CORTEX 71 (2015) 306–322

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Neural networks underlying the metacognitive uncertainty response

Erick J. Paul [a,*], *J. David Smith* [b], *Vivian V. Valentin* [c], *Benjamin O. Turner* [c], *Aron K. Barbey* [a] *and F. Gregory Ashby* [c,**]

[a] *The Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, USA*
[b] *Department of Psychology and Center for Cognitive Science, The University at Buffalo, State University of New York, USA*
[c] *Department of Psychological & Brain Sciences, University of California, Santa Barbara, USA*

ARTICLE INFO

Article history:
Received 16 December 2014
Reviewed 24 February 2015
Revised 19 May 2015
Accepted 20 July 2015
Action editor Gui Xue
Published online 1 August 2015

Keywords:
Metacognition
Categorization
Uncertainty response
Neuroimaging
Psychophysics

ABSTRACT

Humans monitor states of uncertainty that can guide decision-making. These uncertain states are evident behaviorally when humans decline to make a categorization response. Such behavioral uncertainty responses (URs) have also defined the search for metacognition in animals. While a plethora of neuroimaging studies have focused on uncertainty, the brain systems supporting a volitional strategy shift under uncertainty have not been distinguished from those observed in making introspective post-hoc reports of categorization uncertainty. Using rapid event-related fMRI, we demonstrate that the neural activity patterns elicited by humans' URs are qualitatively different from those recruited by associative processes during categorization. Participants performed a one-dimensional perceptual-categorization task in which an uncertainty-response option let them decline to make a categorization response. Uncertainty responding activated a distributed network including prefrontal cortex (PFC), anterior and posterior cingulate cortex (ACC, PCC), anterior insula, and posterior parietal areas; importantly, these regions were distinct from those whose activity was modulated by task difficulty. Generally, our results can be characterized as a large-scale cognitive control network including recently evolved brain regions such as the anterior dorsolateral and medial PFC. A metacognitive theory would view the UR as a deliberate behavioral adjustment rather than just a learned middle category response, and predicts this pattern of results. These neuroimaging results bolster previous behavioral findings, which suggested that different cognitive processes underlie responses due to associative learning versus the declaration of uncertainty. We conclude that the UR represents an elemental behavioral index of metacognition.

* *Corresponding author*. Beckman Institute, 405 N Mathews Ave, MC-251, Urbana, IL 61801-2325, USA.
** *Corresponding author*. Department of Psychological & Brain Sciences, University of California, Santa Barbra, CA 93106, USA.
E-mail addresses: ejpaul@illinois.edu (E.J. Paul), ashby@psych.ucsb.edu (F.G. Ashby).

http://dx.doi.org/10.1016/j.cortex.2015.07.028
0010-9452/

1. Introduction

When the cost of an error is high and the correct response is uncertain, an adaptive choice is to pass rather than guess. Such conscious control of behavior in response to the subjective feeling of uncertainty is a hallmark of metacognition (e.g., Dunlosky & Bjork, 2008; Flavell, 1979; Koriat & Goldsmith, 1994; Metcalfe & Shimamura, 1994; Nelson, 1992; Schwartz, 1994), and is a crucial cognitive capacity that affects every aspect of adaptive human behavior.

Much behavioral research has studied the adaptive nature of the uncertainty response (UR)—sometimes called an "opt-out" response—in which the primary task response may be declined to avoid a (presumably low-confidence) decision. The UR is a simultaneous perceptual and confidence judgment, leveraging uncertainty to modify behavior during a primary perceptual task. In the present study, we adopted a simple perceptual categorization paradigm wherein participants made sparse versus dense categorization judgments about the number of illuminated pixels in a circle. Critically, participants were also allowed to opt out of any trial (i.e., avoid forced, low-confidence categorization judgments) by using the UR.

There are two clear advantages afforded by this paradigm choice. First, it offers the ability to directly compare perceptual categorization responses with a volitional, adaptive opt-out response. Although the brain structures and functions underlying any confidence judgment need not necessarily differ, this paradigm allows us to assess particular brain structures related to a confidence judgment that is implicit, prospective, and adaptive with respect to the primary task goal (i.e., maximizing performance by avoiding errors). Previous research has attributed activity in posterior parietal cortex, dorsal anterior cingulate cortex (dACC), and both anterior and dorsolateral prefrontal cortex (PFC) to explicit, post-decisional, retrospective (not behaviorally adaptive) confidence judgments (Fleming, Huijgen, & Dolan, 2012; Hilgenstock, Weiss, & Witte, 2014; Rounis, Maniscalco, Rothwell, Passingham, & Lau, 2010; Yokoyama et al., 2010). Thus, the present research is positioned to complement those findings.

Second, this paradigm addresses an actively debated controversy regarding the putative metacognitive nature of the UR. The field of comparative psychology has used UR paradigms in order to infer higher-level cognitive capacity in animals through behavioral comparisons to humans. Humans actively report that URs are prompted by conscious control over uncertainty; unfortunately, comparative psychology is at a disadvantage because animals cannot report such feelings. Proponents of a metacognitive theory of uncertainty monitoring have demonstrated remarkable behavioral isomorphisms between humans and animals in UR-based paradigms (Kornell, Son, & Terrace, 2007; Smith, 2009; Smith, Beran, & Couchman, 2012), which supports the view that animals possess putative metacognitive abilities. For example, in an effort to dissociate reinforcement signals with individual stimuli, Smith, Beran, Redford, and Washburn (2006) observed that monkeys continue to adaptively and strategically use the UR when feedback is deferred to the end of multi-trial blocks of stimuli. Using an information-seeking metacognitive paradigm, Basile, Schroeder, Brown, Templer, and Hampton (2015) tested and subsequently discounted seven alternative explanations for monkeys' performance, and instead argued that their empirical evidence is wholly consistent with a metacognitive account of the animals' performance.

Opposing the view that non-human primates possess metacognitive capacity, some argue that the UR is simply a first-order "middle-category response", assigned to stimuli near the category bound (Carruthers, 2008; Jozefowiez, Staddon, & Cerutti, 2009). In the sparse versus dense categorization task used in the present investigation, this "middle-region" essentially divides the category space into three regions—sparse, medium-dense, and dense—separated by two category bounds. In such a scenario, stimuli are associated with the "middle-region" response through the same reinforcement and conditioning mechanisms that guide learning and performance for the sparse and dense stimulus regions. More recent arguments for the low-level, associative nature of the UR seek to account for the apparent metacognitive behavior of animals through reinforcement learning models. These models capture the idea that animals are simply conditioned to respond uncertain to the most difficult stimuli—i.e., those most likely to produce errors (Le Pelley, 2012, 2014). At its core, this theory is argued as being a more parsimonious account of animals' behavior, eschewing the attribution of UR use to metacognition in animals. This alternative to the metacognitive UR explanation necessarily predicts that the same brain structures supporting associative, reinforcement learning processes also underlie the UR. Note that despite the fact that humans appeal to feelings of uncertainty in using the UR, the associative reinforcement account of the UR is readily applicable to humans as an explanatory mechanism for their observed behavior.

There is empirical evidence for and against each account, and the theoretical debate is active (Smith, Couchman, & Beran, 2014a, 2014b). However, neuroscience evidence seeking to disentangle the UR from competing associative accounts is sparse, and two of those few extant studies found conflicting results. Komura, Nikkuni, Hirashima, Uetake, and Miyamoto (2013) found support for dissociable UR brain regions (by selectively affecting UR with muscimol injections), but this result was contested by Kiani and Shalden (2009), who showed that the same neurons in parietal cortex both accumulate evidence in service of a categorization response and trigger a UR when uncertainty is high. Behaviorally, humans and macaques both use the UR in similar ways (Smith et al., 2006; Smith, Coutinho, Church, & Beran, 2013; Zakrzewski, Coutinho, Boomer, Church, & Smith, 2014), suggesting similar underlying function and homologous brain structures. Partly for this reason, we hypothesize that recently evolved networks—and not reinforcement learning or associative learning networks—underlie the UR.

In addition to speaking to this debate, our design offers the ability to elucidate the component processes involved in using the UR. According to the metacognitive theory, the UR is a

complex behavioral adjustment that involves both avoiding errors (uncertainty monitoring) and declining the primary task (task-set switching). Uncertainty monitoring recruits regions that include medial PFC/dorsolateral PFC (DLPFC), posterior parietal cortex, ACC and posterior cingulate cortex (PCC), and anterior insula (AI) (Bach, Hulme, Penny, & Dolan, 2011; Carter et al., 1998; Fleck, Daselaar, Dobbins, & Cabeza, 2006; Grinband, Hirsch, & Ferrera, 2006; Huettel, Song, & McCarthy, 2005; Kable & Glimcher, 2007; MacDonald, Cohen, Stenger, & Carter, 2000; McCoy & Platt, 2005; Platt & Huettel, 2008; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Stern, Gonzalez, Welsh, & Taylor, 2010; Ullsperger, Harsay, Wessel, & Ridderinkhof, 2010; Volz, Schubotz, & von Cramon, 2004). A role in selecting from conflicting action plans has been attributed to dorsal medial frontal cortex (Taylor, Nobre, & Rushworth, 2007) and DLPFC (Cohen et al., 1997; Fleck et al., 2006; MacDonald et al., 2000; Ridderinkhof et al., 2004), so these areas may be involved in choosing the UR. Observing activation during URs in these areas, as well as areas that have been previously associated with retrospective categorization uncertainty judgments, would therefore be in line with the predictions of the metacognitive account.

The structure of results within our task can also serve to disentangle competing accounts of the UR. Choosing the UR is most adaptive near the sparse–dense category bound where categorization difficulty is highest and the probability of a correct response is lowest. However, the UR is necessarily a subjective judgment made when an individual stimulus is deemed too difficult to categorize (e.g., due to perceptual noise), and can occur for stimuli far from the bound as well. The metacognitive account predicts that the neural profile of URs far from the bound will be functionally identical to those made close to the bound because the brain structures driving the decision to decline categorization should be the same regardless of the source of the increased categorization difficulty (e.g., noise from eye-fatigue versus low categorization confidence). That is, conditioned on a UR having occurred, a fixed set of brain regions (distinct from those involved in categorization) is expected to be activated, regardless of difficulty. Although there may nonetheless be gradations difficulty effects, due to gradations in these extra-categorizational brain regions, or mixing effects from multiple processes operating within a single trial, the form of any such effects should be distinct from pure categorization trials. The alternative "middle-category" associative learning account would make a different prediction, because all responses would essentially be categorization responses. Any response in the middle of the "medium-dense" category region (i.e., near the sparse–dense category bound) corresponds to the highest middle-category certainty, making the choice of a "medium-dense" response easier than near the edges of the "medium-dense" region. Therefore, according to the middle-category account, during a "medium-dense-category response" (UR) brain activation intensity should vary with distance-to-bound to exactly the same extent as "regular" categorization trials. Our investigation is thus positioned to address a theoretical controversy, and at the same time contribute to the cognitive neuroscience literature of metacognition and decision making by uncovering the neural correlates of a strategic and adaptive response to uncertainty.

2. Method

2.1. Participants

Thirty-two undergraduates from the University of California, Santa Barbara community were recruited for the imaging experiment after completing a prescreening laboratory experiment as described below. Three participants were unable to complete the entire imaging experiment leaving 29 (14 female and 15 male) participants. All participants had normal or corrected-to-normal vision, and reported no previous neurological injuries or disorders. Participants received course credit for the laboratory session and monetary compensation for their participation in the fMRI session. Prior to data collection, the experiment was reviewed and approved by the UCSB Institutional Review Board.

2.2. Stimuli

The stimuli were circular pixel patterns adapted from Smith et al. (2006). Each stimulus was a black disc subtending about 5° of visual angle presented on a black background with a pre-determined proportion of random pixels illuminated white. Fig. 1 shows four sample stimuli. There were 41 possible stimulus levels and each level had 1.8 percent more pixels illuminated than the previous. The proportion of illuminated pixels in the disc was calculated using the formula: proportion $= .10 \times 1.018$ level (rounded down to the nearest integer). The proportion of pixels illuminated white in the disc stimuli ranged from .1018 (Level 1) to .2078 (Level 41).

Stimulus Levels 1–20 were defined as "sparse" and deserved the Sparse response. Stimulus Levels 22–41 were defined as "dense" and deserved the Dense response. Stimulus Level 21 was never shown to participants and served as the separating boundary between the lower 20 Sparse levels and the upper 20 Dense levels. Nonetheless, the Sparse and Dense trial levels nearest Level 21 (the Sparse–Dense discrimination breakpoint of the task) were most difficult to discriminate as Sparse or Dense for participants, and these trials levels were expected to receive the most Uncertain responses. The number of stimuli sampled from any particular level was calculated using a mirrored, truncated geometric distribution: $G(p = .15, k = 20)$. A histogram of the distribution of stimuli appears in Fig. 2; 498 stimuli were sampled across all 41 observed stimulus levels.

2.3. Apparatus

The experiment was controlled using custom MATLAB scripts and functions from Brainard (1997) Psychophysics Toolbox. During laboratory sessions, participants made Sparse and Dense responses with one hand on keyboard keys specially labeled "S" and "D". They made Uncertain responses with the opposite hand on a key specially labeled "?". Specifically, participants in the laboratory session were assigned to make Uncertain responses either with the left hand or with the right hand, and Sparse/Dense responses with the other hand. Participants selected for the scanning session maintained the same left/right hand-to-button assignment in the scanner.

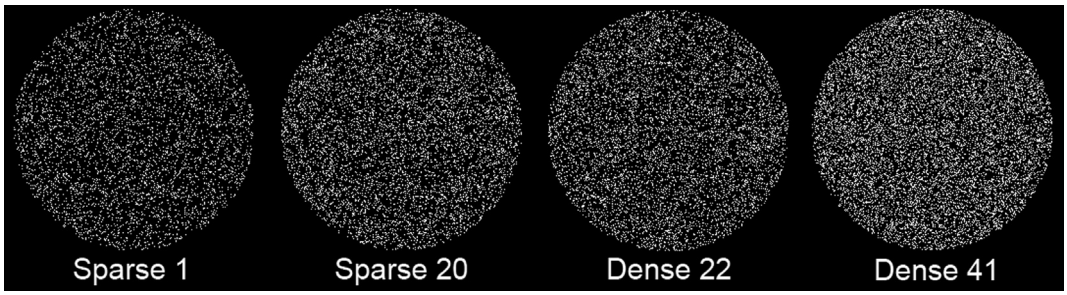

Fig. 1 – Sample stimuli with corresponding category assignment and density level. Level 1 is the sparsest stimulus; Level 41 the densest. Levels 20 and 22 are adjacent to the category boundary (Level 21) that separates sparse from dense.

During scanning sessions, participants emitted responses using the Lumina Response Pad System (model LU400-Pair). For 14 participants (7 female), the right-hand button box was used to make Sparse and Dense responses and the left-hand button box was used to make Uncertain responses; the button assignments were reversed (i.e., right-hand to make Uncertain responses) for the remaining 15 participants (7 female). Note that because there are two buttons per response pad and three total response options, we chose to segregate the Uncertain response from the categorization responses by lateralizing the former to one hand or the other, which also reduces working memory demands associated with maintaining the motor response set. This lateralization is counterbalanced in the present sample. Stimuli and feedback were presented on an LCD monitor during laboratory sessions and on a digital projector and screen viewed through a head-coil-mounted mirror during scanning sessions.

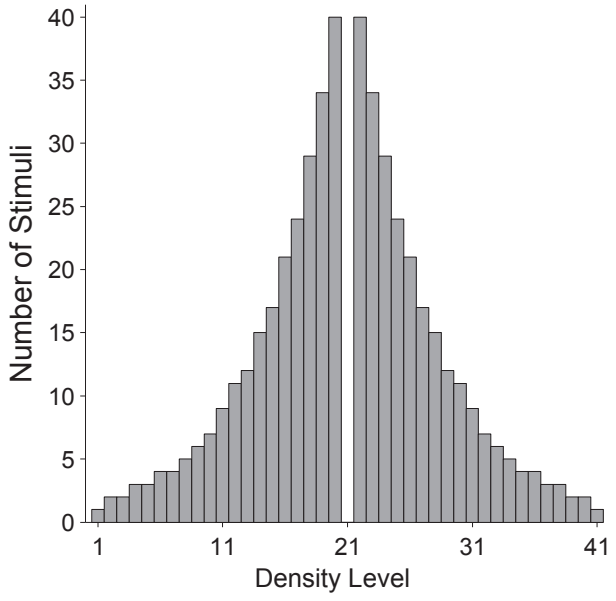

Fig. 2 – Histogram of stimuli at each density level. Stimuli from Level 21 were never shown and served as the category boundary separating sparse stimuli (Levels 1–20) from dense stimuli (Levels 22–41).

2.4. *Procedure*

During the laboratory prescreening session, participants were instructed that they would see black stimuli on a black background with varying numbers of white illuminated pixels. Their job was to categorize each stimulus as Sparse or Dense depending on how many pixels were illuminated. Participants were also informed they could choose to decline a response by using the Uncertain response. In order to motivate the use of this response, participants received +1 point for correct responses, −5 points for incorrect responses, and 0 points for uncertain responses. They were told that their goal was to maximize the accumulated score.

Prior to the experimental trials, participants completed 20 practice trials to become familiar with the procedure. During the practice trials, only stimuli from levels 1 and 41 were shown to participants. On each trial, a crosshair appeared for 500 msec immediately before a single stimulus presentation. The stimulus appeared centered on the screen and was response terminated. Feedback was presented immediately for 1500 msec as the total accumulated score along with "+1" in green text for correct responses, "−5" in red for incorrect responses, and "Don't know: +0" in white for URs. If the response time exceeded 2000 msec, "Too slow: −5" was presented in red. This feedback procedure was repeated on every trial. Points accumulated during the practice session were reset for the experimental session. There were 498 experimental trials and the experiment lasted approximately 60 min.

Participants who initially completed the laboratory version of the Sparse-Uncertain-Dense task were screened for their frequency of URs in order to determine their eligibility for brain imaging during this task. Although we expect the cognitive processes to be common across subjects regardless of their use of the UR, power considerations in the fMRI setting demand a minimum number of events of each type of interest (including, in this case, the UR). The participants in the laboratory session who used the UR on at least 30% of the most difficult trials (i.e., within three stimulus bins of the category boundary) were invited back for the fMRI experiment. Out of 156 participants, 70 met this criterion. The average accuracy and uncertainty-response profiles for the 70 participants who met the criterion and the 86 participants who did not are shown in Fig. 3. Although there is a clear difference in the

frequency of use, the general pattern of uncertainty responding across these two groups is similar and both groups closely match the qualitative performance of the participants in previous research (Smith et al., 2006). In addition, the accuracy of the two groups is functionally identical, especially for those stimuli near the category bound. Thus, in our prescreening sample, the task appears to be equally difficult for both high- and low-uncertainty responders, and the uncertainty-response profile peaks near the category bound.

Of the 70 participants who met the inclusion criterion, 32 were able or willing to participate in the full fMRI experiment. The procedure during the scanning session was identical to the lab session except that timing in the scanner was slightly different in order to synchronize stimulus and feedback with the scanner TRs. In the scanner, stimuli and feedback were presented for 2000 msec each and stimuli were no longer response terminated. Prior to the actual scanning session, participants completed a brief practice block of 40 trials in the laboratory that was matched to the timing of the scanning experiment, to reacquaint them with the task and to familiarize them with the new timing and pace. In the scanner, each participant completed six blocks across six functional runs with 83 stimuli per block. The first 10 TRs of each functional run were left blank to allow adequate time to reach steady-state scanning; similarly, the last 10 TRs were left blank to allow the BOLD response to decay from previous trials. Fig. 4 provides a visual representation of the trial timing. The number of TRs prior to stimulus (maximum of 5) and feedback (maximum of 3) presentation were jittered using a truncated geometric distribution with $p = .5$. Stimulus and feedback each lasted one TR (2000 msec). Whenever more than one TR preceded stimulus presentation, a crosshair was shown for half of the TR (1000 msec) immediately preceding stimulus presentation in order to reorient the participants' attention to the pending stimulus presentation (occurring on 48% of all trials).

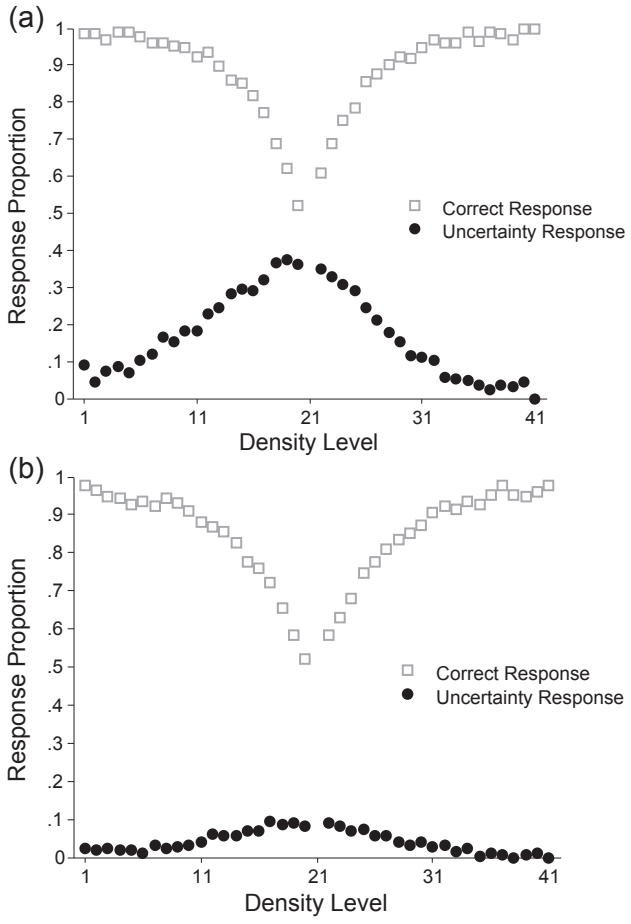

Fig. 3 – Average proportion of correct responses (open gray squares) and uncertainty responses (closed black circles) across all 41 density levels for (a) liberal ($n = 70$), and (b) conservative ($n = 86$) uncertainty responders. Note that density Level 21, which separates sparse and dense stimuli, was never shown to participants.

2.5. *Neuroimaging acquisition*

The scanning sessions were conducted at the UCSB Brain Imaging Center using a 3T Siemens Tim Trio MRI scanner with an 8-channel phased array head coil. Cushions were placed around the head to minimize head motion. Functional runs used a T2* weighted single shot gradient echo, echo-planar sequence sensitive to BOLD contrast (TR: 2000 msec, TE: 30 msec, FA: 90°, FOV: 192 mm) with generalized auto-calibrating partially parallel acquisitions (GRAPPA). Each volume consisted of 33 slices (interleaved acquisition, 3 mm thick with .5 mm gap, 3 mm × 3 mm in-plane resolution, 64 × 64 matrix) acquired at an angle manually adjusted to minimize in-plane artifact susceptibility near orbitofrontal cortex from sinus cavities (Deichmann, Gottfried, Hutton, & Turner, 2003). A localizer, a GRE field map (3 mm thick, FOV: 192 mm, voxel: 3 × 3 × 3 mm, FA = 60°), and a T1-flash structural scan (TR = 15 msec, TE = 4.2 msec, FA = 20°, 192 sagittal slices 3-D acquisition, .89 mm thick, FOV: 220 mm, voxel: .9 × .9 × .9 mm, 256 × 256 matrix) were obtained before the EPI scans, and an additional GRE field-mapping scan was acquired at the end of each scanning session. Slice orientation was identical for all GRE and EPI sequences. Each scanning session lasted about 90 min.

2.6. *Neuroimaging GLM analyses*

Preprocessing and data analysis were conducted using FSL's (www.fmrib.ox.ac.uk/fsl) FEAT (FMRI Expert Analysis Tool, version 6.00) software package. Preprocessing was carried out separately on each EPI scan. Preprocessing included motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002), BET brain extraction (Smith, 2002), spatial smoothing with a 5 mm FWHM kernel, grand-mean intensity normalization, and a high pass filter with a cutoff of .02 Hz. Functional data were assessed for exclusion if excessive head motion (i.e., >3 mm) occurred; no EPI sequences exceeded criterion. The high-resolution structural scan was registered to the MNI152-T1-2 mm standard brain via FLIRT (Jenkinson & Smith, 2001; Jenkinson et al., 2002) and further refined using FNIRT (nonlinear registration; Anderson, Jenkinson, & Smith, 2007). A transformation for each functional scan to the standard brain was generated using a two-step process to improve alignment first by registering each EPI to the T1-flash structural scan and then registering the T1-flash structural scan to

Fig. 4 – Example sequence of one experimental trial with jittering parameters.

the standard brain MNI152 brain template. These transformations were applied at the mid-level analysis (described below).

Low-level analyses were performed using FEAT separately for every BOLD scan (i.e., each block of the experiment). Six explanatory variables (events) were defined post-hoc using each participant's responses. Stimulus trials were redefined in one of three ways depending on a participant's unique responses: "correct categorization stimulus" for every trial on which a participant emitted a correct Sparse/Dense response; "incorrect categorization stimulus" for every incorrect Sparse/Dense response; and "uncertain stimulus" for uncertainty-response trials. All feedback events were redefined in the same way (e.g., correct feedback, etc.). Because our goal was to model the brain response related to the decision processes for each stimulus, the stimulus events were modeled as boxcar functions with a height of one and duration determined by observed response times (RTs)—the behavioral index of a final decision. This procedure was used to capture variation in pre-decisional processing time across different stimulus events. Feedback events had a height of one and duration of 2 sec. To generate a predicted BOLD time course for analysis, the boxcar was convolved with an HRF (gamma function with a standard deviation of 3 sec and a mean lag of 6 sec). All GLM analyses additionally included six unconvolved motion correction nuisance parameters in order to capture signal changes due to subjects' movement.

We performed two different GLM analyses. The first analysis was designed to identify brain regions that respond to uncertainty, categorization, or task difficulty. This analysis included regressors for the three experimental events (as defined above). In addition to those events, three parametric modulator terms (one each for correct, incorrect, and URs) were included to capture any variation in the BOLD signal due to stimulus difficulty as measured by absolute distance to bound. We can use these regressors in two ways: first, as a means of addressing the potential confound between uncertainty and task difficulty (see below for more details), and second, to find further evidence that there are distinct processes underlying uncertain events compared to categorization events. In particular, to the extent that the neural results for distance to bound differ across event types, it strongly argues against the notion that there is a unitary process responsible for the different events.

For the three distance-to-bound regressors, the height of the boxcar was determined by a simple linear function so that stimuli far from the bound were modeled with a tall boxcar (maximum height of 1 for stimulus levels 1 and 41), and stimuli close to the bound were modeled with a short boxcar (minimum height of 0 for stimulus Level 21). Each of these three parametric regressors was mean centered relative to the average non-zero boxcar height for the corresponding event type on a block-by-block basis, guaranteeing the three parametric regressors were uncorrelated with the stimulus EVs. Note that this procedure intentionally leaves differences in the event-wise mean distance-to-bound: uncertainty events had an average distance of .163, correct events .331, and incorrect events .136. However, because the ranges of the distance-to-bound regressors are much larger than the mean differences between events, we can combine these distance-to-bound regressors to identify regions that are exclusively responsive to distance-to-bound (irrespective of trial type), and mask out these regions in our contrasts of interest involving uncertainty events in order to alleviate the difficulty confound mentioned above. A total of nine regressors were included in this model (in addition to the motion parameters): three stimulus events (correct, incorrect, uncertain); three feedback events, one for each stimulus type; and three parametric regressors, one for each stimulus type.

The second GLM analysis was designed to test the hypothesis that uncertainty events—particularly those close to the category bound—are more compatible with a "middle-category" categorization response than a metacognitive judgment. For this analysis, regressors were included for each of the experimental events, but were re-defined to separately estimate "near to bound" events from "far from bound" events. This affords the opportunity to directly assess whether particular brain regions respond differentially to near or far stimulus events. Note that delineating "near" from "far" in this way also splits the stimuli into more or less difficult groups; because the goal of this analysis is to make distance/difficulty comparisons within event types, the parametric distance-to-bound regressors were not included. A total of 12 regressors were included in this model: three stimulus events and three feedback events (correct, incorrect, uncertain) for each distance group.

After low-level analyses were complete, the results were input into mid-level analyses to aggregate the block data for each participant using a fixed effects model. For each participant, the mid-level analyses yielded a statistical z-map for every contrast. Finally, the results of mid-level analyses were transformed into MNI standard space and input into a high-level analysis to generate group maps using a mixed effects model (FLAME 1; Woolrich, 2008). After specifying a minimum z threshold for each voxel, Gaussian random field theory was used to estimate a cluster-size threshold to achieve a specific experiment-wise false positive rate (Ashby, 2011; Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994).

3. Results

The results are reported in several parts. First, the behavioral results from the fMRI session are presented. Next, whole-

brain functional imaging results using a standard general linear model (GLM) approach are described for several contrasts of events and the parametric distance-to-bound modulators. Finally, results of a second GLM analysis are presented to directly test whether uncertainty events are compatible with a "middle" categorization stimulus event. All neuroimaging results are from high-level analyses, which include all 29 participants counterbalanced for left/right hand response assignment.

For all GLM analyses, two different thresholds were used to identify active clusters. A more conservative threshold (initial z-threshold of 3.54 with a cluster significance threshold $\alpha = .01$) was used to match false-positive rates across reported contrasts. A more liberal threshold ($z = 2.58$, $\alpha = .05$) was used to increase (and approximately match) power across reported contrasts at the cost of a greater false-positive rate. We focus on the results of either the liberal or conservative threshold in every case (and we report the results using the other threshold in supplementary materials, where applicable). Our justification for which threshold we use in each case is straightforward: in cases where we favor the alternative hypothesis (e.g., we believe that differences exist between conditions or events), we focus on the results using the conservative threshold (so that any detected differences are statistically conservative, lending more credibility to our results when differences are found), and only focus on the liberal results if nothing survives the conservative correction, in which case we treat the results as more exploratory. In cases where we favor the null hypothesis (e.g., we believe no differences exist between conditions or events), we focus on results using the liberal threshold (so there is statistically enough power to guard against failing to disconfirm the null hypothesis due to a lack of power).

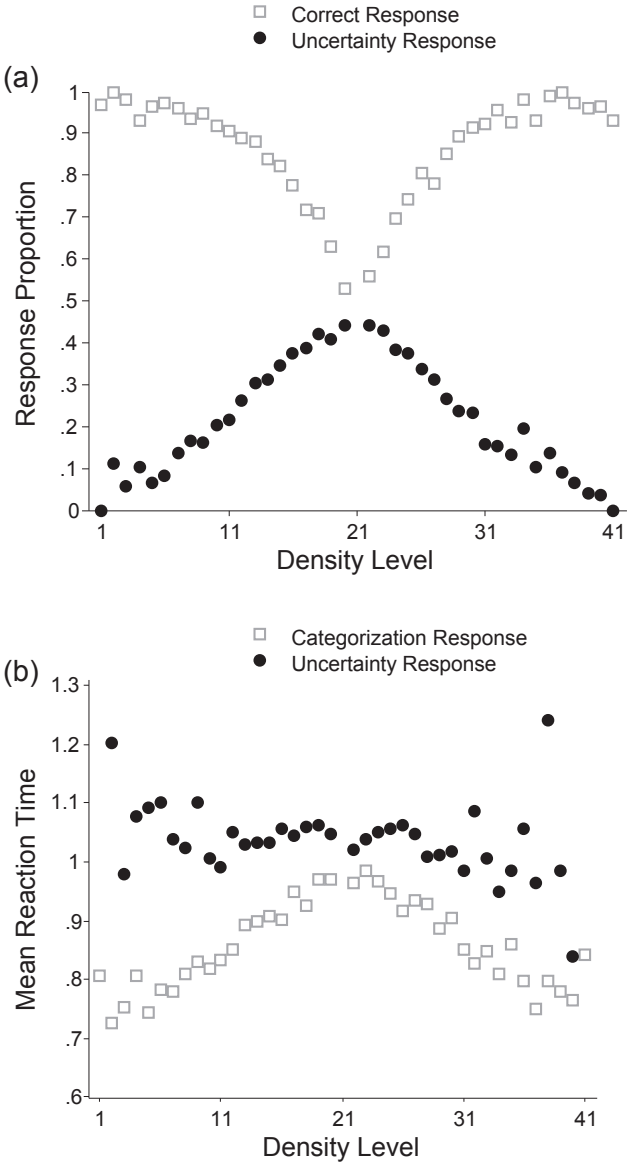

Fig. 5 – (a) Average proportion of correct responses (open gray squares) and uncertainty responses (closed black circles) across all 41 density levels. Note that density Level 21 was never shown to participants and separated sparse and dense stimuli. (b). Average response times for Sparse/Dense categorization (open gray squares) and Uncertainty (closed black circles) responses across all 41 density levels.

3.1. *Behavioral results*

Behavioral performance was assessed by calculating the average proportion of correct responses and URs separately for each density level across participants. Proportion of correct responses was calculated by dividing the number of correct responses by the number of categorization responses (i.e., accuracy was computed only for completed categorization trials); proportion of URs was calculated by dividing the number of URs by the total number of trials. Fig. 5a presents the averaged correct/uncertainty curves across all 29 participants. As expected, accuracy decreased toward chance as the density level moved closer to the Sparse/Dense category boundary, whereas uncertainty responding increased toward its maximum at the category boundary. Similarly, participants' mean RTs for both categorization responses and URs at each density level are shown in Fig. 5b. Whereas categorization RT appears generally to increase for responses as the stimuli become most difficult (e.g., at the Level 21 boundary), mean RT for URs remains uniformly longer than for categorization responses and are relatively constant across stimulus levels.

3.2. *Uncertainty, categorization, and distance-to-bound results*

Whole brain group results were computed for correct responses and for URs, along with the distance-to-bound regressors. We focus first on the latter, which frame our understanding of the results for the former. For the distance-to-bound regressors, we focus on three varieties of contrasts. The first is each regressor by itself (i.e., versus baseline). This is the most basic result, and can give an overview of which event types (of uncertain, correct, or incorrect) evince distance-to-bound effects. The second set of contrasts are direct contrasts between each pair of event types. These contrasts follow up on the patterns of results from the first group, and can confirm whether apparent differences between events hold up in a direct contrast. Finally, the third

contrast combines the three regressors equally, and tests that their sum is nonzero. Because it imposes a constraint that the distance-to-bound effect be event-type agnostic, this contrast is meant to identify voxels whose activity is due only to distance to bound per se. In later contrasts between uncertain and correct events, which differ in mean distance-to-bound (see section 2.6), significant results in any voxel might plausibly be due distance to bound, rather than to the event types. Therefore, we can conservatively use the results of this contrast to mask our primary contrasts of interest.

For analyses focusing on uncertain and correct events, the results are summarized for the contrast "correct categorization > uncertain" and for the contrast "uncertain > correct categorization". In a sense, correct trials represent the purest instances of categorization-related processing, so these contrasts are expected to separate categorization-related regions from UR-related regions (if they are, in fact, separable). Incorrect trials are more closely matched to UR trials in terms of their distribution across density levels. However, our inclusion of distance-to-bound regressors allows us to isolate effects related to such distributional differences. Moreover, errors can occur for multiple reasons, including a failure of normal categorization-related processing, and in that sense, may represent a less pure separation of categorization and UR processes. All results reported for correct trials are qualitatively unchanged when using incorrect trials, and so we focus our analyses on correct categorization trials.

3.2.1. *Parametric distance-to-bound modulators*

Including parametric distance-to-bound modulators in the model provides an opportunity to identify voxels whose BOLD signal is modulated by stimulus difficulty as indexed by stimulus distance-to-bound. For this first analysis, we simply tested the single-parameter estimates against the null hypothesis that they are equal to zero. Positive parameter estimates (i.e., significantly greater than zero) indicate that a voxel's BOLD signal amplitude is higher for stimuli farther from the bound (and vice versa for negative parameter estimates). We tested the parametric distance-to-bound modulators for all three response types (uncertainty, correct, and incorrect) separately. For both the incorrect and uncertainty distance-to-bound regressors, no voxels were significantly different than zero in either direction even at our more liberal threshold ($z = \pm 2.58$, $p < .05$) suggesting distance-to-bound on these trials did not modulate the BOLD response in any voxel. For the correct response distance-to-bound modulator, several significant clusters were identified (using a threshold of $z = 3.54$, $p < .01$; see Fig. 6 and Supplementary Table 1). This result indicates that many voxels' amplitude is higher for correct categorization stimuli far from the boundary relative to those close to the boundary. No voxels survived the converse contrast, even at the more liberal threshold ($z = -2.58$, $p < .05$).

To explore these single-regressor results more closely, we carried out each pairwise comparison between event types. Neither of the contrasts involving incorrect events had any significant voxels in either direction ($z = \pm 2.58$, $p < .05$), possibly due to the relatively small number of events for that event type (69.9 incorrect responses on average per participant over the entire experiment—much less than the 257.7 correct or 165.3 uncertain responses averaged per participant). However, for the direct contrast between the correct and uncertain distance-to-bound regressors, several clusters are present for correct > uncertain at our more liberal threshold ($z = 2.58$, $p < .05$, shown in Fig. 7 and Supplementary Table 2; there were no significant results for the converse contrast). For the most part, these results are a subset of the results for correct distance-to-bound regressor alone (Fig. 6): 83.4% (925/1109) of the voxels in this contrast are present in the correct alone results [and note that this is using the results for the more conservatively thresholded correct distance-to-bound

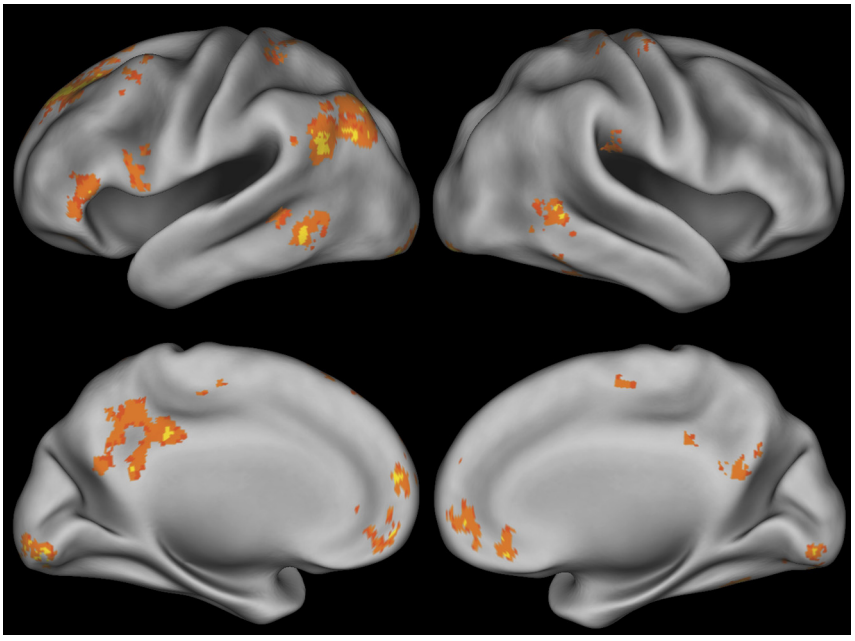

Fig. 6 – Whole-brain results from the correct response distance-to-bound modulator projected on an inflated lateral and medial cortical surface. Images are cluster thresholded (correcting for multiple comparisons) at $z > 3.54$, $p < .01$.

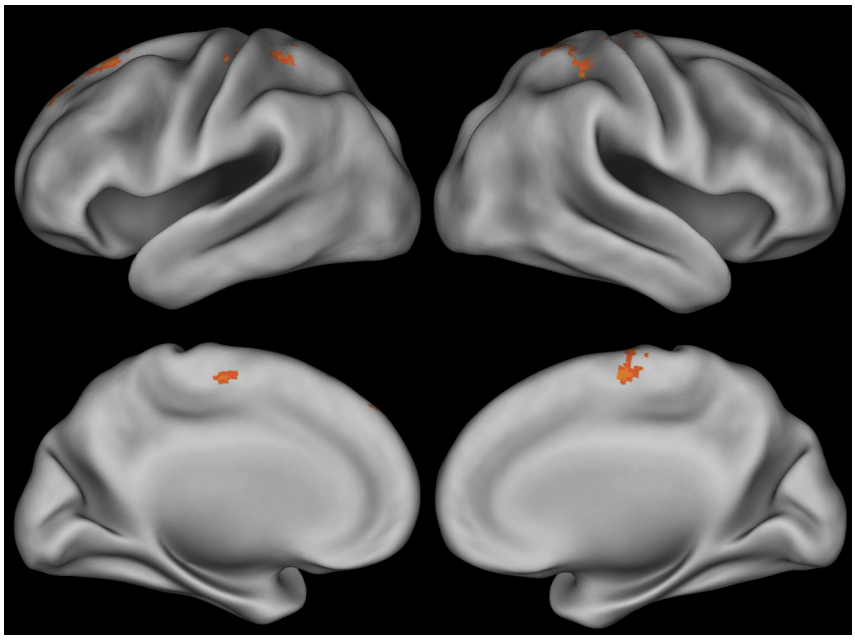

Fig. 7 – Whole-brain results for the direct contrast between correct and uncertain distance-to-bound regressors projected on an inflated lateral and medial cortical surface. Images are cluster thresholded (correcting for multiple comparisons) at $z > 2.58$, $p < .05$.

results; the overlap would necessarily be as large or larger using the more liberal threshold (shown in Supplementary Fig. S1)]. In a sense, these parametric modulator results are compatible with the pattern of behavioral response times (shown in Fig. 5b). Uncertainty RT remains relatively constant across all stimulus distances, as does the amplitude of the BOLD response. On the other hand, correct categorization RT decreases with distance to bound, and BOLD amplitude increases (in at least a small number of regions).

The last result from our analyses of the distance-to-bound regressors is the combined contrast, which compares the average parameter estimate across the three trial-type-specific distance-to-bound regressors against zero. Although there were no significant results in either direction even at our "liberal" threshold of $z = 2.58$, $p < .05$, we relaxed the threshold to $z = 1.96$, $p < .05$. Remember that these results will only be used to identify voxels whose activity in contrasts between correct and uncertain trials (see below) might be due purely to distance to bound, so a liberal threshold here is conservative with respect to our contrast of interest. At this liberal threshold, several significant clusters are evident—shown in Fig. 8 and listed in Table 1. As we present later results, we will refer back to this result to identify regions whose activity may be driven by difficulty.[1]

[1] As a secondary confirmation, we conducted a complementary GLM analysis, which specified a single distance-to-bound regressor (i.e., collapsing across the three event types, each demeaned relative to its own event-specific mean distance to bound). The resulting statistical map from this single regressor was qualitatively very similar to the results of the combined contrast, confirming that this analysis successfully identifies voxels purely responding to distance to bound, irrespective of trial type.

3.2.2. *Correct > uncertain*

The purpose of this contrast was to identify brain regions unique to correct categorization decisions (i.e., distinct from URs). Fig. 9 shows significant clusters on lateral and medial cortical views. Table 2 lists the coordinates in MNI space and z statistic for the peak voxel in each cluster as well as the anatomical region at the peak voxel. Using the conservative threshold ($z = 3.54$ and $p < .01$), four significant clusters—including two clusters spanning regions of the occipital lobe bilaterally, one cluster in left striatum extending from the head of the caudate ventrally to nucleus accumbens (NAcc), and a cluster near the posterior extent of the right middle and inferior frontal gyri—remained after thresholding (see Table 2 and Fig. 9; see Supplementary Fig. S2 for the liberal threshold map). When comparing these maps to the combined distance-to-bound result described above, .8% of the voxels at the conservative threshold (16/1973) are common to both images (2.6% common using the liberal threshold; see Supplementary Fig. S3 for the overlap map).

3.2.3. *Uncertain > correct*

The purpose of this contrast was to identify brain regions uniquely related to uncertainty responding (versus successful categorization). Fig. 10 displays significant clusters using our conservative threshold ($z = 3.54$ and $p < .01$) on lateral and medial cortical views. Significant clusters appeared throughout the brain; a total of 12 clusters survived thresholding (Table 3; see Supplementary Fig. S4 for the liberal threshold map).

The largest significant clusters each spanned the middle frontal gyrus of the PFC bilaterally—largely overlapping with BA9 or DLPFC—and extending near dorsal medial PFC.

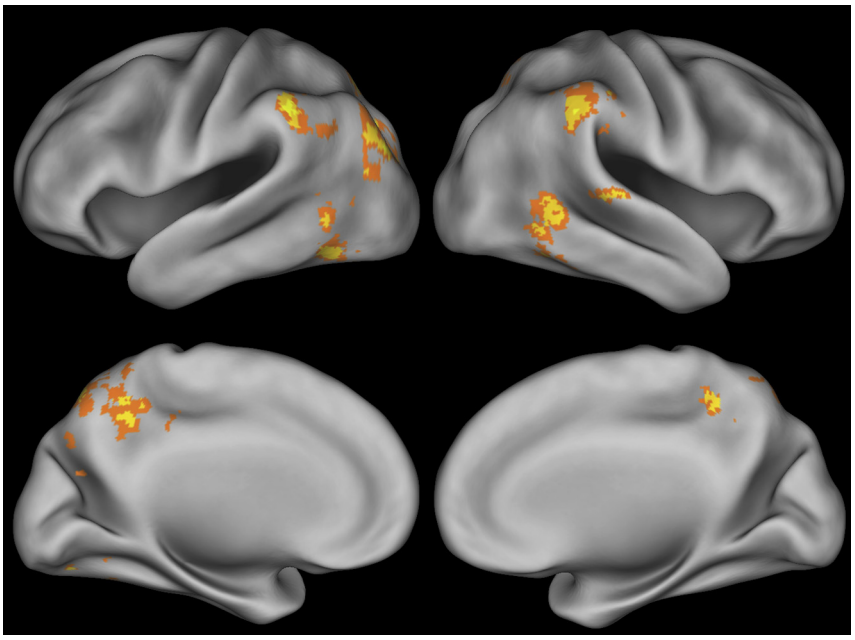

Fig. 8 – Whole-brain results for the combined distance-to-bound contrast projected on an inflated lateral and medial cortical surface. Table 1 lists the coordinates and number of voxels for every significant cluster. Images are cluster thresholded (correcting for multiple comparisons) at $z > 1.96$, $p < .05$.

Table 1 – Combined distance-to-bound contrast.

Cluster	Anatomic regions	# voxels	Max z-stat	x (mm)	y (mm)	z (mm)
5	Superior lateral occipital cortex (L)	1285	3.56	−14	−70	54
4	Posterior supramarginal gurys (L)	1139	3.69	−54	−46	40
3	Posterior supramarginal gurys (R)	804	3.3	54	−40	40
2	Middle temporal gyrus (R)	667	3.47	60	−52	4
1	Middle temporal gyrus (L)	596	3.21	−56	−62	−2

Significant cluster sizes and locations for the combined distance to bound contrast. For each significant cluster, the number of voxels, maximum z-statistic, anatomical region (from the Harvard–Oxford Structural Atlas) and (x, y, z) *mm* coordinates (in MNI standard space) of the max z-statistic are presented.

Significant clusters were also observed in a variety of other regions, including: caudal superior frontal gyrus (bilateral); both ACC and PCC; angular gyrus extending into lateral occipital regions (bilateral); left insular cortex; and the middle temporal gyrus. When comparing these maps to the combined distance-to-bound result described above, 4.0% of the voxels at the conservative threshold (218/5497) are common to both images (4.1% common using the liberal threshold; see Supplementary Fig. S5 for the overlap map).

3.3. *Metacognitive uncertainty versus middle-category response*

The metacognitive and middle-category accounts of the UR make very different psychological assumptions. Even so, at the behavioral level they make many similar predictions. For example, both theories predict that the most frequent and easiest URs will be to stimuli near the boundary between the sparse and dense categories. In the case of the metacognitive theory, this is because these stimuli are the most difficult to categorize and therefore most likely to engender uncertainty and elicit the UR. In the case of the middle-category theory, this is because these stimuli are in the center of the middle (UR) category and therefore are the most prototypical "middle category" exemplars. Despite the similarity between the theories at the behavioral level, they make strikingly different fMRI predictions.

In particular, the metacognitive theory predicts that UR responses rely on a distinct (or supplementary) set of regions compared to categorization responses, regardless of whether they are close to the sparse–dense boundary. The middle-category theory, on the other hand, views near-to-bound URs as correct responses, while far-to-bound URs are incorrect. To test between the metacognitive and middle-category accounts of the UR, we subdivided the stimuli into two sets: 1. stimuli near to the sparse–dense boundary, and 2. stimuli far from this boundary. More specifically, for the following analyses, "near events" were defined as those trials on which

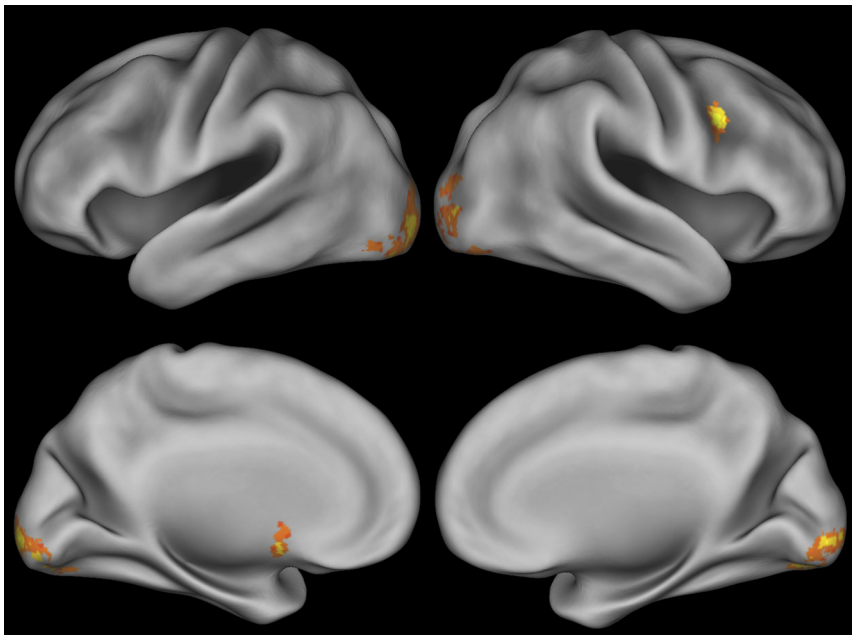

Fig. 9 – Whole-brain results from the correct categorization > uncertainty contrast projected on an inflated lateral and medial cortical surface. Table 2 lists the coordinates and number of voxels for every significant cluster. Images are cluster thresholded (correcting for multiple comparisons) at $z > 3.54$, $p < .01$.

Table 2 – Correct categorization > uncertainty contrast.

Cluster	Anatomic regions	# voxels	Max z-stat	x (mm)	y (mm)	z (mm)
4	Occipital and temporal lobes (L)	965	6.18	−18	−92	−10
3	Occipital and temporal lobes (R)	819	5.17	26	−86	−10
2	Inferior and middle frontal frontal gyrus (R)	112	5.42	40	8	30
1	Caudate/nucleus accumbens (L)	77	4.71	−8	6	−10

Significant cluster sizes and locations for the correct categorization > uncertainty contrast. For each significant cluster, the number of voxels, maximum z-statistic, anatomical region (from the Harvard–Oxford Structural Atlas) and (*x*, *y*, *z*) *mm* coordinates (in MNI standard space) of the max z-statistic are presented.

the stimuli were within ±3 bins from the sparse/dense boundary (bins 18, 19, 20, 22, 23, and 24) and "far events" were defined as those trials on which the stimuli were more than 3 bins from this boundary. Using this definition, the average number of near URs across participants (87.9) was well matched to the average number of far URs (73.8), near correct responses (73.8), as well as to the average number of far correct sparse and dense responses (91.9). Thus, ±3 bins appears to be a reasonable criterion for defining a putative middle-category region, both qualitatively and quantitatively insofar as the average number of responses in each region that would be "correct" for the middle-category account are well-matched.

We computed three contrasts of interest from this design. The first two simply serve to directly test the hypothesis that URs near to the category bound are more compatible with a middle-category, and not a metacognitive judgment. The third contrast leverages the near/far delineation to conduct an additional test as to whether URs are modulated by task difficulty (i.e., distance to bound).

3.3.1. *Near uncertainty > far correct*

The middle-category account of the UR predicts that near URs and far correct responses should be functionally identical because each is essentially a correct categorization response. On the other hand, the metacognitive account of the UR predicts this contrast should produce similar results to the uncertain > correct contrast described in section 3.2.3. Note, however, that any supra-threshold voxels in this contrast argue counter to the middle-category account, regardless of where those voxels are spatially distributed. The results of this contrast show many significant clusters similar to those in Fig. 10 at both our liberal and conservative thresholds (see supplementary Fig. S6). These results are therefore in line with predictions of the metacognitive framework of the UR and counter to the predictions of the middle-category account.

3.3.2. *Far uncertainty > near correct*

The middle-category account of the UR predicts that far URs and near categorization responses are both a kind of

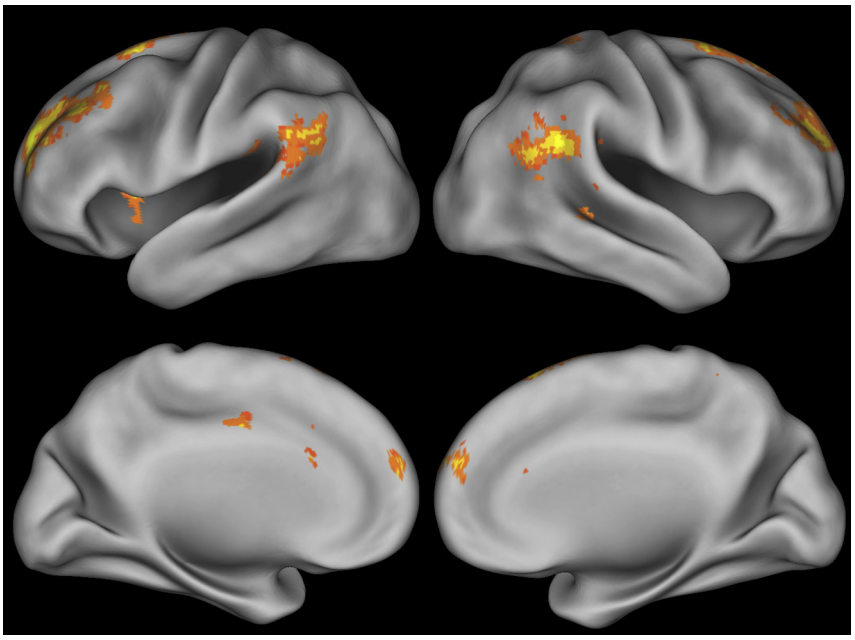

Fig. 10 – Whole-brain results from the uncertainty > categorization contrast projected on an inflated lateral and medial cortical surface. Table 3 lists the coordinates and number of voxels for every significant cluster. Images are cluster thresholded (correcting for multiple comparisons) at $z > 3.54$, $p < .01$.

Table 3 – Uncertainty > correct categorization contrast.

Cluster	Anatomic regions	# voxels	Max z-stat	x (mm)	y (mm)	z (mm)
11	Middle frontal gyrus, DLPFC (L)	1329	5.85	−22	54	24
10	Middle frontal gyrus, DLPFC (R)	1147	5.13	10	62	22
9	Angular gyrus (R)	942	5.03	54	−46	26
8	Angular gyrus/lateral occipital (L)	720	4.78	−58	−60	36
7	SFG (R)	476	5	12	26	58
6	SFG (L)	268	4.89	−18	14	60
5	Insula (L)	206	4.66	−32	8	8
4	Cingulate (anterior)	142	4.67	0	22	26
3	Superior parietal lobule (R)	117	4.3	24	−42	66
2	Middle/superior temporal gurys (R)	80	4.48	46	−36	0
1	Cingulate (posterior)	70	4.47	−6	−8	38

Significant clusters for the uncertainty > correct categorization contrast. For each significant cluster, the number of voxels, maximum z-statistic, anatomical region (from the Harvard–Oxford Structural Atlas) and (*x*, *y*, *z*) *mm* coordinates (in MNI standard space) of the max z-statistic are presented.

incorrect response (because URs should only be made in the near region whereas sparse/dense categorization responses principally reflect responses that should be made in the far region) and so activation to each should be functionally identical. In contrast, the metacognitive account of the UR again predicts results similar to the uncertain > correct contrast reported in section 3.2.2 (and again, any suprathreshold voxels argue against the middle-category account). The results of this contrast also show many significant voxels similar to those in Fig. 10 at both our liberal and conservative thresholds (see supplementary Fig. S7). Thus, these results are again in line with the predictions of the metacognitive framework and counter to the prediction of the middle-category theory.

3.3.3. *Uncertainty and task difficulty*

It is clear that the most difficult stimuli (those close to the sparse/dense boundary) most frequently receive URs. Therefore, it seems possible that the UR is an admission of difficulty rather than a metacognitive judgment. Although our analyses including parametric distance-to-bound modulator terms addresses whether any voxels are modulated by difficulty/distance (which we can then exclude in subsequent analyses, as above), another way to test this is to directly compare the brain activation to URs made to more or less difficult stimuli. Labeling our stimulus bins as near and far serves as a rough segregation of events into more or less difficult stimuli and affords a direct comparison between the most difficult (near) and easiest (far) stimuli.

If the UR is driven by uncertainty rather than task difficulty, then the activity maps for uncertain responses should be the same for near and far events (i.e., unrelated to task difficulty, per se). However, if task difficulty plays an important role in the UR activations, then there should be many voxels that show a stronger response to the near events than to the far events. To test this hypothesis, we generated a contrast directly comparing near uncertain with far uncertain. Even using our liberal cluster-based threshold, no voxels in this analysis were significant. Thus, this analysis produced no evidence that any voxels sensitive to the UR were significantly driven by task difficulty. Similarly, it is unlikely that the differences observed above between correct and uncertainty are due to task difficulty.

4. Discussion

The results of this research support the metacognitive over the middle-category account of the UR. At the broadest level, our results demonstrate that URs are not middle-category responses that simply partition the stimuli into three categories. Instead, as predicted by the metacognitive theory of UR, we found distinct behavioral response profiles and qualitatively different neural activation for URs and categorization responses (Fig. 10). The middle-category theory makes the opposite prediction (i.e., that all responses are "categorization" responses), which was not supported by our data. In addition, the middle-category theory predicts that URs to stimuli near to the sparse/dense category boundary reflect the most confident use of the UR (as these stimuli are most prototypical of the middle category) and should therefore elicit neural activity similar to that elicited during correct responses far from the bound. We did not find support for this prediction.

Our results further demonstrate that brain regions responding to the UR are distinct from those responding to task difficulty. In fact, while it is clear that URs occur most frequently for only the most difficult stimuli (i.e., closest to the categorization bound; Fig. 5a), our analyses identified a consistent collection of brain regions related to the UR, which were generally invariant to various corrections for task difficulty. First, after dividing the stimulus space into "near" and "far" regions from the category boundary, we found that profiles of brain structures particular to the UR did not differ significantly as a function of distance to the sparse/dense bound—in other words, URs near to the bound and URs far from the bound were statistically indistinguishable in our data. Second, in a separate analysis, our inclusion of parametric distance-to-bound modulator terms afforded the ability to specifically look for brain regions whose signal is modulated by stimulus difficulty separately for categorization and URs. We found no support that distance-to-bound for URs modulated the signal of any voxel; however, many voxels were sensitive to distance-to-bound for categorization trials. This suggests distinct processes on these different types of trials.

Note that our analyses include both trial-by-trial duration scaling of the stimulus-event boxcars based on the subjects' observed RT (which varies systematically between response types as in Fig 5), as well as parametric distance to bound modulators. Both methods may be expected to capture some variation in task difficulty: longer RTs occur for trials close to the bound, which are also the least accurate trials. However, in the present application, RT-scaling of the primary stimulus events uniquely (and more effectively) models the trial-by-trial decision processes in contradistinction to modeling RT as its own independent regressor to capture both RT and difficulty/distance variability in one event, as is common in some fMRI analyses. Here, the parametric modulators are designed to capture specific distance to bound effects. We believe the combination of the two simultaneously controls for difficulty differences while also accounting for RT differences in processing times.

The mere fact that there were differences in the distance-to-bound effects between UR and correct trials is enough to strongly suggest that these events are subserved by distinct cognitive processes. However, the spatial pattern of the distance-to-bound effect we observed is interesting in its own right. In particular, our results demonstrating increased activity far from the bound on correct trials (Fig. 6) are intuitively appealing, in that the pattern of activity very closely matches the canonical default mode network (Greicius, Krasnow, Reiss, & Menon, 2003). One possible explanation for this result is that the need for task engagement reduces as category response confidence increases; thus the neural modulation in those confident (easier) categorization trials reflects activation in the default mode network. Likewise, the regions that show differential distance-to-bound effects for correct versus UR trials match our expectations, including somatomotor areas (spanning pre- and post-central gyrus and supplementary motor area) and anterior superior frontal gyrus. The former may reflect something like response preparation or evidence accumulation, while the latter has previously been associated with greater activation to reward versus loss (Cohen, Elger, et al., 2008; Cohen, Fair, et al., 2008; Smith et al., 2010; Xue et al., 2009).

Focusing on our simple contrasts between UR and correct trials, we identified a collection of regions responsive to the UR widely distributed throughout the brain (Fig. 10), suggesting a complex and engaging psychological state. Critically, these brain regions are largely distinct from the regions identified in our investigation of "general" distance-to-bound-(i.e., difficulty-) sensitive processes combined across event types (Fig. 8), sharing a scant 4% of supra-threshold voxels at our most liberal estimate. Despite differences in average distance-to-bound between event types, any attempt to claim that our uncertainty versus categorization results reflect purely difficulty differences must predict that the same regions will appear in the contrast for parametric modulation by distance-to-bound. Instead, we observed minimal overlap between regions that respond to uncertainty and those that are sensitive to distance to bound.

Uncertainty monitoring, whether related to response-conflict, risk, or anticipated errors, correlates with some or all of a group of regions that include medial PFC, posterior parietal, ACC, PCC, and AI (e.g., Botvinick, Cohen, & Carter, 2004; Carter et al., 1998; Fleck et al., 2006; Grinband et al., 2006; Huettel, et al., 2005; Kable & Glimcher, 2007; MacDonald et al., 2000; McCoy & Platt, 2005; Platt & Huettel, 2008; Ridderinkhof et al., 2004; Stern et al., 2010; Ullsperger et al. 2010; Volz et al., 2004). ACC is in the position to relay

this uncertainty information via its strong connections with dorsal and lateral PFC, regions that have been documented to bring about behavioral adjustment or exert top-down control and that overlap with our results (Cohen et al., 1997; Fleck et al., 2006; Huettel et al., 2005; MacDonald et al., 2000; Miller & Cohen, 2001; Miller, Freedman, & Wallis, 2002; Nee, Wager, & Jonides, 2007; Ridderinkhof et al., 2004; Volz et al., 2004).

In our pre-trained task, stimuli near the boundary elicit a known uncertainty resulting from precise probability estimation [p(correct) = .5 at the category boundary]. Uncertainty associated with unknown probabilities is negatively correlated with PCC activity (Bach et al., 2011). Other fMRI studies have found PCC responses representing reward size or subjective value (Ballard & Knutson, 2009; Kable & Glimcher, 2007, 2010; Peters & Büchel, 2009), and prediction of task errors (Li, Yan, Bergquist, & Sinha, 2007). We observed significant voxels spanning the PCC, where activity in our task may reflect the subjective positive valuation of successfully escaping a high probability of error.

One possibility is that the PCC, ACC, insula and the medial PFC provide evidence of uncertainty to the DLPFC, which may then take control and change the strategy from categorization to a UR. Other metacognition neuroimaging research also highlighted dACC and DLPFC in retrospective confidence judgments, ascribing an especially important role to the most anterior aspects of PFC (Fleming et al., 2012; Hilgenstock et al., 2014; Rounis et al., 2010; Yokoyama et al., 2010). The PFC cluster we identified also extends anteriorly to fronto-polar regions, suggesting a metacognitive component of the UR.

Our results partially converge with Grinband et al. (2006) finding that VS, medial PFC and AI correlate with categorization uncertainty. Their task required a guess, while ours required declining to guess. In our task VS activity was specific to categorization trials, which suggests that VS activity in their task was related to a categorization guess, while medial PFC and AI seem to track uncertainty regardless of whether the task requires guessing or opting out. On the surface our tasks are very similar in that they both involve perceptual categorization with increasing uncertainty near the boundary; however, the option of an adaptive UR requires a qualitatively different level of cognitive processing. Not surprisingly, we found large clusters bilaterally, extending to the most anterior aspects of PFC to which evolutionarily nascent functions have been ascribed (Ramnani & Owen, 2004).

In contrast to uncertainty responding regions, correct categorization recruited a relatively small set of regions (Fig. 9) that included the right inferior frontal gyrus pars opercularis (IFGpo), as well as bilateral striatum including the head of the caudate and the NAcc. Both the right IFGpo and striatum are commonly identified in rule-based categorization tasks (Aron, Robbins, & Poldrack, 2004; Ashby, Paul, & Maddox, 2011; Helie, Roeder, & Ashby, 2010). The contribution of ventral striatum (VS) to correct categorization could be because of our design choice to award points for correct responses: several studies have reported NAcc activity related to anticipation of financial gains or rewards (Knutson, Fong, Bennett, Adams, & Hommer, 2003; Knutson & Wimmer, 2007; O'Doherty, et al. 2004; Wächter, Lungu, Liu, Willingham, & Ashe, 2009).

Behaviorally, the response profile observed here matches those reported for humans by Smith et al. (2006), confirming that the present task created sufficient uncertainty and elicited highly engaging and adaptive URs. The distinctive RT distributions across stimulus levels between categorization and URs are also noteworthy: latencies for URs were relatively long and invariant across stimulus levels, but for categorization they increased systematically for more difficult trial levels near the category boundary (Fig. 5b). These longer RTs may indicate that greater cognitive effort is required during URs. Taken together, the behavioral results further corroborate the distinctive psychological role of URs.

One important question is whether our results depended on the particular point scheme we used to encourage frequent URs. Healthy young adults often report dissatisfaction and reluctance in using the UR (Smith, Shields, & Washburn, 2003), even when the task is difficult. FMRI statistical analyses require large samples of responses; the point scheme was adopted in an attempt to maximize UR use. There are several reasons to believe that our results were not unduly influenced by this methodological choice. First, the behavioral results of liberal and conservative "uncertain" responders were qualitatively the same (e.g., compare Fig. 3a and b). Second, the use of different point schemes has not been shown to change the general profile of uncertainty responding (e.g., our behavioral results are qualitatively identical to Smith et al., 2006; wherein a different point scheme was used). Objective point schemes can reduce variability across participants, but there is no reason to believe they change the underlying decision problem in any fundamental way.

The present research provides a particularly compelling map of a UR network because it identified brain regions that underlie the most direct measure of volitional uncertainty responding yet studied. It did so within a psychophysical task that provided a precise titration of trial difficulty. The paradigm also afforded us the ability to observe directly and selectively only those trials exceeding the risk threshold for each individual participant. Finally, the behavioral nature of our paradigm allowed us to study volitional URs shorn of participants' introspected and explicitly declared reports of uncertainty.

The present research also has theoretical implications regarding the animal-metacognition literature (e.g., Kornell, 2009; Smith, 2009; Smith et al., 2012). Different species may have evolved different mechanisms to deal with uncertainty. URs are functionally isomorphic between human and rhesus macaques (Fujita, 2009, p. 575; Roberts, Feeney, McMillan, MacPherson, & Musolino, 2009, p. 130; Sutton & Shettleworth, 2008, p. 266), while capuchin monkeys have repeatedly failed to show the metacognitive performance patterns that macaques show routinely (e.g., Basile, Hampton, Suomi, & Murray, 2008; Beran, Smith, Coutinho, Couchman, & Boomer, 2009; Fujita, 2009). Macaques' URs represent higher-level and controlled decision making (Smith, et al., 2013), while capuchins do not make any URs, but readily learn primary perceptual responses to a middle category (Beran, et al., 2009). Our results may explain this behavioral dissociation. Capuchins may have robustly in place the neural circuits that ground the responses made to primary perceptual inputs like Sparse, Middle, and Dense. But unlike the macaques, they may have only weakly or primitively in place the networks described in this article that ground the adaptive response of

uncertainty. In terms of known neurobiology, macaques possess much greater gyrification in frontal cortex than capuchins, suggesting an overall greater expansion of neocortex in these animals (Rilling & Insel, 1999). Similarly, macaques have a greater proportion of granular PFC tissue than other monkeys as well as a greater density of dendritic spines in granular PFC tissue (Elston et al., 2006). Of course, humans' granular PFC tissue and density of spinous neurons exceeds all other primates. Elston and colleagues suggest that differences in PFC microstructure may underlie fundamental cognitive differences in primates and could be integral to human intelligence. Perhaps these differences also relate to metacognitive capacity in different species.

Electrophysiological research has begun to explore URs in macaques, and those results converge with our human fMRI results. For example, neural responses in the lateral intraparietal cortex of the monkey reflect the degree of uncertainty in an opt-out task similar to ours (Kiani & Shalden, 2009). Furthermore, the pulvinar is inactivated when monkeys escape perceptual categorization due to low confidence (Komura et al., 2013). The pulvinar expanded during primate evolution and receives inhibition from frontal and parietal regions, which were active in our study. In order to better compare whole-brain neural networks between humans and animals, future neuroimaging research of the UR in nonhuman primates will be necessary. Our results provide the groundwork for such comparative research. Elucidating the neural networks of human URs lays the foundations for exploring and comparing neural signatures of URs in other species, which might confirm existing behavioral dissociations, and illuminate the progression by which metacognition and the reflective mind emerged within the primate order. Such work carries with it the potential to strengthen comparative psychology and neuroanatomy as empirical sciences and integrate them more fully within experimental psychology and cognitive neuroscience as well as broadening our understanding of the evolution of humans' remarkable cognitive capabilities.

Acknowledgements

The preparation of this article was supported by the U.S. Army Research Office through the Institute for Collaborative Biotechnologies under Grant W911NF-07-1-0072, by Grant 2R01MH063760 from NIMH, by Grant 1R01HD061455 from NICHD and Grant BCS-0956993 from NSF. The authors declare no competing financial interests.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cortex.2015.07.028

REFERENCES

Anderson, J. L. R., Jenkinson, M., & Smith, S. M. (2007). *Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2*.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8*, 170–177. http://dx.doi.org/10.1016/j.tics.2004.02.010.

Ashby, F. G. (2011). *Statistical analysis of fMRI data*. Cambridge, MA: MIT Press.

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011). COVIS. In E. M. Pothos, & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 65–87). New York: Cambridge University Press.

Bach, D. R., Hulme, O., Penny, W. D., & Dolan, R. J. (2011). The known unknowns: neural representation of second-order uncertainty, and ambiguity. *The Journal of Neuroscience, 31*, 4811–4820. http://dx.doi.org/10.1523/JNEUROSCI.1452-10.2011.

Ballard, K., & Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *NeuroImage, 45*, 143–150. http://dx.doi.org/10.1016/j.neuroimage.2008.11.004.

Basile, B. M., Hampton, R. R., Suomi, S. J., & Murray, E. A. (2008). An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*, 169–180. http://dx.doi.org/10.1007/s10071-008-0180-1.

Basile, B. M., Schroeder, G. R., Brown, E. K., Templer, V. L., & Hampton, R. R. (2015). Evaluation of seven hypotheses for metamemory performance in rhesus monkeys. *Journal of Experimental Psychology: General, 144*, 85–102.

Beran, M. J., Smith, J. D., Coutinho, M. V. C., Couchman, J. C., & Boomer, J. (2009). The psychological organization of "uncertainty" responses and "middle" responses: a dissociation in capuchin monkeys (Cebus apella). *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 371–381. http://dx.doi.org/10.1037/a0014626.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences, 8*, 539–546. http://dx.doi.org/10.1016/j.tics.2004.10.003.

Brainard, D. H. (1997). Psychophysics software for use with MATLAB. *Spatial Vision, 10*, 443–436.

Carruthers, P. (2008). Meta-cognition in animals: a skeptical look. *Mind and Language, 23*, 58–89. http://dx.doi.org/10.1111/j.1468-0017.2007.00329.x.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*, 747–749. http://dx.doi.org/10.1126/science.280.5364.747.

Cohen, M. X., Elger, C. E., & Weber, B. (2008). Amygdala tractography predicts functional connectivity and learning during feedback-guided decision-making. *NeuroImage, 39*, 1396–1407.

Cohen, A. L., Fair, D. A., Dosenbach, N. U. F., Miezin, F. M., Dierker, D., Van Essen, D. C., et al. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage, 41*, 45–57. http://dx.doi.org/10.1016/j.neuroimage.2008.01.066.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386*, 604–608. http://dx.doi.org/10.1038/386604a0.

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage, 19*, 430–441. http://dx.doi.org/10.1016/S1053-8119(03)00073-9.

Dunlosky, J., & Bjork, R. A. (Eds.). (2008). *Handbook of memory and metamemory*. New York: Psychology Press.

Elston, G. N., Benavides-Piccione, R., Elston, A., Zietsch, B., Defelipe, J., Manger, P., et al. (2006). Specializations of the granular prefrontal cortex of primates: implications for cognitive processing. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology, 288*(1), 26–35.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex, 16*, 1623–1630. http://dx.doi.org/10.1093/cercor/bhj097.

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience, 32*, 6117–6125. http://dx.doi.org/10.1523/JNEUROSCI.6489-11.2012.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1*, 210–220.

Fujita, K. (2009). Metamemory in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*, 575–585. http://dx.doi.org/10.1007/s10071-009-0217-0.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences, 100*(1), 253–258.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron, 49*, 757–763. http://dx.doi.org/10.1016/j.neuron.2006.01.032.

Helie, S., Roeder, J. L., & Ashby, F. G. (2010). Evidence for cortical automaticity in rule-based categorization. *The Journal of Neuroscience, 30*, 14225–14234. http://dx.doi.org/10.1523/JNEUROSCI.2393-10.201.

Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You'd better think twice: post-decision perceptual confidence. *NeuroImage, 99*, 323–331. http://dx.doi.org/10.1016/j.neuroimage.2014.05.049.

Huettel, S. A., Song, A. W., & McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *The Journal of Neuroscience, 25*, 3304–3311. http://dx.doi.org/10.1523/JNEUROSCI.5070-04.2005.

Jenkinson, M., Bannister, P. R., Brady, J. M., & Smith, S. M. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage, 17*, 825–841. http://dx.doi.org/10.1006/nimg.2002.1132.

Jenkinson, M., & Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5*, 143–156. http://dx.doi.org/10.1016/S1361-8415(01)00036-6.

Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. (2009). Metacognition in animals: how do we know that they know? *Comparative Cognition and Behavior Reviews, 4*, 29–39. http://dx.doi.org/10.3819/ccbr.2009.40003.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience, 10*, 1625–1633. http://dx.doi.org/10.1038/nn2007.

Kable, J. W., & Glimcher, P. W. (2010). An "as soon as possible" effect in human intertemporal decision making: behavioral evidence and neural mechanisms. *Journal of Neurophysiology, 103*, 2513–2531. http://dx.doi.org/10.1152/jn.00177.2009.

Kiani, R., & Shalden, M. (2009). Representation of confidence associated with a decision by neurons in parietal cortex. *Science, 324*, 759–764. http://dx.doi.org/10.1126/science.1169405.

Knutson, B., Fong, G., Bennett, S., Adams, C., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *NeuroImage, 18*, 263–272. http://dx.doi.org/10.1016/S1053-8119(02)00057-5.

Knutson, B., & Wimmer, G. E. (2007). Splitting the difference: how does the brain code reward episodes? *Annals of the New York Academy of Sciences, 1104*, 54–69. http://dx.doi.org/10.1196/annals.1390.020.

Komura, Y., Nikkuni, A., Hirashima, N., Uetake, T., & Miyamoto, A. (2013). Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nature Neuroscience, 16*, 749–755. http://dx.doi.org/10.1038/nn.3393.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*, 297–315. http://dx.doi.org/10.1037/0096-3445.123.3.297.

Kornell, N. (2009). Metacognition in humans and animals. *Current Directions in Cognitive Science, 18*, 11–15. http://dx.doi.org/10.1111/j.1467-8721.2009.01597.x.

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science, 18*, 64–71. http://dx.doi.org/10.1111/j.1467-9280.2007.01850.x.

Le Pelley, M. E. (2012). Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 686–708. http://dx.doi.org/10.1037/a0026478.

Le Pelley, M. E. (2014). Primate polemic: commentary on Smith, Couchman, and Beran (2014). *Journal of Comparative Psychology, 128*, 132–134. http://dx.doi.org/10.1037/a0034227.

Li, C. S. R., Yan, P., Bergquist, K. L., & Sinha, R. (2007). Greater activation of the "default" brain regions predicts stop signal errors. *NeuroImage, 38*, 640–648. http://dx.doi.org/10.1016/j.neuroimage.2007.07.021.

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835–1838. http://dx.doi.org/10.1126/science.288.5472.1835.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience, 8*, 1220–1227. http://dx.doi.org/10.1038/nn1523.

Metcalfe, J., & Shimamura, A. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: Bradford Books.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202. http://dx.doi.org/10.1146/annurev.neuro.24.1.167.

Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: categories, concepts and cognition. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences, 357*, 1123–1136. http://dx.doi.org/10.1098/rstb.2002.1099.

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 1–17. http://dx.doi.org/10.3758/CABN.7.1.1.

Nelson, T. O. (Ed.). (1992). *Metacognition: Core readings*. Toronto: Allyn and Bacon.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*, 452–454. http://dx.doi.org/10.1126/science.1094285.

Peters, J., & Büchel, C. (2009). Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *The Journal of Neuroscience, 29*, 15727–15734. http://dx.doi.org/10.1523/JNEUROSCI.3489-09.2009.

Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature Neuroscience, 11*, 398–403. http://dx.doi.org/10.1038/nn2062.

Ramnani, N., & Owen, A. M. (2004). Anterior PFC: insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience, 5*, 184–194. http://dx.doi.org/10.1038/nrn1343.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*, 443–447. http://dx.doi.org/10.1126/science.1100301.

Rilling, J. K., & Insel, T. R. (1999). The primate neocortex in comparative perspective using magnetic resonance imaging. *Journal of Human Evolution, 37*(2), 191–223.

Roberts, W. A., Feeney, M. C., McMillan, N., MacPherson, K., & Musolino, E. (2009). Do pigeons (Columba livia) study for a test? *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 129–142. http://dx.doi.org/10.1037/a0013722.

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. C. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience, 1*, 1–11. http://dx.doi.org/10.1080/17588921003632529.

Schwartz, B. L. (1994). Sources of information in metamemory: judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review, 1*, 357–375.

Smith, S. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*, 143–155. http://dx.doi.org/10.1002/hbm.10062.

Smith, J. D. (2009). The study of animal metacognition. *Trends in Cognitive Sciences, 13*, 389–396. http://dx.doi.org/10.1016/j.tics.2009.06.009.

Smith, J. D., Beran, M. J., & Couchman, J. J. (2012). Animal metacognition. In T. Zentall, & E. Wasserman (Eds.), *The oxford handbook of comparative cognition*. Oxford, UK: Oxford University Press.

Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring. *Journal of Experimental Psychology: General, 135*, 282–297. http://dx.doi.org/10.1037/0096-3445.135.2.282.

Smith, J. D., Couchman, J. J., & Beran, M. J. (2014a). Animal metacognition: a tale of two comparative psychologies. *Journal of Comparative Psychology, 128*, 115–131. http://dx.doi.org/10.1037/a0033105.

Smith, J. D., Couchman, J. J., & Beran, M. J. (2014b). A tale of two comparative psychologies: reply to commentaries. *Journal of Comparative Psychology, 128*, 140–142. http://dx.doi.org/10.1037/a0034784.

Smith, J. D., Coutinho, M. V., Church, B. A., & Beran, M. J. (2013). Executive-attentional uncertainty responses by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: General, 142*(2), 458. http://dx.doi.org/10.1037/a0029601.

Smith, D. V., Hayden, B. Y., Truong, T. K., Song, A. W., Platt, M. L., & Huettel, S. A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience, 30*, 2490–2495.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *The Behavioral and Brain Sciences, 26*, 317–373. http://dx.doi.org/10.1017/S0140525X03000086.

Stern, E. R., Gonzalez, R., Welsh, R. C., & Taylor, S. F. (2010). Updating beliefs for a decision: neural correlates of uncertainty and underconfidence. *Journal of Neuroscience, 30*, 8032–8041. http://dx.doi.org/10.1523/JNEUROSCI.4729-09.2010.

Sutton, J. E., & Shettleworth, S. J. (2008). Memory without awareness: pigeons do not show metamemory in delayed matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 266–282. http://dx.doi.org/10.1037/0097-7403.34.2.266.

Taylor, P. C. J., Nobre, A. C., & Rushworth, M. F. S. (2007). Subsecond changes in top–down control exerted by human medial frontal cortex during conflict and action selection: a combined transcranial magnetic stimulation–electroencephalography study. *The Journal of Neuroscience, 27*, 11343–11353. http://dx.doi.org/10.1523/JNEUROSCI.2877-07.2007.

Ullsperger, M., Harsay, H. A., Wessel, J. R., & Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure and Function, 214*, 629–643. http://dx.doi.org/10.1007/s00429-010-0261-1.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *NeuroImage, 21*, 848–857. http://dx.doi.org/10.1016/j.neuroimage.2003.10.028.

Wächter, T., Lungu, O. V., Liu, T., Willingham, D. T., & Ashe, J. (2009). Differential effect of reward and punishment on procedural learning. *The Journal of Neuroscience, 29*, 436–443. http://dx.doi.org/10.1523/JNEUROSCI.4132-08.2009.

Woolrich, M. W. (2008). Robust group analysis using outlier inference. *NeuroImage, 41*, 286–301. http://dx.doi.org/10.1016/j.neuroimage.2008.02.042.

Xue, G., Lu, Z. L., Levin, I. P., Weller, J. A., Li, X., & Bechara, A. (2009). Functional dissociations of risk and reward processing in the medial prefrontal cortex. *Cerebral Cortex, 5*, 1019–1027.

Yokoyama, O., Miura, N., Watanabe, J., Takemoto, A., Uchida, S., Sugiura, M., et al. (2010). Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neuroscience Research, 68*, 199–206. http://dx.doi.org/10.1016/j.neures.2010.07.2041.

Zakrzewski, A. C., Coutinho, M. V., Boomer, J., Church, B. A., & Smith, J. D. (2014). Decision deadlines and uncertainty monitoring: the effect of time constraints on uncertainty and perceptual responses. *Psychonomic Bulletin & Review, 21*, 763–770. http://dx.doi.org/10.3758/s13423-013-0521-1.