Executive-Attentional Uncertainty Responses by Rhesus Macaques (*Macaca mulatta*)

J. David Smith Mariana,
Department of Psychology University at Buffalo, The State University of New York

V. C. Coutinho,
Department of Psychology University at Buffalo, The State University of New York

Barbara Church, and
Department of Psychology University at Buffalo, The State University of New York

Michael J. Beran
Language Research Center, Georgia State University

Abstract

The uncertainty response has been influential in studies of human perception, and it is crucial in the growing research literature that explores animal metacognition. However, the uncertainty response's interpretation is still sharply debated. The authors sought to clarify this interpretation using the dissociative technique of cognitive loads imposed on ongoing discrimination performance. Four macaques (*Macaca mulatta*) performed a Sparse-Dense discrimination within which an uncertainty response let them decline difficult trials or a middle response let them identify middle stimuli. Concurrent memory tasks were occasionally overlain on ongoing discrimination performance. The concurrent tasks disrupted macaques' uncertainty responses far more than their sparse, middle, or dense discrimination responses. This dissociation suggests that the uncertainty response is a higher-level, decisional response that is particularly dependent on working memory and attentional resources. This is consistent with the theoretical possibility that the uncertainty response is an elemental behavioral index of uncertainty monitoring or metacognition.

Keywords

metacognition; uncertainty monitoring; attention; executive function; primate cognition

Metacognition--the monitoring and control of cognitive processes—is one of humans' important and sophisticated cognitive capacities (Balcomb & Gerken, 2008; Benjamin, Bjork, & Schwartz, 1998; Dunlosky & Bjork, 2008; Flavell, 1979; Koriat & Goldsmith, 1994; Nelson, 1992; Scheck & Nelson, 2005; Schwartz, 2008). It affects many aspects of comprehension, learning, and decision-making. It is linked to our self-awareness (Gallup, 1982) and declarative consciousness (Koriat, 2007; Nelson, 1996).

Metacognition's importance and sophistication raises the question of its origins and the question of whether it is uniquely human. Given the importance of these questions, Smith and his colleagues inaugurated research on animal metacognition (Smith, Schull, Strote, McGee, Egnor, & Erb, 1995; Smith, Shields, Allendoerfer, & Washburn, 1998; Smith,

Correspondence can be addressed to J. David Smith, 346 Park Hall, Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY 14260, psysmith@buffalo.edu, 716-645-0229, 716-645-3801 (fax).

Shields, Schull, & Washburn, 1997). Research is actively ongoing in this area (Basile, Hampton, Suomi, & Murray, 2009; Beran, Smith, Coutinho, Couchman, & Boomer, 2009; Beran & Smith, 2011; Beran, Smith, Redford, & Washburn, 2006; Call, 2010; Couchman, Coutinho, Beran, & Smith, 2010; Foote & Crystal, 2007; Fujita, 2009; Paukner, Anderson, & Fujita, 2006; Roberts, Feeney, McMillan, MacPherson, & Musolino, 2009; Smith, Redford, Beran, & Washburn, 2010; Suda-King, 2008; Sutton & Shettleworth, 2008; Washburn, Gulledge, Beran, & Smith, 2010).

Illustrating research in this area, in Smith et al. (1995), a bottlenosed dolphin (*Tursiops truncatus*) was to make a High or Low response, respectively, as the pitch of the tone presented was 2,100 Hz or 1,200–2,099 Hz. The third response paddle, the uncertainty response, let the dolphin decline the present trial and enter a subsequent, easy trial instead. The result was that the dolphin assessed correctly when he was at risk for error in the High-Low discrimination and adaptively declined the difficult trials near threshold. Remarkably, his uncertainty responses peaked near 2,086 Hz, 14 HZ and 0.11 semitones from the standard High tone at 2,100 Hz. The dolphin was performing at his true perceptual limit, and one possible explanation of his uncertainty responding is that he knew (metacognitively) he was at that limit.

Possibly supporting this explanation, the dolphin showed his own ancillary uncertainty behaviors. That is, on difficult trials, he held back from the responses and swam slowly with distinctive head and jaw movements. A factor-analytic study of these behaviors showed that they tracked trial difficulty and peaked at the dolphin's threshold of discrimination. Tolman (1927) called behaviors like these lookings and runnings back and forth. In a remarkable article at the zenith of behaviorism—an article titled *A Behaviorist's Definition of Consciousness*—he argued that these behaviors could operationally define consciousness for behaviorists.

Humans performed similarly to the dolphin, and they reported that their uncertainty responses reflected their metacognitive judgments of knowing and not knowing. There is a strong cross-species isomorphism in the use of the uncertainty response, one that has produced some of the closest existing human-animal performance correspondences (e.g., Shields et al., 1997). This provides a theoretical motivation for seeking common principles of psychological organization behind the uncertainty responses of humans and animals.

However, the attribution of metacognition to the dolphin—or to macaques and apes in related studies (Call & Carpenter, 2001; Hampton, 2001; Kornell et al., 2007; Smith, Beran, Redford, & Washburn, 2006)—bears a heavy burden given comparative psychology's conservatism in interpreting animals' behavior (Morgan, 1906). Animals' performances in uncertainty tasks, though they might seem metacognitive, could nonetheless be underlain by stimulus- and reinforcement-based mechanisms. In fact, the possible low-level basis for uncertainty responses by animals has been the overriding theoretical concern about comparative metacognition research (Couchman et al., 2010; Crystal & Foote, 2009; Hampton, 2009; Jozefowiez, Staddon, & Cerutti, 2009; Smith, Beran, Coutinho, & Couchman, 2008; Smith, Beran, & Couchman, 2012).

This same issue arose in the interpretation of early studies on human perceptual uncertainty. Participants in early perception experiments were sometimes allowed to respond "Uncertain" when they could not classify a stimulus (Angell, 1907; Fernberger, 1914, 1930; Jastrow, 1888; Thomson, 1920; Urban, 1910; Watson, Kellog, & Kawanishi, 1973; Woodworth, 1938). However, some believed that this response was not properly included in psychophysical procedures. They suspected that uncertainty responses were psychologically distinctive because they were so affected by instructional set and participants' personality or

temperament. They suspected that uncertainty responses were qualitatively different from perceptual responses (e.g., Light or Heavy; Short or Long) because they were a non-perceptual, metacognitive comment on the failure of discrimination processes. George (1917) criticized uncertainty responses for introducing "extra-serial" attitudes into tasks that must depend on intra-serial attitudes (i.e., pure reactions to the observed stimulus magnitudes). Boring (1920) criticized uncertainty responses for attitudinally seducing participants away from the series of mental states that should map directly and only to the series of sensory magnitudes.

Watson et al. (1973, pp. 184–185) summarized this debate. Psychologists supported the uncertainty response when their goal in psychophysical research was to establish the determinants of states of consciousness, one of which is the state of uncertainty. In contrast, behaviorists opposed the uncertainty response. For them, the third response was simply a perceptual, middle response, to be included only when the task contained a middle stimulus region. Note that comparative psychologists, seeking an elemental paradigm for studying animal metacognition, converged on the uncertainty-response paradigm that was commonly used within human psychophysics. Likewise, the early perceptual psychologists joined animal-metacognition researchers in considering the appropriate psychological interpretation of uncertainty responses.

So what is the correct interpretation of uncertainty responses? Are they a lower-level, perceptual response (middle) to a middle range of sensory magnitudes? Are they higher-level cognitive assessments about indeterminacy and the failure of the discriminatory process? These are elemental questions about what might be an elemental metacognitive judgment. The present article addresses these questions.

These questions were brought into sharp relief by a recent study on metacognition in capuchin monkeys (*Cebus apella*), a New World primate species. In Beran et al. (2009), capuchin monkeys received two discrimination tasks. In the Sparse-Uncertain-Dense task, difficult stimuli—spanning the breakpoint between the Sparse and Dense stimulus ranges—could be declined through the use of the uncertainty response. In the Sparse-Middle-Dense task, middle responses to *the same stimuli* were rewarded. In both tasks, incorrect responses earned the capuchins a 20 s, trial-less timeout period. Capuchins made almost no uncertainty responses (Figure 1A). They made many middle responses—apparently easily and naturally (Figure 1B).

Given this failure, the researchers more than quadrupled the time cost for errors up to 90 s so that monkeys potentially gave up 30 trials and 30 food rewards during each timeout. Now capuchins had a much more compelling motivation to make uncertainty responses. Five of six capuchins still did not do so, although they made frequent middle responses. This is a striking dissociation between two structurally similar tasks that represent strong controls for one another in a within-subjects design (for related research on capuchin monkeys, see Basile et al., 2009; Fujita, 2009; Paukner et al., 2006).

This dissociation presents a double mystery to cognitive and comparative psychologists. First, why is it that capuchin monkeys freely make middle responses but fail to make uncertainty responses for the same stimuli facing the same strong reinforcement motivation to do so? What is the difference between the psychological signals that prompt the two responses? Second, why do rhesus macaques (*Macaca mulatta*), an Old World primate species, make uncertainty responses (Beran & Smith, 2011; Beran et al., 2006; Couchman et al., 2011; Hampton, 2001; Smith et al., 1997, 1998, 2010; Washburn et al., 2010) when capuchin monkeys do not? What are the differences between the cognitive systems of the

two species that make them differentially able to monitor the psychological signals that prompt uncertainty responses?

The present article's empirical goal is to illuminate these issues psychologically, by exploring the cognitive basis of uncertainty and middle responses using the theoretical constructs of executive functioning and attentional resources. To do so, we measured the effect of concurrent cognitive loads incorporated into macaques' ongoing performance in tasks that allow uncertain or middle responses.

This study may be the first primate-cognition study that has used the methodology of concurrent-task dissociation to assay the different cognitive levels of different aspects of performance by nonhuman primates. It illustrates the many possibilities for theoretical development that could result from using this empirical approach with nonhuman primates. Naturally, though, our study also meets head-on and exemplifies the intrinsic difficulty of this methodology, including the problem of interleaving multiple tasks successfully and the problem of gauging correctly and then sustaining the resource loads of the concurrent tasks.

The results suggest an explanation for the uncertain-middle response dissociation in capuchin monkeys and for the macaque-capuchin species dissociation in uncertainty responding. They provide new insights into the long debate surrounding the interpretation of the uncertainty response in perception and psychophysics. And they begin to suggest the psychological level and psychological organization of uncertainty states, judgments, and responses within humans' and animals' overall cognitive system.

Experiment 1

In Experiment 1, two macaques performed a Sparse-Dense discrimination in which they could also use an uncertainty response to decline any trials they chose. An identity matching-to-sample task was periodically overlain on ongoing discrimination performance. The experiment's hypotheses focus on how the concurrent working-memory load might affect perceptual processing as compared to decisional uncertainty-based processing (if uncertainty-based processing has this decisional character).

Historically, research with humans suggested that primary perceptual processing is not directly affected by a concurrent cognitive load, and that changes in performance under dual-task situations reflect the working memory requirements of the task and bottlenecks during the response processes (e.g. Joliceur, 1999; Logan, 1978; Pashler, 1984; 1989; 1994; Sternberg, 1969). Some recent research has found direct effects of concurrent load on perceptual processing (e.g. Anderson, Mannan, Rees, Sumner, & Kennard, 2005; Liu, Dosher, & Lin, 2009; Peterson, Beck, & Wong, 2008)—for example, when targeted eye-movements are critical to performance (He & McCarley, 2010), or there is direct processing overlap between the perceptual task and the concurrent task (Pannebakker, Band, & Ridderinkhof, 2009), or the stimuli are highly complex (Bourke & Duncan, 2005). These conditions are likely not applicable to the task of Experiment 1. Accordingly, based on the current literature, we expected the sensitive use and accuracy of the primary perceptual responses themselves—that is, the sparse and dense responses—to be mainly unaffected by the working-memory load. In contrast, we expected the adaptive use of the uncertainty response to be strongly affected by the concurrent load—if this response is particularly dependent on executive functioning and attentional resources.

Method

Subjects—Adult male rhesus macaque monkeys (*Macaca mulatta*) Murph and Lou (both 17 years old) were tested. They had been trained, using procedures described elsewhere

(Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989; Washburn & Rumbaugh, 1992), to respond to computer-graphic stimuli by manipulating a joystick. They had both participated in numerous previous computerized experiments (e.g., Beran, 2007, 2008; Beran, Harris, et al., 2008), including experiments with an uncertainty response (e.g., Beran et al., 2006). The macaques were tested individually in their home cages at the *Language Research Center* of *Georgia State University*, with *ad lib* access to the test apparatus, working or resting as they chose during long sessions. The macaques had continuous access to water, and worked for fruit-flavored primate pellets. They received a daily diet of fruits and vegetables independent of their efforts on the task, and thus they were not food deprived for the purposes of this experiment.

Apparatus—The monkeys were tested using the *Language Research Center*'s *Computerized Test System* (Washburn & Rumbaugh, 1992), comprising a computer, a digital joystick, a color monitor, a pellet dispenser, and programming code written in Visual Basic 6.0. Trials were presented on a personal computer with an attached 17-inch color monitor. Joystick responses were made with a Gravis GamePad Pro digital joystick mounted vertically to the cage. Monkeys manipulated the joystick through the mesh of their home cages, producing isomorphic movements of a computer-graphic cursor on the screen. Contacting appropriate computer-generated stimuli with the cursor brought them a 94-mg fruit-flavored chow pellets (Bio-Serve, Frenchtown, NJ) using a Gerbrands 5120 dispenser interfaced to the computer through a relay box and output board (PIO-12 and ERA-01; Keithley Instruments, Cleveland, OH). Correct responses were accompanied by a computer-generated, 2 second melodic chime that bridged the monkeys to their reward. Incorrect responses cleared the screen for 20 seconds and produced a 2 second computer-generated buzz tone. For all tasks reported, the intertrial interval was 1 second after the delivery of food rewards or timeouts.

Density continuum—Macaques completed a sparse-dense discrimination task. On each trial, they saw a 185×185-pixel box in the screen's top center. The box was filled with a varying number of randomly placed lit white pixels on a black background. Sixty stimulus levels could be presented (Levels 1–60). Each level's pixel count was given by pixels=round $(1066 \times 1.018^{Level})$. This formula gave the continuum the appropriate logarithmic character, with each density step a constant percentage increase in pixels.

Response modality—Macaques touched an S or D icon on the screen to classify the boxes as sparse or dense, respectively. These responses produced a 1-pellet reward when correct sparse or dense responses were made. They touched the ? icon to respond uncertain. Choice of that response led to the next trial without providing any feedback or reinforcement to the monkey. Despite this neutral outcome, the uncertainty response is optimally included within the response repertoire for this task. Through its judicious use, the organism will fend off difficult trials, avoid errors and timeouts, and increase its rewards per minute in the task.

Sparse-uncertainty-dense task—The stimulus continuum was divided into sparse and dense regions. Across a session, approximately half of the trials were sparse (Levels 1–30, 1076–1537 pixels) and half were dense (Levels 31–60, 1565–2236 pixels). Approximately 50 percent of the trials were restricted to only the central third of the range (Levels 21–40) whereas the other 50 percent were sampled from the full continuum (Levels 1–60— additional Level 21–40 trials were possible as part of this second sampling method). By oversampling the difficult regions of the stimulus continuum, we increased the difficulty and uncertainty of the task, and also the value of responding uncertain to adaptively mange this difficulty.

During baseline training only, there were forced-escape trials during which the animal had to choose the uncertainty response. These were randomly assigned on 10% of the trials and were randomly assigned to density levels. Thus, monkeys were not trained when the uncertainty response was optimal because they had to make that response on both easy and difficult trials without choice. These trials were simply instructional reminders as to what happened when the uncertainty response was used. We have not observed these reminder trials to support or sustain uncertainty responding by monkeys.

Matching-to-sample task—In the matching-to-sample task, macaques saw a sample image in the center of the screen. These images were nine-pointed white polygons created randomly trial-by-trial for display. The polygons were created according to procedures described in Smith, Redford, and Haas (2008a, b). Monkeys made an observing response to this sample by moving the cursor to touch it, and it disappeared. After a delay of 3, 4, 5, or 6 seconds (randomly chosen across trials), 1 matching and 3 non-matching alternatives were presented randomly in the four corners of the screen—animals had to choose the identical one to the sample most recently shown. Thus, they had to encode the correct sample's visual-feature information at trial's onset, and remember that information during the retention interval. Macaques received two food pellets for correct responses and 20 s timeouts for incorrect responses.

Matching-to-sample conditions—In the matching-to-sample control condition, macaques alternated the completion of matching trials with the completion of sparse-dense trials. The animals saw a sample image that disappeared upon contact, waited through the variable 3 to 6 s delay, and then saw the match choices and could match immediately. After matching, they performed one Sparse-Dense trial. Thus, no information from the matching trial had to be retained through performance on the sparse-dense trial, and so there was no concurrent working-memory load, but there was task switching required between the two tasks.

In the matching-to-sample concurrent-load condition, macaques interleaved the matching and sparse-dense tasks as follows. They saw the sample and made the observing response to it. Next, they completed one sparse-dense trial to which they could respond sparse, dense, or uncertain under the contingencies already described. Following that trial's completion, including the outcome given for the trial (food-pellet delivery for correct responses, clearing of the screen for uncertainty responses, or a 20 s timeout period for incorrect responses), there was a variable 3 to 6 s delay interval during which the screen was blank and then the match choice stimuli appeared. Now, information from the in-progress matching trial had to be retained through performance on the sparse-dense trial, so there was a concurrent working-memory load. In the language of the task-switching literature, now there was task switching and also temporal overlap in task performance—working-memory material had to be maintained during performance of the sparse-uncertain-dense task (Kiesel, Steinhauser, Wendt, Falkenstein, Jost, Philipp, & Koch, 2010).

Matching-to-sample sequence of conditions—Macaques entered the matching conditions from a reacclimation to the sparse-dense discrimination performed alone that served as a baseline for the experiment. Then both animals entered the matching-to-sample concurrent-load condition, in which they performed sparse-dense trials while maintaining their memory for the sample image. Next, they completed the matching-to-sample control condition, in which they completed the matching trial before beginning a sparse-dense trial. Then, they completed a second block of concurrent-load trials and a second block of control trials.

Modeling performance and fitting data—We instantiated a formal model of the sparse-uncertainty-dense task, one grounded in Signal Detection Theory (SDT—MacMillan & Creelman, 2005). SDT assumes that performance in perceptual tasks is organized along an ordered series (a continuum) of psychological representations of changing impact or increasing strength. Here, the continuum of subjective impressions would run from clearly sparse to clearly dense. Given this continuum, SDT assumes that an objective event will create subjective impressions from time to time that vary in a Gaussian distribution around the objective stimulus level presented. This perceptual error is part of what produces errors in discrimination and part of what may foster uncertainty in the task for stimuli near the discrimination breakpoint. Finally, SDT assumes a decisional process by which criterion lines are placed along the continuum so that response regions are organized. Here, by the overlay of Sparse-Uncertain (SU) and Uncertain-Dense (UD) criteria, for example, the stimulus continuum would be divided up into Sparse, Uncertain, and Dense response regions.

Our models took the form of a simulated version of the tasks as macaques in the present studies experienced them. We then placed simulated observers in those task environments for 10,000 trials distributed as the macaques' trials were.

The simulated observers experienced perceptual error (PE). The size of perceptual error— that is, the standard deviation of the Gaussian distribution that governed misperception— was one free parameter in our model. On each trial, given some stimulus (Level 1-Level 60), simulated observers misperceived the stimulus obedient to this Gaussian distribution. Given a standard deviation of 5, for example, they could misperceive a Level 15 stimulus generally in the range of Level 10 to Level 20. This misperceived level became the subjective impression on which the simulated observer based its response choice for that trial.

The simulated observers were also given differently placed criterion points. The placements for criteria SU and UD defined three response regions that determined its response choice to a subjective impression. The placement of SU and UD were two more free parameters that could be adjusted to optimally fit the data.

To fit the data, we moved all three parameters through wide ranges of possibilities. For each parameter configuration, we produced that simulated observer's predicted performance profile, showing its three response levels for 60 stimulus levels. We calculated the sum of the squared deviations (SSD) between the corresponding observed and predicted data points. We minimized this SSD fit measure to find the best-fitting parameter configuration. For this best-fitting configuration, we also calculated a more intuitive measure of fit—the Average Absolute Deviation (AAD). This measure represents the average of the deviations between observed and predicted response levels (with the deviations always signed positively).

Results

Murph—Murph's performance in the concurrent-load, identity-matching task was as follows. For the first concurrent load test, the first control test, the second concurrent-load test, and the second control test, respectively, Murph's performance on the delayed matching task was 79.8% correct (77.21% to 81.25% across the four delay intervals), 91.22% correct (89.56% to 92.2% across delay intervals), 92.49% correct (91.86% to 93.62%), and 94.65% correct (93.29% to 96.14%). There was some performance decrease for the first concurrent condition compared to other conditions. But matching performance was high throughout with no effect of delay interval.

Figure 2,A-E shows Murph's performance in the sparse-uncertainty-dense task during his last baseline session (983 trials) and during all trials of the experiment's four remaining

phases (1116, 1457, 2998, and 2991 trials). Murph's uncertainty responding was nearly eliminated by the addition of the concurrent cognitive load. During baseline performance and his first concurrent-load testing, respectively, Murph declined 64.4% and 7.4% of trials across the difficult trial levels 21–40. $t(1347) = 26.12$, $p < .05$. This change was essentially instantaneous. Comparing the last 40 trials of baseline performance and the first 40 trials of concurrent-load performance, he responded uncertain on 63.0% and 17.4% of trials at levels 21–40, $t(48) = 3.74$, $p < .05$. This suggests that some aspect of the cognitive processing underlying uncertainty responding was immediately blocked, rather than the decrease reflecting gradual extinction or unlearning of the uncertainty response. There is by now a substantial literature on the sources of interference in task mixing and task switching (see, for example, Kiesel et al., 2010). We point out that in the concurrent-load condition for Murph (and Lou), there was task switching and temporal overlap in task performance—working-memory material had to be maintained during performance on the sparse-uncertain-dense task. Either the switching or the temporal overlap, or both factors, could have contributed to the disruption in the uncertainty response we observed in this condition.

Murph's uncertainty responding rebounded when he transitioned to control testing. In this condition, the two tasks alternated. There was task switching but no temporal overlap in the performance of the two tasks. The rebound suggests that a substantial source of the interference in the concurrent-load condition did come from the temporal overlap—that is, from the difficulty of maintaining working-memory information while managing the processing of uncertainty. Murph's uncertainty responding for trial levels 21–40 increased from 7.4% under concurrent-load testing to 42.8% during control testing, $t(1723) = 18.86$, $p < .05$. Through equivalent statistical tests, we confirmed that Murph's uncertainty responding was reduced again when he re-entered concurrent load testing, $t(2972) = 10.78$, $p < .05$, and that it rebounded significantly again when he finally returned to control testing in the experiment's final phase, $t(3958) = 7.44$, $p < .05$.

We used our SDT model to formalize and concretize Murph's first transition in Experiment 1. We fit the model to Murph's data from his baseline condition and his first concurrent-task condition (Figures 2A,B). The respective best-fitting predicted performance profiles are shown in Figure 3A,B.

The model reproduced well Murph's baseline performance and his performances under a concurrent cognitive load. The two SSD measures of fit were 1.012 and 0.246, respectively. The two intuitive fit measures were 0.040 and 0.016, respectively—that is, the model fit with 4.0% or 1.6% error per data point on average. The model estimated that Murph placed his SU and UD criteria, respectively, at Levels 21 and 36 along the 60-Level Sparse-Dense continuum during baseline performance, but at 29 and 31 during concurrent-load performance, defining a very narrow uncertainty-response region that admitted few uncertainty responses. The modeling confirmed that with the concurrent load present—Murph's uncertainty-response region essentially collapsed.

Finally, the modeling estimated Murph's PE to be 4.8 and 4.6 steps along the perceptual continuum during baseline testing and concurrent-load testing, respectively. This convergence shows that Murph's primary perceptual processes, and his sensitive use of the sparse and dense responses, were hardly changed by the presence or absence of the concurrent load. Only his capacity to recruit the uncertainty response was sharply, selectively targeted by the concurrent load. Within the SDT framework, this result implies that the concurrent load left untouched Murph's detection *sensitivity*, while altering his *decisional criteria* to nearly exclude the uncertainty response.

Figure 4A depicts all of Murph's Experiment 1 data. All sessions were adjoined chronologically, with the data summarized in 80-trial blocks according to the percentage of uncertainty responses made on trial levels 21–40. (The final block in each phase could have up to 159 trials, as remaindered trials were included to the phase's end to make a clean block break to the next experimental phase.) Murph entered concurrent-load sessions, went into control testing, re-entered concurrent-load sessions, and returned to control testing at Blocks 13, 26, 44, and 81, respectively. Murph strongly confirmed the hypothesis that the concurrent load would selectively reduce his uncertainty responding, whereas the absence of the concurrent load would let it recover. His results also illustrate the difficulty of holding an animal at a constant level of primary-task skill and concurrent-load effort that are optimal for showing these results. The strength of the declines and rebounds did shift through the experiment, possibly as Murph gradually reduced his reliance on the uncertainty response or as he gained more skill at multi-tasking.

Lou—Lou's performance in the concurrent-load, identity-matching task was as follows. For the first concurrent load test, the first control test, the second concurrent-load test, and the second control test, respectively, Lou's performance on the delayed matching task was 58.49% correct (55.52% to 61.12% across the four delay intervals), 83.15% correct (80.77% to 87.5% correct), 76.23% correct (72.68% to 78.89%). and 88.47% correct (86.2% to 90.23%). Again, there was a decrease in performance for the first concurrent condition compared to the other conditions. Matching performance was otherwise high throughout with no effect of delay length.

Figure 5,A–E shows Lou's performance in the sparse-uncertainty-dense task during his last baseline session (1517 trials) and during all trials of the experiment's four remaining phases (1287, 1482, 2067, and 2991 trials, respectively). Lou's uncertainty responding was nearly eliminated by the addition of the concurrent cognitive load. During baseline performance and his first concurrent-load testing, respectively, Lou declined 30.2% and 4.2% of trials across the trial levels 21–40, $t(1769) = 15.63$, $p < .05$. This change was essentially instantaneous. Comparing the last 40 trials of baseline performance and the first 40 trials of concurrent-load performance, he responded uncertain on 38.9% and 0.0% of trials at levels 21–40, $t(43) = 4.15$, $p < .05$. Again, the uncertainty response in an uncertainty-monitoring task was quickly and strongly affected by the application of a concurrent load. Here, as with Murph, the implication is that the processes underlying uncertainty responding were immediately blocked by the concurrent load. The suddenness of this transition is not consistent with any extinction or gradual unlearning mechanism. Thus, Lou also confirmed the hypothesis that the concurrent load reduces uncertainty responding. However, as Figure 5 shows, Lou's uncertainty responding did not rebound in the control condition that followed his concurrent-task performance or thereafter. With Lou, we only had the opportunity in Experiment 1 to observe once the primary phenomenon under investigation here, whereas with Murph we observed the phenomenon in all four task transitions. Lou did demonstrate the primary phenomenon again in Experiment 2.

Figure 4B summarizes all of Lou's Experiment 1 data in 80-trial blocks as already described. Lou entered concurrent-load sessions, returned to control testing, re-entered concurrent-load sessions, and returned to control testing at Blocks 19, 35, 53, and 78, respectively.

We used our SDT model to formalize and concretize Lou's first transition in Experiment 1. We fit the model to Lou's data from his baseline condition and his first concurrent-task condition. The respective best-fitting predicted performance profiles are shown in Figure 6A and 6B. The model reproduced well Lou's baseline performance and his performances under a concurrent cognitive load. The two SSD measures of fit were 0.6314 and 0.5086,

respectively. The two intuitive fit measures were 0.037 and 0.029, respectively—that is, the model fit with 4% or 3% error per data point on average. The model estimated that Lou placed his SU and UD criteria, respectively, at Levels 28 and 36 along the 60-Level Sparse-Dense continuum during baseline performance, but at 30 and 32 during concurrent-load performance, defining a very narrow uncertainty-response region that admitted few uncertainty responses. Lou's uncertainty-response region also essentially collapsed under the concurrent load.

The modeling estimated Lou's PE to be 5.8 and 5.7 steps along the perceptual continuum during baseline testing and concurrent-load testing, respectively. His sensitive use of the sparse and dense responses was also hardly changed by the presence or absence of the concurrent load. Only his capacity to recruit the uncertainty response was selectively targeted. The concurrent load left untouched Lou's detection *sensitivity*, while altering his *decisional criteria* to exclude the uncertainty response.

Experiment 2

In Experiment 2, macaques performed the sparse-dense discrimination of Experiment 1, but now paired with a new concurrent memory task to which they were unaccustomed. We hoped to replicate the results of Experiment 1, showing again that the application of the cognitive load would reduce uncertainty responding.

Subjects, apparatus, response modality, sparse-dense discrimination—These basic aspects of Experiment 2 were as already described.

Spatial-memory task—In the spatial-memory task, macaques saw a sample image (a colored group of circles on a yellow background) in the screen's center. Simultaneously, they saw four images in the corners, one of which identically matched the sample, and three foil images that were solitary colored squares. However, monkeys did not choose that match image at this point. Instead, they only moved the cursor to the sample image in the center of the screen. Then, they waited through a 2, 3, 4, 5, or 6 s delay interval randomly selected for each trial, during which nothing was present on the screen. Next, four identical white squares appeared in the four corners of the screen, directly over the previous locations of the four match choices, and the monkeys had to move their cursor to the remembered location where the matching image had been seen. Thus, they had to encode the matching item's location at trial's onset, and remember that location, not the matching item's shape. Correct responses led to a reward of three pellets. Incorrect responses led to a 20 second timeout.

Spatial-memory conditions—In the spatial-memory control condition, monkeys alternated completing spatial-memory trials as just described with completing sparse-dense trials that were identical to those presented in Experiment 1. Thus, no information from the spatial-memory trial had to be retained during performance on a sparse-dense trial, so there was no concurrent spatial-memory load, and no temporal overlap of performance of the two tasks, but there was task switching.

In the spatial-memory concurrent-load condition, monkeys interleaved the spatial-memory and sparse-dense tasks as follows. They saw the sample and the four images at the screen's four corners, one image matching the sample. Next, however, they completed one sparse-dense trial to which they could respond sparse, dense, or uncertain under the contingencies already described. During the sparse-dense trial, the sample and the four corner images were removed from the screen. After completing the sparse-dense trial, they waited through the variable 2 to 6 second delay interval. Then, the four identical white squares appeared in the four corners, and the monkeys had to move their cursor to the remembered location where

the matching image had been seen. Now, information from the in-progress spatial-memory trial (specifically, the location to which they later needed to respond) had to be retained through performance on the sparse-dense trial, so there was a concurrent spatial-memory load and temporal overlap between the performance of the two tasks.

Spatial-memory sequence of conditions—Macaques were first trained in the basic spatial-memory task. (There was no need for a sparse-uncertain-dense baseline phase in this experiment, because the animals had already been acclimated to this task through their participation in Experiment 1.) Next, they completed the concurrent-load spatial-memory condition, in which they had to perform Sparse-Dense trials while maintaining their spatial memory for the matching image. Then, they completed the spatial-memory control condition, in which they completed spatial-memory trials before attempting any Sparse-Dense trials. Finally, they completed another phase of testing in the spatial-memory concurrent-load condition.

Results

Murph—Murph's performance in the concurrent-load, spatial-memory task was as follows. For the first concurrent load test, the first control test, and the second concurrent load test, respectively, Murph's matching performance was 34.66% correct (32.27% to 36.09% across the five delay intervals), 60.03% correct (57.21% to 63.57% correct), and 45.54% correct (42.27% to 49.29%). Murph's matching performance was reduced in the concurrent conditions of this experiment, but there was no large effect of delay length within a condition.

Figure 7,A–C shows Murph's performance in the sparse-uncertainty-dense task during all trials of each of the experiment's three phases (2,993, 2,993, and 2,996 trials). Entering the first spatial-memory concurrent-load condition (7A), Murph had recently emerged from a control condition in which he responded uncertain on 33.3% of trials across density levels 21–40. However, with the application of the new concurrent task, his uncertainty response was selectively eliminated leaving his sparse-dense discrimination intact. Now, he only used the uncertainty response on 3.6% of trials across those density levels, $t(2609) = 26.09$, $p < .05$. Once, again, this transition was essentially instantaneous. In the last 40 trials of his previous control condition, Murph had responded uncertain on 40% of trials across density levels 21–40. In the first 40 trials of his performance with an accompanying concurrent load, that uncertain response percentage was 4.5%, $t(45) = 3.29$, $p < .05$. Again we see that the uncertainty response and the primary perceptual responses in an uncertainty-monitoring task were differentially affected by the application of a concurrent load.

In the following spatial-memory control condition (7B), Murph used the uncertainty response significantly more than when the concurrent load had been present, on 14.5% of trials across density levels 21–40, $t(3950) = 12.18$, $p < .05$. Uncertainty responding was reduced again when the spatial-memory concurrent load was applied again (7C). Now Murph used the uncertainty response significantly less, on just 4.4% of trials across density levels 21–40, $t(3942) = 11.00$, $p < .05$.

Once again, the SDT model formalized these results. We fit the model to Murph's performances in Figures 7A and 7B. The model reproduced well Murph's performance with a concurrent load—the SSD fit measure was minimized at 0.1522 with an intuitive fit measure (AAD) of .018 (a 1.8% average prediction error). The model estimated that Murph placed his SU and UD criteria at Levels 30 and 32 along the 60-Level Sparse-Dense continuum, defining a narrow uncertainty region that afforded few uncertainty responses. For the control condition, the SSD fit measure was minimized at 0.1316 with an AAD of .015 (a 1.5% average prediction error). The model estimated that Murph placed his SU and

UD criteria at Levels 29 and 33 along the Sparse-Dense continuum, defining a larger uncertainty region. By the end of the control sessions, the uncertainty-response region had grown to nine steps wide, with SU and UD at Levels 25 and 34.

Given a concurrent load or no concurrent load, Murph's perceptual error was estimated to be 4.9 steps, Again Murph's detection sensitivity was unaffected by the concurrent load. Only his capacity to recruit the uncertainty response was curtailed by the concurrent-task requirement.

Figure 8 summarizes Murph's Experiment 2 data in 80-trial blocks as already described. Murph returned to control testing, and re-entered concurrent-load sessions, at Blocks 38 and 75. A distinctive feature of this summary is that Murph's uncertainty response gradually recovered to high levels during the control-testing blocks 38–74.

Lou—Lou's performance in the concurrent-load, spatial-memory task was as follows. For the first concurrent load test, the first control test, and the second concurrent load test, respectively, Lou's matching performance was 35.0% correct (33.55% to 37.93% across the delay intervals), 68.86% correct (67.97% to 69.53%), and 41.57% (38.81% to 44.88%). Lou's matching performance was reduced in the concurrent conditions of this experiment, but there was no large effect of delay length within a condition.

Figure 9,A–C shows Lou's performance in the sparse-uncertainty-dense task during all trials of the experiment's three phases (2997, 2995, and 2994 trials). During the first spatial-memory concurrent-load condition (9A), Lou responded uncertain on only .2% of trials across Levels 21–40 (4 times in 1,983 trials). This result is not probative, however, because Lou had emerged from a control condition within Experiment 1 during which he had also responded uncertain minimally. Figure 9B shows that his uncertainty responding did rebound a little when the concurrent load was removed. Now Lou responded uncertain on 4.2% of trials across density levels 21–40, significantly more than he had under concurrent-load conditions, $t(3971) = 8.70$, $p < .05$. By the end of Lou's control phase, he was responding uncertain on 15.6% of trials across density levels 21–40. Lou's uncertainty responding was reduced again when the spatial-memory concurrent load was applied again (9C). Now Lou used the uncertainty response significantly less, on 2.1% of trials across density levels 21–40, $t 3983) = -3.93$, $p < .05$.

Once again, the SDT model formalized these results. We fit the model to Lou's performance in Figures 9A and 9B. The model reproduced well Lou's performance with a concurrent load—the SSD fit measure was minimized at 0.0784 with an AAD of .009 (a 1% average prediction error). The model estimated that Lou placed his SU and UD criteria at Levels 30 and 31 along the 60-Level Sparse-Dense continuum, defining the narrowest possible uncertainty region. For the following control condition, the SSD fit measure was minimized at 0.0534 with an AAD of .008 (a .8% average prediction error). The model estimated that Lou placed his SU and UD criteria at Levels 30 and 32 along the Sparse-Dense continuum, defining a larger uncertainty-response region. By the end of the control phase, the uncertainty region spanned Levels 29 to 33.

Given a concurrent load or not, Lou's PE was estimated at 4.4 and 4.3 steps along the continuum. Lou's sensitive use of the Sparse and Dense responses were not affected by the presence or absence of working-memory demands. Only his capacity to recruit the uncertainty response was curtailed by the concurrent-task requirement.

Figure 8B summarizes Lou's Experiment 2 data in 80-trial blocks as already described. Lou returned to control testing, and re-entered concurrent-load sessions, at Blocks 38 and 75. As

with Murph, his uncertainty response substantially but gradually rebounded during the control-testing blocks 38–74. It is an interesting fact that both animals had to relearn to include the uncertainty response when the working-memory load was removed. In this, they might be very different from fully metacognitive humans.

Experiment 3

In Experiments 1 and 2, sparse and dense responses were robust to concurrent loads relative to the uncertainty response, perhaps because these primary perceptual responses required less executive attention or memory resources. One might predict that the primary response middle—if it could be instantiated as a perceptual response like sparse and dense—might also be robust to concurrent loads.

Accordingly, two macaques were tested in a sparse-middle-dense discrimination. They could use the middle response to positively identify stimuli in the continuum's middle range —the same stimuli that generally led macaques to respond uncertain in Experiments 1 and 2. If middle responses were more robust to concurrent loads than uncertainty responses, this would also suggest that uncertainty responses and perceptual responses in psychophysical tasks are psychologically different. It could help explain why Beran et al.'s (2009) capuchins made middle responses easily but uncertainty responses haltingly.

Several considerations attend this experiment's interpretation. First, Experiment 3's subjects had never participated in a three-choice discrimination like sparse-middle-dense. The middle response would be novel and difficult to learn to include in the behavioral repertoire, perhaps leaving it vulnerable to concurrent-load interference. Second, both subjects had had substantial experience in psychophysical paradigms—including sparse-dense discriminations—within which they were to use the uncertainty responses for stimuli near the continuum's middle (e.g., Smith et al., 2010). Based on this experience, the middle response could have been given a mixed middle-uncertainty construal that would also create a vulnerability to interference.

Third, experience/training aside, middle stimuli are inherently more decisionally difficult than sparse and dense stimuli. Consider a subject with a generous middle-response region spanning steps 26–36 along the continuum. Even an ideal middle stimulus (Level 31) would still be only 6 steps from being misperceived into either or two incorrect perceptual classes. Six steps is within the macaques' range of perceptual error based on the extensive modeling in this article. We ran a simulation to illustrate this point. We gave a simulated subject the 26–36 middle-response region and a Gaussian perceptual error of 6 continuum steps. We gave this subject 10,000 trials at each of 60 stimulus levels. We found the proportion of trials on which the subjective impression of the objective stimulus presented crossed over into another perceptual category (e.g., an objectively sparse stimulus perceived through perceptual error as middle or dense, an objectively middle stimulus perceived as sparse or dense, etc.). Figure 10 shows the result. The middle stimuli are nearly as difficult and confusable as the ultimately confusable stimuli right at the sparse-middle and middle-dense boundaries (Levels 26 and 36). They are far more difficult and confusable than the vast majority of Sparse and Dense stimuli. That the middle stimuli are structurally fraught with difficulty and confusability goes with the territory of this kind of psychophysical task.

Taken together, these considerations show that it is impossible to instantiate the middle response as a purely perceptual response exactly equivalent to the sparse and dense responses. We entered Experiment 3 knowing that its test of the perceptual-response hypothesis was highly conservative and that middle responses—themselves requiring substantial decisional and indeterminacy-resolution processes—would probably not be completely invulnerable to a concurrent load. Instead, we approached Experiment 3 as

exploratory, as a way of asking whether we could use a concurrent-task manipulation to illuminate possible psychological differences between uncertainty and middle responding.

Method

Subjects—Male macaques Hank and Gale (both 28 years old at the time of testing) participated. They had been trained as already described to perform computer-based cognitive tasks and had participated in many such tasks previously. In particular, they had had substantial experience in two-response psychophysical paradigms within which they were also allowed to use the uncertainty-response option (e.g., Smith et al., 2010). To our knowledge, these animals had never participated in a three-choice discrimination task like the present sparse-middle-dense task. The macaques were tested, and given food rewards and penalty timeouts, as described for Experiment 1.

Density continuum—The density continuum already described was also used for the sparse-middle-dense task, though the mapping of responses to stimuli was modified now. Levels 1–24, Levels 25–35, and Levels 36–60 of the 60 stimulus levels were assigned to the sparse middle, and dense classifications, respectively. The middle stimulus range was deliberately made narrower, to roughly match the width of the uncertain-response regions in Experiments 1 and 2, and to roughly equate the dominance of the middle and uncertainty responses in their respective tasks. Stimuli were always sampled randomly from the whole 60-step stimulus continuum.

Response modality—Macaques moved their joystick-controlled cursor to touch an S, M, or D icon on the screen to classify the boxes as sparse, middle, or dense, with the M icon appearing in the location the uncertainty icon (?) had occupied in Experiments 1 and 2. Each response produced positive, food-reward feedback (1 pellet) or negative, penalty-timeout feedback (30-second timeout) depending on the objective stimulus presented on that trial. There was no uncertainty-response option in this experiment. The middle response replaced it. Indeed, the middle response was appropriate for the same range of stimuli that had generally led macaques to voluntarily select the uncertainty response in Experiments 1 and 2.

Training the sparse-middle-dense discrimination task—The present task was the first in which the participating monkeys had been trained in a three-choice discrimination task wherein the stimulus continuum had to be divided in thirds using three responses. Accordingly, both monkeys worked on the task until—across three combined test sessions—they had produced middle response curves that approximated the baseline uncertainty-response curves produced by Murph and Lou in Experiment 1. To make this determination, we considered the peak of their middle response curve and its placement along the stimulus continuum. To reach this point, Hank completed 51,618 trials across 25 test sessions. Gale was trained for only about one fourth as many trials—13,920 training trials across 16 test sessions.

Matching-to-sample task—The identity matching to sample concurrent task already described in Experiment 1 was used to provide a concurrent cognitive load to the macaques in some sessions. The concurrent-matching trials and responses and the discrimination trials and responses were interleaved in the same way as described earlier. Correct and incorrect responses, respectively, led to a 2 pellet reward, and a 20 second timeout.

The delay between seeing the sample image in the matching task and when the matching response could be made was titrated within a session and started at 1 second, and then increased by one additional second when the animal was correct on 14 of the most recent 20

matching trials. This matched the training regimen used to teach these monkeys to match using these stimuli. Thus, delays could progress higher and higher throughout the session given performance on the matching task. These delays still occurred immediately after contact with the sample in the control condition, and immediately after the conclusion of the outcome of the discrimination trial in the concurrent-load condition (i.e., immediately after pellet delivery for a correct classification or the end of the timeout period following an incorrect classification response), so in these respects the interleaving of the tasks was identical to the procedure for Experiment 1.

Sequence of conditions—Monkeys first completed baseline training, then a concurrent-load phase of testing, then a control phase of testing, and then a second, concurrent-load phase of testing. Each phase for each macaque lasted about 3,000 trials.

Results

Hank—For the first concurrent load test, the first control test, and the second concurrent load test, respectively, Hank's performance on the delayed matching task was 73.3% correct, 87.97% correct, and 73.85% correct. Hank performed at high levels on the matching task throughout, although his performance was somewhat reduced during the concurrent load phases.

Figure 11,A–D shows Hank's performance in the sparse-middle-dense task during all trials of the experiment's four phases (3139, 2997, 2991, and 3000 trials). Hank performed adaptively in all phases, using the three responses to respond to stimuli in the three stimulus ranges designated sparse, middle, and dense.

In particular, Hank's middle responding appears to have fully survived the transition to the first concurrent-load condition and back again. In his baseline phase, his first concurrent-load phase, his control phase, and his second concurrent-load phase, respectively, he responded middle on 27.1%, 33.1%, 33.4%, and 33.2% of trials across levels 21–40. We also confirmed that Hank responded middle on 13.3% of trials across density levels 21–40 in the last 40 trials of his baseline condition, and on 22.2% of trials in the first 40 trials of his first concurrent-load condition, $t(22) = 0.79$, ns. Overall, Hank actually mildly increased his use of the Middle response during his first concurrent-load phase, $t(2018) = 2.96$, $p < .05$, but after that the rate of middle responding held constant.

Once again we formalized these results using the SDT model that has already been described. The model estimated that Hank in his initial baseline performance placed his SM and MD criteria at Levels 25 and 33, defining a middle-response region that was eight steps wide. For the following concurrent-load phase of performance, the estimated criteria were Levels 29 and 38, defining a slightly wider middle-response region. For the following two phases of performance, the estimated criteria were Levels 25 and 36 (control condition), and Levels 25 and 35 (second concurrent-load condition). On average, Hank's estimated middle-response region was 9.5 steps wide on average whether or not a concurrent load was imposed. There is no evidence that the imposition of the concurrent load affected his capacity to respond middle. Hank confirmed strongly and qualitatively that the psychophysical response middle can be more robust to cognitive loads than the uncertainty response measured in Experiments 1 and 2.

Figure 12A summarizes Hank's Experiment 3 data in 80-trial blocks as already described. Hank moved to concurrent-load testing, to control testing, and re-entered concurrent-load sessions, at Blocks 40, 77, and 114.

Gale—For the first concurrent load test, the first control test, and the second concurrent load test, respectively, Gale's performance on the delayed matching task was 68.06% correct, 79.89% correct, and 73.18% correct. Gale also performed at high levels on the matching task throughout, with performance only slightly reduced during the concurrent load phases.

Figure 13,A–D shows Gale's performance in the sparse-middle-dense task during all trials of the experiment's four phases (2694, 2998, 2993, and 2969 trials). In his baseline phase, his first concurrent-load phase, his control phase, and his second concurrent-load phase, respectively, he responded middle on 46.7%, 17.9%, 33.7%, and 26.4% of trials across levels 21–40. Gale also performed adaptively by responding middle robustly in all phases of testing, an observation not made for the uncertainty response in Experiments 1 and 2. However, there were significant differences in the level of middle responding as follows. Gale's middle responding did decrease during the concurrent-load testing, $t(1872) = 13.90$, $p < .05$, and increased during the control testing that followed, $t(1971) = 8.17$, $p < .05$. We also confirmed that Gale responded middle on 50% of trials across density levels 21–40 in the last 40 trials of his baseline condition, but only 7.7% of trials in the first 40 trials of his first concurrent-load condition, t (23) = 3.68, p < .05. Gale's middle responding also decreased slightly, from 33.7% to 26.4%, during the second period of concurrent-load testing compared to the previous period of control testing, $t(1972) = 3.55$, $p < .05$.

The SDT model estimated that Gale in his baseline, performance, in his first concurrent-load testing, in his control testing, and in his second concurrent-load testing, respectively, placed his SM and MD criteria at Levels 26–38, 29–35, 26–36, and 27–35 along the 60-step stimulus continuum. His middle-response regions were 12, 6, 10, and 8 steps wide. He maintained robust middle-response regions in all cases, as was not seen for the uncertainty response in Experiments 1 and 2, but the modeling still revealed that the concurrent load shrank those regions somewhat.

Figure 12B summarizes Gale's Experiment 3 data in 80-trial blocks as already described. Gale moved to concurrent-load testing, to control testing, and re-entered concurrent-load sessions, at Blocks 34, 71, and 108.

General Discussion

In several cases the concurrent loads sharply reduced uncertainty responding. In contrast, in Experiments 1 and 2, the concurrent load left intact the use of the primary perceptual responses sparse and dense. In Experiment 3, concurrent loads had no effect (Hank) or a relatively small effect (Gale) on the use of the primary perceptual response middle, despite the structural decisional difficulty that attends classifying middle stimuli.

Figure 14 illustrates the article's principal result. It shows the dramatic effect that concurrent-load testing had on uncertainty responding in Experiment 1 (Murph and Lou's performance averaged). It shows the small effect that concurrent-load testing had on middle responding in Experiment 3 (Hank and Gale's performance averaged). The differential effects of concurrent loads on uncertainty and middle responses are important because these responses were associated with the same stimuli along the same stimulus continuum. This makes it difficult to argue that the effect on uncertainty responses in Experiments 1 and 2 arose solely because those stimuli were more difficult than were the sparse and dense trials near the ends of the stimulus continuum. Perceptual difficulty may be part of the whole explanation, as Gale's data in Experiment 3 suggests, but not all of it. The dissociation shown by capuchins between uncertainty and middle responses (Beran et al., 2009) also shows that perceptual difficulty is not the whole story. Instead, the results suggest that uncertainty responses have a different psychological character than do primary, perceptual

responses. The uncertainty response may be more demanding of working-memory resources. It may be more executive and attentional in psychological character.

The present result mirrors human research in which participants performed primary memory tasks while reporting metacognitive tip-of-the-tongue states and judgments of learning (Schwartz, 2008). Schwartz found that concurrent working-memory loads strongly affected both kinds of metacognitive judgment, in particular sharply decreasing reports of tip-of-the-tongue experiences. Schwartz concluded that working memory and metamemory use similar monitoring processes, a conclusion analogous to that made here. The present result also mirrors human research by Paul, Smith, and Ashby (submitted) that used rapid event-related functional magnetic resonance imaging to explore the neural systems underlying humans' uncertainty responses as distinct from their primary perceptual responses. These scanning results indicated that the uncertainty response is not just a middle category, but rather that uncertainty-monitoring is mediated by a distinct network that includes anterior cingulate cortex, prefrontal cortical areas, and insula. Neuroimaging could make an important contribution to the comparative literature on this topic, to the extent it is feasible.

The present study constructively advances the theoretical arc of the animal-metacognition literature. Early studies in this area prompted concerns that associative mechanisms might explain animals' uncertainty responding. Accordingly, extensive further research was conducted to address these associative concerns. Animals transcended these associative explanations (review in Smith et al., 2012). The uncertainty response is preserved when direct rewards for uncertainty responses are omitted (Beran et al., 2006). It is preserved when deferred, rearranged reinforcement defeats the mechanisms of association and conditioning (Smith et al., 2006). These studies suggested that animals' uncertainty responses be given some higher-level, cognitive description. The present study begins to specify this higher-level description.

The present results also help explain a number of animal-metacognition results that were previously difficult to explain. Why is it that animals can invoke the uncertainty response on the first trial of new tasks (Washburn et al., 2006), or when a task's reinforcement contingencies are made invisible to the animal (Smith et al., 2006)? Why was a monkey able to respond uncertain when trans-cranial magnetic stimulation erased his visual working memory post trial (Washburn et al., 2010)? Why are animals able to show uncertainty monitoring while multitasking (Smith et al., 2010)? These results are more naturally explained if animals have an executive-level, decisional uncertainty system that resides actively online in the working-memory system and that is broadly available when the cognitive going gets tough.

On consideration, one can see that there are several reasons why uncertainty responses might plausibly be higher-level and more attentionally demanding?

First, primary perceptual responses like sparse, middle, and dense are directly rewarded when made correctly. Uncertainty responses are never directly rewarded or punished. Their benefit is indirect. Indeed, their benefit could be functionally invisible to a system that depends on the catalysis of dopamine reward signals to update neural connections allowing adaptive response patterns to be stamped in (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Smith, Beran, Crossley, Boomer, & Ashby, 2010).

Second, primary perceptual responses like sparse, middle, and dense are mapped to a particular class of perceptual inputs. The extents and boundaries of those perceptual categories are trained through positive reinforcement and punishment—that is, through the external task contingencies. In contrast, the uncertainty response is a response without a stimulus country. The extent and boundaries of the uncertainty-response region are whatever

the animal's cognitive system decides it should be. That extent and those boundaries must be adjudicated internally and cognitively through the perceiver's own regulatory processes.

Third, the structure of psychophysical uncertainty tasks means that subjects face many indeterminate perceptual representations (e.g., a Level 29 subjective impression). Shiffrin and Schneider (1977) explained why these indeterminate perceptual representations cannot reliably guide the organism's behavioral choices. Because of perceptual error, these impressions could have arisen equally from sparse or dense objective trials so neither response is clearly indicated. Shiffrin and Schneider argued that in this inconsistently mapped situation, the organism must engage higher-level, indeterminacy-resolution processes to fashion adaptive behavioral solutions. They called these higher-level processes controlled cognitive processes, by which they meant that these processes would be slower, serial, more deliberate, and more transparent to conscious awareness. The present study, that manipulated attentional resources and selectively compromised the uncertainty response, encourages the extension of Shiffrin and Schneider's ideas to the domain of discrimination tasks by suggesting that uncertainty responses near the limits of discrimination may be controlled processes in the sense of being attentionally and resource demanding. Of course uncertainty responses by macaques might possess some characteristics of controlled cognitive processing without possessing all of them.

The present results also illuminate the historical debate about uncertainty responses in the literature on human psychophysics. Behavioral theorists treated the uncertainty response as a middle response to middle stimuli. Cognitive theorists treated the uncertainty response as more decisional and reflective and less purely yoked to sensory magnitudes. The demonstration that uncertainty responses and sparse/middle/dense responses can be partially dissociated favors the latter viewpoint, by suggesting that uncertainty responses are more dependent on attentional resources because they are more challenged by a working-memory load.

The present results suggest solutions to some mysteries in the comparative literature on uncertainty monitoring. Beran et al. (2009) found that capuchin monkeys essentially would not make uncertainty responses. But when the task let them label the middle stimulus levels middle instead of responding uncertain to them, they made the middle response immediately at high levels (Figures 1A,B). This behavioral dissociation is predicted if uncertainty responding is a higher-level attentional process as the current results suggest. Thus, capuchin monkeys may have demonstrated in their uncertainty task a cognitive system lacking the attentional resources to respond appropriately to the indeterminacy signals offered by the uncertainty task.

In contrast, macaques make uncertainty responses relatively easily, sometimes nearly identically to human primates (Shields et al., 1997). Perhaps macaques, compared to capuchins, more easily invoke the higher-level decisional process that supports uncertainty responding. This hints at an intriguing divide in the cognitive evolution of the order *Primates*, a divide that is supported by converging lines of evidence from different laboratories using different uncertainty paradigms (Basile et al., 2009; Fujita, 2009; Paukner et al., 2006).

Finally, the present result may also help explain why macaques' and humans' uncertainty responses show the same fragility and changeability. In Smith et al. (1997) and Smith et al. (2006), the uncertainty response—but not the Sparse or Dense response—was behaviorally fragile, so that it fell out of the response repertoire of one macaque. Smith et al. (1997, 2006) observed exactly the same strong individual differences among humans in their use of the uncertainty response. Hampton (2001) found in his Experiment 3 that only one of two

macaques used the uncertainty response adaptively. Here, we observed again the special changeability and fragility of the uncertainty response. Changeability and fragility could confirm in a converging way that the uncertainty response is different from their primary discrimination responses. Indeed, it was those attributes that historically led some early psychophysicists to give humans' uncertainty responses a higher-level theoretical interpretation.

Yet even if one acknowledges that uncertainty tasks are inconsistently mapped, that uncertainty processes are controlled, and that the uncertainty response is executive and attentional, one may still question whether uncertainty responses have full status as metacognitive responses, and whether they are made with conscious awareness and self-knowledge. We endorse the importance of these questions that are sharpened to weaponry when they are asked of animals' uncertainty responses.

Nonetheless, it is a significant step forward to show here using the dissociative framework that macaques' uncertainty responses are controlled and attentionally demanding. Such responses are definitionally not reactive, not reflexive, and not associative in nature. Moreover, executive uncertainty responses may share some features —if not all features— with humans' explicit and conscious forms of on-line cognition and metacognition. The point is that the dissociative approach and complementary approaches are progressively elevating our interpretation of animals' uncertainty responses. Eventually, they might even lead to paradigms that could let researchers ask experimentally about awareness. This progressive elevation is a constructive, gradual form of theoretical development in the field.

In fact, we believe that gradualism in this area has many strengths. The all-or-none philosophical and behaviorist insistence against animal metacognition did not foster theoretical development. One reason is that metacognition may not be all or none. It may have different facets of behavioral expression and cognitive organization. Therefore, the best approach may be to specify which facets are present in which species, instead of always trying to answer the metacognition question Yes or No. In our view, this approach is best for studying the emergence of self-reflective mind in the primates, as it leads to full-fledged human metacognition. It may be best for exploring the earliest developmental roots of metacogniton, too. It could even have utility in other domains of metacognition research, including cognitive-aging and special-education research. For by asking what self-reflective capacities are preserved in different human populations, one might best tailor programs of training and remediation for them.

Acknowledgments

References

Anderson EJ, Mannan SK, Rees G, Sumner P, Kennard C. A role for spatial and nonspatial working memory processes in visual search. Experimental Psychology. 2008; 55:301–312.

Angell F. On judgments of "like" in discrimination experiments. American Journal of Psychology. 1907; 18:253.

Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A neuropsychological theory of multiple systems in category learning. Psychological Review. 1998; 105:442–481. [PubMed: 9697427]

Balcomb FK, Gerken L. Three-year-old children can access their own memory to guide responses on a visual matching task. Developmental Science. 2008; 11:t750–760.

Basile BM, Hampton RR, Suomi SJ, Murray EA. An assessment of memory awareness in tufted capuchin monkeys (*Cebus apella*). Animal Cognition. 2009; 12:169–180. [PubMed: 18712532]

Benjamin AS, Bjork RA, Schwartz BL. The mismeasure of memory: When retrieval fluency is misleading as a metacognitive index. Journal of Experimental Psychology: General. 1998; 127:55–68. [PubMed: 9503651]

Beran MJ. Rhesus monkeys (*Macaca mulatta*) succeed on a computerized test designed to assess conservation of discrete quantity. Animal Cognition. 2007; 10:37–45. [PubMed: 16868737]

Beran MJ. Monkeys (*Macaca mulatta* and *Cebus apella*) track, enumerate, and compare multiple sets of moving items. Journal of Experimental Psychology: Animal Behavior Processes. 2008; 34:63–74. [PubMed: 18248115]

Beran MJ, Harris EH, Evans TA, Klein ED, Chan B, Flemming TM, Washburn DA. Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*) and rhesus monkeys (*Macaca mulatta*): The effects of differential and nondifferential reward. Journal of Comparative Psychology. 2008; 122:52–61. [PubMed: 18298281]

Beran MJ, Smith JD. Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). Cognition. 2011; 120:90–105. [PubMed: 21459372]

Beran MJ, Smith JD, Coutinho MVC, Couchman JC, Boomer J. The psychological organization of "uncertainty" responses and "middle" responses: A dissociation in capuchin monkeys (*Cebus apella*). Journal of Experimental Psychology: Animal Behavior Processes. 2009; 35:371–381. [PubMed: 19594282]

Beran MJ, Smith JD, Redford JS, Washburn DA. Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. Journal of Experimental Psychology: Animal Behavior Processes. 2006; 32:111–119. [PubMed: 16634654]

Boring EG. The control of attitude in psychophysical experiments. Psychological Review. 1920; 27:440–452.

Bourke PA, Duncan J. Effect of template complexity on visual search and dual-task performance. Psychological Science. 2005; 16:208–213. [PubMed: 15733201]

Call J. Do apes know that they could be wrong? Animal Cognition. 201010.1007/s10071-010-0317-x

Call J, Carpenter M. Do apes and children know what they have seen? Animal Cognition. 2001; 4:207–220.

Couchman JJ, Coutinho MVC, Beran MJ, Smith JD. Beyond stimulus cues and reinforcement history: a new approach to animal metacognition. Journal of Comparative Psychology. 2010

Crystal JD, Foote AL. Metacognition in animals. Comparative Cognition and Behavior Reviews. 2009; 4:1–16.

Dunlosky, J.; Bjork, RA., editors. Handbook of memory and metamemory. New York: Psychology Press; 2008.

Fernberger SW. The effect of the attitude of the subject upon the measure of sensitivity. American Journal of Psychology. 1914; 25:538–543.

Fernberger SW. The use of equality judgments in psychophysical procedures. Psychological Review. 1930; 37:107–112.

Foote A, Crystal J. Metacognition in the rat. Current Biology. 2007; 17:551–555. [PubMed: 17346969]

Flavell JH. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. American Psychologist. 1979; 34:906–911.

Fujita K. Metamemory in tufted capuchin monkeys (*Cebus apella*). Animal Cognition. 2009; 12:575–585. [PubMed: 19242741]

Gallup GG. Self-awareness and the emergence of mind in primates. American Journal of Primatology. 1982; 2:237–248.

George SS. Attitude in relation to the psychophysical judgment. American Journal of Psychology. 1917; 28:1–38.

Hampton RR. Rhesus monkeys know when they remember. Proceedings of the National Academy of Sciences. 2001; 98:5359–5362.

Hampton RR. Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? Comparative Cognition and Behavior Reviews. 2009; 4:17–28. [PubMed: 20046911]

Hays, WL. Statistics. New York: Holt, Rinehart, and Winston; 1981.

He J, McCarley JS. Executive working memory load does not compromise perceptual processing during visual search: Evidence from additive factors analysis. Attention, Perception, & Psychophysics. 2010; 72:308–316.

Jastrow J. A critique of psycho-physic methods. American Journal of Psychology. 1888; 1:271–309.

Jolicoeur P. Dual-task interference and visual encoding. Journal of Experimental Psychology: Human Perception & Performance. 1999; 25:596–616.

Jozefowiez J, Staddon JER, Cerutti D. Metacognition in animals: How do we know that they know? Comparative Cognition and Behavior Reviews. 2009; 4:29–39.

Kiesel A, Steinhauser M, Wendt M, Falkenstein M, Jost K, Philipp AM, Koch I. Control and interference in task switching-A review. Psychological Bulletin. 2010; 136:849–874. [PubMed: 20804238]

Koenker, RH. Simplified statistics. Totowa, NJ: Littlefield, Adams, & Co; 1981.

Koriat, A. Metacognition and consciousness. In: Zelazo, PD.; Moscovitch, M.; Thompson, E., editors. The Cambridge handbook of consciousness. Cambridge, UK: Cambridge University Press; 2007. p. 289-325.

Koriat A, Goldsmith M. Memory in naturalistic and laboratory contexts: distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. Journal of Experimental Psychology: General. 1994; 123:297–315. [PubMed: 7931094]

Kornell N. Metacognition in humans and animals. Current Directions in Cognitive Science. 2009; 18:11–15.

Kornell N, Son L, Terrace H. Transfer of metacognitive skills and hint seeking in monkeys. Psychological Science. 2007; 18:64–71. [PubMed: 17362380]

Liu S, Dosher BA, Lu Z. The role of judgment frames and task precision in object attention: Reduced template sharpness limits dual-object performance. Vision Research. 2009; 49:1336–1351. [PubMed: 18817804]

Logan GD. Attention in character-classification tasks: Evidence for the automaticity of component stages. Journal of Experimental Psychology: General. 1978; 107:32–63.

Macmillan, NA.; Creelman, CD. Detection theory: A user's guide. 2. Mahwah, N.J: Lawrence Erlbaum Associates; 2005.

Morgan, CL. An introduction to comparative psychology. London: Walter Scott; 1906.

Nelson, TO., editor. Metacognition: Core readings. Toronto: Allyn and Bacon; 1992.

Nelson TO. Consciousness and metacognition. American Psychologist. 1996; 51:102–116.

Pannebakker MM, Band GPH, Ridderinkhof KR. Operation compatibility: A neglected contribution to dual-task costs. Journal of Experimental Psychology: Human Perception & Performance. 2009; 35:447–460. [PubMed: 19331500]

Pashler HE. Processing stages in overlapping tasks: Evidence for a central bottleneck. Journal of Experimental Psychology: Human Perception & Performance. 1984; 10:358–377. [PubMed: 6242412]

Pashler HE. Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. Cognitive Psychology. 1989; 21:469–514.

Pashler HE. Dual-task interference in simple tasks: Data and theory. Psychological Bulletin. 1994; 116:220–244. [PubMed: 7972591]

Paukner A, Anderson JR, Fujita K. Redundant food searches by capuchin monkeys (*Cebus apella*): A failure of metacognition? Animal Cognition. 2006; 9:110–117. [PubMed: 16184375]

Paul, EJ.; Smith, JD.; Ashby, FG. Neural correlates of uncertainty monitoring in human categorization. Under review

Peterson MS, Beck MR, Wong JH. Were you paying attention to where you looked? The role of executive working memory in visual search. Psychonomic Bulletin & Review. 2008; 15:372–377. [PubMed: 18488654]

Roberts WA, Feeney MC, McMillan N, MacPherson K, Musolino E. Do pigeons (*Columba livia*) study for a test? Journal of Experimental Psychology: Animal Behavior Processes. 2009; 35:129–142. [PubMed: 19364222]

Rumbaugh DM, Richardson WK, Washburn DA, Savage-Rumbaugh ES, Hopkins WD. Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. Journal of Comparative Psychology. 1989; 103:32–38. [PubMed: 2924530]

Scheck P, Nelson TO. Lack of pervasiveness of the underconfidence-with-practice effect: Boundary conditions and an explanation via anchoring. Journal of Experimental Psychology: General. 2005; 134:124–128. [PubMed: 15702968]

Schwartz BL. Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgments. Memory & Cognition. 2008; 36:9–19.

Shields WE, Smith JD, Washburn DA. Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. Journal of Experimental Psychology: General. 1997; 126:147–164. [PubMed: 9163934]

Shiffrin RM, Schneider W. Controlled and automatic human information processing: 11. Perceptual learning, automatic attending and a general theory. Psychological Review. 1977; 84:127–190.

Smith JD. The study of animal metacognition. Trends in Cognitive Sciences. 2009; 13:389–396. [PubMed: 19726218]

Smith JD, Beran MJ, Redford JS, Washburn DA. Dissociating uncertainty states and reinforcement signals in the comparative study of metacognition. Journal of Experimental Psychology: General. 2006; 135:282–297. [PubMed: 16719654]

Smith, JD.; Beran, MJ.; Couchman, JJ. Animal metacognition. In: Zentall, T.; Wasserman, E., editors. Comparative cognition: Experimental explorations of animal intelligence. Oxford, UK: Oxford University Press; 2012. p. 282-304.

Smith JD, Beran MJ, Coutinho MVC, Couchman JC. The comparative study of metacognition: Sharper paradigms, safer inferences. Psychonomic Bulletin and Review. 2008; 15:679–691. [PubMed: 18792496]

Smith JD, Beran MJ, Crossley MJ, Boomer J, Ashby FG. Implicit and explicit category learning by macaques (*Macaca mulatta*) and humans (*Homo sapiens*). Journal of Experimental Psychology: Animal Behavior Processes. 2010; 36:54–65. [PubMed: 20141317]

Smith JD, Redford JS, Beran MJ, Washburn DA. Monkeys adaptively monitor uncertainty while multi-tasking. Animal Cognition. 2010; 13:93–101. [PubMed: 19526256]

Smith JD, Redford JS, Haas SM. Prototype abstraction by monkeys (*Macaca mulatta)*. Journal of Experimental Psychology: General. 2008; 137:390–401. [PubMed: 18473665]

Smith JD, Redford JS, Haas SM. The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). Journal of Experimental Psychology: Animal Behavior Processes. 2008; 34:361–374. [PubMed: 18665719]

Smith JD, Schull J, Strote J, McGee K, Egnor R, Erb L. The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). Journal of Experimental Psychology: General. 1995; 124:391–408. [PubMed: 8530911]

Smith JD, Shields WE, Allendoerfer KR, Washburn WA. Memory monitoring by animals and humans. Journal of Experimental Psychology: General. 1998; 127:227–250. [PubMed: 9742715]

Smith JD, Shields WE, Schull J, Washburn DA. The uncertain response in humans and animals. Cognition. 1997; 62:75–97. [PubMed: 8997171]

Sternberg S. The discovery of processing stages: Extensions of Donders' method. Acta Psychologica. 1969; 30:276–315.

Suda-King C. Do orangutans (*Pongo pygmaeus*) know when they do not remember? Animal Cognition. 2008; 7:239–246.

Sutton JE, Shettleworth SJ. Memory without awareness: Pigeons do not show metamemory in delayed matching-to-sample. Journal of Experimental Psychology: Animal Behavior Processes. 2008; 34:266–282. [PubMed: 18426309]

Thomson GH. A new point of view in the interpretation of threshold measurements in psychophysics. Psychological Review. 1920; 27:300–307.

Tolman EC. A behaviorist's definition of consciousness. Psychological Review. 1927; 34:433–439.

Urban FM. The method of constant stimuli and its generalizations. Psychological Review. 1910; 17:229–59.

Washburn DA, Gulledge JG, Beran MJ, Smith JD. With his memory magnetically erased, a monkey knows he is uncertain. Biology Letters. 2010; 6:160–162. [PubMed: 19864272]

Washburn DA, Rumbaugh DM. Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. Behavior Research Methods, Instruments, and Computers. 1992; 24:157–164.

Washburn DA, Smith JD, Shields WE. Rhesus Monkeys (*Macaca mulatta*) immediately generalize the uncertain response. Journal of Experimental Psychology: Animal Behavior Processes. 2006; 32:85–89.

Watson CS, Kellogg SC, Kawanishi DT, Lucas PA. The *uncertain* response in detection-oriented psychophysics. Journal of Experimental Psychology. 1973; 99:180–185.

Woodworth, RS. Experimental psychology. New York: Holt; 1938.

A. Capuchins
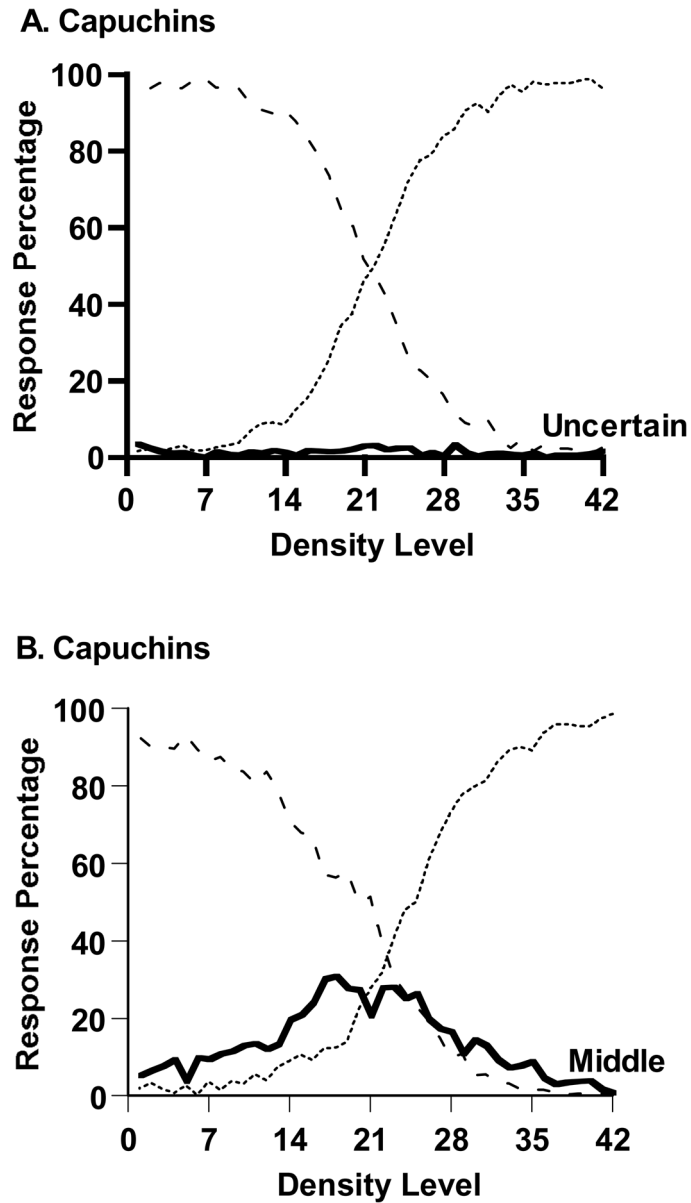
B. Capuchins

Figure 1.
A. Percentage of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by capuchin monkeys (*Cebus apella*) in Beran et al.'s (2009) Sparse-Uncertain-Dense task. *B*. Percentage of middle responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by capuchins in Beran et al.'s Sparse-Middle-Dense task.

A. Murph Baseline

B. Murph Concurrent 1

C. Murph Control 1

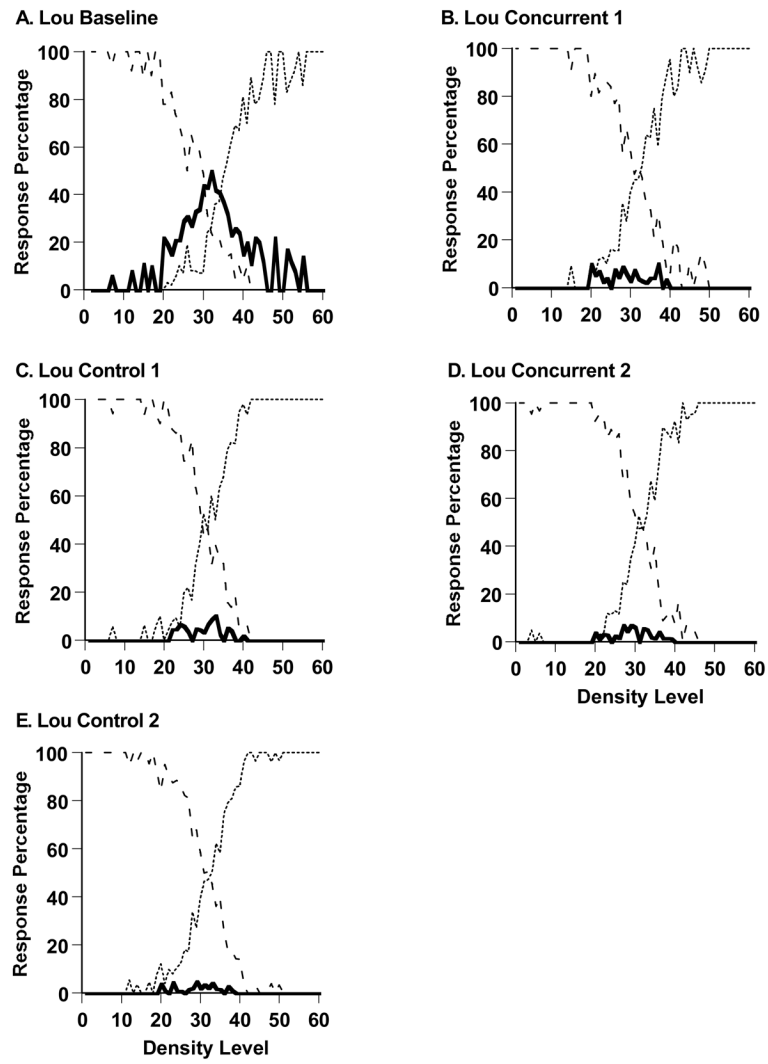### D. Murph Concurrent 2

E. Murph Control 2
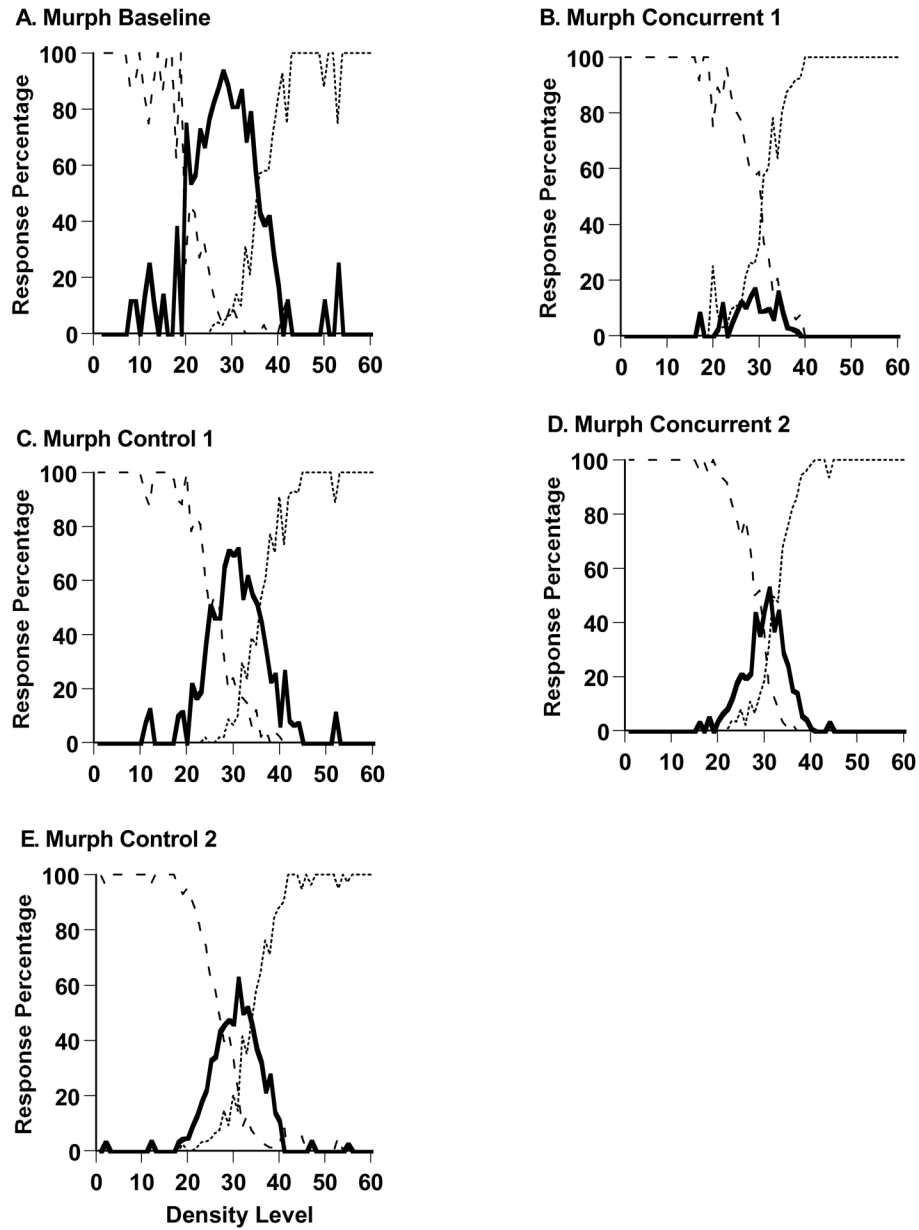
Figure 2.
Percentage of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by the macaque Murph in Experiment 1's Sparse-Uncertain-Dense task. *A*. Baseline performance. *B,D*. Performance with an shape-memory concurrent load. *C,E*. Performance with no shape-memory concurrent load.

A. Murph Baseline (Simulated)

B. Murph Concurrrent 1 (Simulated)
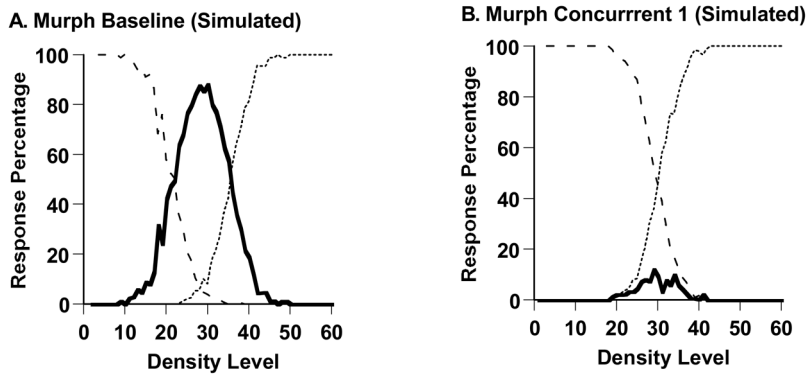
Figure 3.
A. The best-fitting predicted performance profile to the macaque Murph's performance
(Figure 2A) in a Sparse-Uncertain-Dense task with no concurrent shape-memory
requirement. The percentages of uncertainty responses (solid line), sparse responses (dashed
line), and dense responses (dotted line) are shown. The data were fit using formal-modeling
procedures specified in the text. *B.* The best-fitting predicted performance profile to the
macaque Murph's performance (Figure 2B) in a Sparse-Uncertain-Dense task with a
concurrent shape-memory requirement. The data were modeled and are depicted in the same
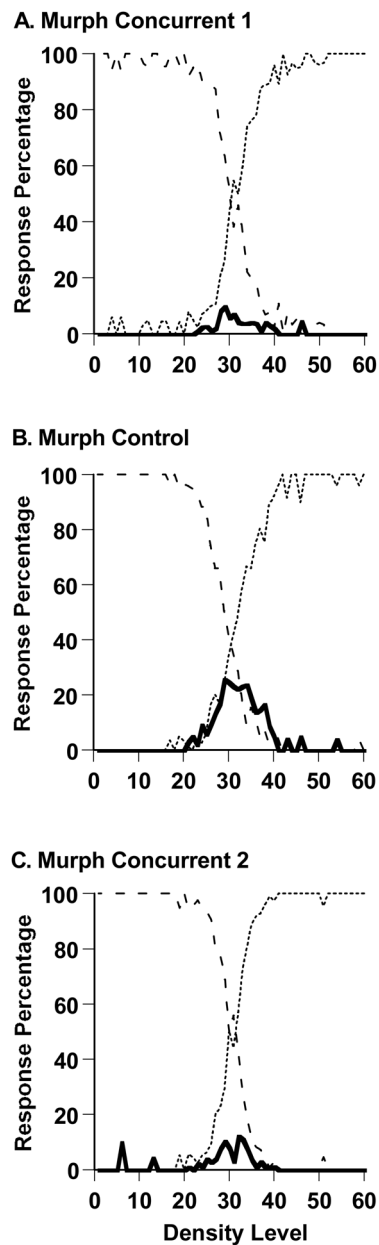way.

A. Murph Experiment 1
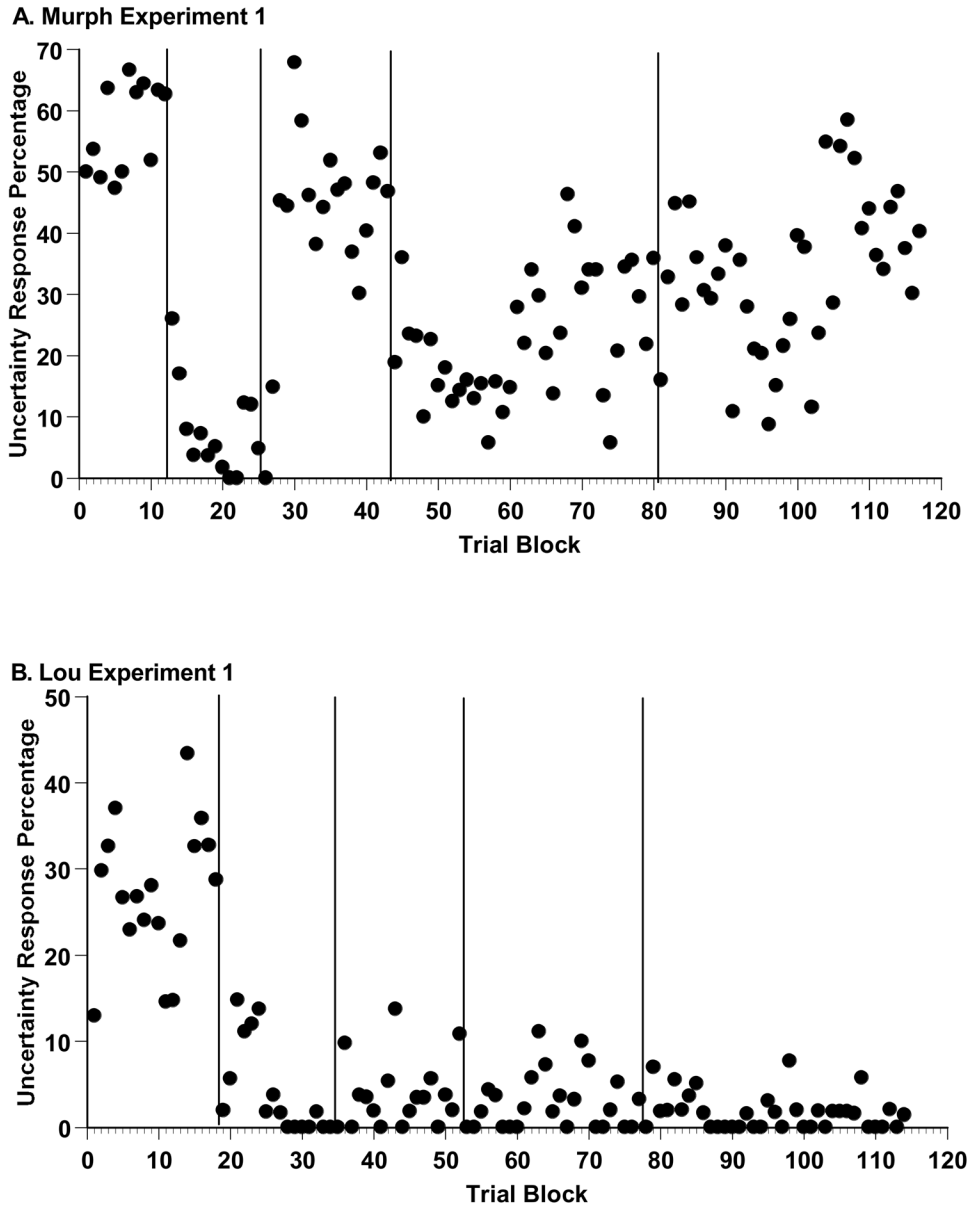
B. Lou Experiment 1

Figure 4.
A. Murph's uncertainty-response percentage for trial levels 21–40, by sequential trial block, in the five phases of Experiment 1 (from left to right: baseline, concurrent-load 1, control 1, concurrent-load 2, and control 2). The Y-axis is scaled to the maximum uncertainty-response percentage. *B.* Lou's uncertainty-response percentages, depicted in the same way.
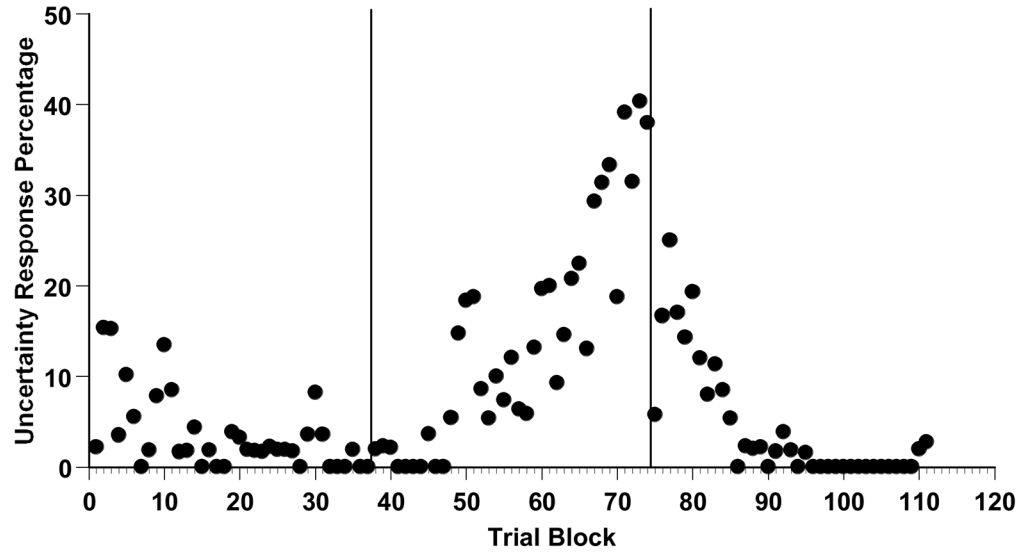
Figure 5.
Percentage of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by the macaque Lou in Experiment 1's Sparse-Uncertain-Dense task. *A.* Baseline performance. *B.* Performance with a shape-memory concurrent load. *C.* Performance with no shape-memory concurrent load.

A. Lou Baseline (Simulated)

B. Lou Concurrent 1 (Simulated)
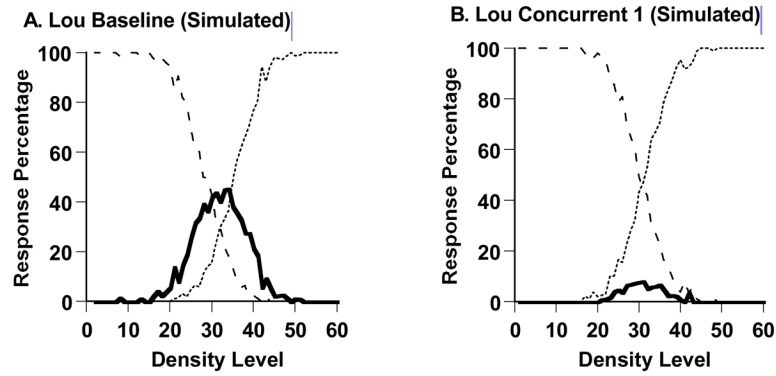
Figure 6.

A. The best-fitting predicted performance profile to the macaque Lou's performance (Figure 5A) in a Sparse-Uncertain-Dense task with no concurrent shape-memory requirement. The percentages of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) are shown. The data were fit using formal-modeling procedures specified in the text. *B*. The best-fitting predicted performance profile to the macaque Lou's performance (Figure 5B) in a Sparse-Uncertain-Dense task with a concurrent shape-memory requirement. The data were modeled and are depicted in the same way.

Figure 7.
Percentage of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by the macaque Murph in Experiment 2's Sparse-Uncertain-Dense task. *A*. Performance with a spatial-memory concurrent load. *B*. Performance with no spatial-memory concurrent load. *C*. Performance with a spatial-memory concurrent load.

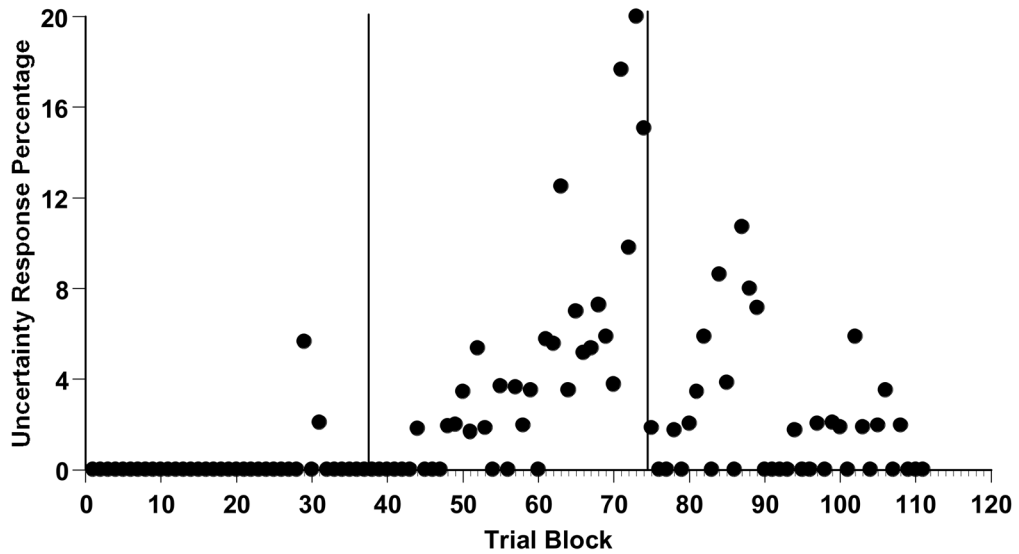**A. Murph Experiment 2**

B. Lou Experiment 2

Figure 8.
A. Murph's uncertainty-response percentage for trial levels 21–40, by sequential trial block in the three phases of Experiment 2 (from left to right: concurrent-load 1, control 1, concurrent-load 2). The Y-axis is scaled to his maximum uncertainty-response percentage. *B.* Lou's uncertainty-response percentages, depicted in the same way.
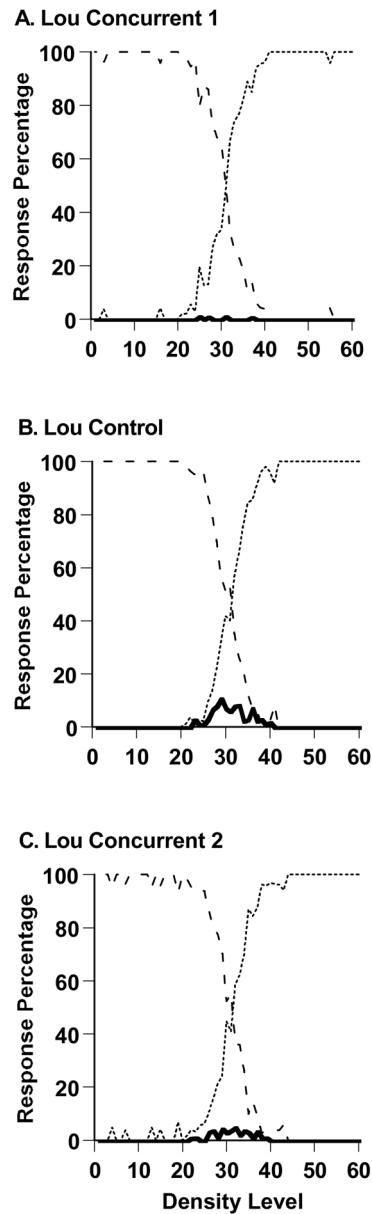
A. Lou Concurrent 1

B. Lou Control

C. Lou Concurrent 2

Figure 9.
A. Percentage of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by the macaque Lou in Experiment 2's Sparse-Uncertain-Dense task. *A.* Performance with a spatial-memory concurrent load. *B.* Performance with no spatial-memory concurrent load. *C.* Performance with a spatial-memory concurrent load.
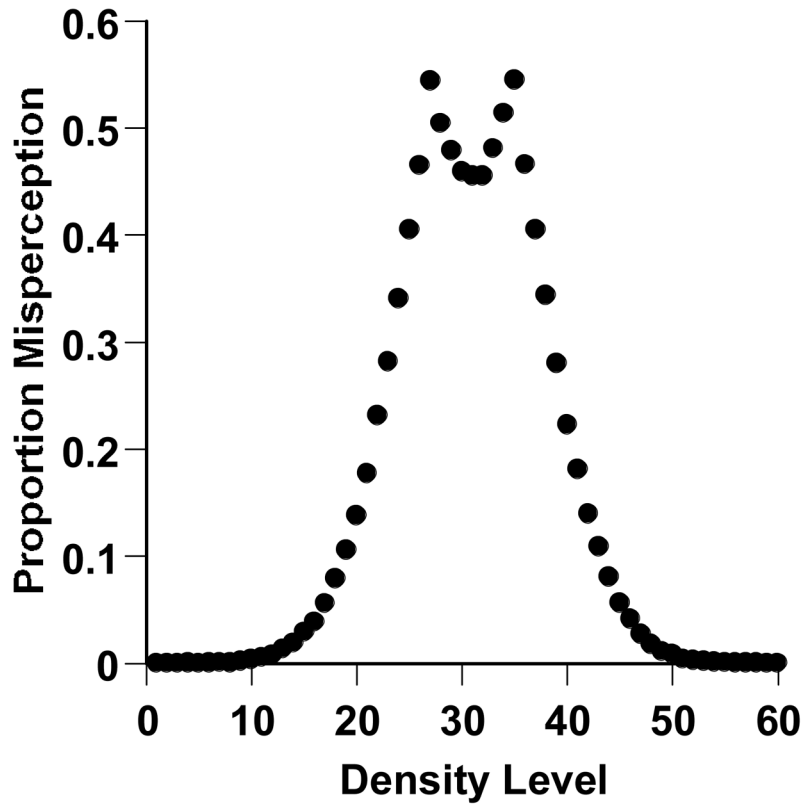
Figure 10.
A formal simulation illuminating the psychophysical situation in the sparse-middle-dense task of Experiment 3. The simulated subject performed 600,000 trials in that task with a Gaussian perceptual error of 6 steps along the continuum and with sparse-uncertain and uncertain-dense criteria, respectively, of 26 and 36. The graph shows the proportion of all trials at each density level on which the stimulus would have been misclassified into the wrong perceptual category (e.g., an objectively middle stimulus perceived and classified as sparse or dense).
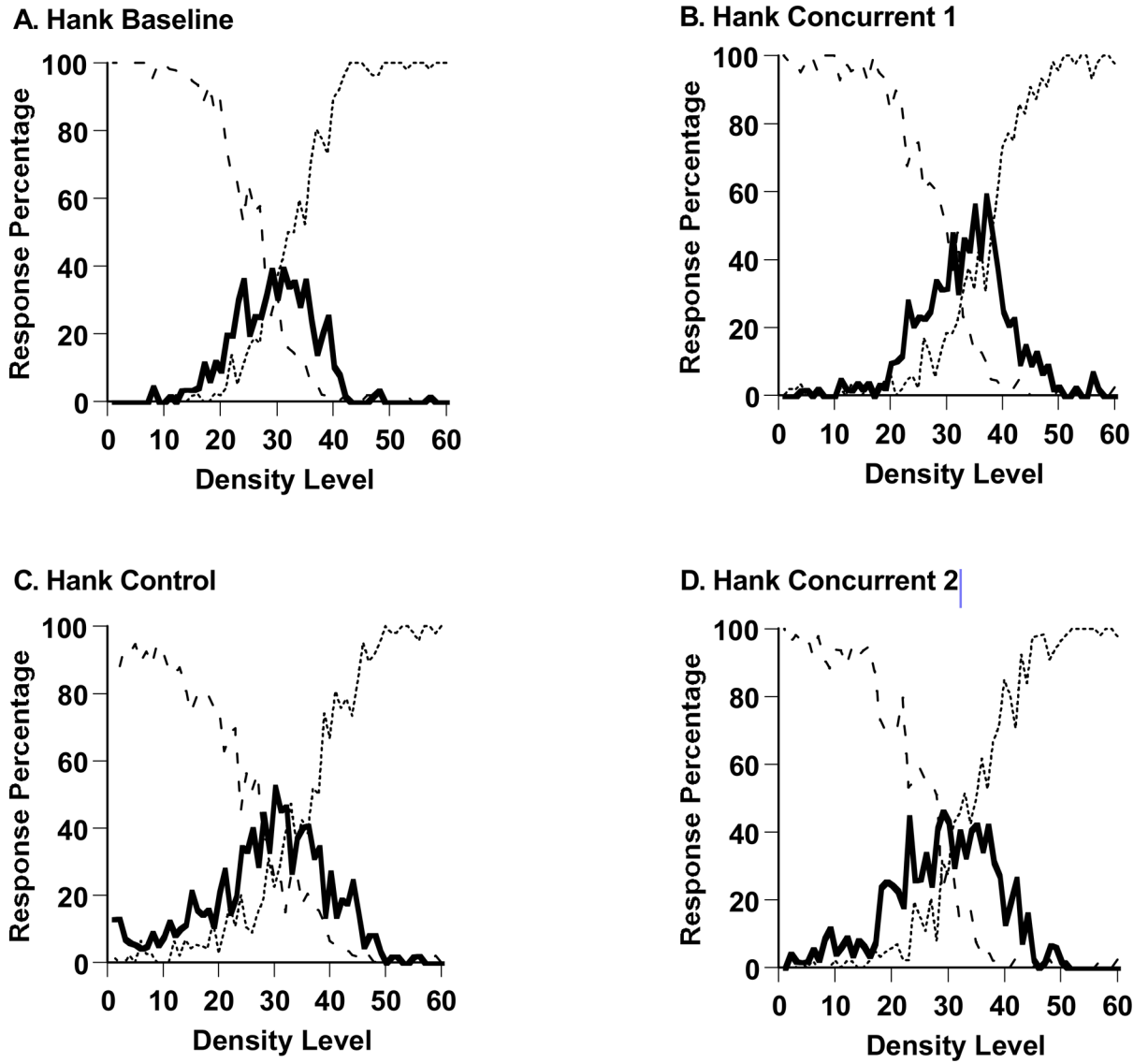
A. Hank Baseline

B. Hank Concurrent 1

C. Hank Control

D. Hank Concurrent 2

Figure 11.
Percentage of middle responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by the macaque Hank in Experiment 3's Sparse-Middle-Dense task. *A,C.* Performance without a concurrent cognitive load. *B,D.* Performance with a concurrent cognitive load.
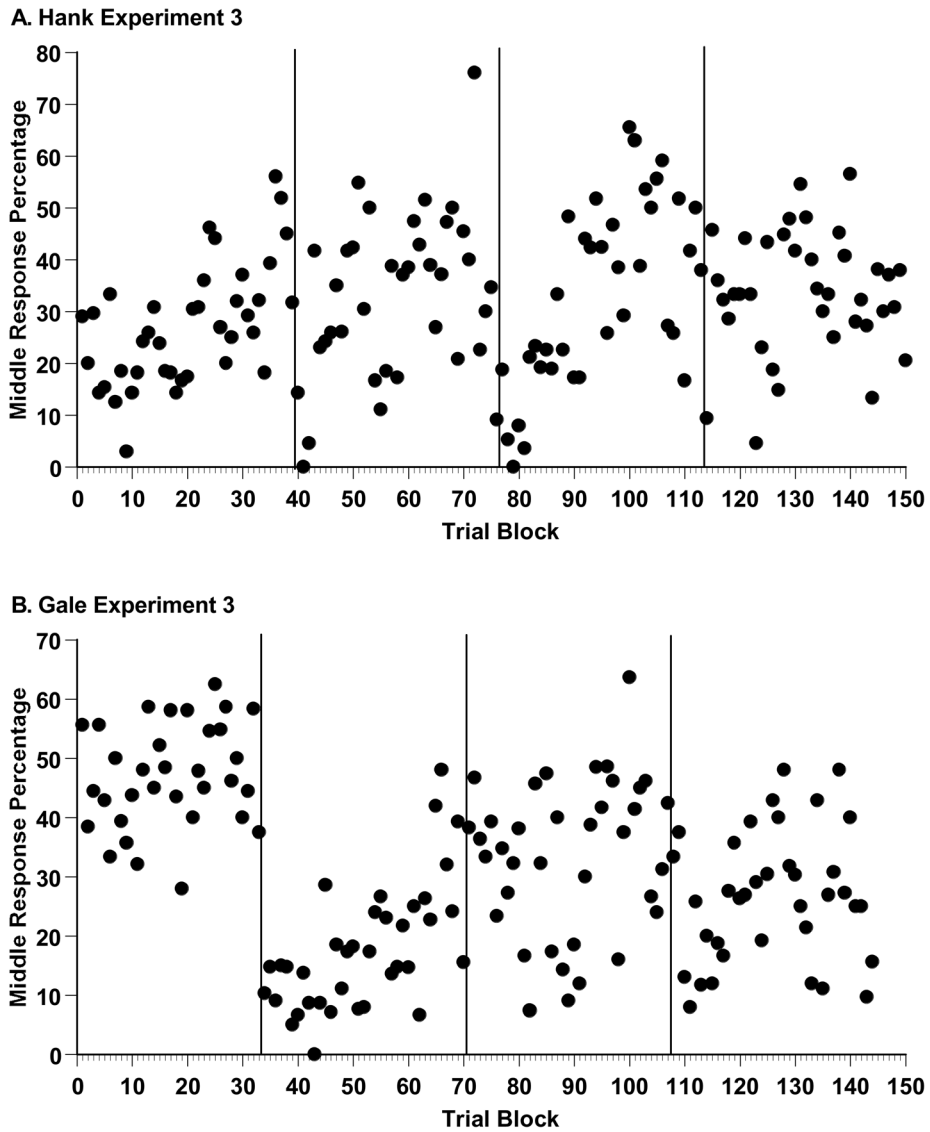
Figure 12.
A. Hank's middle-response percentage for trial levels 21–40, by sequential trial block in the four phases of Experiment 2 (from left to right: baseline, concurrent-load 1, control, concurrent-load 2). The Y-axis is scaled to his maximum middle-response percentage. *B.* Gale's middle-response percentages, depicted in the same way.
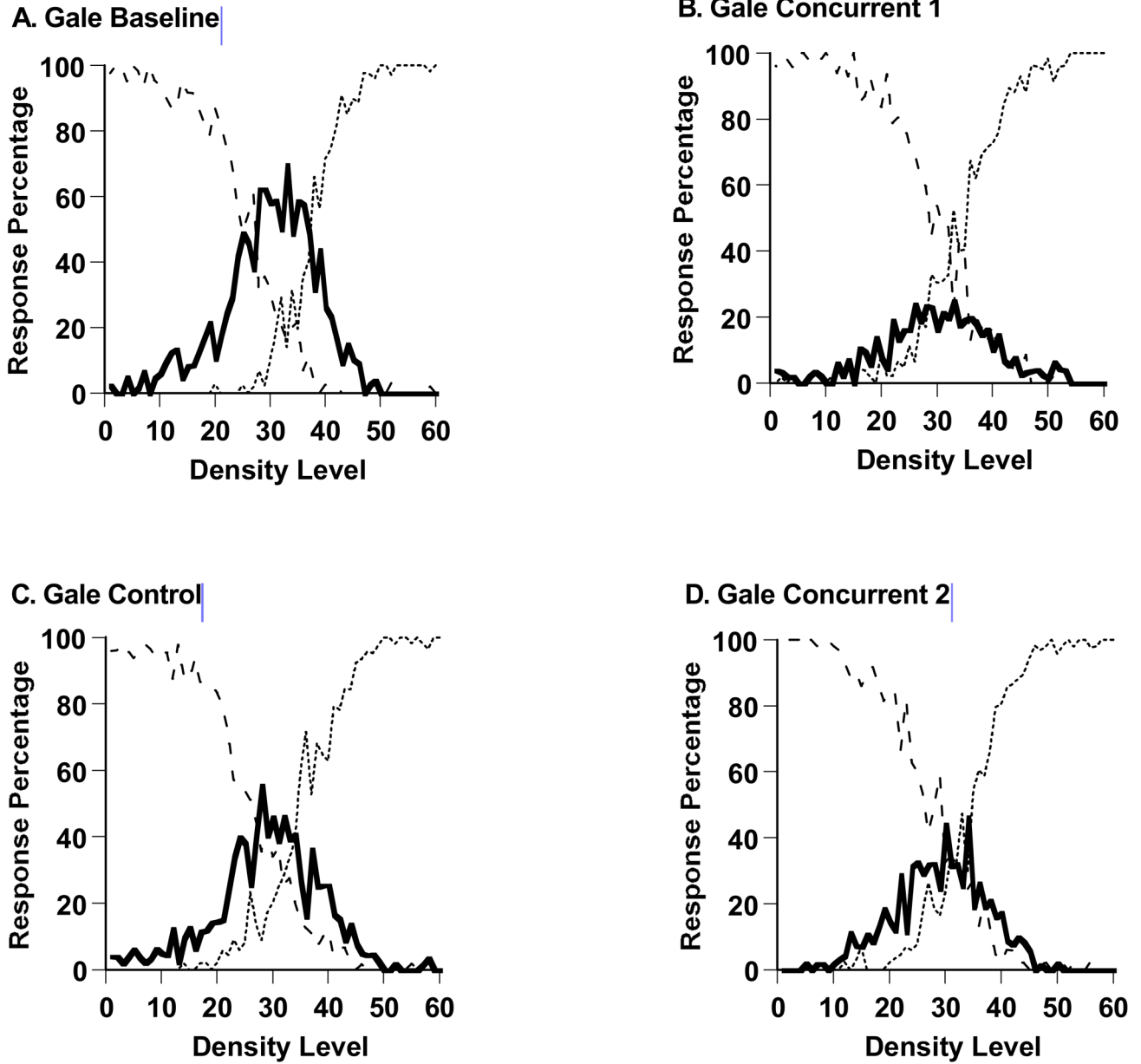
A. Gale Baseline

B. Gale Concurrent 1

C. Gale Control

D. Gale Concurrent 2

Figure 13.
Percentage of middle responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by the macaque Gale in Experiment 3's Sparse-Middle-Dense task. *A,C*. Performance without a concurrent cognitive load. *B,D*. Performance with a concurrent cognitive load.

A. Uncertainty Baseline

B. Uncertainty Concurrent 1

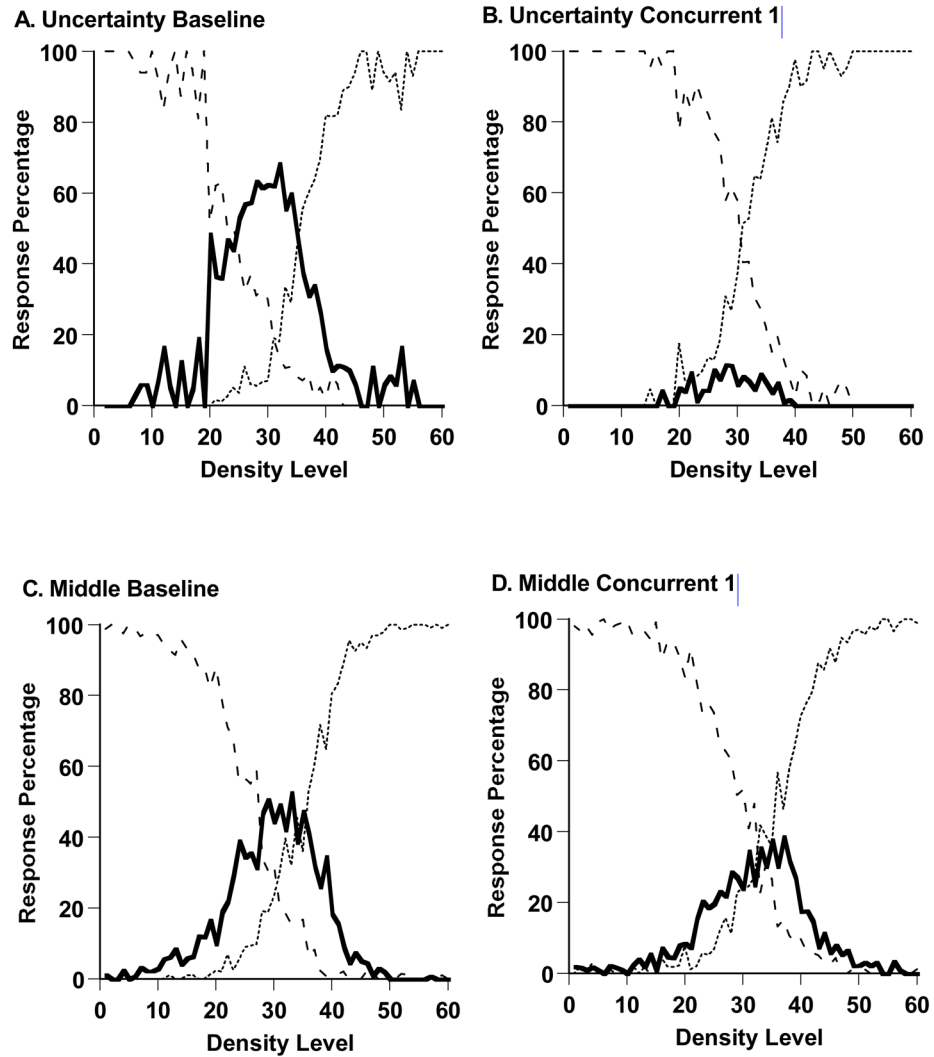**C. Middle Baseline**

D. Middle Concurrent 1

Figure 14.
A,B. Percentage of uncertainty responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by macaques Murph and Lou in their baseline performance and in their first phase of concurrent-load testing. *C,D.* Percentage of middle responses (solid line), sparse responses (dashed line), and dense responses (dotted line) made by macaques Hank and Gale in their baseline performance and in their first phase of concurrent-load testing.