

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

2009

Volume 4, pp 40 -53

Animal Metacognition: Problems and Prospects

J. David Smith

Department of Psychology and Center for Cognitive Science, SUNY Buffalo

Michael J. Beran

Language Research Center, Georgia State University

Justin J. Couchman, Mariana V. C. Coutinho, Joseph B. Boomer

Department of Psychology, SUNY Buffalo

Researchers have begun to evaluate whether nonhuman animals share humans' capacity for metacognitive monitoring and self-regulation. Using perception, memory, numerical, and foraging paradigms, they have tested apes, capuchins, a dolphin, macaques, pigeons, and rats. However, recent theoretical and formal-modeling work has confirmed that some paradigms allow the criticism that low-level associative mechanisms could create the appearance of uncertainty monitoring in animals. This possibility has become a central issue as researchers reflect on existing phenomena and pause to evaluate the area's current status. The present authors discuss the associative question and offer our evaluation of the field. Associative mechanisms explain poorly some of the area's important results. The next phase of research in this area should consolidate the gains achieved by those results and work toward a theoretical understanding of the cognitive and decisional (not associative) capacities that animals show in some of the referent experiments.

Keywords: uncertainty monitoring, metacognition, comparative cognition, decision making.

J. David Smith, Department of Psychology and Center for Cognitive Science, University at Buffalo, the State University of New York; Michael J. Beran, Language Research Center, Georgia State University; Justin Couchman, Mariana Coutinho, and Joseph Boomer, Department of Psychology, University at Buffalo, the State University of New York.

Correspondence concerning this article can be addressed to

J. David Smith, 346 Park Hall, Department of Psychology, SUNY Buffalo, Buffalo, NY, 14260 psysmith@buffalo.edu

Michael J. Beran, Language Research Center, Georgia State University, 3401 Panthersville Rd., Decatur, GA 30034 mjberan@yahoo.com

Justin J. Couchman, 384 Park Hall, Department of Psychology, SUNY Buffalo, Buffalo NY, 14260 jjc38@buffalo.edu

Mariana V. C. Coutinho, 384 Park Hall, Department of Psychology, SUNY Buffalo, Buffalo NY, 14260 mvc5@buffalo.edu

Joseph B. Boomer, 384 Park Hall, Department of Psychology, SUNY Buffalo, Buffalo NY, 14260 jboomer@buffalo.edu

The preparation of this article was supported by Grant HD-38051 from NICHD and by Grant BCS-0634662 from NSF.

Humans feel uncertain. They know when they do not know or remember, and they respond well to uncertainty by deferring response and seeking help or information. Their responses to doubt and uncertainty ground research on metacognition, or thinking about thinking (Benjamin et al., 1998; Flavell, 1979; Koriat, 1995; 2007; Koriat et al., 2006; Metcalfe, 2000; Nelson, 1992; Schwartz, 1994; Serra & Dunlosky, 2005). The idea in this field is that human minds have a cognitive executive that looks in on thought to evaluate its problems and prospects.

When humans behave metacognitively, scientists make intriguing attributions about their minds. Metacognition reveals hierarchical structure in mind (Nelson & Narens, 1990) because the executive oversees the rest of cognition. Metacognition is linked to self-awareness (Gallup, 1982) because doubt is so personal and self-oriented. Metacognition is linked to declarative consciousness (Koriat, 2007; Nelson, 1996) because we can introspect and declare states of knowing. Thus, metacognition is a sophisticated capacity in humans that might be uniquely human (Metcalfe & Kober,

2005). This raises the question of whether nonhuman animals share that capacity.

To address this question, researchers have evaluated whether animals can adaptively monitor and respond to their uncertainty (e.g., Beran et al., 2006; Foote & Crystal, 2007; Hampton, 2001; Kornell et al., 2007; Smith et al., 1995, 1998, 2006; Suda-King, 2008; Sutton & Shettleworth, 2008; Washburn et al., 2006). The paradigms used in this research have two key features. First, they contain a mix of easy and difficult trials, the latter trial type so that researchers may stir up something like an uncertainty state. Second, they grant animals an additional response—beyond the primary discrimination responses—with which they can decline to complete trials of their choosing. This potentially lets them handle uncertainty adaptively. Animals who monitor cognition should recognize difficult trials as error-risking and decline them selectively. Some animals do so, producing data patterns in cognitive-monitoring tasks that are strikingly like those of humans. This additional response is called the uncertainty response, and it is presently interpreted to show some species' capacity for uncertainty monitoring and metacognition.

If this interpretation is correct, then these experiments tap theoretically important cognitive capacities in animals that could bear on their cognitive awareness and consciousness. However, the burden of proof in this area is heavy. Morgan's (1906) Canon established a tradition in comparative psychology of erring toward explaining animals' behavior at the lowest possible psychological level. Therefore, even when animals perform in a way that might demonstrate metacognition, researchers should consider carefully the alternative possibility that these performances can be explained through low-level, associative mechanisms based in stimulus cues and reinforcement contingencies. Morgan's Canon has seldom had a fatter target to shoot at than animal metacognition. This theoretical debate has become the area's central issue as researchers reflect on existing phenomena and pause to evaluate the area's current status (Carruthers, 2008; Proust, 2003; Smith et al., 2008; Staddon et al., 2007). Here, we discuss this debate and offer our evaluation of the field. We appreciate the opportunity for a dialog on this important issue.

Criticism begins at home. We will illustrate the interpretative tension in this field using the paradigm that initiated the comparative study of uncertainty monitoring (Smith et al., 1995). Smith et al. gave a bottlenosed dolphin (Figure 1) a psychophysical threshold task. One response (High) was correct for a repeating tone of 2,100 Hz. The other response (Low) was correct for a repeating tone between 1,200 Hz and 2,099 Hz. In addition to the High and Low responses, the animal could make a trial-decline response that replaced

the current trial with an easier one. Difficulty was varied so as to map the dolphin's threshold for distinguishing High from Low.

Figure 2A shows that High responses predominated on High trials and difficult Low trials; Low responses predominated on easier Low trials. The discrimination was performed at chance where these two response curves cross, and the trial-decline response was used most in this region of maximum uncertainty. The dolphin assessed accurately when he was liable to err and he declined those trials adaptively. That he responded Uncertain most to trials that were 14 Hz (one eighth of a semitone) from true High trials shows that his uncertainty responding focused on his true psychophysical limit.



Figure 1. The dolphin participant in the study of Smith et al. (1995). Photograph Credit: Dolphin Research Center, Inc., Grassy Key, Florida. Used with permission.

The dolphin showed his own brand of auxiliary uncertainty behaviors. On solvable trials, he swam toward the response paddles so fast that his bow wave sometimes soaked our electronics. But he hesitated and wavered when the trial

was uncertain. These behaviors were distributed as his uncertainty responses were (Figure 2B). Tolman (1927) was intrigued by these behaviors that he called *lookings and runnings back and forth*. He thought these behaviors could define animal consciousness for the behaviorist.

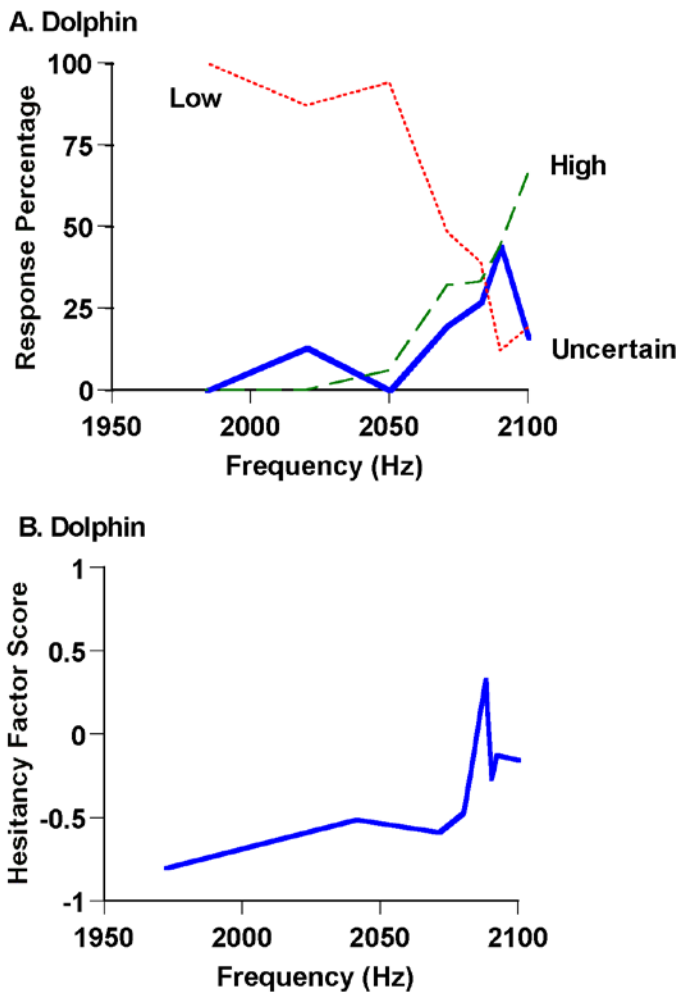


Figure 2. A. Performance by a dolphin in the auditory discrimination of Smith *et al.* (1995). The horizontal axis indicates the frequency (Hz) of the trial. The High response was correct for tones at 2,100 Hz—these trials are represented by the rightmost data point for each curve. All lower-pitched tones deserved the Low response. The solid line represents the percentage of trials receiving the Uncertain response at each trial level. The percentages of trials ending with the High response (dashed line) or Low response (dotted line) are also shown. B. The dolphin's weighted overall Factor 1 behavior (hesitancy, slowing, wavering) for tones of different frequencies (Hz). From “The Uncertain Response in the Bottlenosed Dolphin (*Tursiops truncatus*),” by J. D. Smith, J. Schull, J. Strote, K. McGee, R. Egnor, and L. Erb, 1995, *Journal of Experimental Psychology: General*, 124, p. 391, p. 402. Copyright 1995 by the American Psychological Association. Reprinted with permission.

But against that high-level interpretation are important associative considerations. First, there is the problem of objective stimulus cues—the dolphin was experiencing a class of error-causing threshold stimuli. Second, there is the problem of transparent reinforcement—the dolphin mainly received timeout penalties, not food rewards, for those stimuli. However, when he responded Uncertain to them, the consequence (a new, easy trial) was possibly more positive for him. Perhaps the problematic stimuli and High and Low responses in those stimulus contexts became slightly aversive for the dolphin. Perhaps he was conditioned to respond Uncertain in those stimulus contexts. That trial-decline responses were aversion-avoidance responses would be a very different psychological matter from their being expressions of uncertainty on difficult trials.

A common practice in this wider research area increases the force of associative interpretations. Researchers often reward animals directly for their use of the uncertainty response (e.g., Foote & Crystal, 2007; Kornell *et al.*, 2007; Hampton, 2001; Inman & Shettleworth, 1999; Suda-King, 2008; Sutton & Shettleworth, 2008). This methodology might grant the uncertainty response a positive associative response strength independent of any metacognitive role it plays in a task. It might be used because of its attractive reward properties. Thus, this approach makes it more difficult to disconfirm associative descriptions or affirm metacognitive interpretations.

These three potential associative dimensions to performance—stimulus cues that could occasion aversion and trigger avoidance, transparent reinforcement contingences that could catalyze conditioning processes, and the positive response strength accrued by directly rewarded uncertainty responses—all present significant challenges to this field. In combination, they represent the central interpretative issue facing researchers. The next sections of this article evaluate the explanatory weight borne by these associative mechanisms.

The Stimulus Component of Associative Interpretations

Shields *et al.* (1997). The dolphin experiment focused on primary stimulus qualities (pitch height) and thus encouraged stimulus-based descriptions of performance. It left open the possibility that specific, first-order stimuli trigger uncertainty responses. In fact, Smith *et al.* (2008) used formal modeling approaches to show that some uncertainty-monitoring data patterns can indeed be explained using low-level aversion and avoidance reactions to first-order stimuli. Accordingly, Shields *et al.* (1997) asked whether animals could recruit adaptive uncertainty responses in a psychophysical Same-Different (SD) task. The SD task requires an abstract judgment about the relation between two stimuli that goes

beyond the task's absolute stimuli. This is why SD performance is difficult to train in many species and why SD performance is acknowledged to be cognitively sophisticated (Premack, 1978; Herrnstein, 1990).

Shields et al. (1997) gave rhesus monkeys an SD density discrimination. On each trial, two rectangles filled randomly with lit pixels were shown. Animals were to make the Same or Different response when the two pixel densities were the same or different, respectively. The size of the stimulus differences on Different trials was adjusted dynamically based on recent performance to constantly challenge subjects' discriminative ability. In addition, Same and Different trials at several absolute stimulus levels were intermixed to ensure a true relational performance.

Yet monkeys, despite the difficulty and abstractness of the SD task, still declined adaptively indeterminate stimulus relations at the crux between Same and Different (Figure 3A). Their performance was essentially identical to that of human participants (Figure 3B)—the two performance profiles correlated at 0.98. In this case, uncertainty responses cannot have been triggered by low-level stimulus cues because the relevant cue was abstract-relational in nature. In this case, uncertainty responses had to be prompted by the indeterminacy of the relation that two highly variable stimuli instantiated. This is a performance of psychological complexity and sophistication.

Hampton (2001); Smith et al. (1998). These two studies also confirmed that stimulus-based interpretations of animals' uncertainty-monitoring performances are untenable. Hampton used the Delayed Matching to Sample task—animals had to remember a sample shape and identify it after a forgetting interval. Smith et al. used the Serial-Probe Recognition task—animals saw a list of stimuli and then judged whether a probe stimulus was part of that list or not. Both studies amplified their originating paradigm by giving animals the uncertainty response with which to decline trials of their choosing.

Both studies, too, cycled randomly through a limited set of stimuli that were used as memory probes. All stimuli became memory probes or were the foils to memory probes with the same frequency. All stimuli were rewarded and non-rewarded following both primary responses in the same way. No stimulus cue ever indicated any response. Only the status of the item as a to-be-remembered item (i.e., as a sample or a list member) was relevant. Thus, the psychology of this experiment unfolded along an internal continuum of subjective trace strength, with animals declining memory tests for memories of indeterminate strength. The signal of memory strength is very different from the signals available in traditional operant situations. The monkeys' behavior in these

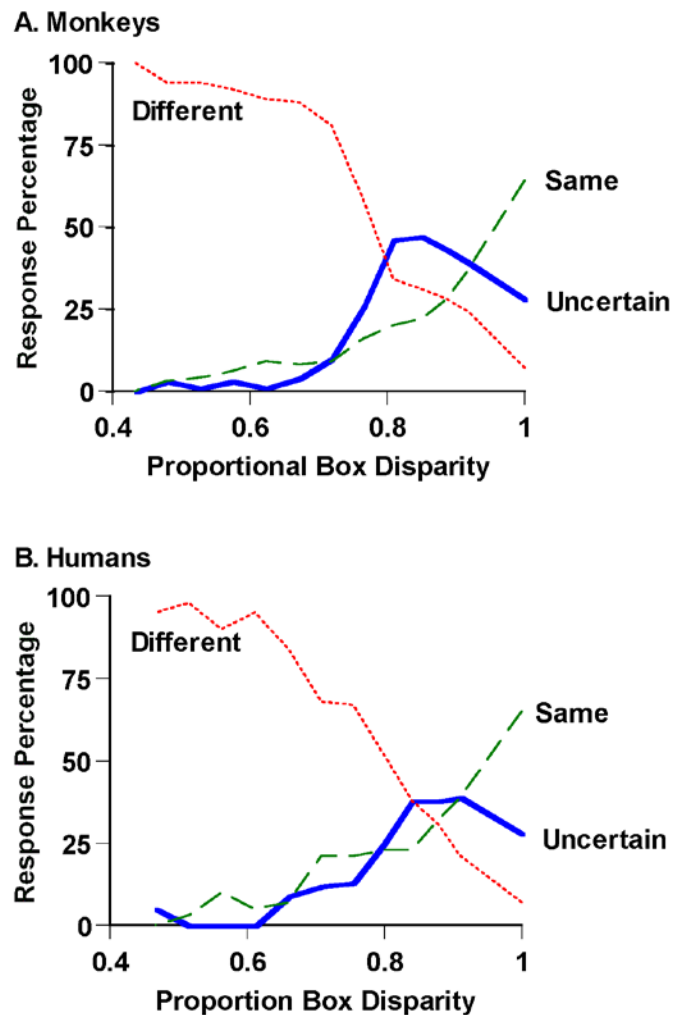


Figure 3. A. Performance by two monkeys in the Same-Different task of Shields et al., 1997. The horizontal axis gives the ratio between the densities of the comparison box and the standard box for trials of different disparities. The Same response was correct for trials at a proportional box disparity of 1.0. These trials are represented by the rightmost data points. All other trials deserved the Different response. The solid line represents the percentage of trials receiving the Uncertain response at each density ratio. The percentages of trials ending with the Same response (dashed line) or Different response (dotted line) are also shown. B. Performance by six humans in the Same-Different task, depicted in the same way. From “Uncertain Responses by Humans and Rhesus Monkeys (*Macaca mulatta*) in a Psychophysical Same-Different Task,” by W. E. Shields, J. D. Smith, and D. A. Washburn, 1997, *Journal of Experimental Psychology: General*, 126, p. 158. Copyright 1997 by the American Psychological Association. Reprinted with permission.

tasks is far from traditional senses of stimulus control. In a limited sense, these animals were showing metamemory—they were monitoring the contents of memory to determine their response.

Associative theorists naturally react sharply to a claim like this. They note that of course animals respond to internal stimuli—but it is still stimulus control. They note that animals' internal representations alter the objective stimulus input. Behavior is controlled by functional stimuli, not nominal, objective stimuli. They might insist that relational cues are still controlling stimulus cues, even if they are abstract or cognitive. They might even insist that states of memory are controlling, internal stimuli, too.

We agree that all stimuli are internal and represented functionally. Neither humans nor animals experience things in themselves. However, the idea of internality misses our point. Our point is that the level of the cognitive performance to which the uncertainty response attaches is critical in evaluating the cognitive level of the uncertainty response itself. If an animal reported uncertainty about whether a philosophical proof of the existence of God had a circularity flaw, no-one would claim that this report was an instance of stimulus control.

Now one may think that uncertainty about an abstract judgment of Sameness-Difference, or uncertainty about a memory, are or are not high enough yet to quiet the stimulus-control argument. That is a matter of one's cognitive or associative bias. Either way, the Same-Different and memory tasks illustrate an approach. They are waypoints toward demonstrations of animal metacognition that cognitive and associative theorists will both endorse.

We do caution, though, against the approach of calling everything a stimulus to preserve the idea of associationism and stimulus control. This does harm for a number of reasons. First, it will demote patent demonstrations of conscious metacognition in humans (or animals) down to instances of behaviorism. Second, there is a theoretical casualness in making everything a stimulus, because this will blur distinctions among performances that are profoundly different in psychological character. In our view, responding to the indeterminacy of a conceptual relation or responding to the dimness of a memory are profoundly different from responding to a present, aversive stimulus. But we understand that this may not be so for everyone.

Finally, we point out that there is a strong chance that the threshold state of indeterminacy between stimulus/response classes is psychologically distinctive and higher-level in its own right. For example, animals respond to internal states of hunger. They press bars for food. They stop pressing on satiation. But, suppose a pigeon made a non-appetitive response to report: I don't know if I'm hungry or not, perhaps a little peckish. This non-appetitive declaration of hunger indeterminacy would have considerable psychological complexity (and this declaration would be consciously metacognitive in

humans even if responding appetitively to hunger is not).

The Reinforcement Component of Associative Interpretations

Smith et al. (2006). The trial-by-trial reinforcement given in many comparative uncertainty-monitoring experiments has also encouraged reinforcement-based descriptions of performance. This leaves open the possibility that reinforcement history allows animals to construct low-level gradients of reward richness that could be used to trigger uncertainty responses *Smith et al.'s (2008)* formal models also illustrated this possibility.

Accordingly, some researchers have worked to show that uncertainty responses can be dissociated from a task's reinforcement structure and history. In one example of this approach, *Smith et al. (2006)* gave monkeys Sparse-Dense psychophysical discriminations, but they also required that the monkeys work for blocks of trials before receiving summary feedback for the entire block. During that summary, all reinforcements for correct responses in the block were delivered first, then all penalty timeouts for errors subsequently. In this way, feedback was rearranged out of trial-by-trial order, leaving monkeys with no way to associate reinforcement signals to particular stimuli or particular stimulus-response pairs. Nonetheless, one monkey of two was still able to make adaptive uncertainty responses (Figure 4A). This subject sustained this ability even when he was moved into new Sparse-Dense tasks (at new absolute pixel-density levels) in which he never experienced difficult or uncertain trials except under the contingency of deferred, rearranged feedback. (He did receive brief periods of trial-by-trial feedback on the easiest, extreme ends of the new Sparse and Dense stimulus continua as he entered new tasks.)

Smith et al. provided strong evidence that the reinforcement structure of this task did not determine the animal's uncertainty-response strategy. For example, Figure 4A shows that the monkey declined trials at the same frequency for Density Bins 6 and 13, even though he was 95% and 24% correct at those two trial levels, respectively, when he sometimes tried them. This is contrary to all associative interpretations. They would hold that animals should decline most those trials with the poorest reinforcement history (i.e., Bin 13 stimuli). Yet Figure 4B shows that there was essentially no relationship between uncertainty responding and reinforcement density in this case. The monkey was not responding to those reinforcement-based, associative cues. His uncertainty responses did not follow the task's reinforcement patterning.

Instead, *Smith et al. (2006)* showed that this monkey's uncertainty responses followed his decisional organization of the task. Figure 4C shows that this task had a subjective

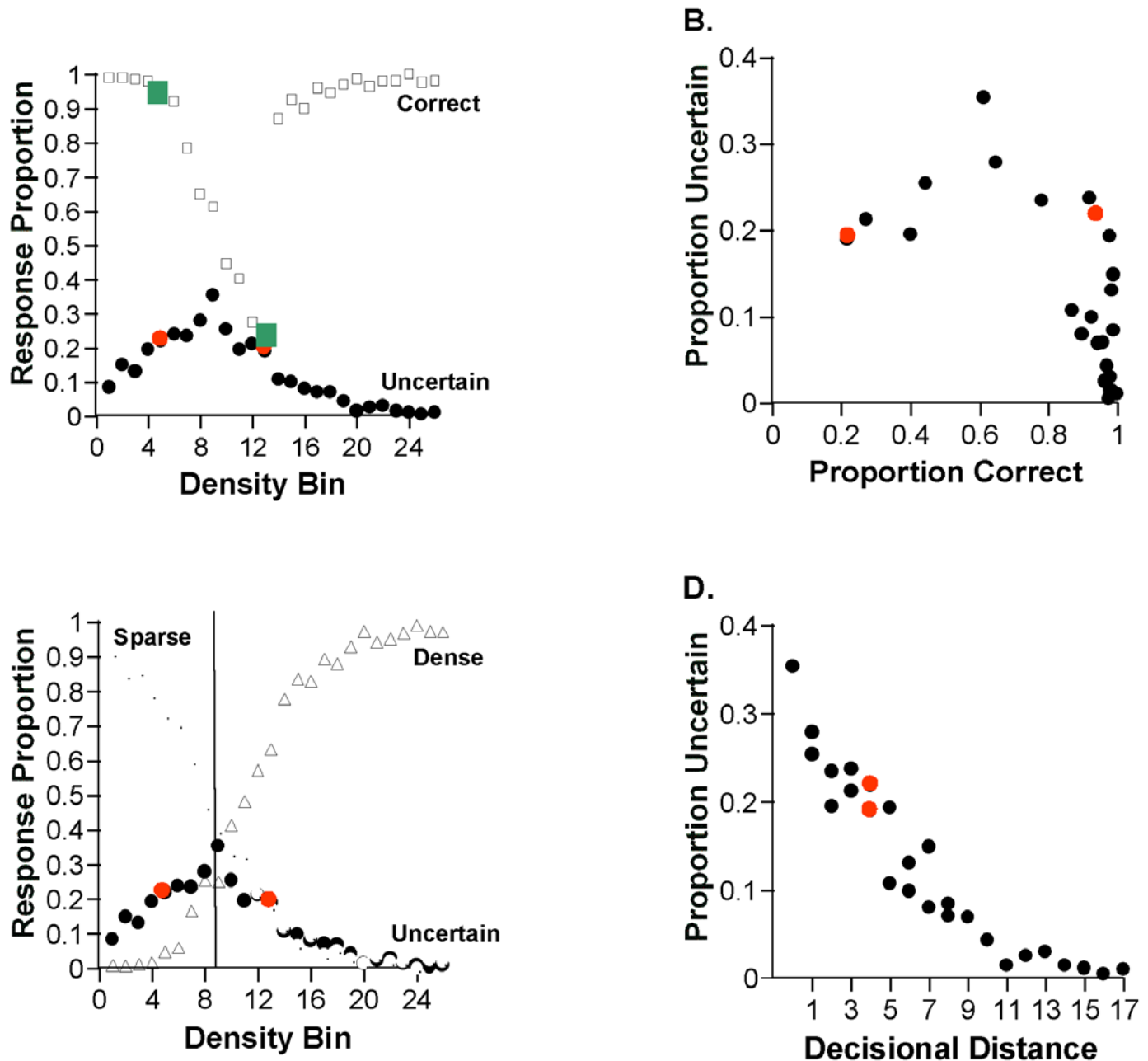


Figure 4. A monkey's performance in the Sparse-Dense discrimination of Smith et al. (2006). A. The horizontal axis indicates the density bin of the box. The Sparse and Dense responses, respectively, were correct for boxes at Density Bins 1-13 and 14-26. The open circles show the proportion of trials attempted that were answered correctly. The dark circles show the proportion of trials receiving the uncertainty response at each density bin. The level of uncertainty responding and correct responding at two trial levels are singled out using red circles and green squares. B. The monkey's performance in the same task, but now showing his proportion of trials declined in each density bin plotted against his proportion of trials answered correctly. Performance at Density Bins 6 and 13 are singled out using red circles. C. The monkey's performance in the same task, but now showing separately his use of the Sparse and Dense responses (open circles and open triangles). The vertical line indicates the monkey's subjective discrimination breakpoint at Density Bin 9. D. The monkey's performance in the same task, but now showing his proportion of trials declined in each density bin plotted against the decisional distance of the bin from his decisional breakpoint (Bin 9 = 0; Bins 8 and 10 = 1; Bins 7 and 11 = 2, etc.). From "Dissociating Uncertainty States and Reinforcement Signals in the Comparative Study of Metacognition," by J. D. Smith, M. J. Beran, J. S. Redford, and D. A. Washburn, 2006, *Journal of Experimental Psychology: General*, 135, p. 292. Copyright 2006 by the American Psychological Association. Reprinted with permission.

breakpoint at Bin 9 for the monkey. This is where his Sparse and Dense curves crossed. This was his discrimination breakpoint. His uncertainty responding was symmetrical about that breakpoint, and he declined fewer trials as they were farther from that breakpoint (note again: reinforcement was not symmetrical about that breakpoint). Figure 4D shows that understanding the monkey's uncertainty responding decisionally in this way explains well his use of the uncertainty response.

Thus, this monkey showed that an animal can organize a decisional framework for a task—with two response criteria defining three response regions—absent trial-by-trial signals of reinforcement. He also showed that the decisional patterning of a task can be dissociated from that task's reinforcement patterning. This is the first dissociation of its kind in the animal uncertainty-monitoring literature, and it speaks against many associative descriptions of animals' performance in those tasks.

Couchman et al. (submitted). Couchman et al. carried this demonstration further. Whereas Smith et al. (2006) had moved monkeys from one Sparse-Dense discrimination to another (e.g., from a Sparse-Dense discrimination involving absolutely sparse stimuli into a Sparse-Dense discrimination involving absolutely dense stimuli), Couchman et al. moved monkeys across qualitatively different task contexts. For example, a monkey trained on a Sparse-Dense discrimination was later moved into discriminations involving shorter-longer lines or rounded-flattened ellipses. Crucially, some of these tasks could not be solved on the basis of pixel illumination or density as in the Sparse-Dense task. Monkeys performed these tasks in blocks of trials without direct reinforcement and were still able to adaptively use the uncertainty response for the most difficult trials. This shows a level of cognitive monitoring that goes beyond low-level associations because associations learned in one task context would not be useful in the next. It also shows that monkeys have at least some capacity for self-learning and self-instruction under conditions of deferred and delayed reinforcement. As part of that capacity, they evidently can recognize difficult and error-causing stimuli, and construct adaptive regions of behavioral uncertainty, absent immediate reinforcement signals.

Reinforcement is a linchpin of associative accounts of learning and behavior. This is true no matter the exact character of the associative account (S-R, S-S, etc.). It is true no matter the exact character of one's theory of discrimination learning (componential, configural, etc.). But in the studies just described, reinforcement—a primary driving force behind the formation of associations—was dissociated away as an interpretative factor. The monkeys' uncertainty-monitoring performances in these tasks are far from traditional sens-

es of associative learning. This is an additional constraint on psychological theorizing about animals' uncertainty monitoring. It shows that animals' uncertainty responses can occur at a cognitive, decisional psychological level that is dissociable from a task's reinforcement structure.

This dissociation is a necessary and constructive step toward the inference that non-human species have something like a metacognitive capacity. However, we caution that it is somewhat artificial to completely dissociate reinforcement cues away from uncertain situations. Not only are errors, punishments, and uncertainty strongly co-occurrent, but also error and punishment are important teachable moments by which life signals uncertainty, doubt, danger, lack of knowledge, and domain non-expertise. A valid, important part of metacognition comes from this instruction and the organism's behavioral response to it. In the longer run of empirical study in this area, these aspects of cognitive monitoring—unfortunately lost in reinforcement's dissociation—should be studied in animals as well as humans.

The Direct Rewards Given to Uncertainty Responses

Finally, researchers have now shown that animals still make adaptive uncertainty responses even when those responses bring no hint, no information, no easy next trial, no direct food reward, and no reward token, but only the next randomly chosen trial. For example, Beran et al. (2006) had two rhesus monkeys judge arrays of dots as having more or fewer dots than a center value that was never presented in trials (Figure 5A). Monkeys were also given an uncertainty response that let them decline to make a numerosity judgment whenever they chose. Across center values (3-7) that shifted across sessions, the most errors occurred for the numerosities nearest the center value. Monkeys also used the uncertainty response most often for those most difficult numerosities (Figure 5B). This research joins that in Smith et al. (2006) in showing monkeys' use of the purest uncertainty response possible, uncontaminated by any secondary positive motivator. Therefore, this third dimension of the associative challenge falls away in the end as well.

A Misconception about Signal-Detection Models of Uncertainty-Monitoring

Many times it is useful to model animals' uncertainty-monitoring performances (e.g., Smith et al., 2003). These models frequently have a signal-detection character (Green & Swets, 1966; MacMillan & Creelman, 1991), though this is a choice of convenience and though other modeling frameworks might serve as well. Signal-detection models assume that performance is organized along an ordered series (a continuum) of psychological representations of changing impact

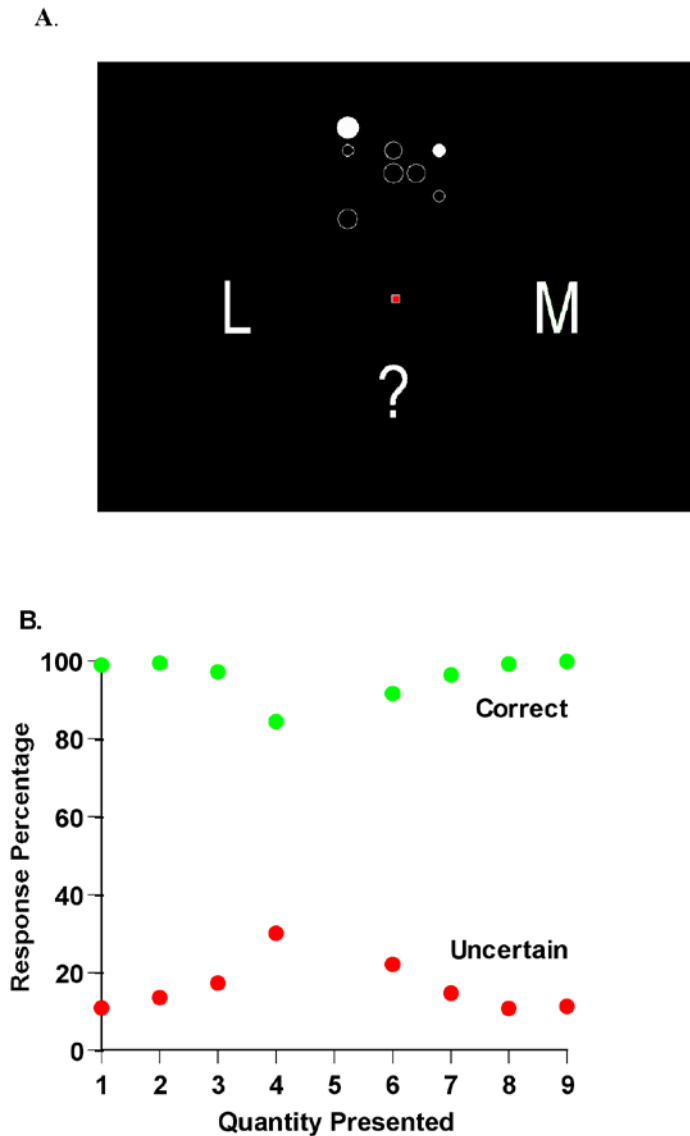


Figure 5. A. The screen from a numerosity-judgment trial in Experiment 2 of Beran et al. (2006). B. Performance by a monkey as a function of array size. Green circles indicate the percentage of trials correct in the primary numerosity discrimination when it was attempted. Red circles indicate the percentage of trials on which the uncertainty response was selected for each array size. From “Rhesus Macaques (*Macaca mulatta*) Monitor Uncertainty during Numerosity Judgments,” by M. J. Beran, J. D. Smith, J. S. Redford, and D. A. Washburn, 2006, *Journal of Experimental Psychology: Animal Behavior Processes*, 32, p. 113, p. 117. Copyright 2006 by the American Psychological Association. Reprinted with permission.

or increasing strength (memory strength, confidence level, etc.). They assume that the subject establishes response regions by placing decision criteria along the continuum of cognitive states. By the usual metacognitive interpretation of the referent experiments, one would assume that there are upper and lower criteria defining three response regions, the

leftmost reserved for one primary discrimination response, the rightmost for the other discrimination response, and the middle region demarcating indeterminate cognitive states that should receive uncertainty responses.

Regarding this descriptive framework, a serious misconception has arisen. In essence it states: if the general framework of the signal-detection model applies to a behavioral situation, then this shows directly that the performance can be, and should be, interpreted in a low-level, first-order, associative manner. This misinterpretation needs to be corrected, because it is a common view, because it is mistaken, and because it does harm to theoretical progress in this area.

Of course in some cases associative descriptions will be appropriate. If the underlying continuum concerns stimulus generalization, response strength, or aversion-avoidance, then, yes, the signal-detection model would be about low-level, associative mechanisms

But, in other cases, associative descriptions will be inappropriate. For the underlying cognitive continuum may also concern higher-level cognitive representations or even metacognitive states, such as degrees of conscious certainty, levels of confidence in retrieved memories, gradations of intensity of tip-of-the-tongue states, and so forth. Note that the signal-detection model will be suitable even when the performance is fully conscious, metacognitive, and reportable through language. That is, any metacognitive performance that humans would ever show could be described using the same signal-detection framework, but by all accounts none of them would deserve to have Morgan’s Canon aimed at them. Likewise, the previous sections of this article have shown that in many cases the primary stimulus basis for association is not there, nor are the reinforcement signals that could fuel association. Without association’s ground and spark, it cannot occur. Another psychological description is required, even though a signal-detection model fits the data.

Perhaps we can partly explain why this misunderstanding arises. Many times, comparative models instantiate mathematically parameters that are strongly grounded in low-level psychological processes in mind or brain. For example, timing/counting models envision duration or event accumulators and comparators that may even reflect the neurophysiology of timing/counting. Perhaps it is natural, then, that modelers assume a model’s mathematics to be instantiating low-level processes. However, this assumption does not apply to the models used in fitting animal’s uncertainty-monitoring performances. The reason for this is that no one knows how the animal or human organizes the decision criteria that define the regions of uncertainty responding in these tasks. This could be done associatively or consciously. Even if one uses Morgan’s Canon to grant the low-level interpretation

provisional priority, this priority evaporates once one discovers that those decision criteria are evidently organized cognitively, decisionally, abstractly, and dissociably from the task's reinforcement signals. Then, a willful suspension of the low-level bias is the appropriate scientific stance.

Meanwhile, this misunderstanding has a significant cost to the field because it artificially equalizes phenomena that may be qualitatively different in psychological level and character. Indeed, it ignores the psychological content of performances (their representations, their processes) to focus on a metaphorical, formal similarity. Worst of all, it lumps together phenomena in order to dismiss them, when what the field actually needs is to parse phenomena and separate them carefully so as to understand them clearly.

We stress that this field needs to keep its focus sharply on the cognitive processes and representations that underlie uncertainty-monitoring performance, not on their suitability for signal-detection modeling. That suitability has no conceivable relationship to or implication for the level or character of a psychological performance.

A Misunderstanding about Reinforcement in Uncertainty-Monitoring Tasks

There is an additional concern about existing associative explanations of uncertainty-monitoring performances. These explanations often contain statements to this effect: Naturally animals use the uncertainty response, for it reduces the delay to reinforcement in subsequent trials because it is never followed by a timeout. Reduction in delay to reinforcement is itself reinforcing, and this is the likely basis for uncertainty responding. This claim is a version of the common claim by associative theorists that reinforcement rate over time is the molar and critical factor determining choice.

But these explanations are also mistaken. This is a case wherein associative theorists have not given sufficient attention to the animal's real situation in an uncertainty-monitoring task. It is not the case that the uncertainty response speeds up the time to the next reward. If the uncertainty response were used on every trial, in many experiments there would be no rewards given, ever—the uncertainty response would produce an indefinite delay until reward. Or, if the animal used the uncertainty response on easy trials that would almost certainly bring reward, it would delay reward by several seconds each time.

Indeed, the use of the uncertainty response will only speed up the arrival of the next reward under narrow circumstances—that is, if the animal can use it adaptively and selectively to decline only the most difficult trials in the discrimination. But, to do so, the animal must have some way

to monitor either difficulty or the risk of error, so that it can decline trials in that range. In essence, the associative claim about speeded-up rewards only stands if the animal has an adaptive uncertainty-monitoring performance at work in the background. But, in that case, the metacognitive dog is wagging the associative tail.

Regarding this misunderstanding, we would caution our field that there is also a theoretical casualness in attributing a wide range of behavioral patterns to the molar maximization of reinforcement. Once again, this is liable to blur distinctions among performances that are profoundly different in psychological character. Put another way, the reward-maximization description is often functionally correct, but psychologically empty. One still needs to come to understand the cognitive representations and processes that allow this maximization to occur (are they conscious and deliberative, or low-level and reactive). On doing so, the psychological description will explain the behavioral pattern more theoretically richly than will the reward-maximization description.

Phylogenetic Studies of Uncertainty Monitoring and Associative Interpretations

Finally, the growing body of cross-species research on uncertainty monitoring also speaks against the possibility that uncertainty-monitoring performances reflect low-level associative mechanisms based in stimuli and reinforcement.

Beran et al. (in press). For example, Beran et al. (in press) gave six capuchins (New World Monkeys, family *Cebidae*) two density discrimination tasks. In the Sparse-Uncertainty-Dense task, difficult trials at the breakpoint between Sparse and Dense could be declined through an uncertainty response. This was a standard uncertainty-monitoring task in which rhesus monkeys (Old World Monkeys, family *Cercopithecidae*) have commonly used the uncertainty response adaptively. In the Sparse-Middle-Dense task, making a middle response to the same intermediate stimuli could be rewarded. This was a standard, three-response discrimination, in which associative mechanisms might encourage middle responses for middle stimuli. Capuchins essentially did not use the uncertainty response at all (Figure 6A). But they used the middle perceptual response easily (Figure 6B).

Beran et al. (in press) then used formal optimality studies to examine the reinforcement landscape of the uncertainty task. They found ways to raise the associative stakes in the task, so that now the capuchins had a strong motivation to learn to use that uncertainty response adaptively. To accomplish this, Beran et al. increased the penalty timeout in the uncertainty task to 90s, more than quadrupling the time cost for each error and increasing the advantage to be gained from declining difficult trials. Now, the capuchins gave up

potentially 30 trials for each penalty timeout (because correctly answered trials lasted about 3 s). In addition, Beran et al. eliminated from the task the easiest trial levels, so that the overall reinforcement rate in the task went down. This ensured that the capuchins were not just satisficing: accepting a decent reinforcement return based on a cognitively easy task strategy that did not incorporate the effortful uncertain-

way, their associative competence would have also let them make uncertainty responses adaptively. If uncertainty responding was just a matter of distributing available responses to globally maximize the rate of rewards, they would have done so as well. But they did not. Taking the uncertainty and middle results together, one sees that each animal acted as its own control, in the sense that uncertainty responding

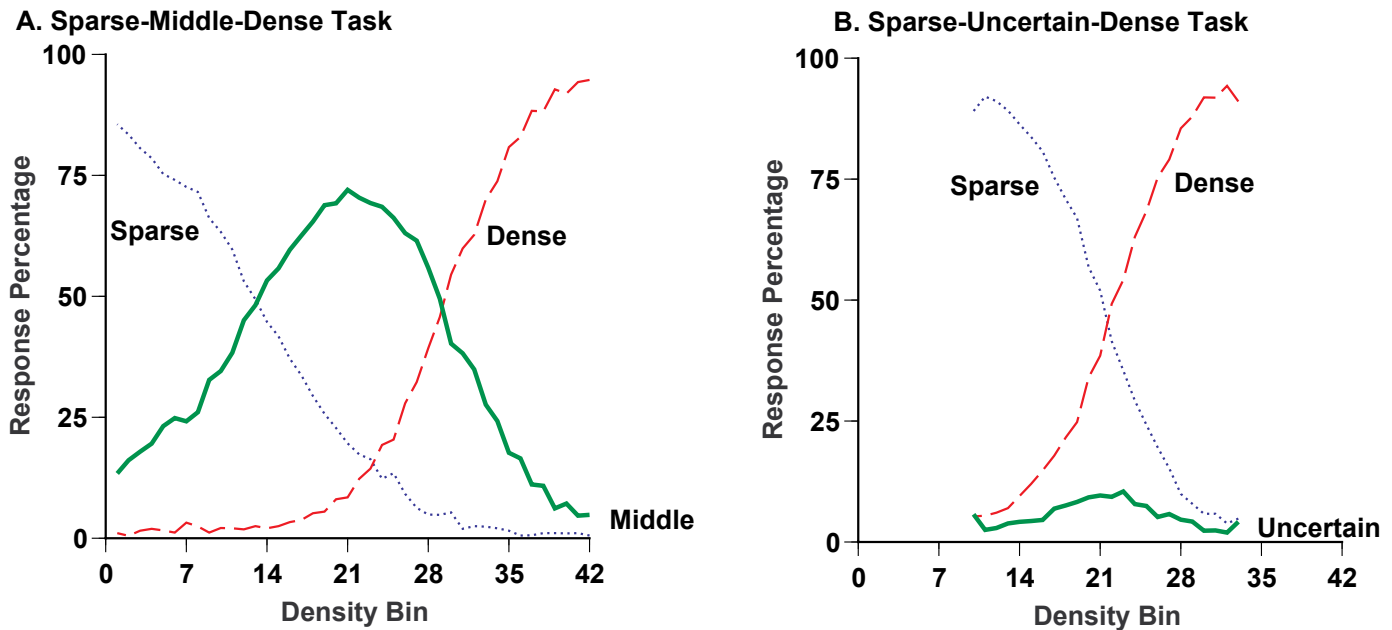


Figure 6. The performance of six capuchin monkeys in the Sparse-Uncertainty-Dense task of Beran et al. (2009). The horizontal axis indicates the density level of the box. The Sparse and Dense responses, respectively, were correct for boxes at Density Levels 1-21 and 22-42. The solid line represents the percentage of trials receiving the Uncertain response at each trial level. The percentages of trials ending with the Dense response (dashed line) or Sparse (dotted line) are also shown. B. The performance of the same capuchin monkeys in the Sparse-Middle-Dense task of Beran et al. (2009), depicted in a similar way. From “The Psychological Organization of ‘Uncertainty’ Responses and ‘Middle’ Responses: A Dissociation in Capuchin Monkeys (*Cebus apella*),” by M. J. Beran, J. D. Smith, M. V. C. Coutinho, J. J. Couchman, and J. B. Boomer, 2009, *Journal of Experimental Psychology: Animal Behavior Processes*, in press. Copyright 2009 by the American Psychological Association. Reprinted with permission.

ty response. Even then, in this second experiment, five of the six capuchins showed essentially no use of the uncertainty response. One animal did use that response now, though he may have done so because the repeated alternation of middle and uncertainty tasks had finally taught him that uncertainty and middle responses were fungible.

These experiments have an important implication. The middle task defined a class of middle discriminative stimuli that animals could use to anchor associative processes. Middle stimuli were transparently reinforced, allowing associative mechanisms to operate and associative connections to form. Clearly, those associative processes operated well in the capuchins. They responded middle sensitively and accurately.

If uncertainty responding was associatively based in a like

and middle responding were manipulated within animals in an alternating and counterbalanced way. Clearly, the psychological mechanism that underlies uncertainty responding is different from the associative mechanism that underlies middle responding. Uncertainty responding is responsive to some psychological signal that capuchin monkeys barely monitored in this task. The performances are qualitatively different. One performance is associatively based, the other is not. One performance maximized reward density, one did not. These animals didn’t reward maximize in the uncertainty task because they didn’t apprehend appropriately that task’s psychological signals. The psychology of the situation was in explanatory control over the animals’ behavioral patterns, not their need to maximize rewards or their ability to establish associative connections.

Inman and Shettleworth (1999); Sutton and Shettleworth (2008). These studies on pigeons made a similar point. Pigeons have consistently refused to express an uncertainty-monitoring capacity, even though in the relevant experiments they have been directly rewarded with food for making uncertainty responses. Shettleworth and her colleagues have concluded that pigeons either lack a capacity for cognitive monitoring, or else they express that capacity with such difficulty that experimenters can hardly show it. Yet no one would say that pigeons are associatively challenged—they are not. If there were patent associative mechanisms and cues available in those tasks, pigeons would find and use them, and reward maximize by producing the uncertainty-monitoring data pattern. Because they do not, one sees again that a different, non-associative psychological description of the uncertainty-monitoring task is required.

Controlled, Decisional Processes in Uncertainty-Monitoring Tasks

A crucial conclusion from this discussion is that animals have transcended the basic associative mechanisms that have been proposed. Not every time, but sufficiently for the relevant existence proofs. Other approaches to associative interpretations contain misunderstandings as discussed above. In some cases, therefore, animals in their performances have gone to the next psychological level. Humans in their theorizing have not. Another crucial conclusion is that theoretical development in this area lags behind empirical development. Researchers are still focused on asserting or refuting associative interpretations—one sees this focus in the present article. Consequently, the field has not found the new cognitive-decisional description of animals' performances that is warranted and necessary for further theoretical advancement in this area. In this section, we outline some aspects of this description.

Consider a traditional, auditory-threshold task. Silent intervals deserve the No Sound response. Intervals containing a faint, 50%-detectable (threshold) sound deserve the Sound response. In this situation, only two events can occur. Two responses map exhaustively to those events—all events deserve one response or the other. There is no intermediate stimulus class that could ground associative processing. This is not a Red-Yellow discrimination with the possibility of attaching a third response to Orange in between. Here there is no "orange," because the whole psychological range of the threshold task plays out within a single JND. So, there is nothing between Sound and No Sound except Sound-No Sound indeterminacy. This is why—logically—one knows that uncertainty responding in a task of this kind is about resolving indeterminacy. This is why even associative theorists acknowledge that threshold states are psychologically unique. The rules of stimulus control are different there,

animals have difficulty finding adaptive solutions there, animals become minimally informed observers there, and there are not stimuli there on which to ground operants or conditioned reflexes (Boneau & Cole, 1967; Commons et al., 1991; Davison et al., 1985; Miller et al., 1980; Terman & Terman, 1972). The classical psychophysicists agreed that the threshold state is psychologically complex (Boring, 1920; Fernberger, 1914; George, 1917; Thomson, 1920; Watson et al., 1973; Woodworth, 1938). A threshold event is not the clear stimulus signal that can ground a reflexive response. A threshold event is about the failure to assign an event to a stimulus class. It is about the failure to know which reactive response to make. The laws of association and conditioning break down near threshold, and this helps explain why the psychology of threshold is still poorly understood.

However, one can describe the psychological structure of uncertainty-monitoring tasks, or indeterminacy-resolution tasks, in a way that advances us toward understanding these performances across tasks and species. Shiffrin and Schneider (1977) analyzed the information-processing consequences of cognitive indeterminacy. Indeterminate mental representations necessarily map inconsistently and unreliably onto behavioral responses. This makes those representations inadequate behavioral indicators. Consequently, the organism must invoke higher levels of controlled cognitive processes to adjudicate and resolve the indeterminacy. An important theoretical statement follows from this. All the uncertainty-monitoring tasks are inconsistently mapped in Shiffrin and Schneider's sense. All would benefit from controlled processing (slow, deliberate, etc.). This is true of the dolphin caught at threshold, 14 HZ from the standard at 2,100 HZ. It is true for monkeys needing to handle ambiguous memory traces (e.g., Hampton, 2001; Hampton et al., 2004; Smith et al., 1998). Atkinson and Juola (1974) also suggested that indeterminate memory traces require qualitatively different, secondary information-processing strategies.

To grasp intuitively the situation animals face in uncertainty-monitoring tasks, consider your cognitive processes in a world in which traffic lights gradually morphed, pixel by pixel, between red and green, and you had to decide whether your light was Red Enough to Stop or Green Enough To Go. Gone would be all the associative clarity of the Stop and Go operants and their conditioned emotional and response contents. Instead, approaching a light, you would set in motion an elaborate criterion-setting process that would include—sadly—your schedule, the speed of opposing traffic, guilt, ethics, the proximity of police vehicles, the points already on your license, whether your children were watching from the back seat, and so forth. Criterion-setting at threshold is a qualitatively different thing from associative responding, and this field needs sophisticated theories of this kind of performance to apply to animals' uncertainty-monitoring per-

performances.

In short, psychophysical procedures ensure indeterminate stimulus-response mappings and encourage controlled decision-making no matter the species. Uncertainty responses to threshold states probably represent a controlled decision, on the threshold of perception or memory, to decline the trial. This realization grants animals' uncertainty responses some of the cognitive sophistication that is due them, without loading them down with consciousness or other human concomitants of metacognition.

Conclusion

In this new area of comparative inquiry, researchers have found innovative ways to ask animals difficult psychological questions concerning their metacognitive capacities. It is natural that associative concerns have arisen as part of this research program. It is natural that the field would pause to evaluate its problems and prospects. On doing so, it is clear that this field shows encouraging signs of progress, including this special discussion forum, the many insightful scholars who have engaged this research, and the crucial fact that animals sometimes transcend reasonable associative interpretations to show genuine uncertainty-monitoring performances.

Accordingly, we believe it is a good time for theory in this area to ratchet up a level to meet animals' uncertainty-monitoring performances, and to consider with careful models and psychological descriptions the cognitive-decisional performances that animals show. What kind of controlled processing do animals demonstrate in these tasks? Are these processes executive, explicit, declarative—possibly even conscious? These and related questions of psychological interpretation have been held in abeyance while the associative issue dominated. However, these new questions will open up along with the field's theoretical horizon, and new insights will emerge as they become the focus of this growing field.

References

- R. Luce, and P. Suppes (Eds.), *Contemporary Developments in Mathematical Psychology*, Vol. 1. San Francisco: W.H. Freeman.
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metacognitive index. *Journal of Experimental Psychology: General*, **127**, 55-68.
- Beran, M. J., Smith, J. D., Coutinho, M. V. C., Couchman, J. J., Boomer, J. (in press). The psychological organization of "uncertainty" responses and "middle" responses: A dissociation in capuchin monkeys (*Cebus apella*). In press, *Journal of Experimental Psychology: Animal Behavior Processes*.
- Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *Journal of Experimental Psychology: Animal Behavior Processes*, **32**, 111-119.
- Boneau, C. A., & Cole, J. L. (1967). Decision theory, the pigeon, and the psychophysical function. *Psychological Review*, **74**, 123-135.
- Boring, E. G. (1920). The control of attitude in psychophysical experiments. *Psychological Review*, **27**, 440-452.
- Carruthers, P. (2008). Meta-cognition in animals: a skeptical look. *Mind and Language*, **23**, 58-89.
- Commons, M. L., Nevin, J. A., & Davison, M. C. (Eds.) (1991). *Signal detection: mechanisms, models, and applications*. Hillsdale, NJ: Erlbaum.
- Couchman, J. J., Coutinho, M. V. C., Beran, M. J., & Smith, J. D. (2008, submitted). Beyond stimulus cues and reinforcement signals: A new approach to animal metacognition. Manuscript submitted for publication.
- Davison, M., McCarthy, D., & Jensen, C. (1985). Component probability and component reinforcer rate as biasers of free-operant detection. *Journal of the Experimental Analysis of Behavior*, **44**, 103-120.
- Fernberger, S. W. (1914). The effect of the attitude of the subject upon the measure of sensitivity. *American Journal of Psychology*, **25**, 538-543.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, **34**, 906-911.
- Foot, A., & Crystal, J. (2007). Metacognition in the rat. *Current Biology*, **17**, 551-555.
- Gallup, G. G. (1982) Self-awareness and the emergence of mind in primates. *American Journal of Primatology*, **2**, 237-48.
- George, S. S. (1917). Attitude in relation to the psychophysical judgment. *American Journal of Psychology*, **28**, 1-38.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, **98**, 5359-5362.
- Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition*, **7**, 239-246.
- Herrnstein, R. J. (1990). Levels of stimulus control—A functional approach. *Cognition*, **37**, 133-166.
- Inman, A., & Shettleworth, S. J. (1999). Detecting metamemory in nonverbal subjects: A test with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **25**, 389-395.
- Koriat, A. (1995). Dissociating knowing and the feeling of

- knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, **124**, 311-333.
- Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 289-325). New York, NY: Cambridge University Press.
- Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, **135**, 36-69.
- Kornell, N., Son, L., & Terrace, H. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, **18**, 64-71.
- MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- Metcalf, J. (2000). Metamemory: Theory and data. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197-211). New York: Oxford University Press.
- Metcalf, J., & Kober, H. (2005). Self-reflective consciousness and the projectable self. In H. S. Terrace and J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 57-83). New York: Oxford University Press.
- Miller, J. T., Saunders, S. S., & Bourland, G. (1980). The role of stimulus disparity in concurrently available reinforcement schedules. *Animal Learning & Behavior*, **8**, 635-641.
- Morgan, C. L. (1906). *An introduction to comparative psychology*. London: Walter Scott Publishing.
- Nelson, T. O. (Ed.) (1992). *Metacognition: Core readings*. Toronto: Allyn and Bacon.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, **51**, 102-116.
- Nelson, T. O. & Narens, L. (1990) Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, **26**, 125-41.
- Premack, D. (1978). On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 423-451). Hillsdale, NJ: Erlbaum.
- Proust, J. (2003). Does metacognition necessarily involve metarepresentation? *The Behavioral and Brain Sciences*, **26**, 352.
- Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review*, **1**, 357-375.
- Serra, M. J., & Dunlosky, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 1258-1266.
- Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology: General*, **126**, 147-164.
- Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, **84**, 127-190.
- Smith, J. D., Beran, M. J., Couchman, J. J., & Coutinho, M. V. C. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review*, **15**, 679-691.
- Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating uncertainty states and reinforcement signals in the comparative study of metacognition. *Journal of Experimental Psychology: General*, **135**, 282-297.
- Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: General*, **124**, 391-408.
- Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, W. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General*, **127**, 227-250.
- Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *The Behavioral and Brain Sciences*, **26**, 317-373.
- Staddon, J. E. R., Jozefowicz, J., & Cerutti, D. (2007). Metacognition: A problem not a process. *PsyCrit*, 1-5.
- Suda-King, C. (2008). Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Animal Cognition*, **11**, 21-42.
- Sutton, J. E. & Shettleworth, S. J. (2008). Memory without awareness: Pigeons do not show metamemory in delayed matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, **34**, 266-282.
- Terman, M., & Terman, J. (1972). Concurrent variation of response bias and sensitivity in an operant-psychophysical test. *Perception and Psychophysics*, **11**, 428-432.
- Thomson, G. H. (1920). A new point of view in the interpretation of threshold measurements in psychophysics. *Psychological Review*, **27**, 300-307.
- Tolman, E. C. (1927). A behaviorist's definition of consciousness, *Psychological Review*, **34**, 433-439.
- Washburn, D. A., Smith, J. D., & Shields, W. E. (2006). Rhesus Monkeys (*Macaca mulatta*) immediately generalize the uncertain response. *Journal of Experimental Psychology: Animal Behavior Processes*, **32**, 85-89.
- Watson, C. S., Kellogg, S. C., Kawanishi, D. T. & Lucas, P. A. (1973). The uncertain response in detection-oriented psychophysics. *Journal of Experimental Psychology*, **99**,

180-185.

Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.