other factors can make people interesting as well. Targets in the present study were two live individuals whom the participants had just met in the scanning environment—with whom participants had shaken hands, conversed—and, critically, with whom the participants expected to work immediately after scanning. These two people may have therefore been more relevant and interesting to participants than those represented by face databases or invented names. The processing demands may be quite different in such situations. Specifically, because of constraints on perceptual and cognitive processing capacity (Todd et al., 2005), people tend to dynamically shift processing strategies, continually re-allocating cognitive resources to process the people and information that is most functionally important (e.g., Ackerman et al., 2006; Rodin, 1987; Sporer, 2001).

Thus, the processing demands of a more personally-involving paradigm provide one a theoretically acceptable explanation as to why the present study yielded a different pattern of results from those described above. However, direct comparisons are complicated by the fact that, in these previous studies, half of targets were expectancy-consistent and the other half were inconsistent; whereas in the present experiment, both targets were consistently inconsistent—that is they sometimes behaved as expected, and other times they did not. Relatedly, the expectations being violated in the previous studies were general categories (liberal/conservative) or valence (good/bad); while the expectations about the two teachers would seem to be more specific, personal, and relevant to participants, since these

expectations were about (a) people with whom participants were about to work, (b) the very thing that they were about to work on, and (c) the perceived likelihood of obtaining a desired outcome.

None of this is to ignore the fact that people do sometimes quickly form impressions of large numbers of people in succession, based on only a picture, a few words, and perhaps a category (e.g., members of a Facebook group, characters on a TV show, researchers on a departmental website). This observation underscores the importance of these previous studies. However, impression formation is sometimes also more personal, directed at a smaller number of people whom one has actually met, and has something to do with anticipating behavior, or developing some sense of the people with whom one might interact. We would suggest that the proportion of experiments that presently map onto these two kinds of situations may not reflect the proportions of actual human experience. In fact, despite the large number of neuroimaging experiments addressing social impression formation, we know of no other experiment on this topic in which participants formed impressions of real people whom they had actually met in person.

Integration with existing behavioral research—The BOLD modulations observed in this study align with previous behavioral research on the topic of outcome dependency. These studies demonstrated that outcome dependency can increase the time people spend looking at unexpected social information and the number of dispositional statements offered when participants are asked to think aloud into a tape recorder (Erber and Fiske, 1984; Neuberg and Fiske, 1987; Ruscher and Fiske, 1990). However, direct attempts to measure impression formation have been largely inconclusive to this point. Several experiments in this line of research attempted to explore this link by examining the effect of outcome dependency on liking for a target individual. These liking effects proved inconsistent across studies however. Moreover, when these results were found, the predicted mediating relationship between social attention and changes in liking did not emerge. This pattern of results is problematic for the prospect of inferring impression formation processes from liking ratings. As we have already noted however, there is more to social impressions than liking or disliking someone; thus “amount of liking” may simply be too coarse a measure for testing some hypotheses concerning impression formation. Likewise, while think-aloud protocols are not limited to assessing valence, such procedures force the participant to repeatedly break away from the task to report on their ongoing cognition (a process likely to interfere with normal processing strategies).

Perhaps the most frequently employed behavioral measure of “more” or “less” impression-formation is recall for information about the target. The typical inference is that if one remembers more about person A than person B, one is engaging in “more impression formation” for person A than for person B. However, evidence that impression-formation actually produces increased recall is, at best, mixed (Dabady et al., 1999; Erber and Fiske, 1984; Hamilton et al., 1980; Klein and Loftus, 1990; Ma et al., 2011; Marmurek, 1990; Mitchell et al., 2004; Sedikides et al., 1991). For example, Dabady et al. (1999) randomly assigned participants either to form an impression of a person using a set of information, or to remember the information about the person. At test, impression formation participants remembered no more information than did memorization participants.

The foregoing reviews previous behavioral approaches to linking outcome dependency to impression formation, and suggests possible methodological limitations to this research. A more general observation is that behavioral investigations concerning this topic typically aim to retroactively infer the extent to which participants *were* forming impressions—by using a summary statistic such as “number of items remembered” or “average liking for target.” Such measures either require the participant to repeatedly break away from the task (as with think-aloud protocols), or more commonly, are collected offline—after the hypothesized period of impression formation has concluded. Such methods, though ideally suited for many questions, make it difficult to capture the rapid and dynamic nature of the impression formation process, which can change rapidly as a continuous stream of information is presented. An alternative approach therefore is to measure neural responses in the brain online, as the process of interest (impression formation) is occurring. While the large number of previous studies implicating dmPFC activity in impression formation (see Achieving consistency by discounting the inconsistent section) provides the primary empirical basis for our interpretation of dmPFC activity in the context of this experiment, we also localize this function a-priori in our participants' brains using an independent task and provide a novel brain-behavior correlation to support this functional interpretation within the localized region (see Main analysis section).

However, this study does nothing to resolve the problems described in this section concerning the dearth of online *behavioral* measures that index “extent of engagement of impression formation processing.” As such, we can adduce no such measure from this experiment to further support the case for dmPFC's involvement in impression formation, and this is a limitation of the present experiment.

Conclusion

Operating from the premise that thinking is for doing (Fiske, 1992)—this experiment aimed to study how participants *thought* about other people when they were preparing to *do something* with those people, and when the doing had meaningful consequences for the participants. For this reason, we examined participants learning about real people whom they met under personally involving circumstances, and with whom they expected to work immediately following scanning.

Further integration of socially ecological paradigms with neuroscience is desirable, and this study serves in part as a response to recent calls for more ecologically valid paradigms in cognitive neuroscience, with an eye toward building more ecologically valid theories (Zaki and Ochsner, 2009). However, the fact that the predicted effects were observed within the region defined by a canonical impression formation task also offers some support for the validity of traditional tasks. Specifically, given that the dmPFC is by far the region most often implicated in traditional fMRI investigations of impression formation, it is encouraging that activity in this region does indeed comport with theory-driven predictions about impression formation and cognitive resource allocation when the targets are real people.

This work identifies two ways in which the neural resources subserving impression formation are allocated in the face of mixed information, and specifies social circumstances under which each is observed. The study combines various methodological approaches, incorporating a functional localizer, individual difference measures correlated with brain function (PNS), the use of actual human targets, and observation of subtle but ecologically meaningful behavior (stolen glances) to address what is currently a major topic of interest in cognitive neuroscience.

We speculate that the tactic of shifting impression formation resources away from expectancy-disconfirming information and toward expectancy-confirming information may be something of a default strategy for maintaining coherence (Fiske and Neuberg, 1990). However the suggestion that outcome dependency can elevate the importance of expectancy-disconfirming information may be encouraging for those concerned with contemporary issues in social welfare. An inflexible policy of discounting inconsistent information can thwart professional advancement (“black people are lazy”), personal relationships (“those kinds of men can’t commit”), and diplomatic efforts (“dictatorships are immune to reason”), while simultaneously causing perceivers to write off the very information that would undermine their preconceptions (a sterling résumé, a loving gesture, a request for peace talks). If depending on another person for something increases some kinds of processing for expectancy-disconfirming information, it may be possible to combat potentially blinding social expectancies (e.g., stereotypes) by making people (either apparently or actually) dependent on one-another for their outcomes. Though not all social situations can (or should) be restructured to create networks of interdependence, it may be possible to make interdependence more salient where it already exists. After all, people depend on others, either directly or indirectly, for nearly all of their desired outcomes, both social and material.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Ackerman J, Shapiro J, Neuberg S, Kenrick, Vaughn Becker D, Griskevicius V, Maner J, Schaller M. They all look the same to me (unless they're angry) from out-group homogeneity to out-group heterogeneity. *Psychol. Sci.* 2006; 17(10):836–840. [PubMed: 17100781]
- Ambady N, Rosenthal R. Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *J. Personal. Soc. Psychol.* 1993; 64(3):431–441.
- Ames, DL.; Fiske, ST.; Todorov, A. Impression formation: a focus on others' intents. In: Decety, J.; Cacioppo, J., editors. *The Handbook of Social Neuroscience*. Oxford University Press; 2011. p. 419-433.
- Anderson JL. Embracing uncertainty: the interface of Bayesian statistics and Cognitive Psychology. *Conserv. Ecol.* 1998; 2(1):2.
- Asch S. Forming impressions of personality. *J. Abnorm. Soc. Psychol.* 1946; 41:258–290.
- Baker CL, Tenenbaum JB, Saxe R. Bayesian models of human action understanding. *Adv. Neural Inf. Process. Syst.* 2006; 18:99–106.
- Bartlett, FC. *Remembering: A Study in Experimental and Social Psychology*. Cambridge, England: Cambridge University Press; 1932.

- Battaglia PW, Jacobs RA, Aslin RN. Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A.* 2003; 20(7):1391–1397.
- Bennett C, Wolford G, Miller M. The principled control of false positives in neuroimaging. *Soc. Cogn. Affect. Neurosci.* 2009; 4(4):417–422. [PubMed: 20042432]
- Binder J, Desai R, Graves W, Conant L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex.* 2009; 19(12):2767–2796. [PubMed: 19329570]
- Cantor, N.; Mischel, W. Prototypes in person perception. In: Berkowitz, L., editor. *Advances in Experimental Social Psychology*. Vol. 12. New York: Academic; 1979. p. 3-52.
- Choi I, Nisbett RE, Norenzayan A. Causal attribution across cultures: variation and universality. *Psychol. Bull.* 1999; 125(1):47–63.
- Cloutier J, Gabrieli JDE, O’Young D, Ambady N. An fMRI study of violations of social expectations: when people are not who we expect them to be. *NeuroImage.* 2011; 57(2):583–588. [PubMed: 21569855]
- Corbetta M, Shulman G. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 2002; 3(3)
- Dabady M, Bell M, Kihlstrom JF. Person memory: organization of behaviors by traits. *J. Res. Personal.* 1999; 33:369–377.
- Delia JG, Crockett WH. Social schemas, cognitive complexity, and the learning of social structures. *J. Personal.* 1973; 41:413–429.
- Denny B, Kober K, Wager T, Ochsner K. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* 2012; 24(8):1742–1752. [PubMed: 22452556]
- Devine P. Stereotypes and prejudice: their automatic and controlled components. *J. Personal. Soc. Psychol.* 1989; 56:5–18.
- Erber R, Fiske ST. Outcome dependency and attention to inconsistent information about others. *J. Personal. Soc. Psychol.* 1984; 47:709–726.
- Fiske ST. Thinking is for doing: portraits of social cognition from daguerreotype to laserphoto. *J. Personal. Soc. Psychol.* 1992; 63:877–889.
- Fiske ST, Linville PW. What does the schema concept buy us? *Personal. Soc. Psychol. Bull.* 1980; 6:543–557.
- Fiske, ST.; Neuberg, SL. A continuum of impression formation, from category- based to individuating processes: influences of information and motivation on attention and interpretation. In: Zanna, M., editor. *Advances in Experimental Social Psychology*. San Diego, CA: Academic Press; 1990. p. 1-74.
- Fiske, ST.; Taylor, SE. *Social Cognition: From Brains to Culture*. London: Sage; 2013.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved assessment of significant change in functional magnetic resonance imaging (fMRI): use of a cluster size threshold. *Magn. Reson. Med.* 1995; 33:636–647. [PubMed: 7596267]
- Freeman JB, Schiller D, Rule NO, Ambady N. The neural origins of superficial and individuated judgments about ingroup and outgroup members. *Hum. Brain Mapp.* 2010; 31:150–159. [PubMed: 19618409]
- Frith U, Frith C. The social brain: allowing humans to boldly go where no other species has been. *Philos. Trans. R. Soc. B.* 2010; 365:165–176.
- Gilbert DT, Malone PS. The correspondence bias. *Psychol. Bull.* 1995; 117:21–38. [PubMed: 7870861]
- Good IJ. 46656 varieties of Bayesians. *Am. Stat.* 1971; 25:62–63.
- Hamilton DL, Sherman SJ. Perceiving persons and groups. *Psychol. Rev.* 1996; 103:336–355. [PubMed: 8637962]
- Hamilton DL, Katz LB, Leirer VO. Cognitive representation of personality impressions: organizational processes in first impression formation. *J. Personal. Soc. Psychol.* 1980; 39(6):1050–1063.
- Harris LT, Todorov A, Fiske ST. Attributions on the brain: neuroimaging dis- positional inferences, beyond theory of mind. *NeuroImage.* 2005; 28(4):763–769. [PubMed: 16046148]

- Hassabis D, Spreng N, Rusu A, Robbins C, Mar R, Schacter D. Imagine all the people: how the brain creates and uses personality models to predict behavior. *Cereb. Cortex*. 2013 in press.
- Hasson U, Nusbaum HC, Small SL. Brain networks subserving the extraction of sentence information and its encoding to memory. *Cereb. Cortex*. 2007; 17:2899–2913. [PubMed: 17372276]
- Jones EE, Harris VA. The attribution of attitudes. *J. Exp. Soc. Psychol.* 1967; 3(1):1–24.
- Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. *Annu. Rev. Psychol.* 2004; 55:271–304. [PubMed: 14744217]
- Klein SB, Loftus J. Rethinking the role of organization in person memory: an independent trace storage model. *J. Personal. Soc. Psychol.* 1990; 59(3):400–410.
- Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 2004; 27(12):712–719. [PubMed: 15541511]
- Körding KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature*. 2004; 427(6971):244–247. [PubMed: 14724638]
- Kuhl P. Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 2004; 5(11):831–843. [PubMed: 15496861]
- Levy SR, Stroessner SJ, Dweck CS. Stereotype formation and endorsement: the role of implicit theories. *J. Personal. Soc. Psychol.* 1998; 74:1421–1436.
- Ma N, Vandekerckhove M, Van Overwalle F, Seurinck R, Fias W. Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: spontaneous inferences activate only its core areas. *Soc. Neurosci.* 2011; 6(2):123–138. [PubMed: 20661837]
- Ma N, Vandekerckhove M, Baetens K, Van Overwalle F, Seurinck R, Fias W. Inconsistencies in spontaneous and intentional trait inferences. *Soc. Cogn. Affect. Neurosci.* 2012; 7(8):937–950. [PubMed: 22006990]
- Malle B. The actor-observer asymmetry in attribution: a (surprising) meta-analysis. *Psychol. Bull.* 2006; 132(6):895–919. [PubMed: 17073526]
- Marmurek HHC. The dissociation of impression formation and person memory: the effects of processing resources and trait favorableness. *J. Res. Personal.* 1990; 24(2):191–205.
- Martín-Loeches M, Casado P, Hernández-Tamames J, Álvarez-Linera J. Brain activation in discourse comprehension: a 3 T fMRI study. *NeuroImage*. 2008; 41(2):614–622. [PubMed: 18394923]
- Mende-Siedlecki P, Cai Y, Todorov A. The neural dynamics of updating person impressions. *Soc. Cogn. Affect. Neurosci.* in press.
- Mitchell JP. Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex*. 2008; 18(2):262–271. [PubMed: 17551089]
- Mitchell JP. Inferences about other minds. *Philos Trans. R. Soc. B.* 2009; 364:1309–1316.
- Mitchell JP, Macrae CN, Banaji MR. Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *J. Neurosci.* 2004; 24(21):4912–4917. [PubMed: 15163682]
- Mitchell JP, Macrae CN, Banaji MR. Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. *NeuroImage*. 2005; 26:251–257. [PubMed: 15862225]
- Mitchell JP, Cloutier J, Banaji MR, Macrae CN. Medial prefrontal dissociations during processing of trait diagnostic and nondiagnostic person information. *Soc. Cogn. Affect. Neurosci.* 2006; 1(1):49–55. [PubMed: 18985100]
- Neuberg SL, Fiske ST. Motivational influences on impression formation: outcome dependency, accuracy driven attention, and individuating processes. *J. Personal. Soc. Psychol.* 1987; 53:431–444.
- Neuberg SL, Newsom JT. Personal need for structure: individual differences in the desire for simple structure. *J. Personal. Soc. Psychol.* 1993; 65:113–131.
- Pelphrey KA, Singerman JD, Allison T, McCarthy G. Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*. 2003; 41:156–170. [PubMed: 12459214]
- Rayner K. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 1998; 124:372–422. [PubMed: 9849112]
- Roberts B, Donahue E. One personality, multiple selves: integrating personality and social roles. *J. Personal.* 1994; 62(2):199–218.

- Rodin MJ. Who is memorable to whom: a study of cognitive disregard. *Soc. Cogn.* 1987; 5:144–165.
- Rosenberg S, Sedlak A. Structural representations of implicit personality theory. *Adv. Exp. Soc. Psychol.* 1972; 6:235–297.
- Ross, L.; Nisbett, R. *The Person and the Situation: Perspectives of Social Psychology*. New York: McGraw-Hill; 1991.
- Ruscher JB, Fiske ST. Interpersonal competition can cause individuating impression formation. *J. Personal. Soc. Psychol.* 1990; 58:832–842.
- Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science.* 1996; 274(5294):1926–1928. [PubMed: 8943209]
- Saxe R, Carey S, Kanwisher N. Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 2004; 55:87–124. [PubMed: 14744211]
- Schank, RC.; Abelson, RP. *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum; 1977.
- Schiller D, Freeman JB, Mitchell JP, Uleman JS, Phelps EA. A neural mechanism of first impressions. *Nat. Neurosci.* 2009; 12:508–514. [PubMed: 19270690]
- Schoenfeld MA, Tempelmann C, Martinez A, Hopf J-M, Sattler C, Heinze H-J, Hillyard SA. Dynamics of feature binding during object-selective attention. *Proc. Natl. Acad. Sci.* 2003; 100(20):11806–11811. [PubMed: 12960369]
- Sedikides C, Devine PG, Fuhrman RW. Social perception in multi-target settings: effects of motivated encoding strategies. *Personal. Soc. Psychol. Bull.* 1991; 17(16):625–632.
- Snyder M, Swann WB. Hypothesis-testing processes in social interaction. *J. Personal. Soc. Psychol.* 1978; 36:1202–1212.
- Sporer SL. Recognizing faces of other ethnic groups: an integration of theories. *Psychol. Publ. Policy Law.* 2001; 7:36–97.
- Sporns O, Chialvo D, Kaiser M, Hilgetag C. Organization, development and function of complex brain networks. *Trends Cogn. Sci.* 2004; 8(9):418–425. [PubMed: 15350243]
- Spunt RP, Satpute AB, Lieberman MD. Identifying the what, why, and how of an observed action: an fMRI study of mentalizing and mechanizing during action observation. *J. Cogn. Neurosci.* 2011; 23(1):63–74. [PubMed: 20146607]
- Srull TK, Wyer RS. Person memory and judgment. *Psychol. Rev.* 1989; 96:58–83. [PubMed: 2648446]
- Todd, PM.; Hertwig, R.; Hoffrage, U. Evolutionary cognitive psychology. In: Buss, DM., editor. *The Handbook of Evolutionary Psychology*. New York: Wiley; 2005. p. 776-802.
- Todorov A, Pakrashi M, Oosterhof NN. Evaluating faces on trustworthiness after minimal time exposure. *Soc. Cogn.* 2009; 27:813–833.
- Tononi G, Edelman G, Sporns O. Complexity and coherency integrating information in the brain. *Trends Cogn. Sci.* 1998; 2:474–484. [PubMed: 21227298]
- Treisman A, Gelade G. A feature-integration theory of attention. *Cogn. Psychol.* 1980; 12(1):97–136. [PubMed: 7351125]
- Van Duynslaeger M, Van Overwalle F, Verstraeten E. Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Soc. Cogn. Affect. Neurosci.* 2007; 2(3):174–188. [PubMed: 18985139]
- Van Overwalle F. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 2009; 30(3): 829–858. [PubMed: 18381770]
- Wagner DD, Haxby JV, Heatherton TF. The representation of self and person knowledge in the medial prefrontal cortex. *Wiley Interdiscip. Rev. Cogn. Sci.* 2012; 3(4):451–470. [PubMed: 22712038]
- Wang S-H, Morris R. Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annu. Rev. Psychol.* 2010; 61:49–79. [PubMed: 19575620]
- Whitson J, Galinsky A. Lacking control increases illusory pattern perception. *Science.* 2008; 322(5898):115–117. [PubMed: 18832647]
- Zaki J, Ochsner K. The need for a cognitive neuroscience of naturalistic social cognition. *Ann. N. Y. Acad. Sci.* 2009; 1167:16–30. [PubMed: 19580548]

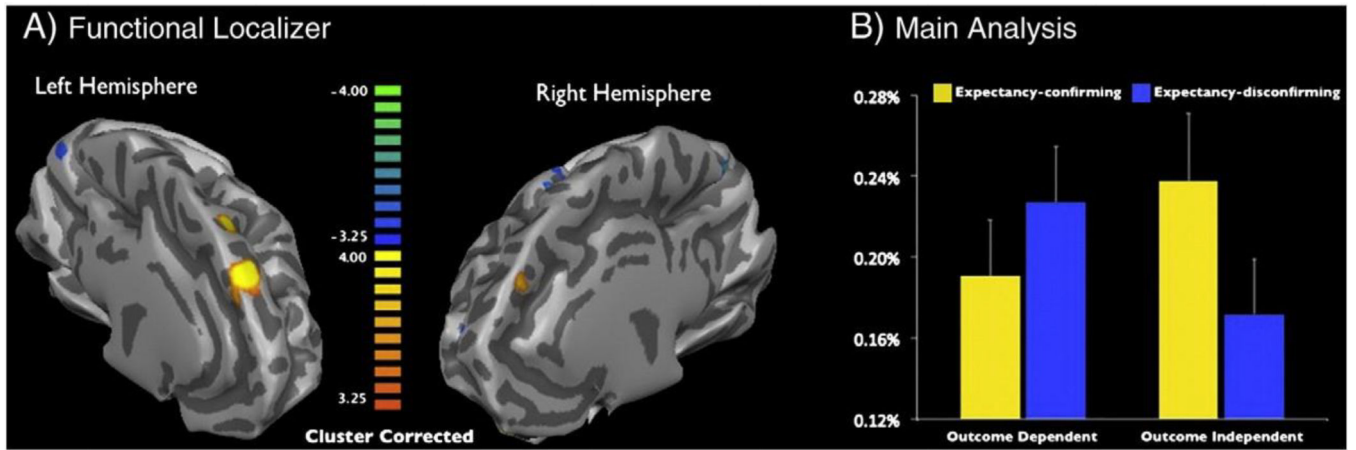


Fig. 1.

Average BOLD response of dmPFC as a function of perceivers' dependence on target person and expectancy-consistency of information about that person. Consistent with previous research, a region of dorsomedial prefrontal cortex emerged from the contrast *form impression* > *remember order* in the functional localizer (A). This a-priori selected region revealed the predicted interaction in the main task, with greater activation to expectancy-confirming (vs. expectancy-disconfirming) information for the outcome-independent target, and the opposite pattern for the outcome-dependent target (B).

Table 1

Center of mass coordinates and cluster sizes for localizer task.

Contrast / region	Talairach coordinates			Cluster Size
	x	y	z	
Form impression > remember order				
mPFC	-4	46	39	87
	-12	25	54	22
Right anterior temporal lobe	46	8	-23	164
Remember order > form impression				
Right anterior insula	37	13	7	29
Right superior frontal gyrus	25	3	59	40
Right superior parietal lobule	31	69	42	114
Precuneus	-1	-68	48	52

Coordinates refer to center of mass in Talairach space. Cluster size is shown in functional voxels ($3 \times 3 \times 3$ mm).

Table 2

Center of mass coordinates and cluster sizes for main task.

Contrast / region	Talairach coordinates			Cluster Size
	x	y	z	
Dependent > independent				
Left cerebellum	-52	-49	-30	41
Independent > dependent				
No significant regions				
Consistent > inconsistent				
No significant regions				
Inconsistent > consistent				
No significant regions				
Dependence × consistency interaction				
Left inferior frontal gyrus	-46	37	-2	75
Left superior parietal lobule	-43	-33	41	23

Coordinates refer to center of mass in Talairach space. Cluster size is shown in functional voxels ($3 \times 3 \times 3$ mm).