Social Cognition as Reinforcement Learning: Feedback Modulates Emotion Inference

Jamil Zaki[1], Seth Kallman[2], G. Elliott Wimmer[3], Kevin Ochsner[2]*, and Daphna Shohamy[2]*

Abstract

■ Neuroscientific studies of social cognition typically employ paradigms in which perceivers draw single-shot inferences about the internal states of strangers. Real-world social inference features much different parameters: People often encounter and learn about particular social targets (e.g., friends) over time and receive feedback about whether their inferences are correct or incorrect. Here, we examined this process and, more broadly, the intersection between social cognition and reinforcement learning. Perceivers were scanned using fMRI while repeatedly encountering three social targets who produced conflicting visual and verbal emotional cues. Perceivers guessed how targets felt and received feedback about whether they had guessed correctly. Visual cues reliably predicted one target's emotion, verbal cues predicted a second target's emotion, and neither reliably predicted the third target's emotion. Perceivers successfully used this information to update their judgments over time. Furthermore, trial-by-trial learning signals—estimated using two reinforcement learning models—tracked activity in ventral striatum and ventromedial pFC, structures associated with reinforcement learning, and regions associated with updating social impressions, including TPJ. These data suggest that learning about others' emotions, like other forms of feedback learning, relies on domain-general reinforcement mechanisms as well as domain-specific social information processing. ■

INTRODUCTION

"Perceivers" (individuals focusing on another person) often draw inferences about the experiences of social "targets" (individuals who are the focus of perceivers' attention) based on complex and sometimes conflicting social cues (Zaki, 2013; Freeman & Ambady, 2011; Gilbert, 1998). Neuroscientific studies of social cognition typically employ paradigms in which perceivers draw single-shot inferences about strangers and do not learn whether these inferences are correct (Frith & Frith, 2012; Zaki & Ochsner, 2011). Although this approach provides experimental control, it fails to capture two key features of social cognition (Zaki & Ochsner, 2009; Neisser, 1980). First, perceivers outside laboratory contexts often draw inferences about familiar targets such as friends (Stinson & Ickes, 1992). Second, perceivers often receive feedback about when their inferences are correct, for instance, when targets correct perceivers after those perceivers make incorrect guesses about targets' experiences (Swann, 1984).

Feedback allows perceivers to learn "rules" for understanding social targets and, crucially, to learn which types of information they should use when drawing inferences about particular others. Consider a perceiver who meets two targets. One (Bob) expresses his emotions through facial expressions: When he grins, he is typically happy. Another target (Joe) produces misleading facial expressions, such as smiles that hide anxiety or politely hidden boredom. However, Joe expresses emotions verbally. Initially, a perceiver might make guesses about each target based on visual and verbal information. Upon receiving feedback, a perceiver might "tune" her inferences toward cues that are useful for understanding each target (visual information for Bob and verbal information for Joe) and away from the use of misleading information. This same perceiver might further develop, through experience and feedback, increasingly strong assumptions about the cues that accurately signal a given target's affect. Finally, as a perceiver learns to reliably understand a target's experience, they might imbue a familiar target—and their own ability to predict the target's internal states—with value.

At least two mechanisms could support such learning, depending on the extent to which social cognition is supported by processes distinct from those engaged by nonsocial phenomena (Ostrom, 1984). First, social learning could rely on domain-general reinforcement learning mechanisms, through which individuals update the expected value with which they imbue options (i.e., the extent they believe that option will result in reward) based on previous experiences that were better or worse than expected (prediction errors; Sutton & Barto, 1998).

[1]Stanford University, [2]Columbia University, [3]University Medical Center Hamburg-Eppendorf
*These authors contributed equally to this research.

Expected value and prediction errors track engagement of dopaminergic targets, especially the ventral striatum (VS) and ventromedial pFC (vMPFC; see Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008). Learning how to "read" social targets over time might thus depend on individuals' ability to update their sense of how informationally valuable, or likely to result in correct inferences, a particular class of social cue is. This learning would occur based on feedback they have received after using this type of cue in the past. This would be quite similar to learning how reliably other stimuli, such as decks of cards, produce valuable outcomes. If this is the case, social learning should behaviorally follow patterns associated with reinforcement learning in nonsocial settings and recruit activity in VS and vMPFC.

Second, learning about social targets could recruit domain-specific mechanisms associated with social cognition. Neuroscientists have documented consistent engagement of a circumscribed set of brain regions—including medial pFC, TPJ, and posterior cingulate cortex (PCC)—when perceivers consider the minds of social targets (Frith & Frith, 2012). This system differentiates processing of social, as opposed to nonsocial, information in several domains including long-term memory (Macrae, Moran, Heatherton, Banfield, & Kelley, 2004), working memory (Meyer, Spunt, Berkman, Taylor, & Lieberman, 2012), and categorization (Contreras, Banaji, & Mitchell, 2012). Parts of this system—especially PCC and TPJ—also support the updating of social impressions, for instance, perceivers' sense for the morals or qualities of social targets based on information about those targets' behaviors (Bhanji & Beer, 2013; Mende-Siedlecki, Baron, & Todorov, 2013; Mende-Siedlecki, Cai, & Todorov, 2012; Schiller, Freeman, Mitchell, Uleman, & Phelps, 2009). To the extent that learning about others' emotions through reinforcement is likewise domain specific, it should recruit these regions, especially when feedback indicates that a perceiver needs to update their use of a given cue when inferring how a target feels.

Finally, social learning could represent a hybrid phenomenon, encompassing both domain-general and domain-specific mechanisms (Zaki, Hennigan, Weber, & Ochsner, 2010). Here, we tested these possibilities and more broadly explored the intersection between reinforcement learning and social cognition.

METHODS

Overview

We scanned participants with fMRI while they repeatedly encountered one of three social targets. On each trial, the target produced two social cues—one visual and one verbal—that suggested opposite (positive and negative) emotional valence. Participants used this information to infer targets' emotion and received feedback indicating whether they were correct. Over the course of the

study, feedback indicated that targets differed with respect to the cues that accurately tracked their emotions. Specifically, one target (the "visual-correct" target) produced ~88% correct feedback when perceivers inferred her or his emotions using visual cues. A second target (the "caption-correct" target) produced ~88% correct feedback when perceivers used verbal cues, and a third, "unpredictable" target produced 50% correct feedback regardless of the cues on which targets relied (see Figure 1 for task schematic). The particular social targets paired with visual-correct, caption-correct, and unpredictable feedback varied across participants.

We fit learning about each target using two models. In a standard Q-learning model (Daw, 2009; Sutton & Barto, 1998), we assumed that individuals learned about the correctness or incorrectness of only the cue (visual or caption) that they used on a given trial and did not learn about the informational value of nonchosen cues. As each trial included two cues, we were also interested in whether participants updated the expected value from both chosen and nonchosen cues (see below for more discussion of this possibility). To examine this, we employed an augmented Q-learning model that allowed for individuals to also learn about the nonchosen cue based on feedback.

We used the fit parameters from each model to (i) test whether people indeed learn about both chosen and nonchosen cues during social learning and (ii) generate
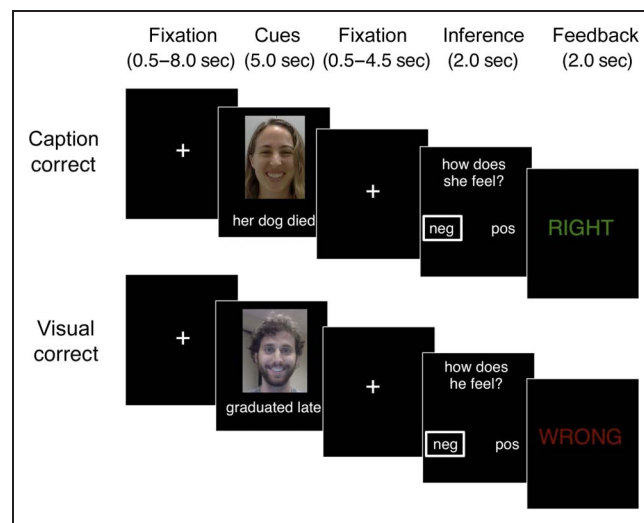

Figure 1. Task schematic. Participants encountered silent videos and captions from three social targets, each of whom produced conflicting social cues (here, two positive videos accompanied by two negative captions). Participants then drew inferences about whether the target in fact felt positively or negatively and received feedback about whether their inferences were correct. For one target (caption correct or video correct), inferences based on verbal cues produced correct feedback 87.5% of the time. For another target (video correct or caption correct), inferences based on visual cues produced correct feedback 87.5% of the time (pictures here, a caption-based inference receives incorrect feedback). For a third target (not pictured), neither visual nor verbal social cues reliably produced correct feedback.

trial-by-trial regressors through which to examine brain activity associated with expected value and reward prediction errors during social learning. As described below, we fit different sets of parameters to the different cue conditions (visual correct, caption correct, and control) to examine learning in each condition.

Participants

Participants ($n = 25$, 13 women, age = 23.1 ± 6.7 years) completed informed consent in accordance with the guidelines of the Columbia University institutional review board and were remunerated for their participation. Data from two participants were lost because of computer error, data from one participant were lost because of excessive movement, and four additional participants either failed to respond on >25% of trials or used only one informational channel (e.g., visual cues) in judging targets' affective states on >90% of trials, rendering their data unusable in subsequent analyses. Thus, our final sample consisted of 18 participants (11 women, age = 23.7 ± 7.9 years).

Stimuli and Procedure

Participants were scanned with fMRI while they viewed and drew inferences about three social targets. On each trial, they viewed two simultaneously presented cues: a silent, 5-sec video of a target talking and a short (<15-word) caption describing an emotional event. Participants were told that the captions summarized the event about which the target in the video spoke. Videos and captions were drawn from a stimulus library described elsewhere (Zaki et al., 2010). Videos and captions were normed to ensure that, viewed individually, they were judged as either positive or negative emotion on a 9-point Likert scale (1 = *very negative*, 9 = *very positive*; ratings for positive and negative videos: 6.03 ± 0.97 and 4.05 ± 0.78, respectively; ratings for positive and negative captions: 7.42 ± 0.44 and 2.69 ± 0.51, respectively). Positive and negative captions did not differ with respect to word count (9.81 ± 2.38 and 9.31 ± 2.91, respectively; $t < 1.0$, $p > .25$). It is worth noting that, overall, videos were viewed as less affectively valenced than captions. This is consistent with the idea that naturalistic emotion expressions like the ones used here are rarely as clear or exaggerated as posed expressions (Russell, Bachorowski, & Fernandez-Dols, 2003). Nonetheless, our norming data indicate that these visual cues were reliably identifiable as positive and negative (i.e., observers rarely rated a positive video below the midpoint of the scale, and vice versa).

Each trial began with a jittered fixation of 500–8500 msec, optimized for rapid presentation using Optseq (Dale, 1999). After this, participants encountered a combination of visual and verbal cues, each of which suggested opposite interpretations of targets' likely affective valence. On a given trial, a participant might encounter a video in which a target appeared to feel negatively, paired with a positive caption, or vice versa. After a jittered ISI of 500–4500 msec, participants were given 2000 msec to decide whether they believed the target felt positive or negative. After drawing each inference, participants received feedback for 2000 msec, indicating whether their choice had been correct or incorrect (Figure 1). Participants viewed and made judgments about each target 48 times—24 including a positive video and a negative caption and 24 including a negative video and a positive caption—for a total of 144 trials.

Critically, feedback varied across the three targets. Specifically, when viewing one target (the visual-correct target), the "correct" answer corresponded to visual information on 87.5% of trials (42/48). Note that this source of information typically produced "correct" feedback irrespective of the affective valence of each cue. That is, when the visual-correct target was paired with a positive video and a negative caption, participants who answered "positive" received feedback that they were correct on 87.5% of trials. When the visual-correct target instead was paired with a negative video and a positive caption, participants who answered "negative" were correct on 87.5% of trials. When viewing a second target (the caption-correct target), captions likewise provided the "correct" answer on 87.5% of trials. When viewing a third target (the unpredictable or U target), neither videos nor captions reliably predicted the "correct" valence. The cue type that produced "correct" feedback for each target was constant across the session, such that an optimal strategy for correctly understanding a given target (e.g., caption correct) would be to rely on their associated cue type (e.g., verbal information) on all trials.

Each video was shown to each perceiver only once. All perceivers saw 48 unique videos of each social target, paired with a pseudorandomized set of captions that were opposite in valence from videos (i.e., 24 positive videos from each target paired with 24 negative captions, and vice versa). As mentioned above, we also randomized, across perceivers, whether each target was paired with video-correct, caption-correct, or unpredictable feedback. Therefore, although all participants saw the same videos of each target, (i) these videos did not repeat within participant, and (ii) the pairing of videos with captions and the feedback paired with videos for a given target varied across participants. Trials were fully randomized across targets, such that a given trial was equally likely to include any of the three targets.

Behavioral Analysis

To visualize participants' learning over time, we split the 48 trials on which participants viewed each target into six "bins" of eight trials apiece and assessed the proportion of trials in each bin during which participants relied on the correct cue type to assess each target's state. Note that this, in some cases, constituted assigning a response as correct even if it resulted in incorrect feedback,

provided that the participants' guess reflected the type of information they should have used for each target. For instance, on 12.5% of trials including the visual-correct target, perceivers would receive "incorrect" feedback for making judgments based on visual information although this was the correct general strategy; these trials would be counted as correct in our analysis. We used repeated-measures one-way ANOVAs within condition (visual correct, caption correct, and unpredictable) to assess increases in the proportion of correct judgments across blocks and repeated-measures two-way ANOVAs (of Block × Condition) for each pair of conditions to assess differences in improvement as a function of the type of correct information provided by targets.

Modeling

To formally assess learning, we employed modeling approaches from reinforcement learning that describe the value participants ascribe to different cues, the choices they make, and how they learn based on feedback. At any given point, a perceiver holds a belief about the expected value of each social cue. Here, this value is informational, reflecting the value of this cue in allowing the perceiver to correctly guess how the target feels. For instance, an expected value of 0.8 for a visual cue indicates that a perceiver believes that relying on that cue—and inferring a target's emotional state based on visual information—will give the perceiver 80% of the maximum reward value associated with being correct about that target's state.

As perceivers make choices about a target's emotion and receive feedback about whether they are correct, the model updates the values assigned to these parameters, based on the difference between the expected values and the feedback experienced as a consequence of each choice (a prediction error). For instance, if a perceiver relies on visual information to infer a target's state but learns that this inference is incorrect, the perceiver will reduce the informational value they expect such cues to have in subsequent judgments. We used Q-learning models to examine how participants updated the value of chosen information (see Sutton & Barto, 1998).

First, we applied a variant of a Q-learning model that is often used to study instrumental learning based on feedback (Daw, 2009; Sutton & Barto, 1998). This "standard" model focuses on, and updates, only the informational value of the chosen cue—for instance, visual cues on trials when perceivers based their inference on visual information. Option values were initialized at 0.5. When an individual chooses to follow a particular information source i (visual or caption information) on trial t and receives a reward (0 or 1), the expected value of that cue in correctly predicting the target's emotion is updated according to

$$V_{i,t} = V_{i,t-1} + \alpha \cdot \delta_t \tag{1}$$

where α is a free parameter controlling the learning rate—or how quickly the perceiver updates their valuation based on feedback—and δ_t is the prediction error associated with choosing visual information on that trial. Prediction error, in turn, is specified by

$$\delta_t = r_t - V_{i,t-1} \tag{2}$$

the difference between the actual reward associated on that trial (r_t) and the expected value of the chosen cue on the prior trial, $V_{i,t-1}$.

Given value estimates on a particular trial, individuals are assumed to choose between the options stochastically with probabilities P_{vis} and P_{cap} according to a softmax distribution (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006):

$$P_{i,t} \approx \exp(\beta(V_{i,t})) / \sum\nolimits_{i=1}^{k} \exp(\beta(V_{i,t})) \tag{3}$$

The free parameter β represents the softmax inverse temperature, which refers to an individual's tendency to reliably select the option with the higher expected value on any given trial.

To provide an example of how social learning might operate under this model, consider the case where, at the start of the task, a perceiver believes a verbal cue to have an initial value of 0.5. The perceiver nonetheless uses this cue to infer the target's emotion and learns that this inference is correct. To calculate updated next-trial values, the prediction error will be computed as in Equation 2: 1 (feedback) − 0.5 (expectation) = 0.5. Then, with an example learning rate of 0.25, as per Equation 1, (0.5 * 0.25) would be added to the initial value of verbal cues, yielding an estimated value of 0.625. On the next trial in which the perceiver encounters this same target, the perceiver's choice probability of the verbal and visual cues, respectively, would be updated according to this new value. For instance, under an example β (softmax) of 4, these choice probabilities would be 62.25% and 37.75%. For further elaboration on reinforcement learning models, see Daw (2009).

All four parameters were fit to participants' behavior and task feedback using MATLAB's (The MathWorks, Natick, MA) *fmincon*, which included 20 starting points to avoid local minima. The resulting model provides a single estimate of each participant's learning rate (α) and inverse temperature (β) and trial-by-trial estimates of the value associated with visual and caption cues ($V_{t(\text{vis})}$ and $V_{t(\text{cap})}$) and prediction error (δ_t) at each time point during the task. In calculating trial-by-trial estimates, we followed common practice in the field by using group mean learning rates, as individual learning rates are typically too unstable to produce robust regressors for neuroimaging (following previous work: Wimmer, Daw, & Shohamy, 2012; Daw, 2009; Schönberg, Daw, Joel, & O'Doherty, 2007; Daw et al., 2006).

Consistent with prior work, we modeled brain activity that tracked both positively and negatively with prediction error (Abler, Walter, Erk, Kammerer, & Spitzer,

2006; McClure, Berns, & Montague, 2003). Prediction error signals were "signed," meaning that negative prediction errors occurred when a perceiver was less accurate than they would have expected on a given trial based on our model (i.e., the inverse of a reward response). As such, neural correlates of negative prediction errors in this task were meant to isolate a learning signal associated with a need to correct one's inferences after errors.

A standard Q-learning model assumes that participants learn about and update only the value of the chosen cue. For instance, if a perceiver draws an inference about a target based on visual information and learns that he or she was right, his or her expected value for subsequent visual cues will increase, but the value he or she places on caption cues (which he or she did not choose) will remain unchanged. However, it is possible that individuals in our task simultaneously learned about both chosen and nonchosen cues. For instance, if a perceiver relies on a visual cue to infer how a target feels and receives feedback that they are incorrect, he or she might simultaneously decrease the expected informational value of visual cues and increase the expected informational value of later verbal cues. This is especially possible given the fact that, in our task, visual and caption cues always provided conflicting information about targets' valence, such that, whenever a chosen cue was wrong, a nonchosen cue was also correct.

To examine the possibility that perceivers simultaneously update the value associated with chosen and nonchosen cues, we also examined an "augmented" Q-learning model that allows for learning about the nonchosen option in addition to the chosen option.[1] By adding a secondary learning rate ($\alpha2$), the model tests whether agents simultaneously learn about multiple cues—here, both the chosen and nonchosen cues.

In the augmented Q-learning model, the value for the attended ("chosen") cue is updated according to Equation 1. The value for the nonattended ("nonchosen") informational cue is updated according to

$$\delta_{non,t} = r_t - V_{non,t-1} \qquad (4)$$

$$V_{non,t} = V_{non,t-1} + \alpha2 \cdot \delta_t \qquad (5)$$

The value of the nonchosen option was updated with an additional learning parameter, $\alpha2$. A negative $\alpha2$ produces a prediction error in the opposite direction as the chosen option, for example, increasing the value of the chosen information source and decreasing the value of the nonchosen information source. This kind of updating may be expected if participants assume that there is only one correct information source. As $\alpha2$ decreases to zero, the prediction error applies only to the chosen ("attended") cue. Thus, if $\alpha2 = 0$, perceivers only update their expected value based on prediction errors over the chosen cue, and the augmented Q-learning model learning update rule "reduces" to the Q-learning model described above. Finally, as $\alpha2$ increases above zero, learning applies positively to both the chosen and nonchosen options. In cases where the $\alpha2$ parameter is negative, to allow for negative updating of the nonchosen option, the update rule for the nonchosen cue uses the absolute value of $\alpha2$ in Equation 6 and replaces the r term with the absolute value of $(r - 1)$ in Equation 5. Separate learning models were fit for the different cue types: visual correct, caption correct, and control. Choices were modeled using the softmax function as described above in Equation 3, where the free parameter β fits were optimized as described above.

Standard and augmented Q-learning models produced very similar fits to the behavioral learning data, and participants did not show significant learning about the nonchosen cue (see Results below). This suggests that the standard Q-learning model likely serves as a more parsimonious descriptor of social learning in this paradigm. As such, we restricted neuroimaging analyses to a model using variables derived from the base Q-learning model.

fMRI Data Acquisition and Analysis

Imaging data were collected on a 3.0-T Phillips scanner using a gradient-echo echo-planar pulse sequence (38 contiguous axial slices, repetition time = 2000 msec, echo time = 20, 3 × 3 × 3 mm voxel size). A high-resolution T1-weighted structural scan (magnetization prepared rapid gradient echo) was collected before three functional runs of 336 repetition times each. Each functional run began with five volumes before the first stimulus onset; these volumes were discarded to allow for magnetic field equilibration. Stimuli were presented onto a screen at the end of the magnet bore using Presentation Software (www.neurobs.com). Participants viewed the screen via a mirror mounted on the head coil, and a pillow and foam cushions were placed inside the coil to minimize head movement.

MRI data were preprocessed and analyzed using custom batch scripts that interfaced with SPM2 (Wellcome Department of Cognitive Neurology, London, UK). Functional data were time-corrected for differences in acquisition time between slices for each whole-brain volume and realigned to correct for head movement. Data were then transformed into a standard anatomical space (3-mm isotropic voxels) based on the ICBM 152 brain template (Montreal Neurological Institute). Normalized data were then spatially smoothed (6-mm FWHM) using a Gaussian kernel. Statistical analyses were performed using general linear models in which the events were convolved with a canonical hemodynamic response function, its temporal derivative. General linear models included a choice value regressor modeled for the 5-sec duration of the cue period and a prediction error regressor modeled at the time of feedback with 0-sec duration as well as same-duration control regressors for these trial periods. We further included additional covariates of no interest—a session mean and a

linear trend—to ensure that any patterns in brain activity we saw do not reflect scanner drift over time (either across or within runs). This analysis was performed individually for each participant, and contrast images for each participant were subsequently entered into a second-level analysis treating participants as a random effect.

To assess common activity in response to both visual-correct and caption-correct targets, we first isolated activity associated with value signals in response to each target in separate imaging analyses (by including positive contrast weights for the value regressor in response to each target alone). We then computed a two-way conjunction analysis combining activation maps reflecting (i) the expected value regressor in response to the visual-correct target alone and (ii) the same regressor in response to the caption-correct target alone. We computed a similar conjunction of activation maps reflecting prediction error-related engagement in response to the visual-correct target and the caption-correct target. These conjunction maps employed the minimum statistic approach described by Nichols, Brett, Andersson, Wager, and Poline (2005), which requires each individual map in the conjunction (as opposed to either one) to meet a chosen statistical criterion.

To isolate preferential brain activity associated with learning about our two predictable targets, we directly contrasted brain activity associated with the value regressor for the visual-correct, as compared with the caption-correct, target and with the caption-correct, as compared with the visual-correct, target. We computed similar direct contrasts of prediction error-related activity across visual-correct and caption-correct targets.

Brain regions fitting single contrast models (for preferential activity related to learning from one cue type) were identified using a threshold of 32 or more contiguous voxels at a voxelwise threshold of $p < .005$. These height and extent thresholds were selected on the basis of a Monte Carlo simulation implemented in MATLAB (similar to Monte Carlo simulation in AFNI and SPM), to correspond with an overall false-positive rate of <5%, corrected for multiple comparisons (Slotnick, Moo, Segal, & Hart, 2003). To compute appropriate thresholds for conjunctions between two contrast maps (assessing common learning-related activity across cue types), we used Fisher's (1932) methods, which combine probabilities of multiple hypothesis tests using the following formula:

$$\chi^2 = -2\sum_{i=1}^{k} \log e(p_i) \qquad (6)$$

where p_i is the p value for the ith test being combined, k is the number of tests being combined, and the resulting test has a χ^2 distribution with $2k$ degrees of freedom. This equation reveals that thresholding each parametric map of value-related and prediction-error-related activity at a p value of .02 in a two-way conjunction corresponds to a combined threshold of $p < .003$. We then combined

this threshold with an extent threshold of 30 contiguous voxels—again using Monte Carlo simulations—consistent with a false discovery rate of <0.05.

RESULTS

Behavioral Results

Participants demonstrated significant learning across the session in response to the visual-correct target (accuracy by eight-trial "bins": $F(5, 85) = 3.55$, $p < .01$) and marginal learning in response to the caption-correct target ($F(5, 85) = 2.12$, $p = .07$; see Figure 2 for accuracy across trial bins in all conditions). This likely reflects an initial bias toward using caption cues. Perceivers initially relied on visual cues in response to the visual-correct target at chance (mean use of visual cues in Trial bin 1 = 46%, in comparison with 50%: $t = -1.16$, $p > .25$) but relied on captions in response to the caption-correct target significantly more than chance (mean use of verbal information in Trial bin 1 = 60%, in comparison with 50%: $t = 2.18$, $p < .05$). This initial pattern of inference left less "room for improvement" in response to the caption-correct, as compared with visual-correct, target. An initial bias toward captions in social inference is consistent with prior demonstrations that perceivers preferentially utilize contextual, as compared with facial, cues when drawing inferences about others' emotions (Zaki et al., 2010; Aviezer et al., 2008).

By contrast, participants demonstrated no learning in response to the unpredictable target across trial bins.[2] Direct Time × Condition contrasts revealed that participants learned more quickly (i.e., adjusted their behavior more quickly across trials) from the visual-correct target than the unpredictable target ($F(5, 85) = 4.49$, $p < .001$)
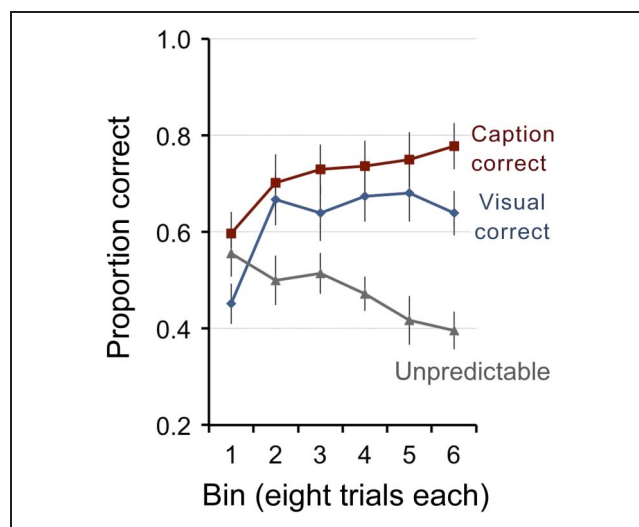

Figure 2. Participants successfully learned about visual-correct and caption-correct targets, improving inferential accuracy across trial bins. By contrast, accuracy about unpredictable targets did not improve over time.
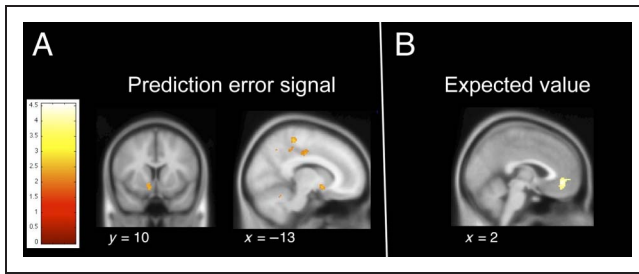
Figure 3. (A) Conjunction analysis of brain activity predicted by prediction error signals in response to visual-correct and caption-correct targets. (B) Conjunction analysis of brain activity predicted by expected value signals in response to visual-correct and caption-correct targets. Scale represents *t* values for each contrast inputted into the conjunction map.

and from the caption-correct target than from the unpredictable target ($F(5, 85) = 3.78, p < .005$) but did not differ in the speed at which they learned about visual-correct and caption-correct targets ($F(5, 85) = 0.49, p > .75$).

Modeling of Reinforcement Learning

Standard Q-learning model. The Q-learning model produced learning rates in response to visual-correct and caption-correct targets that did not differ from each other (mean $\alpha_{vis} = 0.22$, $SD = 0.25$; mean $\alpha_{cap} = 0.26$, $SD = 0.30$; $t = 0.42$, $p > .65$) and approximated learning rates documented in prior studies of nonsocial reinforcement learning (Gläscher, Daw, Dayan, & O'Doherty, 2010; O'Doherty et al., 2004).

Augmented Q-learning model. The augmented Q-learning model, which allowed for learning about the nonchosen as well as chosen cues, produced learning rates for the visual-correct and caption-correct targets that did not differ from each other (mean $\alpha_{vis} = 0.32$, $SD = 0.34$; mean $\alpha_{cap} = 0.31$, $SD = 0.40$; $t = 0.01$, $p > .99$). For the secondary learning rate, controlling generalization of feedback to the nonchosen option, a negative value would indicate counterfactual learning, whereas a positive value would update the nonchosen option in the same direction as the chosen option. We found that the secondary learning rate was not different than zero and also similar across target conditions (mean $\alpha2_{vis} = 0.05$, $SD = 0.52$, $p > .75$; mean $\alpha2_{cap} = 0.04$, $SD = 0.56$, $p > .69$; caption vs. video: $t = 0.04$, $p > .97$). Furthermore, no participants were better fit by the augmented, as compared with standard, Q-learning model (likelihood ratio test, $p < .05$ significance level). The nonsignificant secondary learning rate suggests that, in accordance with a simpler Q-learning model, perceivers strongly tracked the informational value of the cue that they chose on a given trial but might not have attended to, or updated, the value of the nonchosen cue.

Neuroimaging Results

Learning Signals Common across Cue Type

Conjunction analysis of reward prediction error signals common to caption-correct and visual-correct targets revealed engagement of the left nucleus accumbens, along with the precuneus, precentral gyrus, and a number of clusters in visual cortex (Figure 3A, Table 1). A similar analysis probing brain activity negatively correlated with prediction errors revealed engagement of dorsal ACC, anterior insula, and TPJ (Table 2). Conjunction analysis of expected value signals across targets revealed activity only in vMPFC and precuneus (Figure 3B, Table 1).

Cue-specific Learning Signals

Expected value signals in response to the visual-correct, as compared with caption-correct, target preferentially engaged regions associated with processing visual movement and faces, including the fusiform gyrus, supra-marginal gyrus, and inferior parietal lobe (Figure 4A, Table 3). Direct comparison of prediction error signals in response to the visual-correct, as compared with caption-correct, target also revealed engagement of visual cortex

Table 1. Brain Areas Commonly Responding to Learning Signals in Response to Both Visual-correct and Caption-correct Targets

| Region | Coordinates | | | | Volume (Vox) |
	x	*y*	*z*	*t*	
Expected value					
vMPFC	2	42	−6	3.19	157
Precuneus	−8	−36	76	3.11	38
Prediction error					
Nucleus accumbens	−10	12	−12	2.82	34
vMPFC	4	48	12	2.52	29
Precuneus	12	−76	46	2.91	51
Precuneus	12	−40	42	3.72	107
Precentral gyrus	−32	−24	56	3.85	447
Precentral gyrus	24	−28	62	2.7	38
Superior parietal lobe	−18	−72	44	4.56	87
Fusiform gyrus	28	−78	−10	2.78	56
Lingual gyrus	−4	−72	−4	2.92	119
Superior occipital gyrus	24	−90	30	3.85	397
Superior occipital gyrus	−24	−84	22	4.12	474
Inferior occipital gyrus	−26	−92	−8	4.35	450

Coordinates are in stereotaxic space of the Montreal Neurologic Institute. *t* values reflect the statistical difference between conditions, as computed by SPM.

Table 2. Brain Areas Negatively Correlated with Prediction Error Signals

Region	Coordinates				Volume (Vox)
	x	y	z	t	
Dorsal ACC/dMPFC	−2	34	38	4.46	188
ACC/SMA	−10	16	56	3.8	50
Inferior frontal gyrus (IFG)	−30	50	10	4.98	296
IFG/dorsal anterior insula	34	22	12	4.55	128
IFG/dorsal anterior insula	−36	22	14	3.79	31
Anterior insula	−38	18	−8	4.2	88
Anterior insula	26	20	−12	3.85	46
Superior frontal gyrus	−54	20	30	3.83	36
TPJ/inferior parietal lobe	54	−50	26	3.66	35
Superior temporal gyrus	44	−40	2	4.29	202

Coordinates are in stereotaxic space of the Montreal Neurologic Institute. t values reflect the statistical difference between conditions, as computed by SPM.

Table 3. Brain Areas Preferentially Responding Learning Signals in Response to Visual-correct > Caption-correct Targets

Region	Coordinates				Volume (Vox)
	x	y	z	t	
Expected value					
Cuneus	16	−90	8	6.91	730
Cerebellum	42	−54	−40	4.84	69
Inferior parietal lobe	−30	−58	38	4.3	176
Lingual/fusiform gyrus	−18	−86	0	4.29	182
Supramarginal gyrus	−38	−34	28	3.63	77
Calcarine gyrus	−22	−58	12	3.53	63
Prediction error					
Inferior temporal gyrus	−34	−4	−36	5.24	121
Cuneus	−12	−82	28	4.16	95

Coordinates are in stereotaxic space of the Montreal Neurologic Institute. t values reflect the statistical difference between conditions, as computed by SPM.

(Figure 4B, Table 3). Direct comparisons of value and prediction error signals associated with the caption-correct, as compared with visual-correct, target produced no activation. At a relaxed extent threshold of 20 voxels, we did identify two clusters that responded more to prediction errors for caption-correct, as compared with visual-correct, targets (peak coordinates and statistics: −14, 12, 24; t = 4.53, cluster size = 28 voxels; 20, 10, 32; t = 3.48, extent = cluster size = 24 voxels) and one cluster that responded more to value signals over caption-correct, as compared with visual-correct, targets (peak: 64, 0, 22; t = 3.47, cluster size = 35 voxels). Interestingly, the cluster associated with uniquely verbal value, in the precentral gyrus, is associated with listening to and producing speech (Wilson, Saygin, Sereno, & Iacoboni, 2004). However, we hesitate to interpret these data further given their exploratory nature.
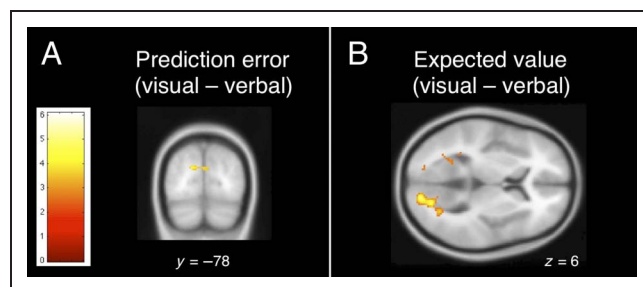

Figure 4. (A) Contrast analysis isolating activity related to prediction errors in response to the visual-correct, as compared with caption-correct, target. (B) Contrast analysis isolating activity related to expected value in response to the visual-correct, as compared with caption-correct, target. Scale represents t values.

DISCUSSION

The vast majority of everyday social cognition does not involve people making single-shot judgments about strangers. Instead, perceivers engage in a dynamic, interactive process in which they learn about individual targets through trial, error, and feedback (Schilbach et al., 2012; Neisser, 1980). Here, we provide evidence that such learning approximates other forms of reinforcement, in that observers use successful and unsuccessful social inferences to update the perceived value of social cues, a process accompanied by activity in brain regions associated with domain-general feedback learning. We also found that updating the value of social cues produced activity in regions associated with domain-specific social information processing.

With respect to domain generality, providing perceivers with feedback about the accuracy of their social inferences produced activity in VS and vMPFC, in a manner consistent with reinforcement learning. First, as the expected value of each target—in essence, the probability that perceivers could correctly assess the target's emotions using the proper cue—increased across the experimental session, encountering that target also produced stronger signals in vMPFC, but not in VS. Second, prediction errors signifying feedback that was better than the current expected value of a target tracked activity in VS more consistently than in vMPFC. These data dovetail with a growing consensus about the complementary roles these regions play in reinforcement learning: with the striatum encoding a "teaching signal" that allows organisms to update value representations based on feedback and

the vMPFC encoding an integrative measure of stimulus value in the service of decision-making (Grabenhorst & Rolls, 2011; Rangel & Hare, 2010). Importantly, the VS and vMPFC did not display preferential activity to learning based on visual, as compared with verbal social, cues, or vice versa. Finally, prediction error signals negatively tracked activity in dorsal ACC and insula, regions broadly associated with negative feedback (Barch, Braver, Sabb, & Noll, 2000; Carter et al., 2000).

These data connect in interesting ways with the small but growing literature on disjunctions between social and nonsocial information processing. Our paradigm closely mimicked other tasks examining brain responses to informational value (Tricomi & Fiez, 2012; Behrens, Woolrich, Walton, & Rushworth, 2007), with the only difference being that we employed social, as compared with nonsocial, information. This might make it unsurprising that we found patterns of activation similar to these past studies. However, at least a few studies have found that other tasks—such as working memory or categorization—produce qualitatively different patterns of brain activity when performed over social, versus nonsocial, information (Contreras et al., 2012; Meyer et al., 2012). As such, it is useful to document which cognitive operations produce brain activity that varies, versus remaining unchanged, across these stimulus types.

With respect to domain specificity, both expected value and prediction errors tracked engagement of the precuneus, posterior cingulate, and (negatively) TPJ. These regions are associated with self-projection (Buckner & Carroll, 2007), social cognition (Mitchell, 2009a), and updating social impressions (Schiller et al., 2009). These data are consistent with an account under which reinforcement learning signals in the social domain involve both a "value component" encoded in VS and vMPFC and a "social updating" component encoded in PCC, precuneus, and TPJ.

Finally, neural correlates of learning signals differed in response to visual-correct and caption-correct targets. Extrastriate visual cortex—including the fusiform gyrus—demonstrated even tighter domain specificity, tracking prediction errors and expected value only when social feedback favored visual cues. It is interesting that learning from verbal cues (caption-correct targets), as compared with visual cues, was accompanied by weaker brain activity. One possible explanation for this discrepancy surrounds participants' overall reliance on each cue type. Consistent with prior work (Aviezer et al., 2008; Carroll & Russell, 1996), perceivers in our task exhibited an initial bias toward using verbal information in drawing inferences about targets. In the context of learning, high initial reliance on verbal cues could restrict the range with which participants could update the value of such cues. This effect was likely compounded by the fact that our visual cues, when viewed alone, appeared less strongly valenced than the captions alone (see Methods above). Future work should examine whether manipulations that decrease ini-

tial reliance on verbal information, or increase initial reliance on visual cues, also intensify the behavioral and neural indices of learning based on verbal information.

Broadly speaking, these data support a conceptual bridge between feedback learning and social cognition. Under this model, perceivers initially make guesses about novel social targets, and targets provide feedback about whether perceivers have correctly understood them. This feedback drives activity in dopaminergic targets associated with domain-general reinforcement learning signals as well as regions more specifically associated with social updating, including TPJ and PCC. These mechanisms, in turn, allow perceivers to update the value of particular forms of inference (e.g., using facial expressions to understand targets) and "tune" their inferences about others (e.g., learning that facial expressions are a strong guide to one target's emotions, but not another's).

The TPJ's response to negative prediction errors draws interesting connections with past work. This region is associated both with inferences about others' beliefs and internal states and more generally with redirecting attention in response to exogenous (e.g., unexpected) feedback. In particular, recent theoretical work has suggested that TPJ is associated with generating predictions about the social world and testing those predictions against evidence about others' internal states (Koster-Hale & Saxe, 2013). Consistent with this framework, one interpretation for TPJ activity in our paradigm is that participants generated, over time, increasingly confident estimates of the cues that tracked each target's internal states. As such, feedback that conflicted with their expectations—thus generating negative prediction errors—may have prompted participants to allocate additional attention to targets' internal states, with TPJ activity accompanying this allocation. This interpretation is of course highly speculative. Future work should more directly examine whether other situations featuring unexpected social feedback also elicit TPJ activity consistent with this idea.

The presence of activity in a given brain structure does not allow for strong inference about specific underlying psychological processing (Poldrack & Yarkoni, 2015; Moran & Zaki, 2013). For instance, the PCC and precuneus are involved in nonsocial processes, such as mental imagery (Cavanna & Trimble, 2006), "stimulus independent thought" (Gusnard, Akbudak, Shulman, & Raichle, 2001), and risk perception (McCoy & Platt, 2005). TPJ similarly responds not only to social inference but also to reorienting of attention more generally (Corbetta, Patel, & Shulman, 2008). The TPJ cluster identified here (peak MNI coordinates: 54, −50, 26) does exhibit a strong association with social processing. For instance, when evaluated meta-analytically using Neurosynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011), this peak exhibited strong associations with the terms "theory of mind" and "belief" (posterior probabilities for each = .89). However, we cannot conclusively assert that activity in PCC, precuneus,

or TPJ during social learning reflects domain-specific social cognitive mechanisms. Future work should more directly test this prediction by manipulating social context, for instance, targets' group membership (Cikara & Van Bavel, 2014) or observers' motives to understand targets (Zaki, 2014; Pickett, Gardner, & Knowles, 2004), and examining the effects of such manipulations on social learning and accompanying engagement of these regions.

Our data extend a rich and growing literature on the intersection between social cognition and reinforcement learning. A number of innovative studies have examined how individuals learn about the world through social experiences. For instance, Behrens, Hunt, Woolrich, and Rushworth (2008) found that individuals update the expected value of options in a two-armed bandit task as a function of others' instructions and also accounted for the reliability of other social agents when making such judgments. Similar neural mechanisms also appear to underlie individuals' updating of value representation based on others' opinions during social influence (Nook & Zaki, 2015; Klucharev, Hytonen, Rijpkema, Smidts, & Fernandez, 2009). A group of studies has focused on the related construct of vicarious learning, through which individuals update their representation of value based on others' experience of reinforcement (Braams et al., 2013; Burke, Tobler, Baddeley, & Schultz, 2010) or punishment (Olsson & Phelps, 2007). Here, we complement this approach by examining the role that reinforcement plays in changing, updating, and shaping the inferences that individuals draw not only because of social targets but also specifically about targets and their internal states. Interestingly, these forms of social learning appear driven by strongly overlapping neural mechanisms.

Furthermore, the models applied in our data provide preliminary insight into the structure of learning in social contexts. In particular, we fit participants' behavior to two models that make different predictions about how social learning operates. The first "standard" learning model assumed that participants respond to feedback by updating the value of only the cue they used to make their decision (e.g., trusting visual information less or more only after having relied on that cue type to draw inferences about targets). The second "augmented" model allowed for the possibility that perceivers also simultaneously update the value they imbue to a nonchosen option (e.g., learning to trust verbal cues after incorrectly relying on visual cues to draw an inference). We found that the learning rate for nonchosen options did not differ from zero in our data set. This suggests that participants focused mainly on the value of chosen options in this particular learning paradigm and paid little attention to the nonchosen option. It is important to note, however, that, in our paradigm, chosen and nonchosen options provided inverse—and therefore redundant—information about targets' states. For instance, if a visual cue offered accurate information about a target, then the accompanying textual cue was necessarily inaccurate. Therefore,

the nonchosen option does not offer unique information about the target. As such, low learning rates over nonchosen options in our paradigm might reflect an artifact of our design, rather than a general feature of social learning.

In particular, there are likely social learning contexts in which participants learn simultaneously from multiple cues, and future work should explore such situations. In our paradigm, targets produced visual and verbal cues that supported opposing inferences, for instance, a negative visual display paired with a positive caption. As such, perceivers who learned that visual cues signaled a target's states simultaneously learned that verbal cues negatively predicted the same target's state. In many learning contexts, cue types such as verbal and visual information might instead be uninformative (neither stably correct nor incorrect), be informative to varying degrees (e.g., 60% vs. 80% correct), or be informative at some times and not others (Boorman, O'Doherty, Adolphs, & Rangel, 2013; Rushworth & Behrens, 2008; Behrens et al., 2007). Furthermore, in most real-world social encounters, targets likely produce congruent information about their internal states, for instance, smiling while verbally describing positive events. As such, although our paradigm allowed us to isolate brain systems involved in one form of social learning, it might fail to approximate learning in other social settings.

Future work should vary the informational value of multiple cues to more systematically assess the brain systems involved in social learning and test how these systems vary as a function of social contexts, for instance, when integrating across congruent versus incongruent social cues. Using cues with varying informational value could also allow future studies to decouple the value of chosen and unchosen cues. This would allow for more accurate modeling of such updating, expanding on our use of an augmented Q-learning model.

Another important direction for future work will be exploring how differences across observers affect learning during social inference. The current study used a relatively small sample size of 18 participants. This afforded us enough power to detect learning effects in our repeated-measures within-participant design but nonetheless constitutes an important limitation. In particular, our study was quite underpowered to detect between-participant effects (18 participants, for instance, would only allow us to reliably detect correlations with r values over .60). Future work should examine brain correlates associated with learning about others' affective states in larger samples. This would allow for exploration of at least two interesting individual differences in social learning. First, observers might differ in the type of social cues they initially employ when drawing social inferences. For instance, individuals with alexithymia, who have difficulties labeling emotional states (Mayer, DiPaolo, & Salovey, 1990), especially based on facial expressions (Nook, Lindquist, & Zaki, 2015), might rely more strongly on verbal cues. Second,

individuals might vary in their ability to efficiently update social inferences based on feedback. Individuals with ventromedial prefrontal damage, for instance, might exhibit difficulties in adjusting inferences after negative feedback.

In addition to bridging research traditions across social cognition and feedback learning, our data make key points for work in each domain.

Social Informational Value

Although the vast majority of research on reinforcement learning employs tangible rewards such as food (Daw & Doya, 2006) or money (Delgado, Miller, Inati, & Phelps, 2005) as feedback, we found that social feedback—even when divorced from any tangible reward—drove learning and activity in vMPFC and VS. This is consistent with the idea that abstract stimuli, and not only material gains, drive learning and activity in dopaminergic targets. Two such abstract phenomena are knowing that one has made a correct response (Foerde & Shohamy, 2011; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006) and receiving information needed to complete a task (Bromberg-Martin & Hikosaka, 2009). To wit, people value being right and knowing about the world around them, and feedback learning processes track not only reward narrowly construed but any feedback that helps to structure behavior (Bromberg-Martin, Matsumoto, & Hikosaka, 2010). Under this framework, it follows that social information should be especially rewarding. Other people stand out in the environment both in their motivational relevance (Zaki, 2014; Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012; Baumeister & Leary, 1995) and their unpredictability (Mitchell, 2009b). As such, information that allows perceivers to reliably track others' internal states should be experienced as highly valuable.

The Breadth of "Social Reward"

These data also add to a growing trend documenting the breadth of social phenomena that are experienced as reinforcing. Thus far, such "social rewards" include learning that one has been judged positively (Izuma, Saito, & Sadato, 2008), sharing opinions with others (Nook & Zaki, 2015; Klucharev et al., 2009), watching others receive rewards (Zaki, Lopez, & Mitchell, 2014; Mobbs et al., 2009), and acting generously toward others (Dawes et al., 2012; Zaki & Mitchell, 2011; Harbaugh, Mayr, & Burghart, 2007). Here, we document that another social phenomenon—correctly understanding others—similarly drives learning and reinforcement-related brain activity. This separates our paradigm from prior work by emphasizing the extent to which people imbue even abstract social information with value.

The current study further emphasizes the breadth of social reward by decoupling such reward from positive social cues. Typically, social rewards are operationalized using positive cues such as smiling faces (Lin, Adolphs, & Rangel, 2012; Spreckelmeyer et al., 2009). Here, we document value signals in vMPFC and VS even in response to negative social cues—and feedback about targets' negative emotions—provided that perceivers accurately perceived those emotions. This highlights the idea that negative information can nonetheless provide value if it helps perceivers predict and adaptively interact with their social environment.

Conclusion

Neuroscientific examinations of social cognition typically focus on perceivers drawing inferences about strangers in the absence of any feedback. In real-world social encounters, perceivers instead become familiar with specific social targets and learn rules about how to correctly understand these targets through repeated feedback. Here, we model these key social cognitive processes and find that learning how to optimally infer targets' emotions behaviorally and neuroscientifically approximates feedback learning more broadly. These data provide a link between social cognition on the one hand and computational and physiological models of reinforcement learning on the other. Just as importantly, they provide a novel and mechanistic window into the phenomenon of familiarity and social learning: central features of real-world social cognition.

Notes

1. Note that this kind of learning can also be accomplished in a multicue variant of the Rescorla–Wagner model (Rescorla & Wagner, 1972). However, a Q-learning model more simply models instrumental choice behavior than a Rescorla–Wagner model.
2. The only trend identified for U targets was for a numerical, but nonsignificant, tendency toward decreased accuracy over time (accuracy by bin: $F = 1.95$, $p < .09$), because of performance lower than chance in the final block. Given that there was no pattern to the cues deemed "correct" in response to the U target, we assume this is an artifact.

REFERENCES

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage, 31,* 790–795.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid?

Studies on the malleability of emotion perception. *Psychological Science, 19,* 724–732.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12,* 298–309.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117,* 497–529.

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

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10,* 1214–1221.

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

Boorman, E. D., O'Doherty, J. P., Adolphs, R., & Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. *Neuron, 80,* 1558–1571.

Braams, B. R., Güroğlu, B., de Water, E., Meuwese, R., Koolschijn, P. C., Peper, J. S., et al. (2013). Reward-related neural responses are dependent on the beneficiary. *Social Cognitive and Affective Neuroscience, 9,* 1030–1037.

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron, 63,* 119.

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron, 68,* 815–834.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences, 11,* 49–57.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 14431–14436.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70,* 205–218.

Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., et al. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 1944–1948.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain, 129,* 564–583.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences, 16,* 231–239.

Cikara, M., & Van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science, 9,* 245–274.

Contreras, J. M., Banaji, M. R., & Mitchell, J. P. (2012). Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Social Cognitive and Affective Neuroscience, 7,* 764–770.

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

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping, 8,* 109–114.

Daw, N. D. (2009). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning:*

Attention and performance XXIII (pp. 3–38). New York: Oxford University Press.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology, 16,* 199–204.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441,* 876–879.

Dawes, C. T., Loewen, P. J., Schreiber, D., Simmons, A. N., Flagan, T., McElreath, R., et al. (2012). Neural basis of egalitarian behavior. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 6479–6483.

Delgado, M. R., Miller, M. M., Inati, S., & Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *Neuroimage, 24,* 862–873.

Fisher, R. (1932). *Statistical methods for research workers.* London: Oliver & Boyd.

Foerde, K., & Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *Journal of Neuroscience, 31,* 13157–13167.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118,* 247–279.

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology, 63,* 287–313.

Gilbert, D. (1998). Ordinary personology. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 89–150). New York: McGraw Hill.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron, 66,* 585–595.

Grabenhorst, F., & Rolls, E. T. (2011). Value, pleasure and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences, 15,* 56–67.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 4259–4264.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science, 316,* 1622–1625.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience, 28,* 5623–5630.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron, 58,* 284–294.

Klucharev, V., Hytonen, K., Rijpkema, M., Smidts, A., & Fernandez, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron, 61,* 140–151.

Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron, 79,* 836–848.

Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience, 7,* 274–281.

Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex, 14,* 647–654.

Mayer, J. D., DiPaolo, M., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment, 54,* 772–781.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38,* 339–346.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience, 8,* 1220–1227.

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

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2012). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience, 8,* 623–631.

Meyer, M. L., Spunt, R. P., Berkman, E. T., Taylor, S. E., & Lieberman, M. D. (2012). Evidence for social working memory from a parametric functional MRI study. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 1883–1888.

Mitchell, J. P. (2009a). Inferences about mental states. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 364,* 1309–1316.

Mitchell, J. P. (2009b). Social psychology as a natural kind. *Trends in Cognitive Sciences, 13,* 246–251.

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., et al. (2009). A key role for similarity in vicarious reward. *Science, 324,* 900.

Moran, J., & Zaki, J. (2013). Functional neuroimaging and psychology: What have you done for me lately? *Journal of Cognitive Neuroscience, 25,* 834–842.

Neisser, U. (1980). On "social knowing". *Personality and Social Psychology Bulletin, 6,* 601–605.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage, 25,* 653–660.

Nook, E. C., Lindquist, K. A., & Zaki, J. (2015). A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion, 15,* 569–578.

Nook, E. C., & Zaki, J. (2015). Social norms shift behavioral and neural responses to foods. *Journal of Cognitive Neuroscience, 27,* 1412–1426.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304,* 452–454.

Olsson, A., & Phelps, E. A. (2007). Social learning of fear. *Nature Neuroscience, 10,* 1095–1102.

Ostrom, T. (1984). The sovereignty of social cognition. In *Handbook of social cognition.* Princeton, NJ: Lawrence Erlbaum.

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality & Social Psychology Bulletin, 30,* 1095–1107.

Poldrack, R. A., & Yarkoni, T. (2015). From brain maps to cognitive ontologies: Informatics and the search for mental structure. *Annual Review of Psychology, 67,* 587–612.

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology, 20,* 262–270.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black, & W. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 65–99). New York: Appleton-Century-Crofts.

Rushworth, M. F., & Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience, 11,* 389–397.

Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54,* 329–349.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2012). Toward a second-person neuroscience. *Behavioural Brain Research, 36,* 393–414.

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

Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience, 27,* 12860–12867.

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J., Jr. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Research. Cognitive Brain Research, 17,* 75–82.

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad, K., et al. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social Cognitive and Affective Neuroscience, 4,* 158–165.

Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology, 62,* 787–797.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91,* 457–477.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience, 18,* 1029–1043.

Tricomi, E., & Fiez, J. A. (2012). Information content and reward processing in the human striatum during performance of a declarative memory task. *Cognitive, Affective & Behavioral Neuroscience, 12,* 361–372.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience, 7,* 701–702.

Wimmer, G. E., Daw, N. D., & Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. *European Journal of Neuroscience, 35,* 1092–1104.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8,* 665–670.

Zaki, J. (2013). Cue integration: A common framework for physical perception and social cognition. *Perspectives on Psychological Science, 8,* 296–312.

Zaki, J. (2014). Empathy: A motivated account. *Psychological Bulletin, 140,* 1608–1647.

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

Zaki, J., Lopez, G., & Mitchell, J. (2014). Activity in ventromedial prefrontal cortex covaries with revealed social preferences: Evidence for person-invariant value. *Social Cognitive and Affective Neuroscience, 9,* 464–469.

Zaki, J., & Mitchell, J. (2011). Equitable decision making is associated with neural markers of subjective value. *Proceedings of the National Academy of Sciences, U.S.A., 108,* 19761–19766.

Zaki, J., & Ochsner, K. (2009). The need for a cognitive neuroscience of naturalistic social cognition. *Annals of the New York Academy of Sciences, 1167,* 16–30.

Zaki, J., & Ochsner, K. (2011). You, me, and my brain: Self and other representations in social cognitive neuroscience. In A. Todorov, S. T. Fiske, & D. Prentice (Eds.), *Social neuroscience: Toward understanding the underpinnings of the social mind* (Vol. 26–48, pp. 25–52) New York: Oxford University Press.