Social
Neuroscience

How unexpected events are processed in theory of mind regions: A conceptual replication

Journal:	*Social Neuroscience*
Manuscript ID	SNS-RP 81.22.R1
Manuscript Type:	Research Paper
Date Submitted by the Author:	14-Apr-2023
Complete List of Authors:	McManus, Ryan; Boston College, Dungan, James; University of Chicago Booth School of Business Jiang, Kevin; Boston College Young, Liane; Boston College, Psychology
Keywords:	theory of mind, prediction error, social neuroscience

How Unexpected Events are Processed in Theory of Mind Regions: A Conceptual Replication

Ryan M. McManus[1*]

James A. Dungan[2]

Kevin Jiang[1]

Liane Young[1]

[1]Department of Psychology and Neuroscience, Boston College, Boston, MA, USA

[2]Booth School of Business, University of Chicago, IL, USA

* Corresponding author email:

mcmanurd@bc.edu

Abstract

Recent research in social neuroscience has postulated that Theory of Mind (ToM) regions play a role in processing social prediction error (PE: the difference between what was expected and what was observed). Here, we tested whether PE signal depends on the type of prior information people use to make predictions—an agent's prior mental states (e.g., beliefs, desires, preferences) or an agent's prior behavior—as well as the type of information that confirms or violates such predictions. That is, does prior information about mental states (versus behavior) afford stronger predictions about an agent's subsequent mental states or behaviors? Additionally, when information about an agent's prior mental states or behavior is available, is PE signal strongest when information about an agent's subsequent mental state (vs behavior) is revealed? In line with prior research, results suggest that DMPFC, LTPJ, and RTPJ are recruited more for unexpected than expected outcomes. However, PE signal does not seem to discriminate on the basis of prior or outcome information type. These findings suggest that ToM regions may flexibly incorporate any available information to make predictions about, monitor, and perhaps explain, inconsistencies in social agents.

Keywords: Theory of Mind, Prediction Error, Social Neuroscience

Introduction

Consider making the following prediction: "My father is going to pick me up from the airport when I fly home to visit tomorrow." Given this prediction, you may be extremely surprised if he forgets to pick you up. Conversely, if your prediction was that he would forget to pick you up (perhaps because you just phoned him and he did not mention your travel plans), you may be especially surprised when he shows up at the airport right on time. In either case, you would experience social prediction error (i.e., a difference between what was expected and what was observed, within a social context).

There is a variety of information on which people base their predictions of others. This information can include but is not limited to: an agent's prior behavioral history (Dungan, Stepanovic, & Young, 2016; Heil et al., 2019), their prior mental states (Dungan et al., 2016), and descriptive or prescriptive norms (see Theriault, Young, & Feldman Barrett, 2021). These sources can be used to predict not only an agent's future behavior, but also their future mental states. For example, you might predict that your father will pick you up (or simply wants to pick you up) from the airport based on: his calling you to talk about your travel plans, explicit knowledge of his desire to pick you up (perhaps because your mother told you this), or the idea that parents tend to pick up their children from the airport. The current paper zeroes in on prior behavior and prior mental states as sources of social prediction, investigating whether the type of information on which people base their predictions, as well as the type of information that is being predicted, affects how surprising an outcome is. Specifically, we examine the brain regions implicated in processing agents' unexpected behavior and mental states as a function of their prior behavior and mental states.

Research on social prediction error has been steadily increasing, due to a seminal review article linking predictive coding models to theory of mind (ToM) tasks in neuroscience (Koster-Hale & Saxe, 2013). Predictive coding, put simply, is the idea that neuronal activity contains not only information about sensory input, but also information about the difference between expected and actual sensory input (e.g., Fiorillo, Tobler, & Schultz, 2003; Wacongne, Changeux, & Dehaene, 2012). The crucial idea behind the marriage of predictive coding and ToM is that "most experiments on ToM depend on predictions based on prior expectations and an internal model of human behavior" (Koster-Hale & Saxe, 2013). As in the opening example, you may make predictions about your father's mental states (or behavior) based on his prior mental states (or behavior). In line with Koster-Hale & Saxe's (2013) theorizing, if your predictions are not borne out, your experience of prediction error may be due to ToM regions (i.e., DMPFC, PC, LTPJ, and RTPJ) preferentially responding to information that is inconsistent with your predictions.

Extant research supports this notion, showing that predictions about a person's mental states or behaviors are influenced by social knowledge. For example, Saxe and Wexler (2005) presented participants with stories about social agents whose mental states were consistent or inconsistent with expectations that follow from social norms within the target agent's culture (e.g., your friend from high school, who has a happy marriage, confides in you that he [really hates the idea that/ would find it fun if] his wife might ever have an affair). For participants, and for their imagined target friend, the cultural norm was that happily married people do not find it exciting if their partner wants a relationship with another person. RTPJ activity was stronger when the target's cultural norms were violated, and this result held even when participants' own cultural norms differed from the target's. Similarly, bilateral TPJ and MPFC were recruited more when participants learned about politicians whose political desires were incongruent (versus congruent) with their party identity (e.g., a democrat who wants a [smaller/larger] government), suggesting that these regions preferentially respond to expectation-violating information

(Cloutier et al., 2011). Even knowledge of an agent's ability (e.g., a novice versus an experienced bowler) is enough to produce prediction error in mentalizing regions when performance is inconsistent with that ability (Heil et al., 2019).

In total, it seems that ToM regions flexibly encode different information types to make predictions and process expectation violations. This flexibility account finds additional support in recent work on metaethical judgments. Theriault et al. (2020) demonstrated that moral statements judged as more preference-like elicit greater ToM activity, whereas moral statements judged as more fact-like elicit less ToM activity. In each case, ToM activity was related not to an agent, but to social consensus regarding metaethics. However, even though past research is generally consistent with an account in which ToM regions flexibly encodes and uses the available types of social information to make predictions and process their outcomes, individual studies have tended to focus on only a single type of information when engendering predictions or revealing outcomes.

More targeted research has attempted to isolate expectation-based ToM effects that result from ~~various~~ non-social and social knowledge sources. Dungan et al. (2016) presented participants with stories that varied the source of an expectation (i.e., prior information about a non-social object, and prior information about a social agent ’~~s~~[-prior mental state vs prior behavior]), and whether subsequent behavior was consistent with that expectation. For non-social objects, results suggested that ToM regions were recruited less overall when compared to social agents. Moreover, ToM regions did not show effects at the univariate level for non-social objects’ violation of expectations, nor did these regions discriminate between non-social objects’ unexpected versus expected behavior in multi-voxel pattern analyses (MVPA). For social agents, however, r~~R~~esults showed that DMPFC and bilateral TPJ were -preferentially recruited for unexpected over expected behaviors at the univariate level, and that these regions discriminated between unexpected and expected behavior in MVPA. Overall, these results suggest that ToM regions are selectively recruited for prediction error that is social in nature. However, simple effects analyses revealed that the social prediction error effects in bilateral TPJ were~~as~~ evident only when prior information was behavior.

Current Research

Because Dungan et al. (2016) varied only the prior information type (i.e., behavior versus mental state) and not the outcome information (i.e., all outcomes were behaviors), it was impossible to know whether ToM regions would also preferentially respond to unexpected mental state outcomes. Similarly, participants might have experienced especially strong prediction error if they had instead learned about a mental state that was inconsistent with a prior mental state. Additionally, because Dungan et al's expectedness effect in bilateral TPJ occurred only when prior information was about behavior, it was impossible to know whether the lack of an effect in the other prior information condition (i.e., mental state) was simply due to a mismatch between prior and outcome types. For example, if your father did not show up at the airport to pick you up, and your prior information was that he simply *wanted* to pick you up, you may have considered the fact that even though people sometimes want to do something, they often do not (or cannot) do it. Therefore, you may not find your father's behavior particularly surprising. For these reasons, the current paper uses a novel paradigm to examine whether social prediction error occurs across brain regions typically implicated in thinking about others' minds, as well as whether these regions' sensitivity to social prediction error differs as a function of the type of information that is used to make predictions and the type of information that confirms or violates them.

fMRI Experiment

Method

Participants

Participants were 25 right-handed adults were recruited from the Greater Boston Area. One participant, unable to remain still, was removed from the scanner partway through the study, resulting in a final $N = 24$ (age: M=24.08, SD=4.11; 50% female). All participants were native English speakers, had normal or corrected-to-normal vision, and gave written informed consent in accordance with the Boston College Internal Review Board. Additionally, participants reported no psychiatric disorders or history of learning disabilities. Sample size was determined by available resources at the time data were collected; we note that although this sample size is small, it is typical of the time these data were collected in 2015.

Procedure and Materials

Participants were scanned while reading and responding to 64 vignettes, learning and making predictions about 64 different agents (see Supplemental Online Materials [SOM]). Each story was presented in three sequential segments: Initial Info, Prediction, and Final Info (see Figure 1). During the Initial Info segment, participants read background information to establish an expectation about how an agent would likely think or behave in the future. During the Prediction segment, participants were presented with a multiple-choice question asking them to make a prediction about the agent's future thoughts or behavior. Four options were provided: one that was expected based on the Initial Info segment, and three others that would be relatively unexpected. Participants responded to this question by using a button-box. During the Final Info segment, the vignette's outcome was presented, which corresponded to one option from the Prediction segment's multiple-choice question.

Critically, we varied the type of information presented in each segment. The Initial Info segment consisted of either the agent's 1) prior behavior, or 2) prior mental states. Similarly, during the Prediction segment, participants made a prediction about either the agent's 1) subsequent behavior, or 2) subsequent mental state. Last, the Final Info segment presented participants with either the agent's 1) actual subsequent behavior, or 2) actual subsequent mental state, which was the same type of information queried in the Prediction segment. Also, the Final Info segment presented was either expected or unexpected based on the information provided in the Initial Info segment. For the Final Info segment, an unexpected ending was shown on half of all trials, whereas an expected ending was shown on the remaining trials. Crossing these dimensions (Initial Info, Final Info, and Expectedness) yielded 8 conditions in a 2 (Initial Info: behavior, mental state) x 2 (Final Info: behavior, mental state) x 2 (Expectedness: expected, unexpected) design. These three experimental factors were also manipulated within stimuli, constituting a fully within-subject/within-stimulus design. The order of conditions and pairing of conditions and vignettes were randomized across participants. An online behavioral sample, that completed the same task as the fMRI participants did, verified that we successfully manipulated expectedness (see SOM).

The vignettes were presented in a pseudo-randomized order in white font on a black background via an Apple Macbook Pro running MATLAB 8.5 (2015) with Psychophysics Toolbox. The Initial Info segment was presented on-screen for 10 seconds, the Prediction segment for 8 seconds, and the Final Info segment for 4 seconds. To analyze Initial Info and Final Info segments separately, 0, 2, or 4 seconds of jittered fixation were included between each story segment. Stimulus presentation was divided into 8 equal runs (8 stimuli per run, 1 per condition) lasting 4 minutes and 4 seconds each.

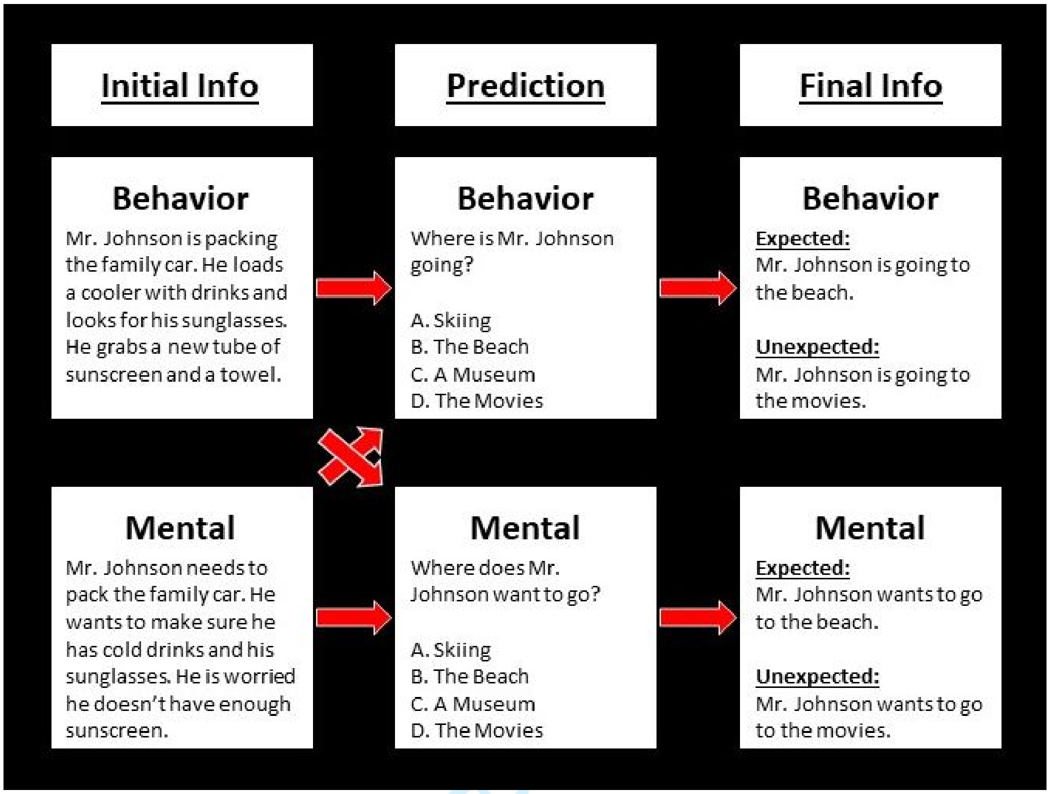

Figure 1. Example experimental stimulus and its variants. Participants first read the initial information segment of the stimulus (Behavior or Mental), then made a prediction (Behavior or Mental), and last learned the final information. Importantly, final information type (Behavior or Mental) always matched prediction type. Participants never saw the same stimulus across conditions; different participants saw different experimental variations of the same stimulus.

Functional Localizer

Participants also completed a theory of mind (ToM) functional localizer task (Dodell-Feder et al., 2011) consisting of 10 stories about mental states (e.g., false-belief condition) and 10 stories about physical representations (e.g., false-photograph condition). The task was presented in two 4.5-minute runs, interleaved between experiment runs.

fMRI Data Acquisition and Preprocessing

The fMRI data were collected using a 16-channel head coil in a 3T Siemens scanner at the Athinuoula A. Martinos Imaging Center, Massachusetts Institute of Technology. Data were acquired in 36 near-axial slices (3mm isotropic voxels, 0.54mm gap). Standard gradient echo planar imaging (EPI) procedures were used (TR=2000ms; TE=30ms; flip angle=90°; FOV=216 x 216; interleaved acquisition). Anatomical data were collected with T1-weighted multi-echo magnetization prepared rapid acquisition gradient echo image sequences (MEMPRAGE) using the following parameters: TR=2530ms; TE=1.64ms; FA=7˚; 1mm isotropic voxels; 0.5mm gap between slices; FOV=256 x 256. Data processing and analysis were performed using fMRIPrep (Esteban et al. (2019); see Supplementary Materials p. 1 for details), SPM12 (https://www.fil.ion.ucl.ac.uk/spm/software/ spm12/), and custom software. The functional data were realigned, coregistered to the anatomical image, normalized onto a common brain space (Montreal Neurological Institute, MNI, template), spatially smoothed using a Gaussian filter (fullwidth half-maximum = 8 mm kernel), and high-pass filtered (128 Hz). Neural responses were modeled in an event-related design using a general linear model (GLM), with conditions modeled as boxcar functions convolved with a canonical hemodynamic response function (HRF). The GLM included the six components of the anatomical CompCor variant (aCompCor) as nuisance regressors (Behzadi, Restom, Liau, Liu, 2007).

Analytic Approach

Whole-brain and regions of interest (ROI) analyses were conducted. For whole-brain analyses, we first conducted a whole-brain random effects analysis (voxel-wise threshold: p < .001, uncorrected; k > 16; cluster-wise threshold: p < .05, FWE-corrected) of behavior over mental conditions during the Initial Info segment. Second, we conducted a whole-brain random effects analysis of mental over behavior conditions during the Initial Info segment. Third, we conducted a whole-brain random effects analysis of expected over unexpected conditions during the Final Info segment. Fourth, we conducted a whole-brain random effects analysis of unexpected over expected conditions during the Final Info segment. For all whole-brain analyses, assignments of coordinates to brain regions were aided by use of the *label4MRI* package in R (Chuang & Yun-Shiuan, 2020), which performs automatic anatomic labeling (AAL) based on the most recently updated atlas, AAL3 (Rolls et al., 2020; see our OSF page for an RMarkdown file). Next~~For brevity~~, we describe ~~only~~ the ROI analyses ~~here (see SOM for whole-brain results)~~.

A whole-brain contrast of false-belief versus false-photograph stories in the ToM localizer task (Dodell-Feder et al., 2011) was used to identify ROIs implicated in ToM: DMPFC (N = 20), PC (N = 23), LTPJ (N = 22), and RTPJ (N = 23). ROIs were selected for each participant individually and defined as contiguous voxels within a 9-mm radius of the peak voxel that passed contrast threshold (see Table 1 for by-ROI coordinate information). Within each ROI, the average percent signal change (PSC) relative to runwise baseline (PSC=100×raw

BOLD magnitude for (condition−fixation)/raw BOLD magnitude for fixation) was calculated for each condition at each time point (averaging across all voxels in the ROI and all blocks of the same condition). Initial Info and Final Info segments were modeled separately. Timepoints were shifted by 6 seconds to account for hemodynamic lag. The Prediction segment was not analyzed.

Table 1. Average Peak MNI coordinates for ToM in Functional Localizer Task

		MNI Coordinates				
ROI	**N**	X	Y	Z	Voxels	t-value
RTPJ	23	53	-55	23	93	8.07
LTPJ	22	-49	-55	24	75	6.58
PC	23	-2	-58	34	85	6.40
DMPFC	20	3	51	23	61	5.47

Importantly, for all ROI analyses, we analyzed only trials on which participants selected the correct outcome prediction (i.e., the option that we determined, a priori, was the most likely to occur given the information that participants received in the Initial Info segment). This was done to ensure that our results were uncontaminated by the possibility of participants' lack of attention, random responding, or their own expectations being different from the paradigm's intended expectations. The frequency of incorrect predictions was similar across Initial Info x Final Info conditions (i.e., BB, BM, MB, MM): 16%, 12%, 15%, and 13% of each condition's total trials were incorrect, respectively. Further, we removed images according to the following criteria: individual scans, along with their two temporally adjacent scans, were excluded if framewise displacement (FD) (Power et al., 2012) exceeded 0.5mm; individual runs were excluded if either (1) FD for more than two-thirds of the scans in that run exceeded 0.5mm or (2) FD for any scan in that run exceeded 3mm (see results' table notes for final analyzable Ns). With the above data transformations/exclusions in mind, we next explain how we arrived at analyzable data. Specifically, for each participant's stimuli, multiple segments were presented (e.g., participant 1's Initial Info segment for vignette 1). Within each participant's stimulus, we averaged across time course activity during the segment of interest (accounting for hemodynamic lag). Finally, we used those averages in linear mixed effects models, meaning that each participant, for each stimulus, had a single PSC value for a specific segment of the experimental design.

For all ROI analyses, linear mixed effects models were constructed in R (R Core Team, 2021) to simultaneously account for variability across participants and stimuli (Judd, Westfall, & Kenny, 2012). Within each ROI, we attempted to fit a maximal model that allowed all main effects and interactions to vary over participants and stimuli. If the maximal model failed to converge or yielded a singular fit, we followed guidelines to avoid false positives (see Barr, 2013; Barr et al., 2013; Singmann & Kellen, 2019). First, we simplified the random effects structure by removing all correlations between random effects. Next, if the zero-correlation model failed to converge or converged with a singular fit, we further simplified the random effects structure by removing variance components that were estimated as zero. If this further reduced model resulted in non-convergence, or convergence with a singular fit, we repeated the second step. If there were no remaining variances estimated as zero, we removed the smallest variance components, one at a time, until the model converged with a non-singular fit. Last, when appropriate, we attempted to add random effects' correlations back into the model (see Bates et al., 2018). If this extended model converged with a non-singular fit, we retained it as our final model. However, if this extended model did not converge, or converged with a singular fit, we retained the non-extended model as our final model. We report only our final models here. The entire model selection process (i.e., specifications and simplifications) can be found on our OSF page: https://osf.io/tf852/.

Whole-Brain Results

As described above, we first conducted a whole-brain random effects analysis of behavior over mental conditions during the Initial Info segment. This contrast revealed only one peak cluster in the right angular gyrus [48, -64, 31]. Second, we conducted a whole-brain random effects analysis of mental over behavior conditions during the Initial Info segment. However, no peak clusters emerged for this contrast. Third, we conducted a whole-brain random effects analysis of expected over unexpected conditions during the Final Info segment. Fourth, we conducted a whole-brain random effects analysis of unexpected over expected conditions during the Final Info segment. See Table 2 for peak clusters revealed for these final two contrasts. We note that none of the ToM regions passed threshold in whole-brain contrasts for unexpected over expected outcomes.

Table 2. Regions passing threshold in whole-brain analysis during Final Info segment

Contrast	Region	MNI Coordinates X	Y	Z	*t*-Value	Cluster Size
Expected > Unexpected						
	L middle occipital gyrus	-48	-79	1	9.22	704
	L superior temporal gyrus	-63	-31	13	7.83	446
	R precuneus	12	-76	55	7.36	1864
	R superior frontal gyrus (dorsolateral)	21	-1	73	5.43	84
	R inferior temporal gyrus	51	-67	-5	5.33	210
	R lobule VI cerebellar hemisphere	21	-64	-14	5.00	169
	L lingual gyrus	-12	-88	-11	4.93	116
Unexpected > Expected						
	L supplementary motor area	-6	17	64	11.34	3567
	R inferior frontal gyrus pars orbitalis	48	26	-11	9.14	1019
	R thalamus	12	-7	7	7.67	129
	L lobule VI/V cerebellar hemisphere	-18	-37	-29	6.94	166
	R middle temporal gyrus	48	-22	-11	5.90	157
	L middle temporal gyrus	-57	-34	-8	5.82	224
	R lobule VI cerebellar hemisphere	33	-61	-29	5.31	192
	L angular gyrus	-45	-55	31	4.90	147

ROI Results

Initial Info

Here, we investigated the effect of the Initial Info manipulation on neural activity during the Initial Info segment. We note that model reduction within each ROI sometimes led to different random effects structures across ROIs. We chose this strategy of conservatism (rather than an anti-conservative strategy which held random effects structures constant across ROIs) ~~in order~~ to avoid false positives in some ROIs. Within each ToM ROI, there was no effect of the Initial Info manipulation on neural activity during the Initial Info segment. See Tables 3~~1~~ – 6~~4~~ for detailed information about all final models (i.e., coding scheme, random effects structure, random effects estimates, and fixed effects estimates). These patterns held when analyzing neural activity averaged across the entire ToM network (see Table 7~~5~~).

Table 3~~1~~. DMPFC activity during Initial Info segment

Final Model:
PSC ~ Initial + (1 | Item) + (1 | Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)

Random Effects	**Var.**	**SD**			
Item					
Intercept	.003	.052			
Subject					
Intercept	.015	.124			
Residual					
	.116	.341			
Fixed Effects			***b (SE)***	***t* (df)**	***p-value***
Intercept			-.03 (.03)	-1.01 (21)	.325
Initial			.01 (.02)	0.53 (997)	.593

Note. Analysis included 1063/1280 observations from 20 subjects and 64 items. Degrees of freedom were Satterthwaite-approximated and rounded to the nearest integer for all analyses.

Table 4~~2~~. PC activity during Initial Info segment

Final Model:
PSC ~ Initial + (1 + Initial || Item) + (1 | Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)

Random Effects	**Var.**	**SD**				
Item						
Intercept	.002	.048				
Initial	.001	.031				
Subject						
Intercept	.006	.079				
Residual						
	.103	.322				
Fixed Effects			***b (SE)***	***t* (df)**	***p-value***	
Intercept			-.04 (.02)	-2.07 (24)	.049	*
Initial			-.00 (.02)	-0.01 (60)	.992	

Note. Analysis included 1234/1472 observations from 23 subjects and 64 items.

Table 5~~3~~. LTPJ activity during Initial Info segment

Final Model:
PSC ~ Initial + (1 + Initial | Item) + (1 | Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)

Random Effects	**Var.**	**SD**	**Correlations**		
Item					
Intercept	.003	.051	-		
Initial	.005	.067	.23	-	
Subject					
Intercept	.049	.222	-		
Residual					
	.087	.294			

Fixed Effects	***b (SE)***	***t* (df)**	***p-value***	
Intercept	.12 (.05)	2.54 (22)	.019	*
Initial	-.01 (.02)	-0.67 (65)	.508	

Note. Analysis included 1189/1408 observations from 22 subjects and 64 items.

Table 6~~4~~. RTPJ activity during Initial Info segment

Final Model:
PSC ~ Initial + (1 | Item) + (1 | Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)

Random Effects	**Var.**	**SD**
Item		
Intercept	.001	.036
Subject		
Intercept	.012	.108
Residual		
	.071	.267

Fixed Effects	***b (SE)***	***t* (df)**	***p-value***
Intercept	-.02 (.02)	-0.68 (23)	.504
Initial	.01 (.02)	0.61 (1161)	.539

Note. Analysis included 1234/1472 observations from 23 subjects and 64 items.

Table 7~~5~~. ToM Network activity (averaged across ROIs) during Initial Info segment

Final Model:
PSC ~ Initial + (1 | Item) + (1 | Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)

Random Effects		**Var.**	**SD**			
Item						
	Intercept	.002	.039			
Subject						
	Intercept	.009	.095			
Residual						
		.047	.216			
Fixed Effects				***b (SE)***	***t* (df)**	***p-value***
	Intercept			.01 (.02)	0.30 (24)	.768
	Initial			.00 (.01)	0.13 (1158)	.897

Note. Analysis included 1234 observations from 23 subjects and 64 items.

Final Info

Here, we investigated the effect of the Initial Info, Final Info, and Expectedness manipulations on neural activity during the Final Info segment. See Figure 2 for results plotted by ROI and Tables 8~~6~~ – 11~~9~~ for detailed information about all final models. We note, here too, that model reduction within each ROI sometimes led to different random effects structures across ROIs. We chose this strategy of conservatism (rather than an anti-conservative strategy which held random effects structures constant across ROIs) ~~in order~~ to avoid false positives in some ROIs.

In DMPFC, LTPJ, and RTPJ, there were main effects of Expectedness, such that neural activity was higher when the final information was unexpected compared to expected (DMPFC: b = .10, SE = .03, p = .004; LTPJ: b = .06, SE = .02, p = .015; RTPJ: b = .06, SE = .02, p = .004). This main effect of Expectedness held when analyzing neural activity averaged across the entire ToM network (see Table 12~~0~~). To investigate the robustness of these unexpectedness effects, we investigated how many participants showed them, finding that most participants showed these effects within each ROI (DMPFC = 16/20; LTPJ = 13/22~~1~~; RTPJ = 17/23). Additionally, in DMPFC only, there was a main effect of Initial Info, such that neural activity during the Final Info segment was higher when Initial Info was behavior (compared to mental), b = .05 (SE = .02), p = .033. No other main effects or interactions were observed.

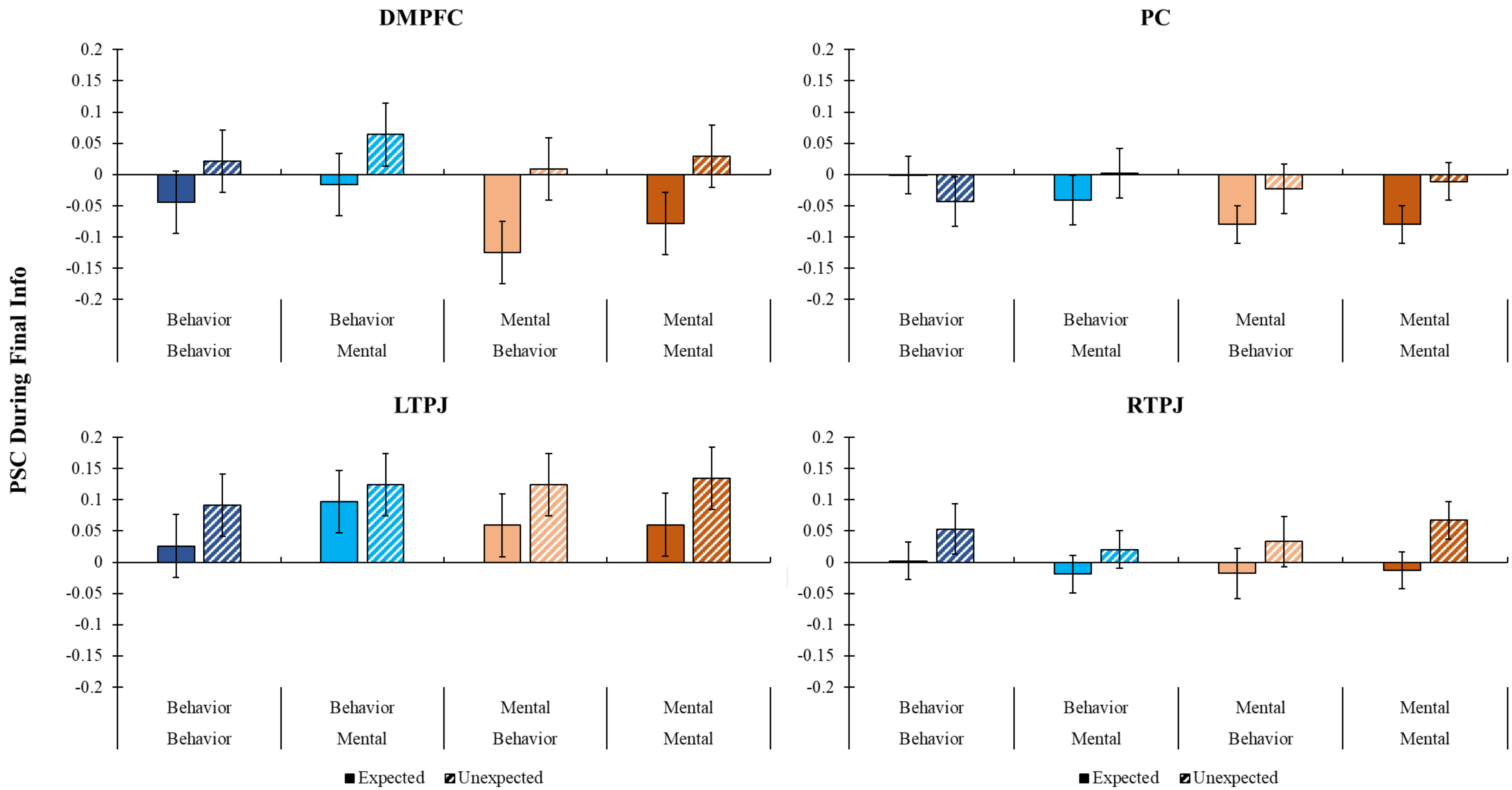

Figure 2. PSC during Final Info segment within each ROI. On the x-axis, the top factor (Behavior vs Mental) refers to the Initial Info manipulation, whereas the bottom factor (Behavior vs Mental) refers to the Final Info manipulation. Solid bars show expected outcome ~~responses~~estimates, whereas patterned bars show unexpected outcome estimates~~responses~~. Estimates are predicted means from each ROI's linear mixed effects model. Error bars represent +/- 1 SE for the predicted mean.

Table ~~6~~8. DMPFC activity during Final Info segment

Final Model:
PSC ~ Initial*Final*Expected +
(1 + Final + Initial:Final + Initial:Expected + Final:Expected || Item) +
(1 + Expected + Initial:Final:Expected || Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Final Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Expectedness: Contrast coded (Expected = -0.5; Unexpected = +0.5)

Random Effects	**Var.**	**SD**			
Item					
Intercept	.002	.045			
Final	.000	.094			
Initial:Final	.014	.119			
Initial:Expected	.008	.091			
Final:Expected	.013	.113			
Subject					
Intercept	.027	.163			
Expected	.007	.081			
Initial:Final:Expected	.004	.064			
Residual					
	.123	.350			
Fixed Effects			***b (SE)***	***t* (df)**	***p-value***
Intercept			-.02 (.04)	-0.45 (20)	.659
Initial			.05 (.02)	2.13 (869)	.033 *
Final			-.03 (.03)	-1.36 (60)	.178
Expected			.10 (.03)	3.35 (17)	.004 **
Initial:Final			-.00 (.05)	-0.04 (63)	.965
Initial:Expected			-.05 (.05)	-1.06 (60)	.294
Final:Expected			.01 (.05)	0.12 (62)	.903
Initial:Final:Expected			-.04 (.09)	-0.46 (14)	.653

Note. Analysis included 1021/1280 observations from 20 subjects and 64 items.

Table 9~~7~~. PC activity during Final Info segment

Final Model:
PSC ~ Initial*Final*Expected +
(1 + Expected + Initial:Final + Initial:Expected + Initial:Final:Expected || Item) +
(1 + Initial:Expected || Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Final Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Expectedness: Contrast coded (Expected = -0.5; Unexpected = +0.5)

Random Effects	**Var.**	**SD**			
Item					
Intercept	.002	.044			
Expected	.000	.001			
Initial:Final	.016	.128			
Initial:Expected	.005	.069			
Initial:Final:Expected	.043	.209			
Subject					
Intercept	.010	.100			
Initial:Expected	.002	.046			
Residual					
	.104	.323			
Fixed Effects			***b (SE)***	***t* (df)**	***p-value***
Intercept			-.03 (.02)	-1.47 (24)	.156
Initial			.03 (.02)	1.51 (991)	.140
Final			-.00 (.02)	-0.21 (989)	.829
Expected			.03 (.02)	1.68 (63)	.100
Initial:Final			.00 (.04)	0.10 (61)	.924
Initial:Expected			-.06 (.04)	-1.55 (17)	.139
Final:Expected			-.05 (.04)	-1.30 (979)	.194
Initial:Final:Expected			-.07 (.08)	-0.91 (59)	.368

Note. Analysis included 1179/1472 observations from 23 subjects and 64 items.

Table 108. LTPJ activity during Final Info segment

Final Model:
PSC ~ Initial*Final*Expected +
(1 + Initial + Expected + Initial:Final + Initial:Expected + Final:Expected +
Initial:Final:Expected || Item) +
(1 + Expected || Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Final Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Expectedness: Contrast coded (Expected = -0.5; Unexpected = +0.5)

Random Effects	**Var.**	**SD**
Item		
Intercept	.002	.040
Initial	.002	.047
Expected	.005	.069
Initial:Final	.029	.169
Initial:Expected	.013	.113
Final:Expected	.022	.148
Initial:Final:Expected	.052	.228
Subject		
Intercept	.037	.192
Expected	.001	.027
Residual		
	.093	.306

Fixed Effects	***b (SE)***	***t* (df)**		***p-value***	
Intercept	.09 (.04)	2.10	(21)	.048	*
Initial	-.01 (.02)	-0.48	(58)	.631	
Final	-.03 (.02)	-1.54	(894)	.125	
Expected	.06 (.02)	2.72	(17)	.015	*
Initial:Final	-.05 (.04)	-1.08	(61)	.284	
Initial:Expected	-.02 (.04)	-0.62	(61)	.550	
Final:Expected	.01 (.04)	0.35	(59)	.726	
Initial:Final:Expected	.05 (.08)	0.60	(60)	.553	

Note. Analysis included 1134/1408 observations from 22 subjects and 64 items.

Table 11~~9~~. RTPJ activity during Final Info segment

Final Model:

PSC ~ Initial*Final*Expected +
(1 + Initial:Final + Initial:Final:Expected || Item) +
(1 + Expected + Initial:Expected + Final:Expected + Initial:Final:Expected || Subject)

Coding:

Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Final Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Expectedness: Contrast coded (Expected = -0.5; Unexpected = +0.5)

Random Effects	**Var.**	**SD**
Item		
Intercept	.001	.024
Initial:Final	.008	.087
Initial:Final:Expected	.002	.069
Subject		
Intercept	.014	.117
Expected	.000	.015
Initial:Expected	.005	.068
Final:Expected	.001	.030
Initial:Final:Expected	.026	.162
Residual		
	.083	.229

Fixed Effects	***b (SE)***	***t* (df)**		***p-value***	
Intercept	.02 (.03)	0.60	(22)	.552	
Initial	-.00 (.02)	-0.21	(986)	.828	
Final	.00 (.02)	0.21	(988)	.836	
Expected	.06 (.02)	3.24	(21)	.004	**
Initial:Final	.05 (.04)	1.31	(60)	.195	
Initial:Expected	-.02 (.04)	-0.55	(18)	.590	
Final:Expected	-.01 (.03)	-0.24	(22)	.813	
Initial:Final:Expected	.04 (.08)	0.52	(17)	.611	

Note. Analysis included 1179/1472 observations from 23 subjects and 64 items.

Table 12~~0~~. ToM Network activity (averaged across ROIs) during Final Info segment

Final Model:
PSC ~ Initial*Final*Expected +
(1 + Final + Expected + Initial:Final + Final:Expected +
Initial:Final:Expected || Item) +
(1 | Subject)

Coding:
Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Final Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Expectedness: Contrast coded (Expected = -0.5; Unexpected = +0.5)

Random Effects	**Var.**	**SD**				
Item						
Intercept	.001	.035				
Final	.001	.032				
Expected	.000	.016				
Initial:Final	.007	.083				
Final:Expected	.007	.082				
Initial:Final:Expected	.021	.146				
Subject						
Intercept	.012	.109				
Residual						
	.052	.227				
Fixed Effects			***b (SE)***	***t* (df)**	***p-value***	
Intercept			.01 (.02)	0.40 (23)	.696	
Initial			.01 (.01)	1.07 (953)	.284	
Final			-.01 (.01)	-1.06 (61)	.294	
Expected			.06 (.01)	4.36 (61)	< .001	***
Initial:Final			.00 (.03)	0.08 (61)	.938	
Initial:Expected			-.03 (.03)	-1.25 (949)	.212	
Final:Expected			-.01 (.03)	-0.41 (61)	.684	
Initial:Final:Expected			-.00 (.06)	-0.02 (64)	.985	

Note. Analysis included 1179 observations from 23 subjects and 64 items.

General Discussion

The current work investigated whether social prediction error (i.e., the difference between what was expected and what was observed) occurs across brain regions typically implicated in thinking about others' minds, as well as whether these regions' sensitivity to social prediction error differs as a function of the type of information that is used to make predictions and confirm or violate them. When people learned about an agent's unexpected mental states or behavior, DMPFC, LTPJ, and RTPJ activity was higher than when people learned about an agent's expected mental states or behaviors. But no region showed differential effects of expectedness based on the type of information that was used to make predictions or confirm/violate them (i.e., mental states versus behaviors), suggesting that these regions may flexibly incorporate any available social information to make predictions about and monitor or explain social inconsistencies. These findings add to a growing literature (e.g., Cloutier et al., 2011; Dungan et al., 2016; Heil et al., 2019; Theriault et al., 2020; Saxe & Wexler, 2005) supporting the link between predictive coding and ToM activity (Koster-Hale & Saxe, 2013). However, some of our results are consistent with prior research, whereas others are not.

In the current experiment, effects of expectedness in DMPFC, LTPJ and RTPJ are consistent with prior work. Specifically, such effects have been found in investigations of unexpected behaviors (Dungan et al., 2016; Heil et al., 2019) and unexpected mental states (Cloutier et al., 2011). Additionally, that we did not find an effect of expectedness in PC is also consistent with prior research (Dungan et al., 2016). However, we caution readers that this lack of detection (as well as other null effects) could be interpreted in the following ways: (1) no effect occurs in PC, (2) a theoretically meaningful but small effect occurs in PC but could not be detected, or (3) any effect that occurs in PC is too small to be theoretically meaningful. Because there will be disagreement about what is theoretically meaningful to all researchers, we do not take a firm position on this issue. We do note the following: in the current experiment, standardized effect sizes for expectedness in DMPFC (d = 0.21), LTPJ (d = 0.11), and RTPJ (d = 0.15) ranged from more than 1.5x – 3x the standardized effect size for expectedness in PC (d = 0.07). (All d's were calculated using variance estimates, as described in Brysbaert & Stevens, 2018; Westfall, Judd, & Kenny, 2014). These estimates, combined with Dungan et al.'s reported null effect, suggests that PC may diverge from other ToM regions in monitoring expectation violations.

Other findings in the current experiment are somewhat surprising considering prior research. In particular, expectedness effects in ToM regions were not moderated by the type of information that engendered or violated predictions. On one popular account of cultural learning, people attend to "credibility-enhancing" displays to determine their degree of confidence in someone else's beliefs (Henrich, 2009), being more confident in someone else's beliefs when a costly behavior reflects the purported belief. Applying this logic to the current data, one possibility is that prior behaviors could have served as stronger predictors of subsequent behaviors/mental states because behaviors are interpreted as better signals of one's current beliefs or intentions. For example, community organizers who had themselves installed solar panels were more effective in recruiting new residents to install solar panels than organizers who had not done so, as the former engaged in costly behavior which was inferred as an honest signal of their belief in the technology's benefit (Kraft-Todd et al., 2018).

This possibility is also consistent with research by Dungan et al. (2016) in which an expectedness effect in RTPJ was driven by one condition (see Figure 2 of Dungan et al.). In the outcome segment, Dungan et al. reported that RTPJ showed an expectedness effect only when prior information was behavior (but not mental), suggesting that RTPJ may be recruited specifically when an agent's subsequent behavior is inconsistent with their prior behavior, perhaps in order to generate mental states that explain the behavioral inconsistency (see Decety & Lamm, 2007). In the current experiment, however, when outcome information was behavior (as was always true in Dungan et al., 2016), we did not replicate this finding. More specifically, both expectedness simple effects were identical in magnitude in our data (initial behavior d = 0.14; initial mental d = 0.14). Using all of our data, we also failed to detect a three-way interaction among prior information, outcome information, and expectedness, suggesting that Dungan et al.'s (2016) findings in RTPJ may not be due to mismatches between prior information and outcome information.

On the other hand, on the logic that people can never be certain about the inferred mental states of others, explicit mental states could have served as stronger predictors of subsequent behaviors/mental states. That participants might make the most reliable predictions about future mental states based on prior mental states is also consistent with recent theoretical work arguing that people can track the transitional probabilities between mental states (Tamir & Thornton, 2018). Even though people do not come with thought bubbles above their heads in the real world, participants in the current experiment were given explicit access to others' mental states in a thought-bubble-like way. Therefore, the current data's lack of prior-by-outcome information moderation on prediction error signals is surprising. However, it has also been argued that social outcomes are inherently much less predictable than non-social outcomes (FeldmanHall & Shenhav, 2019), which may explain why social information type does not moderate the effects here. Further experimentation, and much larger (and therefore higher-powered) fMRI paradigms are needed to better understand if, when, and how social prediction error interacts with the type of social information that predictions and violations were based on, as the observed null interactions could simply be false negatives.

There are multiple methodological features that may explain inconsistencies between the current work and Dungan et al.'s findings specifically. First, a strength of the current work is that stimuli were constructed for both behavioral and mental conditions of the prior information segment (see Figure 1~~2~~ for an example), whereas stimuli in Dungan et al. were nested in a particular prior information condition (i.e., completely different stimuli constituted behavior versus mental conditions). Therefore, Dungan et al. 's effects of initial behavior versus initial mental states may have been driven by stimulus differences rather than a true distinction between information types. Second, data in the current experiment were analyzed with linear mixed effects models as opposed to traditional repeated-measures ANOVAs. For multilevel data (e.g., responses nested within participants/stimuli), linear mixed effects models better control Type I error rates by retaining the true variability in the data and adjusting standard errors of test statistics to account for the possibility that some participants/stimuli will respond (to an experimental manipulation) differently than other participants/stimuli. Therefore, some of Dungan et al.'s effects may have been due to specific participants or stimuli behaving in ways that led to a group-level effect which was not representative of most participants or stimuli. These methodological changes can also explain another discrepancy between the present data and Dungan et al.'s. In the prior information segment, Dungan et al. detected an effect in which RTPJ activity was higher when initial information was behavior (versus mental), suggesting that RTPJ may play a special role in mental state *inference* based on witnessed behavior rather than processing mental states directly. However, in the current experiment, we found no evidence of this effect (d = 0.03). Sample size issues notwithstanding, we believe that the methods of the current experiment offer the best tests of the ideas under investigation thus far. Therefore, that we successfully replicated prior research's effects of expectedness lends especially strong evidence to the idea that DMPFC, LTPJ, and RTPJ are ToM regions coding for discrepancies between social predictions and their outcomes.

Limitations and Future Directions

Although the current experiment improved on prior work, it has important limitations. First, our design focused on people's predictions about unknown others based on a single prior behavior or mental state. However, people more typically interact with and make predictions about agents they know. Social prediction error may occur more strongly or weakly depending on one's relationship to the agent who is the object of prediction. For example, people believe that there are stronger obligations to help family members compared to non-family members (Marshall et al., 2021; McManus, Kleiman-Weiner, & Young, 2020; McManus, Mason, & Young, 2021), leading to neglect of family members being judged as more unexpected (see SOM of McManus, Kleiman-Weiner, & Young, 2020). Additionally, people seem to experience stronger prediction error when they imagine witnessing close (versus distant) others commit crimes (Berg, Kitiyama, & Kross, 2021), which may be due to their having stronger positive priors about close others (see Hughes, Ambady, & Zaki, 2017; Hughes, Zaki, & Ambady, 2017; Kim, Park, & Young, 2020). Future research can shed more light on the neural mechanisms involved in using relationship information to make and monitor predictions about an agent's mental states and behavior.

Second, the current experiment did not systematically vary the social context in which expectations were confirmed or violated, which may alter if and how social prediction error occurs. For example, imagine that an agent thinks to herself, "I want to speak up the next time I hear a sexist joke about women." However, the next time she hears a sexist joke is at her workplace where all of her colleagues are men. When she fails to speak up, you may be unsurprised, perhaps because you understand that she might experience additional negativity at her workplace in the future. Conversely, if at least half of her colleagues were women, you may be extremely surprised when she fails to speak up, perhaps because you believe she would have wanted to alleviate the possible discomfort experienced by her same-gendered colleagues. Such an example suggests that there are many potential features of the social context that can affect predictions and therefore what is considered unexpected, such as an agent's reputational concerns, the demographic composition of surrounding others, and more.

Additionally, this example demonstrates that the time point at which a prediction is made, and whether this prediction is updated, is a crucial factor. If you were to predict the woman's behavior at the exact time a sexist joke occurred, you might make starkly different predictions based on her social context. If, however, you predicted the woman's behavior at an earlier time (e.g., the time at which you first learned she imagined speaking up), you may or may not update your prediction based on the context in which the sexist joke occurs. While recent work has documented that people update their (moral) impressions of others through learning more about the agents' past behaviors (see Brambilla et al., 2019; Kim et al., 2020; Mann & Ferguson, 2017), less is known about how people update their predictions of an agent's single future behavior over time. Future research could investigate the conditions under which predictions are updated and how this relates to prediction error (see Bach & Schenke, 2017, for a detailed discussion).

Third, although we failed to detect interaction effects based on the sources of information used to make and violate predictions, it is unlikely that such effects would never occur. For example, imagine learning that someone thought to himself, "I really hate my coworkers and my role in this company. I want to quit." When he continues to go into work, you may not be very surprised because you realize that people often think and want to do things that they do not do. However, if he instead sent out a company-wide e-mail stating, “I really hate my coworkers and my role in this company. I’m going to quit,” you might be extremely surprised when he continues to go into work. The critical difference between these cases is that the agent’s prior behavior seems to leave no doubt about his near-future intentions. We may not have seen such effects in the current data because most of the agents’ prior behaviors did not yield near-certain predictions. Interestingly, follow-up behavioral and behavior-brain analyses of our data suggest that vignette-level prediction confidence indeed varied substantially and was positively related to the difference in vignette-level neural activity during expectation confirmation/violation (see SOM). Unfortunately, we did not have enough condition-specific data to address whether prior information moderated these relations.

I~~Relatedly, i~~t is possible that our experimental paradigm was responsible for the observed null interactions. That is, after participants saw the first few stimuli, they would have become aware that they would have to continually make predictions about future behavior or future mental states based on both prior behavior and prior mental states. Once this awareness set in, it is possible (even likely) that participants were generating future behavior and future mental state predictions for each stimulus. This could have led to non-interactions between prior and future information on unexpectedness-related neural activity. Therefore, although our experimental paradigm was designed to address potential alternative explanations for prior research (i.e., Dungan et al. 2016), it may have fundamentally altered the psychological experience that we intended to study. Moreover, perhaps our paradigm exaggerated the effect of expectation violations. Specifically, in the real world, people may not engage in explicit predictions in the way that our participants did. Therefore, perhaps prompting explicit predictions led to artificial but equally strong predictions across information types, leading to the observed null interactions.

~~Moreover~~We opted~~,~~ for the explicit prediction methodology for two reasons: first to ensure that participants indeed engaged in prediction, and second to constrain the kinds of predictions made so that different participants would not be making different predictions. However, even though we attempted to control the experimental environment and constrain the types of predictions made as much as possible (i.e., behavior vs mental state), participants were likely to spontaneously generate mental state inferences regardless of the information presented. The reason for this, we propose, is that our experimental task (i.e., social prediction) may fundamentally rest on mental state inferences, as trying to predict a target's future mental states or behavior requires inferring the beliefs of the target. As an additional example of an experimental task that supports this point, prior research shows that people engage in spontaneous mental state inference when making moral judgments (Young & Saxe, 2009), as making moral judgments demands inferring the beliefs and intentions of a transgressor. This possibility results in two falsifiable predictions. Specifically, if participants spontaneously generate mental state inferences, then neural activity should have been similar for behavior and mental state stimuli immediately following the Initial Info segment. Additionally, participants should be equally good at choosing "correct" mental state outcomes during the Prediction segment, regardless of the Initial Info type. Both predictions were supported in our data. First, in the Initial Info segment, ToM recruitment was similar in magnitude for behavioral versus mental state information. Second, in the Prediction segment, participants were similarly likely to make the "correct" mental state prediction following Initial Info that was behavioral versus mental state (i.e., being correct on 88% vs 87% of trials, respectively). Therefore, our data suggest one of ~~a~~ two theoretically plausible possibilities: ToM regions flexibly incorporate any available social information in order to make predictions and monitor their violation/confirmation, or, regardless of what type of social information is perceived, ToM regions spontaneously generate mental state inferences to use for prediction and then monitor their violation/conformation. ~~as noted elsewhere, even if our experimental paradigm was unproblematic, null interactions could have been a function of low statistical power. they~~Consequently, future, more targeted, research is needed to better understand if and when prior behavior (or mental state) information leads to stronger predictions and its consequences on prediction error.

Last, there have been continued calls to communicate "constraints on generalizability" in psychology and neuroscience research (Simons, Shoda, & Lindsay, 2017; Yarkoni, 2020). In addition to the above limitations, it is unclear if our recruited fMRI participants are representative of most people. Recent work suggests that fMRI research suffers from generalizability issues. Specifically, fMRI samples tend to be lower in trait anxiety than behavioral samples (Charpentier et al., 2021), suggesting a self-selection bias. Since past research has linked anxiety to ToM abilities (Washburn et al., 2016) and difficulty in understanding/completing ToM tasks (Lenton-Brym et al., 2018), future research on social prediction error would benefit from considering the role of anxiety and other individual differences.

Conclusion

The current research found that brain regions implicated in theory of mind (ToM: DMPFC, LTPJ, and RTPJ) are especially responsive to an agent's unexpected behavior or mental states based on knowledge of their prior behavior or mental states. These findings also suggest that ToM regions may not discriminate in their sensitivity to expectedness based on information type, though additional research is needed to conclusively provide evidence for or against this possibility. Overall, these findings replicate recent research consistent with a predictive coding account of the neural computations underlying ToM (Koster-Hale & Saxe, 2013), and lay the foundation for future research investigating when, how, and for whom certain kinds of prior social knowledge give rise to robust predictions and therefore shape prediction error signals.

Acknowledgments

We thank current and former members of the Morality Lab at Boston College for feedback on an original version of the manuscript.

References

Bach, P., Schenke, K. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass, 11*(7), e12312.

Barr, D.J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4*, 328. https://doi.org/10.3389/fpsyg.2013.00328

Barr, D.J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D., Reinhold, K., Vasishth, S., & Baayen, R.H. (2018). Parsimonious mixed models. 1506.04967v2.pdf (arxiv.org)

Behzadi, Y., Restom, K., Liau, J., & Liu, T.T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage, 37*, 90-101.

Berg, M. K., Kitayama, S., & Kross, E. (2021). How relationships bias moral reasoning: Neural and self-report evidence. *Journal of Experimental Social Psychology*, *95*, 104156. https://doi.org/10.1016/j.jesp.2021.104156

Brambilla, M., Carraro, L., Castelli, L., & Sacchi, S. (2019). Changing impressions: Moral character dominates impression updating. *Journal of Experimental Social Psychology, 82*, 64-73.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), Article 9. Doi:10.5334/joc.10

Charpentier, C. J., Faulkner, P., Pool, E. R., Ly, V., Tollenaar, M. S., Kluen, L. M., Fransen, A., Yamamori, Y., Lally, N., Mkrtchian, A., Valton, V., Huys, Q. J. M., Sarigiannidis, I., Morrow, K. A., Krenz, V., Kalbe, F., Cremer, A., Zerbes, G., Kausche, F. M., … O'Doherty, J. P. (2021). How representative are neuroimaging samples? Large-scale evidence for trait anxiety differences between fMRI and behaviour-only research participants. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsab057

~~Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), Article 9. Doi:10.5334/joc.10~~

Chuang & Yun-Shiuan (2020). *label4MRI: MRI-labeling.*

Cloutier, J., Gabrieli, J. D. E., O'Young, D., & Ambady, N. (2011). An fMRI study of violations of social expectations: When people are not who we expect them to be. *NeuroImage*, *57*(2), 583–588. https://doi.org/10.1016/j.neuroimage.2011.04.051

Decety, J., & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *The Neuroscientist, 13*(6), 580-593.

Dungan, J. A., Stepanovic, M., & Young, L. (2016). Theory of mind for processing unexpected events across contexts. *Social Cognitive and Affective Neuroscience*, *11*(8), 1183–1192. https://doi.org/10.1093/scan/nsw032

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behavior, 3,* 426-435.

Ferguson, M.J., & Mann, T.C. (2017). Reversing implicit first impressions through reinterpretation after a two-day delay. *Journal of Experimental Social Psychology, 68,* 122 – 127.

Fiorilla, C.D., Tobler, P.N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*, 1898–1902.

Heil, L., Colizoli, O., Harstra, E., Kwisthout, J., van Pelt, S., van Rooij, I., & Bekkering, H. (2019). Processing of predictions errors in mentalizing areas. *Journal of Cognitive Neuroscience*, *31*(6), 900–912. https://doi.org/10.1162/jocn

Henrich, J. (2009). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior, 30*, 244-260.

Hughes, B.L., Ambady, N., & Zaki, J. (2017). Trusting outgroup, but not ingroup members, requires control: neural and behavioral evidence. *Social Cognitive Affective Neuroscience*, *12*(3), 372–381.

Hughes B.L., Zaki, J., Ambady, N. (2017). Motivation alters impression formation and related neural systems. *Social Cognitive Affective Neuroscience*, *12*(1), 49–60.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *Journal of personality and social psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kim, M., Park, B., & Young, L. (2020). The psychology of motivated versus rational impression updating. *Trends in Cognitive Sciences*, *24*(2), 101-111. doi: 10.1016/j.tics.2019.12.001

Kim, M., Mende-Siedlecki, P., Anzellotti, S., & Young, L. (2020). Theory of mind following the violation of strong and weak prior beliefs. *Cerebral Cortex, 31*(2), 884-898.

Koster-Hale, J., & Saxe, R. (2013). Theory of Mind: A neural prediction problem. *Neuron*, *79*(5), 836–848. https://doi.org/10.1016/j.neuron.2013.08.020

Kraft-Todd, G.T., Bollinger, B., Gillingham, K., Lamp, S., & Rand, D.G. (2018). Credibility-enhancing displays promote the provision of non-normative public goods. *Nature, 563*, 245-248.

Lenton-Brym, A. P., Moscovitch, D. A., Vidovic, V., Nilsen, E., & Friedman, O. (2018). Theory of mind ability in high socially anxious individuals. *Anxiety, stress, and coping*, *31*(5), 487–499. https://doi.org/10.1080/10615806.2018.1483021

Marshall, J., Gollwitzer, A., Mermin-Bunnell, N., Shinomiya, M., Retelsdorf, J., & Bloom, P. (in press). How development and culture shape intuitions about prosocial obligations. *Journal of Experimental Psychology: General*.

McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science*, *31*(3), 227–242. doi: 10.1177/0956797619900321

McManus, R.M., Mason, J.E., & Young, L. (2021). Re-examining the role of family relationships in structuring perceived helping obligations, and their impact on moral evaluation. *Journal of Experimental Social Psychology, 96*, 104182.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage, 59*(1), 2142-54. doi: 10.1016/j.neuroimage.2011.10.018

Rolls, E.T., Huang, C., Lin, C., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *NeuroImage, 206*, 116189.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Simons, D.J., Shoda, Y., & Lindsay, D.S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123-1128.

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. *New Methods in Cognitive Psychology*, 4–31. https://doi.org/10.4324/9780429318405-2

Tamir, D.I., & Thornton, M. (2018). Modeling the predictive social mind. *Trends in Cognitive Science, 22*(3), 201 – 212.

Theriault, J. E., Young, L., & Barrett, L. F. (2021). The sense of should: A biologically-based framework for modeling social pressure. *Physics of Life Reviews*, *36*, 100–136. https://doi.org/10.1016/j.plrev.2020.01.004

Theriault, J., Waytz, A., Heiphetz, L., & Young, L. (2020). Theory of Mind network activity is associated with metaethical judgment: An item analysis. Neuropsychologia, 107475.

Wacongne, C., Changeux, J.P., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience, 32*, 3665 – 3678.

Washburn, D., Wilson, G., Roes, M., Rnic, K., & Harkness, K.L. (2016). Theory of mind in social anxiety disorder, depression, and comorbid conditions. *Journal of Anxiety Disorders, 37*, 71-77.

Westfall, J., Kenny, D.A., & Judd, C.M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*(5), 2020-2045.

Yarkoni, T. (2020). The generalizability crisis. *Behavioral and Brain Sciences.* https://doi. org/10.1017/S0140525X20001685.

Young, L., & Saxe, R. (2009). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Nueroscience, 21,* 1396 – 1405.

Supplemental Online Materials for:

How Unexpected Events are Processed In Theory of Mind Regions: A Conceptual Replication

1. **Manipulation Check**

~~2.~~ ~~**Whole-Brain Contrasts**~~

~~3.~~2. **Item-Analysis Studies**

~~4.~~3. **Experimental Stimuli**

~~5. **Additional References**~~

Manipulation Check

To ensure that the fMRI experiment successfully manipulated expectations, we collected data from an additional behavioral sample.

Methods

Participants

Behavioral data were collected from 60 people recruited through Amazon's Mechanical Turk. After excluding participants who failed in-task attention checks, our final $N = 59$.

Procedure and Materials

Participants were presented with the fMRI experiment's 64 vignettes in randomized order. Participants saw all three segments of each vignette, just like the fMRI participants did. They read the Initial Info segment, made a prediction, and finally, reported how (un)expected the outcome was (0 = *Extremely expected* to 100 = *Extremely unexpected*).

No participant saw the same scenario across conditions. For example, Participant A saw Item 1 in the "Behavior" condition of the Initial Info segment, made a prediction about future behavior, saw an unexpected outcome, and then rated its unexpectedness. On the other hand, Participant B saw almost the same variation but instead saw an expected outcome, and so on.

Analytic Approach

To analyze these behavioral data, we employed mixed effects models that allowed all factors (Initial Info, Final Info, Expectedness) to vary randomly over participants and scenarios. We used the same model simplification strategy as reported in the main text to reach final models that converged with non-singular fits. We report only the final model here. All models excluded responses that followed incorrect predictions (3,356 of 3,776 or 89% of all trials). Importantly, predictions were not differentially incorrect by Initial*Final conditions (i.e., BB, BM, MB, MM): 10%, 10%, 11%, and 12% of each condition's total trials were incorrect, respectively.

Results

There was a main effect of Expectedness, such that unexpected outcomes were judged as more unexpected than expected outcomes ($b = 70.99$, $SE = 2.79$, $p < .001$), suggesting that the fMRI manipulation was successful. There were no other main effects or interactions. See SOM Table 1 for detailed information about the final model.

SOM Table 1. Unexpectedness judgments in behavioral sample.

Final Model:

PSC ~ Initial*Final*Expected +
(1 + Initial + Expected + Final:Expected || Item) +
(1 + Expected + Initial:Final + Final:Expected + Initial:Final:Expected || Subject)

Coding:

Initial Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Final Info: Contrast coded (Mental = -0.5; Behavior = +0.5)
Expectedness: Contrast coded (Expected = -0.5; Unexpected = +0.5)

Random Effects	**Var.**	**SD**
Item		
Intercept	16.01	4.00
Initial	0.28	0.52
Expected	64.27	8.02
Initial:Expected	2.56	1.60
Subject		
Intercept	29.36	5.42
Expected	386.06	19.65
Initial:Final	20.28	4.50
Final:Expected	2.55	1.60
Initial:Final:Expected	3.80	1.95
Residual		
	191.95	13.85

Fixed Effects	***b (SE)***	***t* (df)**	***p-value***	
Intercept	41.66 (0.90)	46.31 (87)	< .001	***
Initial	- 0.52 (0.49)	- 1.08 (55)	.283	
Final	0.37 (0.48)	0.76 (2910)	.447	
Expected	70.99 (2.79)	25.43 (70)	< .001	***
Initial:Final	1.24 (1.13)	1.10 (48)	.278	
Initial:Expected	- 0.45 (0.98)	- 0.46 (57)	.650	
Final:Expected	0.54 (0.99)	0.55 (48)	.587	
Initial:Final:Expected	2.62 (1.94)	1.35 (49)	.183	

Note. Analysis included 3356 observations from 59 subjects and 64 items.

~~Whole-Brain Contrasts~~

~~First, we conducted a whole-brain random effects analysis (voxel-wise threshold: p < .001, uncorrected; k > 16; cluster-wise threshold: p < .05, FWE-corrected) of behavior over mental conditions during the Initial Info segment. This contrast revealed only one peak cluster in the right angular gyrus [48, -64, 31]. Second, we conducted a whole-brain random effects analysis of mental over behavior conditions during the Initial Info segment. However, no peak clusters emerged for this contrast. For all whole-brain analyses, assignments of coordinates to particular brain regions were aided by use of the *label4MRI* package in R (Chuang & Yun-Shiuan, 2020), which performs automatic anatomic labeling (AAL) based on the most recently updated atlas, AAL3 (Rolls et al., 2020; see our OSF page for an RMarkdown file).~~

~~Third, we conducted a whole-brain random effects analysis (voxel-wise threshold: p < .001, uncorrected; k > 16; cluster-wise threshold: p < .05, FWE-corrected) of expected over unexpected conditions during the Final Info segment. Fourth, we conducted a whole-brain random effects analysis of unexpected over expected conditions during the Final Info segment. See Table 6 for peak clusters revealed for each contrast.~~

~~**SOM Table 1.** Regions passing threshold in whole-brain analysis during Final Info segment~~

		~~**MNI Coordinates**~~				
~~**Contrast**~~	~~**Region**~~	~~**X**~~	~~**Y**~~	~~**Z**~~	~~***t*-Value**~~	~~**Cluster Size**~~
~~**Expected > Unexpected**~~						
	~~L middle occipital gyrus~~	~~-48~~	~~-79~~	~~1~~	~~9.22~~	~~704~~
	~~L superior temporal gyrus~~	~~-63~~	~~-31~~	~~13~~	~~7.83~~	~~446~~
	~~R precuneus~~	~~12~~	~~-76~~	~~55~~	~~7.36~~	~~1864~~
	~~R superior frontal gyrus (dorsolateral)~~	~~21~~	~~-1~~	~~73~~	~~5.43~~	~~84~~
	~~R inferior temporal gyrus~~	~~51~~	~~-67~~	~~-5~~	~~5.33~~	~~210~~
	~~R lobule VI cerebellar hemisphere~~	~~21~~	~~-64~~	~~-14~~	~~5.00~~	~~169~~
	~~L lingual gyrus~~	~~-12~~	~~-88~~	~~-11~~	~~4.93~~	~~116~~
~~**Unexpected > Expected**~~						
	~~L supplementary motor area~~	~~-6~~	~~17~~	~~64~~	~~11.34~~	~~3567~~
	~~R inferior frontal gyrus pars orbitalis~~	~~48~~	~~26~~	~~-11~~	~~9.14~~	~~1019~~
	~~R thalamus~~	~~12~~	~~-7~~	~~7~~	~~7.67~~	~~129~~
	~~L lobule VI/V cerebellar hemisphere~~	~~-18~~	~~-37~~	~~-29~~	~~6.94~~	~~166~~
	~~R middle temporal gyrus~~	~~48~~	~~-22~~	~~-11~~	~~5.90~~	~~157~~
	~~L middle temporal gyrus~~	~~-57~~	~~-34~~	~~-8~~	~~5.82~~	~~224~~
	~~R lobule VI cerebellar hemisphere~~	~~33~~	~~-61~~	~~-29~~	~~5.31~~	~~192~~
	~~L angular gyrus~~	~~-45~~	~~-55~~	~~31~~	~~4.90~~	~~147~~

Item-Analysis Studies

To better understand possible influences on neural activity during the Final Info segment, we collected behavioral data from additional online samples, with three primary investigation goals: (1) the relationship between by-stimulus confidence judgments during the Prediction segment and by-stimulus unexpectedness judgments following the Final Info segment; (2) the relationship between by-stimulus confidence judgments during the Prediction segment and neural activity during the Final Info segment; and (3) the relationship between by-stimulus unexpectedness judgments following the Final Info segment and neural activity during the Final Info segment. The logic of these tests was to investigate whether certain stimuli elicited more or less confidence in predictions about the outcome, and therefore, whether certain stimuli generated more or less neural activity as a function of the eventual outcome being more or less surprising.

Methods

Participants

Behavioral data were collected from two independent samples. Sample 1 (confidence judgments) contained responses from 33 people recruited through Amazon's Mechanical Turk. Sample 2 (unexpectedness judgments from above "Manipulation Check" section) contained responses from 60 people also recruited through Amazon's Mechanical Turk; we doubled our sample size here to have roughly equal numbers of observations per each stimulus' unexpected and expected outcomes as we did for each stimulus' confidence judgments. After excluding participants who failed in-task attention checks, Sample 1's final *N* = 31, and Sample 2's final *N* = 59.

Procedure and Materials

Participants were presented with the fMRI experiment's 64 vignettes in randomized order. In Sample 1, participants never learned the outcome of the vignette. They read the Initial Info segment, selected which of the four possible outcomes they predicted would occur, and finally, indicated how confident they were in their prediction (0 = *Not at all confident* to 100 = *Extremely confident*). In Sample 2, participants saw all three segments of each vignette, just like the fMRI participants did. They read the Initial Info segment, made a prediction, and finally, reported how (un)expected the outcome was (0 = *Extremely expected* to 100 = *Extremely unexpected*).

Across behavioral samples, as was the case in the fMRI experiment, each participant only saw one variant of each stimulus. For example, in Sample 1 (confidence judgments), Participant A saw Item 1 in the "Behavior" condition of the Initial Info segment and then made a prediction about future behavior, whereas Participant B saw Item 1 in the "Mental" condition of the Initial Info segment and then made a prediction about future behavior, and so on. In Sample 2 (unexpectedness judgments), Participant A saw Item 1 in the "Behavior" condition of the Initial Info segment, made a prediction about future behavior, saw an unexpected outcome, and then rated its unexpectedness. On the other hand, Participant B saw almost the same Item 1 variation but instead saw an expected outcome.

Analytic Approach

First, to create by-stimulus confidence and unexpectedness judgments, we removed all judgments that were associated with incorrect predictions. This was again done to ensure that our results were uncontaminated by the possibility of participants' lack of close attention, random responding, or their own expectations being different from the intended expectations of our paradigm. For confidence judgments, this resulted in 1,596/1,984 trials being retained. For unexpectedness judgments, this resulted in 3,356/3,776 trials being retained. Second, for confidence judgments, we averaged across all remaining judgments for each stimulus to get a single estimate per stimulus. For unexpectedness judgments, we first divided the remaining judgments into unexpected outcomes and expected outcomes, and then averaged across judgments for each stimulus to get a single estimate per stimulus per outcome type. Importantly, because predictions were more often incorrect for some stimuli than others, this led to different numbers of observations contributing to each stimulus' estimate, ranging from 16 to 29 observations for each stimulus' confidence judgments, 14 to 30 observations for each stimulus' unexpectedness judgment in the expected condition, and 15 to 30 observations for each stimulus' unexpectedness judgment in the unexpected condition. For confidence judgments, predictions were not differentially incorrect by Initial*Final conditions (i.e., BB, BM, MB, MM): 19%, 20%, 21%, and 18% of each condition's total trials were incorrect, respectively. For unexpectedness judgments, predictions were not differentially incorrect by Expectedness conditions (i.e., E, U): 11% and 11% of each condition's total trials were incorrect, respectively. Third, we divided neural responses from the fMRI experiment into unexpected outcomes and expected outcomes and created by-stimulus PSC estimates by averaging over neural responses for each stimulus within each outcome type. This was done for each ROI separately, as well as for a combined ToM network that averaged across ROIs. Last, we conducted correlation analyses to understand the associations between by-stimulus confidence judgments, by-stimulus unexpectedness judgments, and by-stimulus neural activity. For completeness, we report both Pearson's r and Spearman's ρ to show that the magnitude of any association is not heavily influenced by extreme data points.

Results

Unexpectedness Judgments ~ Confidence Judgments

Surprisingly, there was no relationship between confidence judgments and unexpectedness judgments following an unexpected outcome. However, there was a negative relationship between confidence judgments and unexpectedness judgments following an expected outcome (i.e., more confident predictions were associated with lower ratings of unexpectedness [or, more accurately describing the data, ratings closer to “extremely expected” on the bidirectional scale]). There was no relationship between confidence judgments and the difference in judgments between unexpected and expected outcomes (see SOM Table 2).

PSC ~ Confidence Judgments

Across ROIs, there was no relationship between confidence judgments and neural activity during unexpected outcomes. However, in all ROIs except for RTPJ, there were negative correlations between confidence judgments and neural activity during expected outcomes (i.e., more confident predictions were associated with less neural activity). Additionally, in all ROIs except for RTPJ, there were positive correlations between confidence judgments and the difference in neural activity between unexpected and expected outcomes (i.e., more confident predictions were associated with more discrimination in neural activity between conditions). These patterns held when analyzing neural activity averaged across the entire ToM network (see SOM Table 3).

PSC ~ Unexpectedness Judgments

Unexpectedly, across ROIs, there was no relationship between unexpectedness judgments and neural activity during unexpected outcomes. There was also no relationship between unexpectedness judgments and neural activity during expected outcomes. And there was no relationship between differences in unexpectedness judgments and differences in neural activity between unexpected and expected conditions. These patterns held when analyzing neural activity averaged across the entire ToM network (see SOM Table 4).

SOM Table 2. Correlations between Confidence and Unexpectedness Judgments

	r		*ρ*	
Unexpected	.07		.05	
Expected	-.29	*	-.38	**
Diff (U – E)	.14		.14	

Note. Confidence judgments were measured on a 0 – 100 scale (*Not at all confident* to *Extremely confident*). Unexpectedness judgments were measured on a 0 – 100 scale (*Extremely expected* to *Extremely unexpected*). Therefore, within each Expectedness condition, positive correlations indicate that more confidence is associated with higher judgments of unexpectedness, whereas negative correlations indicate that more confidence is associated with lower judgments of unexpectedness. (Because unexpectedness judgments were on a bipolar scale, another way to interpret the within-condition correlations is that positive correlations indicate that more confidence is associated with lower judgments of expectedness, whereas negative correlations indicate that more confidence is associated with higher judgments of expectedness). Differences in unexpectedness judgments were calculated by subtracting the average item-level unexpectedness judgments following expected outcomes from the average item-level unexpectedness judgments following unexpected outcomes. Therefore, positive correlations indicate that more confidence is associated with more discrimination in unexpectedness judgments between conditions, whereas negative correlations indicate that more confidence is associated with less discrimination in unexpectedness judgments between conditions.

SOM Table 3. Correlations between Confidence Judgments and Final Info Segment PSC

		r		*ρ*	
DMPFC					
	Unexpected	.06		.07	
	Expected	-.30	*	-.27	*
	Diff (U – E)	.25	*	.27	*
PC					
	Unexpected	.01		-.09	
	Expected	-.30	*	-.27	*
	Diff (U – E)	.25	*	.12	
LTPJ					
	Unexpected	.09		.11	
	Expected	-.32	**	-.33	**
	Diff (U – E)	.31	*	.30	*
RTPJ					
	Unexpected	.18		.17	
	Expected	-.13		-.13	
	Diff (U – E)	.23	†	.22	†
ToM Network					
	Unexpected	.12		.12	
	Expected	-.38	**	-.34	**
	Diff (U – E)	.36	**	.35	**

Note. Confidence judgments were measured on a 0 – 100 scale (*Not at all confident* to *Extremely confident*). Therefore, within each Expectedness condition, positive correlations indicate that more confidence is associated with more neural activity, whereas negative correlations indicate that more confidence is associated with less neural activity. Differences in neural activity were calculated by subtracting the average item-level neural activity following expected outcomes from the average item-level neural activity following unexpected outcomes. Therefore, positive correlations indicate that more confidence is associated with more discrimination in neural activity between conditions, whereas negative correlations indicate that more confidence is associated with less discrimination in neural activity between conditions.

SOM Table 4. Correlations between Unexpectedness Judgments and Final Info Segment PSC

		r	*ρ*
DMPFC			
	Unexpected	.05	.11
	Expected	.18	.22 †
	Diff (U – E)	-.05	.01
PC			
	Unexpected	.00	.03
	Expected	-.06	-.09
	Diff (U – E)	-.07	-.07
LTPJ			
	Unexpected	-.02	.04
	Expected	.10	.12
	Diff (U – E)	-.04	-.07
RTPJ			
	Unexpected	-.04	-.02
	Expected	.03	.08
	Diff (U – E)	-.03	.00
ToM Network			
	Unexpected	.00	.06
	Expected	.10	.13
	Diff (U – E)	-.07	-.03

Note. Unexpectedness judgments were measured on a 0 – 100 scale (*Extremely expected* to *Extremely unexpected*). Therefore, within each Expectedness condition, positive correlations indicate that more unexpected outcomes are associated with more neural activity, whereas negative correlations indicate that more unexpected outcomes are associated with less neural activity. Positive difference score correlations indicate that more discrimination between unexpected and expected judgments (with higher unexpected judgments for unexpected than expected outcomes) is associated with more discrimination in neural activity between conditions, whereas negative difference score correlations indicate that more discrimination between unexpected and expected judgments is associated with less discrimination in neural activity between conditions.

Experimental Stimuli

All stimuli are formatted in the following way. The first two sets of sentences correspond to the two different Initial Info conditions (i.e., behavior versus mental state, respectively). The second two sets correspond to the behavior prediction type and the expected versus unexpected (top versus bottom) outcome for Final Info conditions in which the outcome was behavior. The last two sets correspond to the mental prediction type and the expected versus unexpected (top versus bottom) outcome for Final Info conditions in which the outcome was a mental state. As an illustration, each of scenario #1's variations are communicated in parentheses.

1.
Mr. Johnson is packing the family car. He loads a cooler with drinks and looks for his sunglasses. He grabs a new tube of sunscreen and a towel. *(Initial Info = Behavior)*

Mr. Johnson needs to pack the family car. He wants to make sure he has cold drinks and his sunglasses. He is worried he doesn't have enough sunscreen. *(Initial Info = Mental State)*

Where is Mr. Johnson going? *(Prediction about Behavior)*
A. Skiing B. The Beach C. A Museum D. The Movies

Mr. Johnson is going to the beach. *(Final Info = Behavior, Expected)*
Mr. Johnson is going to the movies. *(Final Info = Behavior, Unexpected)*

Where does Mr. Johnson want to go? *(Prediction about Mental State)*
A. Skiing B. The Beach C. A Museum D. The Movies

Mr. Johnson wants to go to the beach. *(Final Info = Mental State, Expected)*
Mr. Johnson wants to go to the movies. *(Final Info = Mental State, Unexpected)*

2.
Mrs. Ellsbury is attending a fancy dinner party at a friend's estate. In preparation, she gets her hair styled and puts on her new white velvet gloves.

Mrs. Ellsbury wants to go to a fancy dinner party at a friend's estate. She is excited to get her hair done and happy to wear her new white velvet gloves.

What will Mrs. Ellsbury order for dinner?
A. Buffalo wings B. Pizza C. Grilled Salmon D. A Cheeseburger

Mrs. Ellsbury orders grilled salmon for dinner.
Mrs. Ellsbury orders a cheeseburger for dinner.

What does Mrs. Ellsbury want for dinner?
A. Buffalo wings B. Pizza C. Grilled Salmon D. A Cheeseburger

Mrs. Ellsbury wants grilled salmon for dinner.
Mrs. Ellsbury wants a cheeseburger for dinner.

3.
Seth is inviting guests over for a party tonight. He picks out his clothes for the evening and starts to get cleaned up before his guests arrive.

Seth wants to have guests over for a party tonight. He is eager to get cleaned up and look his best before his guests arrive.

What will Seth do before the party?
A. Shower B. Sleep C. Exercise D. Paint

Seth takes a shower before the party.
Seth exercises before the party.

What does Seth feel like doing before the party?
A. Shower B. Sleep C. Exercise D. Paint

Seth feels like taking a shower.
Seth feels like exercising.

4.
Olivia is allergic to something in her home. She sneezes every time she pets her dog and when she cleans its bed. She goes to the doctor to get an allergy test.

Olivia thinks she is allergic to something in her home. She feels like sneezing every time she pets her dog. She decides to visit the doctor to get an allergy test.

What will the test results say that Olivia is allergic to?
A. Penicillin B. Pet hair C. Perfume D. Peanuts

The test results say that Olivia is allergic to pet hair.
The test results say that Olivia is allergic to peanuts.

What does Olivia think she is allergic to?
A. Penicillin B. Pet hair C. Perfume D. Peanuts

Olivia thinks she is allergic to pet hair.
Olivia thinks she is allergic to peanuts.

5.
Bob has to catch a 6:30am flight tomorrow. He is a heavy sleeper and usually sleeps for at least 8 hours a night. He always sets several alarms to wake him up.

Bob is worried about catching a 6:30am flight tomorrow. He thinks he needs 8 hours of sleep a night. He thinks it takes several alarms to wake him up.

The night before his flight, when will Bob go to sleep?
A. 3 pm B. 12 am C. 9 pm D. 2 am

Bob goes to sleep at 9 pm.
Bob goes to sleep at 2 am.

The night before his flight, when does Bob want to sleep?
A. 3 pm B. 12 am C. 9 pm D. 2 am

Bob wants to sleep at 9 pm.
Bob wants to sleep at 2 am.

6.
Madeline is visiting the United States. She goes on several tours to experience American culture. She goes to restaurants to try classic American food.

Madeline is visiting the United States. She wants nothing more than to soak up American culture. She is especially excited to try classic American food.

What does Madeline order at the restaurants?
A. Lasagna B. Hamburger C. Falafel D. Sushi

Madeline orders a hamburger at the restaurant.
Madeline orders sushi at the restaurant.

What food does Madeline want to try?
A. Lasagna B. Hamburger C. Falafel D. Sushi

Madeline wants to try a hamburger.
Madeline wants to try sushi.

7.
Ryan is a dentist. He is trying to clean some food out from between the patient's teeth. He brushes the teeth but the food is still stuck.

Ryan is a dentist. He wants to clean some food out from between the patient's teeth. He brushes the teeth but thinks the food is stuck.

What tool does Ryan reach for?
A. Mirror B. Mouthwash C. A drill D. Floss

Ryan reaches for the floss.
Ryan reaches for a mirror.

What tool does Ryan need?
A. Mirror B. Mouthwash C. A drill D. Floss

Ryan needs the floss.
Ryan needs a mirror.

8.
Laura is picking out a card for her friend. She picks up several cards with pictures of cakes and gifts on them. She buys one with a cat holding a balloon.

Laura needs to get a card for her friend. She looks at several cards with pictures of cakes and gifts on them. She likes one with a cat holding a balloon.

What occasion is Laura buying a card for?
A. Birthday B. Funeral C. Thank you D. Get Well Soon

Laura buys a birthday card for her friend.
Laura buys a thank you card for her friend.

What occasion does Laura want a card for?
A. Birthday B. Funeral C. Thank you D. Get Well Soon

Laura wants a birthday card for her friend.
Laura wants a thank you card for her friend.

9.
Mark is taking his girlfriend on a romantic date. He reserves a table at the best Italian restaurant in town. He buys some liquor for them to drink after dinner.

Mark wants to treat his girlfriend to a romantic date. He plans on going to the best Italian restaurant in town. He wants to buy some liquor to enjoy after dinner.

What liquor will Mark buy for after dinner?
A. Bud Light B. Tequila C. Wine D. Red Bull

Mark buys wine for after dinner.
Make buys Bud Light for after dinner.

What liquor does Mark want for after dinner?
A. Bud light B. Tequila C. Wine D. Red Bull

Mark wants wine for after dinner.
Make wants Bud Light for after dinner.

10.
Leland is one of the smartest kids in his class. He answered all the questions on a math test and finished it before any of his classmates. He will get his grade tomorrow.

Leland thinks he is one of the smartest kids in class. He felt good about his math exam and finished it before any of his classmates. He is eager to see his grade tomorrow.

What grade will Leland get on the exam?
A. 55 B. 74 C. 80 D. 95

Leland gets a 95 on the exam.
Leland gets a 74 on the exam.

What grade does Leland think he got?
A. 55 B. 74 C. 80 D. 95

Leland thinks he got a 95 on the exam.
Leland thinks he got a 74 on the exam.

11.
Marta was first to sign up for a 5K race next weekend. She is an experienced runner and competes often. The weather will be very hot and humid on race day.

Marta eagerly signs up for a 5K race next weekend. She loves running and knows everything it takes to compete. She thinks the weather will be very hot and humid on race day.

What clothes will Marta wear to the race?
A. Synthetic t-shirt B. Long-sleeve shirt C. Jeans D. Sweater

Marta wears a synthetic shirt as she runs the race.
Marta wears jeans as she runs the race.

What clothes will Marta want for the race?
A. Synthetic t-shirt B. Long-sleeve shirt C. Jeans D. Sweater

Marta wants a synthetic t-shirt for the race.
Marta wants jeans for the race.

12.
Thomas is trying out for his High School sports team. In preparation, he lifts weights every morning. He has also started eating extra protein so he can bulk up.

Thomas hopes to join his High School sports team. In preparation, he likes to lift weights every morning. He also believes he should eat extra protein so he can bulk up.

What sports team is Thomas trying out for?
A. Cross-country B. Football C. Swimming D. Tennis

Thomas is trying out for the football team.
Thomas is trying out for the swim team.

What sports team does Thomas hope to join?
A. Cross-country B. Football C. Swimming D. Tennis

Thomas hopes to join the football team.
Thomas hopes to join the swim team.

13.
Madison took up the hobby of carpentry years ago. She started with small projects but has now progressed to making much larger pieces.

Madison's favorite hobby is carpentry. Years ago, she enjoyed making small projects but now wants to make much larger pieces.

What piece will Madison make next?
A. Dining room table B. Bird house C. Picture frame D. Wine rack

Madison makes a dining room table.
Madison makes a picture frame.

What piece does Madison want to make?
A. Dining room table B. Bird house C. Picture frame D. Wine rack

Madison wants to make a dining room table.
Madison wants to make a picture frame.

14.
Tim is a member of the Seattle Bicycling Association. He owns both road bikes and mountain bikes. He often chooses to ride his bicycle over driving a car.

Tim hopes to join the Seattle Bicycling Association. He admires the road bikes and mountain bikes he owns. He prefers riding his bicycle over driving a car.

How does Tim usually commute to work?
A. Car B. Scooter C. Jet pack D. Bicycle

Tim uses his bicycle to commute to work.
Tim uses a scooter to commute to work.

How does Tim prefer to commute to work?
A. Car B. Scooter C. Jet pack D. Bicycle

Tim prefers commuting with his bicycle.
Tim prefers commuting with a car.

15.
Billy runs a successful apple orchard in Virginia, but he always has left over apples that he can't sell in the market. He uses these leftover apples to cook with.

Billy loves the apple orchard he runs, but he thinks he has more apples than he can sell in the market. He wants to use the leftover apples to cook with.

What will Billy make with the apples?
A. Chocolate B. Bread C. A Stew D. A Pie

Billy uses the apples to make a pie.
Billy uses the apples to make bread.

What does Billy want to make with the apples?
A. Chocolate B. Bread C. A Stew D. A Pie

Billy wants to make a pie.
Billy wants to make bread.

16.
Maria is buying a musical instrument. She avoids instruments that are loud and obnoxious. She has a harp at home and often buys classical music CDs.

Maria wants to buy a musical instrument. She thinks loud instruments are obnoxious. She loves the harp and wants something that sounds soft and beautiful.

What instrument will Maria buy?
A. Trumpet B. Drum set C. Violin D. Electric guitar

Maria buys a violin.
Maria buys a trumpet.

What instrument will Maria like?
A. Trumpet B. Drum set C. Violin D. Electric guitar

Maria likes the violin.
Maria likes the trumpet.

17.

Mike has an early morning presentation for his boss. On the way to work, he stops by a store to get a drink that will keep him awake through his presentation.

Mike needs to impress his boss at an early morning presentation. He thinks he'll need a drink that will keep him awake through his presentation.

What drink will Mike buy that morning?
A. Root Beer B. Coffee C. Water D. Whiskey

Mike buys coffee in the morning.
Mike buys root beer in the morning.

What drink does Mike need that morning?
A. Root Beer B. Coffee C. Water D. Whiskey

Mike needs coffee in the morning.
Mike needs root beer in the morning.

18.
Maria wakes up early every morning. Her daily routine consists of drinking coffee and reading the newspaper before heading into work.

Maria loves waking up early every morning. Her favorite part of the day is drinking coffee and browsing the news before going into work.

How does Maria start her Tuesday morning?
A. Sleeping in B. Waking up early C. Jogging D. Going to church

Maria wakes up early on Tuesday.
Maria sleeps in on Tuesday.

How does Maria want to start Tuesday morning?
A. Sleeping in B. Waking up early C. Jogging D. Going to church

Maria wants to wake up early.
Maria wants to sleep in.

19.
Leah has a hard time lacing up her ice skates. She puts on a helmet and elbow pads and carefully wobbles over to the ice rink.

Leah is anxious about trying her new ice skates. She wants a helmet and elbow pads and wonders whether she'll fall down a lot.

What will Leah do on the ice?

A. Race B. Tricks C. Hold the side D. Skate backwards

Leah holds the side on the ice.
Leah performs tricks on the ice.

What does Leah want to do on the ice?
A. Race B. Tricks C. Hold the side D. Skate backwards

Leah wants to hold the side.
Leah wants to perform tricks.

20.
Jean is making her lunch. She goes to her garden and picks a ripe tomato, cucumber, and some lettuce. She washes them and cuts them up.

Jean wants to make lunch. She goes to her garden and happily sees a ripe tomato, cucumber, and some lettuce. She eagerly picks them.

What will Jean make for lunch?
A. Pizza B. Salad C. Chicken D. Pasta

Jean makes a salad for lunch.
Jean makes pasta for lunch.

What does Jean want for lunch?
A. Pizza B. Salad C. Chicken D. Pasta

Jean wants a salad for lunch.
Jean wants pasta for lunch.

21.
Joe often makes a hearty breakfast in the morning. He always makes a large meal which usually includes meat along with three eggs.

Joe's favorite meal is a full hearty breakfast. He thinks it is the most important meal of the day and loves having meat and eggs.

What will Joe make for breakfast tomorrow?
A. Bacon omelet B. Coffee B. Toast C. One egg

Joe makes a bacon omelet for breakfast.
Joe makes toast for breakfast tomorrow.

What will Joe decide to eat tomorrow?
A. Bacon omelet B. Coffee B. Toast C. One egg
Joe decides to eat a bacon omelet.

Joe decides to eat toast for breakfast.

22.
Janice is ordering ice cream. She always gets chocolate instead of vanilla, even though her sister never does. Janice usually gets dark chocolate fudge.

Janice is ordering ice cream. She prefers chocolate and can't understand why her sister gets vanilla. Janice's favorite flavor is dark chocolate fudge.

What flavor ice cream will Janice order?
A. Vanilla B. Fudge Brownie C. Mint D. Hazelnut

Janice orders fudge brownie ice cream.
Janice orders hazelnut ice cream.

What flavor ice cream does Janice feel like?
A. Vanilla B. Fudge Brownie C. Mint D. Hazelnut

Janice feels like fudge brownie ice cream.
Janice feels like hazelnut ice cream.

23.
Ally is meeting up with friends for a hike tomorrow. She checks the weather on her phone. It says that it will be cloudy with a chance of rain.

Ally wants to go for a hike with friends tomorrow. She wants to check the weather beforehand. She thinks it will be cloudy and might rain.

What does Ally bring with her for the hike?
A. Sunscreen B. Hat C. T-shirt D. Rain coat

Ally brings a rain coat with her.
Ally brings sunscreen with her.

What does Ally think she needs for the hike?
A. Sunscreen B. Hat C. T-shirt D. Rain coat

Ally thinks she needs a rain coat.
Ally thinks she needs sunscreen.

24.
As summer approaches, Sam gets ready to be outdoors. He buys a tent and sleeping bag. He asks a friend to recommend a national park.

As summer approaches, Sam wants to be outdoors. He thinks it's a good time to try out a new tent and sleeping bag and looks for a national park.

What will Sam do during his summer?
A. School B. Travel abroad C. Camp D. Ride roller coasters

Sam will camp this summer.
Sam will travel abroad this summer.

What does Sam want to do this summer?
A. School B. Travel abroad C. Camp D. Ride roller coasters

Sam wants to camp this summer.
Sam wants to travel abroad this summer.

25.
Tyler avoids store crowds during Christmas. He often skips the sales at stores during the shopping season. He stays at home instead.

Tyler hates store crowds during Christmas. He finds the whole shopping experience overwhelming. He wishes he could stay at home.

Where will Tyler buy his gifts this season?
A. The mall B. Online C. Department store D. Goodwill

Tyler will buy his gifts online.
Tyler will buy his gifts at Goodwill.

Where does Tyler prefer to buy his gifts?
A. The mall B. Online C. Department store D. Goodwill

Tyler prefers to buy his gifts online.
Tyler prefers to buy this gifts at Goodwill.

26.
Kristin's body is achy and tired. She constantly reaches for tissues. She stays in bed all day and watches TV. She makes herself chicken soup for lunch.

Kristin feels achy and tired. She thinks she is running out of tissues. She wants to stay in bed all day and watch TV. She feels like having chicken soup.

What will Kristin do tomorrow?
A. Go to work B. Rest in bed C. Visit friends D. Exercise

Kristin will rest in bed tomorrow.
Kristin will visit friends tomorrow.

What does Kristin want to do tomorrow?
A. Go to work B. Rest in bed C. Visit friends D. Exercise

Kristin wants to rest in bed tomorrow.
Kristin wants to visit friends tomorrow.

27.
Francesca is having her boyfriend's parents over for dinner. She usually makes them simple meals like meat and potatoes. They don't really eat spicy food.

Francesca wants to cook dinner for her boyfriend's parents. She knows they love simple meals like meat and potatoes. She doesn't think they like spicy food.

What does Francesca serve the family for dinner?
A. Steak B. Pad Thai C. Curry D. Chipotle Chicken

Francesca serves steak for dinner.
Francesca serves pad thai for dinner.

What does Francesca think she'll make for dinner?
A. Steak B. Pad Thai C. Curry D. Chipotle Chicken

Francesca thinks she'll make steak.
Francesca thinks she'll make pad thai.

28.
Georgina is an avid gardener. She buys plants that grow well in strong sunlight. She bought a house where the yard gets plenty of sun on its east side.

Georgina wants to start gardening. She believes her new plants will grow well in strong sunlight. She thinks her yard gets plenty of sun on its east side.

On which side of the house will Georgina plant?
A. North B. South C. East D. West

Georgina plants her garden on the east side of her house.
Georgina plants her garden on the west side of her house.

Which side of the house does Georgina plan to garden?
A. North B. South C. East D. West

Georgina plans on gardening on the east side of her house.
Georgina plans on gardening on the west side of her house.

29.

Frederick is shopping for glasses. He stays updated on fashion trends and usually wears smooth-shaped glasses that balance out his square face.

Frederick wants new glasses. He tries to be fashionable and thinks he should wear smooth-shaped glasses because he hopes to balance out his square face.

What shape glasses will Frederick buy?
A. Rectangular B. Square C. Star-shaped D. Oval

Frederick buys oval shaped glasses.
Frederick buys rectangular glasses.

What shape glasses does Frederick want?
A. Rectangular B. Square C. Star-shaped D. Oval

Frederick wants oval shaped glasses.
Frederick wants rectangular glasses.

30.
Darrell is having trouble sleeping. He has tried several different pillows to improve his sleep. Many pillows he uses are too soft and squishy.

Darrell feels like he can't sleep. He believes that his pillow is the cause. He thinks his current pillow is too soft and wants something firmer.

What kind of pillow will Darrell buy?
A. Memory foam B. Cotton C. Firm foam D. Feathers

Darrell buys a pillow filled with firm foam.
Darrell buys a pillow filled with feathers.

What kind of pillow will Darrell like?
A. Memory foam B. Cotton C. Firm foam D. Feathers

Darrell likes a pillow filled with firm foam.
Darrell likes a pillow filled with feathers.

31.
Trish never sits on the beach with her friends. She doesn't sit in one place for very long. Instead, she signs up for watersport lessons. She tries to do stunts.

Trish finds sitting on the beach boring and wants to learn a watersport to keep her occupied. She would like to learn something fast-paced with stunts.

Which water sport will Trish participate in?
A. Kayaking B. Water skiing C. Snorkeling D. Tubing

Trish participates in water skiing.
Trish participates in snorkeling.

Which water sport does Trish want to learn?
A. Kayaking B. Water skiing C. Snorkeling D. Tubing

Trish wants to learn water skiing.
Trish wants to learn snorkeling.

32.
Grace's students perform well because she has all the latest technology in her classroom. Each year she updates all the old materials in her classroom.

Grace believes that students learn best when classrooms are equipped with the latest technology. She would like to keep her classroom updated.

What will Grace buy for her classroom next?
A. Computer tablets B. Crayons C. A Chalk board D. Poster boards

Grace will buy computer tablets for her class.
Grace will buy a new chalk board for her class.

What does Grace want for her classroom?
A. Computer tablets B. Crayons C. A Chalk board D. Poster boards

Grace wants computer tablets for her class.
Grace wants a new chalk board for her class.

33.
Nick constantly takes videos of his cat with his phone. He makes several videos a day. His phone quickly runs out of hard drive space so he buys a new phone.

Nick loves taking fun videos of his cat with his phone. He likes to make several videos a day but his phone runs out of hard drive space. He wants a new phone.

What feature will Nick add to his phone?
A. Better Wi-Fi B. Bluetooth C. More apps D. Larger hard drive

Nick adds a larger hard drive to his phone.
Nick adds Bluetooth capability to his phone.

What feature does Nick want in his new phone?
A. Better Wi-Fi B. Bluetooth C. More apps D. Larger hard drive

Nick wants a larger hard drive on his phone.

Nick wants Bluetooth capability on his phone.

34.
Alicia does not usually go to amusement parks. She easily gets motion sickness on intense rides. Going on fast rides makes her throw up and ruins her day.

Alicia is dreading going to the amusement park. She believes that she will get motion sickness if she goes on intense rides. She doesn't want to ruin her day.

What attraction will Alicia avoid going on?
A. The go-karts B. The Ferris wheel C. The rollercoaster D. The 3D movie

Alicia will avoid going on the roller coaster.
Alicia will avoid going on the Ferris wheel.

What attraction does Alicia want to avoid?
A. Boat ride B. Ferris wheel C. Roller coaster D. The 3D movie

Alicia wants to avoid the roller coaster.
Alicia wants to avoid the Ferris wheel.

35.
Jessie lives alone. She goes to the pet store to buy a pet that will keep her company. She looks for pets than she can cuddle with on the couch.

Jessie is tired of living alone. She thinks she needs a pet that will help keep her company. She enjoys pets that she can cuddle with on the couch.

What kind of pet will Jessie buy?
A. Dog B. Snake C. Hamster D. Parrot

Jessie buys a dog as a pet.
Jessie buys a hamster as a pet.

What kind of pet does Jessie want?
A. Dog B. Snake C. Hamster D. Parrot

Jessie wants a dog for a pet.
Jessie wants a hamster for a pet.

36.
Tommy competes in the annual sailing regatta. Last year, his boat lagged behind the others. He tried adjusting the sail but it did not catch enough wind.

Tommy wants to win the annual sailing regatta. Last year, he was unhappy with how he finished. He thinks the sail did not catch enough wind.

What will Tommy change about his boat?
A. New hull B. Larger sail C. Bigger rudder D. New paint

Tommy will get his boat a larger sail.
Tommy will get his boat a new hull.

What does Tommy think his boat needs?
A. New hull B. Larger sail C. Bigger rudder D. New paint

Tommy thinks his boat needs a larger sail.
Tommy thinks his boat needs a new hull.

37.
Craig is applying for a job. He quit his office job and is applying for jobs that utilize his creative talents. He is good with fine detailed work.

Craig wants a new job. He hated his last office job and wants to do something that is more creative. He likes fine detailed work.

Where will Craig apply for a job?
A. Bank of America B. Post office C. Graphic design firm D. Wall Street

Craig applies for a job at a graphic design firm.
Craig applies for a job at the post office.

Where is Craig thinking about applying?
A. Bank of America B. Post office C. Graphic design firm D. Wall Street

Craig is thinking about a job at a graphic design firm.
Craig is thinking about a job at the post office.

38.
Sally is giving a presentation at a regional business meeting. Her presentations always use bright vibrant colors. She makes all her slides standout.

Sally is thinking about her presentation for a regional business meeting. She thinks it needs bright vibrant colors to make her slides standout.

What color will Sally use in her presentation?
A. Navy blue B. Bright red C. Forest green D. Steel gray

Sally uses bright red in her presentation.
Sally uses steel gray in her presentation.

What color is Sally thinking of using?

A. Navy blue B. Brick red C. Forest green D. Steel gray

Sally is thinking of using bright red.
Sally is thinking of using steel gray.

39.
Allison is going on vacation. She keeps her vacation budgets very small. She looks up cheap vacation spots that are cold and mountainous.

Allison wants to go on vacation but she doesn't want to spend much money. She thinks she can find a nice spot that is cold and mountainous.

Where will Allison travel on vacation?
A. Miami B. The Appalachian Trail C. Las Vegas D. New York

Allison hikes the Appalachian Trail on vacation.
Allison travels to Las Vegas, Nevada on vacation.

Where is Allison thinking of traveling?
A. Miami B. The Appalachian Trail C. Las Vegas D. New York

Allison is thinking of hiking the Appalachian Trail.
Allison is thinking about going to Las Vegas.

40.
Hank buys tickets whenever NASCAR is in town. He works at a car repair shop. He reads up on the latest car parts and models.

Hank loves NASCAR. He also enjoys working at a car repair shop and thinks he should keep up-to-date on the latest car news.

Which magazine will Hank subscribe to?
A. Cosmopolitan B. Time C. Car and Driver D. Scientific American

Hank subscribes to Car and Driver magazine.
Hank subscribes to Time magazine.

What is Hank's favorite magazine?
A. Cosmopolitan B. Time C. Car and Driver D. Scientific American

Hank loves Car and Driver magazine.
Hank loves Time magazine.

41.
Scott is getting ready for an office Christmas party. The parties in the past have been very formal. Scott's boss will be attending and wearing a suit.

Scott is planning what to wear for his office Christmas party. Scott thinks the parties have been formal and wants to look good in front of his boss.

What shoes will Scott wear to the party?
A. Dress shoes B. Flip-flops C. Cowboy boots D. Old sneakers

Scott wears dress shoes at the party.
Scott wears old sneakers at the party.

What shoes does Scott want to wear to the party?
A. Dress shoes B. Flip-flops C. Cowboy boots D. Old sneakers

Scott wants to wear dress shoes to the party.
Scott wants to wear old sneakers to the party.

42.
Teddy hunts with his uncle to spend quality time together. Today he thinks that his uncle is going to teach him how to track wild game. They hunted for large deer the last couple times they met and today Teddy believes that his uncle will want to hunt something different.

Teddy has never fired a gun before. He goes to the shooting range to practice. He is not very strong. A forceful blast would cause him to fall over.

Teddy wants to learn how to shoot a gun so he goes to the shooting range. He thinks he's weak and is anxious that a forceful blast will knock him over.

What gun will Teddy start shooting with?
A. Magnum B. Shotgun C. Assault rifle D. Pistol

Teddy starts shooting with a pistol.
Teddy starts shooting with a shotgun.

What gun does Teddy intend to shoot?
A. Magnum B. Shotgun C. Assault rifle D. Pistol

Teddy intends to shoot a pistol.
Teddy intends to shoot a shotgun.

43.
Justin keeps his teeth incredibly clean. He rarely misses brushing his teeth at night, but he arrives home very late after a movie and falls straight to sleep.

Justin hates when his teeth are dirty. He feels gross when he doesn't brush his teeth, but he arrives home very late after a movie and falls straight to sleep.

What will Justin do first thing in the morning?
A. Go jogging B. Brush his teeth C. Laundry D. Read a book

Justin brushes his teeth when he wakes up.
Justin reads a book when he wakes up.

What will Justin want to do first in the morning?
A. Go jogging B. Brush his teeth C. Laundry D. Read a book

Justin wants to brush his teeth when he wakes up.
Justin wants to read a book when he wakes up.

44.
Meg takes a class on French baking. She practices making dozens of pastries each day. When the class ends, she cooks a large brunch for her friends.

Meg enjoys taking a class on French baking. She thinks she is great at making pastries. When the class ends, she wants to make brunch for her friends.

What does Meg cook for her friends?
A. Croissants B. Bacon C. Churros D. Grits

Meg makes croissants for her friends.
Meg makes bacon for her friends.

What does Meg plan on making?
A. Croissants B. Bacon C. Churros D. Grits

Meg plans on making croissants.
Meg plans on making bacon.

45.
John joined the orchestra after college. He has been playing stringed instruments for twenty years. He is starting to teach private lessons.

John decided to join the orchestra after college. He knows he is skilled at playing stringed instruments. He wants to teach private lessons.

What instrument will John offer lessons for?
A. Clarinet B. Drums C. Xylophone D. Violin

John offers private violin lessons.
John offers private clarinet lessons.

What instrument does John want to teach?
A. Clarinet B. Drums C. Xylophone D. Violin

John wants to teach the violin.
John wants to teach the clarinet.

46.
Susan is moving next week. Her closet is full of vintage dresses. She frequently goes to estate sales to look for valuable old designer clothing.

Susan is worried about moving. She is thinking about her closet full of vintage dresses. She finds many of her old designer clothing very valuable.

Which items will Susan pack most carefully?
A. Her books B. Her furniture C. Her dresses D. Her jackets

Susan packs her dresses the most carefully.
Susan packs her books the most carefully.

Which items are Susan concerned about moving?
A. Her books B. Her furniture C. Her dresses D. Her jackets

Susan is concerned about packing her dresses.
Susan is concerned about packing her books.

47.
Frank works at the arboretum. He checks the summer weather. It will be unusually wet. Frank always has to deal with harmful bugs after the rain.

Frank works at the arboretum. He realizes the weather this summer will be unusually wet and he figures that harmful bugs will come after the rain.

What will Frank get to protect the trees this summer?
A. Pesticide B. Sprinkler systems C. Warning signs D. Trunk cages

Frank gets pesticide for the trees this summer.
Frank gets sprinklers for the trees this summer.

What does Frank think the trees need this summer?
A. Pesticide B. Sprinkler systems C. Warning signs D. Trunk cages

Frank thinks the trees need pesticide.
Frank thinks the trees need sprinklers.

48.
When her friends visit, Abby usually shows them what her city has to offer. She takes them to eat at fancy restaurants and goes dancing at local clubs.

When her friends visit, Abby wants to show them what her city has to offer. She thinks they will like fancy restaurants and have fun dancing at local clubs.

Where will Abby take her friends out to?
A. Downtown B. The park C. The beach D. The amusement park

Abby takes her friends downtown.
Abby takes her friends to the park.

Where will Abby decide to take her friends?
A. Downtown B. The park C. The beach D. The amusement park

Abby decides to take her friends downtown.
Abby decides to take her friends to the park.

49.
Sean bought a new kitten. He does not get it declawed. When he comes home from work he often finds scratch marks on his couch and chairs.

Sean loves his new kitten. He thinks it's cruel to declaw cats, but when he comes home from work he notices scratches on his couch and chairs.

What will Sean buy for his kitten?
A. Litter box B. Scratching post C. Food bowl D. New bed

Sean gets his kitten a scratching post.
Sean gets his kitten a new food bowl.

What does Sean think his cat needs?
A. Litter box B. Scratching post C. Food bowl D. New bed

Sean thinks his kitten needs a scratching post.
Sean thinks his kitten needs a new food bowl.

50.
David laughs as often as possible. He avoids sadness at all cost and never puts himself in situations that are intense and emotional.

David believes that people should laugh more. He feels uncomfortable around sadness and wants to avoid intense situations.

What kind of movies does David go to?
A. Dramas B. Horror C. Comedies D. Thrillers

David goes to comedies at the movies.
David goes to thrillers at the movies.

What kind of movies does David like?
A. Dramas B. Horror C. Comedies D. Thrillers

David likes to see comedies.
David likes to see thrillers.

51.
Felicia works at NASA. She bought herself a telescope and the entire Star Trek series for her birthday. Most weekends she plays computer games.

Felicia wants a tour of NASA and a telescope for her birthday. Her favorite TV series is Star Trek. She likes to spend her weekends playing computer games.

What genre of books does Felicia collect?
A. Science Fiction B. Romance C. Mystery D. Short stories

Felicia collects science fiction books.
Felicia collects mystery novels.

What book genre does Felicia think is the best?
A. Science Fiction B. Romance C. Mystery D. Short stories

Felicia thinks science fiction is the best.
Felicia thinks mystery books are the best.

52.
Luca keeps his pizzas simple. He never eats overcomplicated pizzas that have a million ingredients. He usually adds just one or two toppings.

Luca likes to keep his pizza simple. He doesn't understand people who complicate things with a million ingredients. He prefers one or two per pizza.

What kind of pizza does Luca order?
A. Hawaiian B. Tomato and basil C. Meat lover's D. Stuffed crust

Luca orders a tomato and basil pizza.
Luca orders a meat lover's pizza.

What kind of pizza does Luca prefer?
A. Hawaiian B. Tomato and basil C. Meat lover's D. Stuffed crust

Luca prefers tomato and basil pizza.
Luca prefers meat lover's pizza.

53.
Owen runs a large grocery store. He holds employees to strict standards. He yells at employees when they are not on-task. He watches everyone closely.

Owen runs a large grocery store. He wants his employees to stay on-task and gets angry when they do not meet his strict standards.

What does Owen make his employees do?
A. Take vacations B. Wear uniforms C. Use cellphones D. Take lunch breaks

Owen makes his employees wear uniforms.
Owen makes his employees take lunch breaks.

What does Owen prefer his employees do?
A. Take vacations B. Wear uniforms C. Use cellphones D. Take lunch breaks

Owen prefers that his employees wear uniforms.
Owen prefers that his employees take lunch breaks.

54.
Robert just finished High School. His lowest grade was in math. He never took an advanced science class. He writes often and published a short story.

Robert just finished High School. He didn't understand math very well and never liked his science classes. He loves to write as often as he can.

What will Robert choose as a major in college?
A. English B. Economics C. Physics D. Chemistry

Robert chooses English as his major.
Robert chooses Economics as his major.

What does Robert want to major in for college?
A. English B. Economics C. Physics D. Chemistry

Robert wants to be an English major.
Robert wants to be an Economics major.

55.

Rebecca quit her last corporate job. She argued daily with businessmen who took advantage of people. She volunteers at a homeless shelter every Sunday.

Rebecca hated working at a corporate job. She thought businessmen were taking advantage of people. She loves volunteering at a homeless shelter.

What does Rebecca do for a living now?
A. Banker B. Consultant C. Corporate Lawyer D. Social worker

Rebecca has a job as a social worker.
Rebecca has a job as a banker.

What job does Rebecca want to have now?
A. Banker B. Consultant C. Computer programmer D. Social worker

Rebecca wants to be a social worker.
Rebecca wants to be a banker.

56.
Lily bought a camera with a very fast shutter speed. She takes pictures of fast moving action. She brings a telephoto lens to capture people far away.

Lily needs a camera with a fast shutter speed. She likes taking pictures of fast moving action. She prefers using a telephoto lens to capture people far away.

What event does Lily usually photograph?
A. Sports games B. Weddings C. Landscapes D. Portraits

Lily usually photographs sports games.
Lily usually photographs weddings.

What event does Lily like photographing?
A. Sports games B. Weddings C. Landscapes D. Portraits

Lily likes to photograph sports games.
Lily likes to photograph weddings.

57.
Michelle breaks in a new pair of pointe shoes. She buys tights and a tutu. She practices every day for the upcoming audition for the Nutcracker.

Michelle hopes her new pair of pointe shoes fit. She almost forgets her tights and tutu. She hopes she nails her audition for the Nutcracker.

What kind of dance does Michelle often perform?
A) Jazz B) Ballet C) Hip hop D) Salsa

Michelle often performs ballet.
Michelle often performs hip hop.

What kind of dance does Michelle like performing?
A) Jazz B) Ballet C) Hip hop D) Salsa

Michelle likes performing ballet.
Michelle likes performing hip hop.

58.
Jordan makes a list of his work experience. He asks three previous employers if he can list them as references. He reads the classified section for open positions.

Jordan tries to remember his past work experience. He hopes he can list his previous employers as references. He likes the positions he sees in the classified section.

What is Jordan working on?
A. Term paper B. Diary C. Resume D. Grocery list

Jordan is working on his resume.
Jordan is working on his diary.

What is Jordan trying to write?
A. Term paper B. Diary C. Resume D. Grocery list

Jordan is trying to write a resume.
Jordan is trying to write his diary.

59.
Bethany is getting ready for a party. She finds the rattle that she bought and wraps it in pink paper. She also brings a bag of hand-me-down clothes.

Bethany is going to a party. She looks for the rattle she bought and decides to wrap it in pink paper. She wants to bring a bag of hand-me-down clothes.

What kind of party is Bethany attending?
A) Bachelorette B) Retirement C) Anniversary D) Baby Shower

Bethany is attending a baby shower.
Bethany is attending a bachelorette party.

What kind of party does Bethany think it is?
A) Bachelorette B) Retirement C) Anniversary D) Baby Shower

Bethany thinks it's a baby shower.

Bethany thinks it's a bachelorette party.

60.
Justin has lived with his girlfriend for four years. He goes to the mall to look at rings and buys one. He brings her to the spot where they first met.

Justin has loved his girlfriend for four years. He decides on a ring he likes at the mall and buys it. He wants to take her to the spot where they first met.

What will Justin ask his partner to do?
A) Cook for him B) Marry him C) Sleep with him D) Hire him

Justin asks his partner to marry him.
Justin asks his partner to sleep with him.

What does Justin want his partner to do?
A) Cook for him B) Marry him C) Sleep with him D) Hire him

Justin wants his partner to marry him.
Justin wants his partner to sleep with him.

61.
Elizabeth eats sugary candy all day. Now she has trouble chewing. She puts ice packs on her jaw to ease the pain. She avoids very hot or cold liquids.

Elizabeth loves eating sugary candy. Now she feels a sharp pain when she tries to chew. She likes to keep ice on her jaw to try to soothe the pain.

What work does Elizabeth get done at the dentist?
A) Gum surgery B) Whitening C) Cavity filling D) Braces

Elizabeth gets a cavity filled.
Elizabeth gets braces put on.

What work does Elizabeth want at the dentist?
A) Gum surgery B) Whitening C) Cavity filling D) Braces

Elizabeth wants a cavity filled.
Elizabeth wants to get braces.

62.
Karina is getting ready for a wedding. She places an order for flowers. She picks out vases in the right colors and then changes the table centerpieces.

Karina is thinking about a wedding. She knows what flowers she wants to order. She looks for vases in the right colors and wants new centerpieces.

What role is Karina doing at the wedding?
A) The caterer B) The decorator C) An usher D) A bridesmaid

Karina is the decorator at the wedding.
Karina is a bridesmaid at the wedding.

What role is Karina thinking about doing?
A) The caterer B) The decorator C) An usher D) A bridesmaid

Karina is thinking about being the decorator.
Karina is thinking about being a bridesmaid.

63.
Kelly keeps a strict budget. She saves every penny and rarely buys luxury items. She saves coupons and looks for bargain deals.

Kelly keeps a strict budget. She wants to save every penny and dislikes luxury items. She loves coupons and discovering bargain deals.

Where does Kelly usually buy clothes?
A) Thrift stores B) Boutique shops C) Gucci D) Nieman Marcus

Kelly usually buys her clothes at thrift stores.
Kelly usually buys her clothes at Gucci.

Where does Kelly like to buy clothes?
A) Thrift stores B) Boutique shops C) Gucci D) Nieman Marcus

Kelly likes to buy clothes at thrift stores.
Kelly likes to buy her clothes at Gucci.

64.
Alex reads books about environmentalism. He made a compost and recycles. He limits his green house gas emissions as much as possible.

Alex knows a lot about environmentalism. He thinks people should compost and recycle. He wants to eliminate green house gas emissions.

What kind of motor vehicle will Alex purchase?
A) Minivan B) Motorcycle C) Electric D) Hummer

Alex purchases an electric car.
Alex purchases a new minivan.

What kind of vehicle does Alex want to purchase?

A) Minivan B) Motorcycle C) Electric D) Hummer

Alex wants an electric car.
Alex wants a new minivan.

~~Additional References~~

~~Chuang & Yun-Shiuan (2020). *label4MRI: MRI labeling.*~~

~~Rolls, E.T., Huang, C., Lin, C., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *NeuroImage, 206*, 116189.~~