



Published in final edited form as:

Anim Cogn. 2010 January ; 13(1): 93–101. doi:10.1007/s10071-009-0249-5.

Rhesus Monkeys (*Macaca mulatta*) Adaptively Monitor Uncertainty While Multi-Tasking

J. David Smith,

University at Buffalo, the State University of New York

Joshua S. Redford,

University at Buffalo, the State University of New York

Michael J. Beran, and

Language Research Center, Georgia State University

David A. Washburn

Language Research Center, Georgia State University

Abstract

As researchers explore animals' capacity for metacognition and uncertainty monitoring, some paradigms allow the criticism that animal participants—who are always extensively trained in one stimulus domain within which they learn to avoid difficult trials—use task-specific strategies to avoid aversive stimuli instead of responding to a generalized state of uncertainty like that humans might use. We addressed this criticism with an uncertainty-monitoring task environment in which four different task domains were interleaved randomly trial by trial. Four of five macaques (*Macaca mulatta*) were able to make adaptive uncertainty responses while multi-tasking, suggesting the generality of the psychological signal that occasions these responses. The findings suggest that monkeys may have an uncertainty-monitoring capacity that is like that of humans in transcending task-specific cues and extending simultaneously to multiple domains.

Keywords

metacognition; uncertainty monitoring; primate cognition; comparative psychology; monkeys

Introduction

Humans are metacognitive (Flavell, 1979; Koriat, 1993; 2007; Koriat, Ma'ayan, and Nussinson, 2006; Metcalfe, 2000; Nelson, 1992; Schwartz, 1994). They make judgments of their level of learning and competence; they reflect upon their knowledge states. They also feel uncertain when they do not know or remember, and they respond adaptively to these feelings by declining behavioral choices and seeking help, hints, or information instead. This sophisticated cognitive capacity is linked to humans' self-awareness (Gallup, 1982) and declarative consciousness (Koriat, 2007; Nelson, 1996). Typically, it is a late developing trend in human children (e.g., Ghetti, Lyons, and Lazzarin, 2008; Kuhn, 2000; Lockl and Schneider, 2002; Roebers, von der Linden, Schneider, and Howie, 2007). Therefore, it is an intriguing question whether nonhuman animals (hereafter animals) have a similar capacity.

Researchers are exploring this question using uncertainty-monitoring paradigms that are suitable for research with nonverbal animals (Beran, Smith, Redford, and Washburn, 2006; Call and Carpenter, 2001; Foote and Crystal, 2007; Hampton, 2001; Kornell, Son, and Terrace, 2007; Smith, Beran, Redford, and Washburn, 2006; Smith, Shields, and Washburn, 2003; Suda-King, 2008; Sutton and Shettleworth, 2008; Washburn, Smith, and Shields, 2006). To do so, they present to animals a mix of easy and difficult trials. They give animals an additional response—beyond the primary discrimination responses—that lets them decline to complete any trials they choose. If animals monitor their cognition, they should recognize difficult trials as error-risking and decline them. Some animals do so, producing human-like data patterns in cognitive-monitoring tasks. This additional response is called the *uncertainty response*—it is interpreted to show some species' capacity for uncertainty monitoring and metacognition.

If this interpretation is correct, these studies tap theoretically important cognitive capacities in animals and raise interesting issues about animal minds, awareness, and consciousness. They could help reveal the evolutionary roots of human metacognition. They also raise theoretical questions about human metacognition (e.g., how dependent on language is human metacognition if nonverbal animals share this capacity). Finally, these studies might also help reveal the developmental roots of human metacognition. The simple, perceptual tasks that are suitable for animals would also be appropriate for exploring the earliest metacognitive achievements of young human children.

However, there are important concerns about the appropriate psychological interpretation of animals' uncertainty responses. One concern is that many studies reward uncertainty responses directly with food (Foote and Crystal, 2007; Hampton, 2001; Kornell et al., 2007; Suda-King, 2008). This gives the uncertainty response an attractiveness independent of any uncertainty role it plays in a task. Under these circumstances, animals might be drawn to make this response as a default, avoidance response when their other response tendencies are weak, not to cope with an uncertainty state. Indeed, Smith, Beran, Couchman, and Coutinho (2008) used formal-modeling approaches to confirm that metacognitive-like data patterns could be recreated under these circumstances using low-level associative mechanisms. Therefore, these direct rewards make it difficult or impossible to dismiss behaviorist interpretations of uncertainty responses and difficult to affirm metacognitive interpretations. The present research addressed this concern by using a pure uncertainty response that had no consequence except to end one trial and begin the next.

Another concern is that animals have always used uncertainty responses in one task at a time. Animals are highly trained in that task environment for weeks or months, and thousands of trials. During this training, they have ample time to become sensitive to some task-specific cue to stimulus difficulty, perhaps one not even known to the experimenters. Uncertainty responses could be occasioned by the presence of this cue, not necessarily by a generalized state of uncertainty. This possibility remains one of the dominant theoretical issues within the animal metacognition literature (Smith et al., 2008a,b). Alternatively, animals could build up reinforcement histories for different specific stimuli in a single, highly trained task, experience the lean rewards and negative outcomes that attend certain stimuli, become averse to those stimuli, and use the trial-decline response just as a low-level aversion-avoidance response when those stimuli appear. Of course such a response would not be equivalent to the flexible, multi-functional use of uncertainty that humans show. Humans are uncertain about many situations daily, constantly, and in parallel. To give just one example, students taking examinations are capable of recognizing their own uncertainty about answers to questions in many formats (e.g., multiple choice, matching, fill in the blank) as they proceed through the exam. So, the question remains whether animals can show a more flexible and general uncertainty-monitoring capacity that is applicable to

several tasks and situations simultaneously. This would indicate that, beyond the task-specific cues available in previous single tasks, they can also monitor and respond to some broad-spectrum signal of uncertainty that is made available by difficult and indeterminate trials of many kinds. Answering this question was the primary purpose of the present article. If animals can do so, the demonstration would be a step toward showing that they have a generalized capacity for uncertainty monitoring.

Methods

Participants

Five male rhesus monkeys (*Macaca mulatta*) were tested: Hank (22 years old), Murph (12 years old), Lou (12 years old), Gale (22 years old), and Willie (20 years old). The animals had been trained using established procedures to respond to computer-graphic stimuli by manipulating a joystick (Rumbaugh Richardson, Washburn, Savage-Rumbaugh, and Hopkins, 1989). Murph and Lou had participated previously in an experiment involving uncertainty monitoring in a dense-sparse discrimination, one of the tasks used here. Murph and Lou both were presented with blocks of trials in which feedback was deferred. Both of these monkeys learned to make the four different discriminations given in the present experiment. However, Lou learned only the most extreme levels within each task (i.e., the two easiest levels of each task), and he was never presented the uncertainty response in that pilot study. Murph, however, was given some sessions with all four tasks intermixed and all levels within each task presented. However, he showed no ability to use the uncertainty response appropriately, and he also showed poor performance in the four discrimination tasks. The monkey Gale also had very brief exposure to only the easiest trials on Task 1 (Dense-Sparse) but without the uncertainty response. Given the difficulties experienced by Murph and Lou, we abandoned the attempt to have monkeys work four simultaneous tasks with deferred feedback because this appeared to be too difficult. To summarize, the pilot test afforded almost no experience of relevance for Gale, only experience with the easiest levels of the four tasks for Lou (but not the uncertainty response), and more extensive experience for Murph in the sense of his seeing all trial types in each of the four tasks and having an uncertainty response available although he never used it properly.

Apparatus

The monkeys were tested in their home cages with *ad lib* access to the test apparatus, working or resting as they chose. They were not food deprived for purposes of testing. No matter how many trials they completed, they always received their usual evening meal of fruit and vegetables, and they had continuous access to water.

Apparatus—The monkeys were tested using the Language Research Center's Computerized Test System (Washburn and Rumbaugh, 1992). Contacting appropriate computer-generated stimuli with a cursor controlled by their hand movements on a joystick brought them 94-mg fruit-flavored chow pellets (Bioserve, Frenchtown, NJ) using a Gerbrands 5120 dispenser interfaced through a relay box and output board (PIO-12 and ERA-01; Keithley Instruments, Cleveland, OH). Incorrect responses brought them a 20 s timeout during which the animals could not make any responses.

Procedure

Psychophysical tasks—The monkeys performed four psychophysical-judgment tasks involving a 201 × 101 pixel box presented in the top left of the screen.

In the Sparse-Dense task (Figure 1A), the box was filled with a variable number of randomly placed blue pixels, from Level 1 (1,425 pixels) to Level 41 (2,909 pixels).

Animals discriminated Sparse and Dense boxes that had fewer pixels or more pixels than the midpoint (Level 21 2,036 pixels).

In the Top-Bottom task (Figure 1B), the box had light red and light green pixels above and below. The dividing line between the colors varied vertically on the screen, from Level 1 (26 red and 75 green rows of pixels) to Level 50 (75-26 red-green). Animals discriminated boxes with smaller or larger top halves.

In the Left-Right task (Figure 1C), the box had light cyan and light magenta pixels to left and right. The dividing line between the colors varied horizontally on the screen, from Level 1 (51 cyan and 150 magenta columns of pixels) to Level 100 (150-51 cyan-magenta). Animals discriminated boxes with smaller or larger left halves.

In the Cyan-Yellow task (Figure 1D), the box contained varying numbers of randomly placed cyan and yellow pixels. For each color, 41 stimulus levels could be presented, from Level 1 (713 pixels) to Level 41 (1,455 pixels). Animals discriminated boxes that were more cyan or more yellow. For this task, trial difficulty is given by the difference (Level Yellow–Level Cyan). This difficulty continuum ran from -40 (easy cyan) to +40 (easy yellow).

Though the four tasks used the same stimulus box for simplicity's sake, they were distinctively different in the cues and relations that indicated a trial's category membership. The Sparse-Dense task presented a cue of box density (or, alternatively, of the sizes and distribution of unfilled black areas in the box). No other task made this cue relevant, and so responses made to those tasks on the basis of density would not produce above-chance performance. Two of the tasks presented a vertical or horizontal edge in pixel color that had to be localized and judged relative to the midline of the box. None of the other tasks made this cue relevant, though there could have been some abstract transfer between the vertical and horizontal edge tasks. Finally, in the cyan-yellow task, the animal had to judge the relational balance between two pixel colors that were fully distributed throughout the box. No other task made this kind of relational color cue relevant across the full width and height of the stimulus box.

For all four tasks, the levels nearest the midpoint of each task continuum were not presented: Level 21 (Sparse-Dense task), Levels 25-26 (Top-Bottom task), Levels 50-51 (Left-Right task), and Difficulty Level -1, 0, +1 (Cyan-Yellow task).

Tasks 3-4 had about twice the levels of Tasks 1-2, and therefore half the trial counts at each level. To better equate these trial counts, we binned the data in Tasks 3 and 4. In Task 3, Levels 1-3 became Bin 1, Levels 4-97 became Bins 2-48 (2 levels per bin), and Levels 98-100 became Bin 49. In Task 4, Levels -40 to -5 became Bins 1 to 18 (two levels per bin), Levels -4 to -2 became Bin 19, Levels -1 to 1 became Bin 20, Levels 2-4 became Bin 21, and Levels 5-40 became Bins 22-39 (2 levels per bin).

Responses—The monkeys touched the box with the cursor to make Dense, SmallTopHalf, SmallLeftHalf, or Yellow discrimination responses. They touched an N icon in the top right of the screen to make Sparse, BigTopHalf, BigLeftHalf, or Cyan discrimination responses. They touched a ? icon at the screen's bottom center to respond Uncertain. In all tasks, the uncertainty response operated in exactly the same way. It ended a trial and cleared the screen, and then the inter-trial interval began. Its selection offered no auditory feedback, punishment, food reward, or hint as to the correct response. Its use did not guarantee an easier or harder next trial. It did not guarantee that the next trial would be from the same task or from a different task type.

Training—Training began with a single task and with the most extreme and discriminable stimuli along the stimulus continuum (e.g., Levels 1 and 41 for Sparse-Dense). The first 20 training trials of a task were guaranteed-win trials in the sense that the cursor could only be moved to the correct response icon. After those trials, both primary responses were available on all trials. When performance on the extreme stimuli exceeded 85% correct over the last 60 trials, training on the next single task began, and so forth through four tasks. Then, animals were tested with the four discriminations intermixed (extreme stimuli only, still no uncertainty response available). When performance on all 4 tasks exceeded 80% correct over the last 60 trials, mature testing was initiated using the double-rectangular distributions described next and the third, uncertainty response became available. Note that this training procedure ensured that animals only experienced the difficult trials in each task after the tasks had been fully interdigitated—when the task of stimulus memory and of stimulus-response association had become extremely difficult. In the present case, animals never had the concentrated exposure to one task’s contingencies and reinforcement histories across the full stimulus continuum that has been available to animals in every other instance of comparative uncertainty-monitoring research.

This sequence of brief training phases was repeated at the beginning of each session, in a re-training phase that lasted about 500 trials. These re-training phases still ensured that the animals experienced the difficult trials in each task only after the tasks had been interdigitated. We analyzed only the mature testing trials (for Hank, Murph, Lou, Gale, and Willie, respectively, 7,280, 6,449, 6,751, 6,352, and 8,046 trials).

Focused trial difficulty—During mature testing, trials were presented in a double-rectangular distribution that let us increase the overall difficulty of the task by over-sampling the most difficult and uncertain trial levels. On 75% of trials, trials were drawn randomly from the whole stimulus range. On 25% of trials, trials were drawn from the most difficult levels as follows: Levels 11-20 and 22-31 (Sparse-Dense), Levels 20-24 and 27-31 (Top-Bottom), Levels 40-49 and 52-61 (Left-Right), and Levels 11-20 and 22-31 for cyan and yellow colors (Cyan-Yellow).

Results

For each task, we conducted two Repeated Measures Analyses of Variance with proportion correct and proportion of uncertainty responses as the dependent measures and trial level or bin level as the within-subjects independent measure. For these analyses, four of the five monkeys were included: Murph, Hank, Gale, and Willie. There was a significant effect of level on the proportion of correct responses in all 4 tasks, $F_{39,117} = 11.21$, $p < 0.001$, $\eta_p^2 = 0.79$ for Sparse-Dense, $F_{47,141} = 8.73$, $p < 0.001$, $\eta_p^2 = 0.74$ for Top-Bottom, $F_{47,141} = 13.16$, $p < 0.001$, $\eta_p^2 = 0.81$ for Left-Right, and $F_{37,111} = 4.11$, $p < 0.001$, $\eta_p^2 = 0.58$ for Yellow-Cyan. Figure 2 (A-D) shows these classic psychophysical data patterns, with excellent performance for easy trial levels at the ends of the stimulus continua but poor performance for levels near the discriminations’ breakpoints.

We conducted two additional analyses to confirm this relationship between level/bin and proportion correct. First, a test of within-subject contrast indicated that the best fit to the primary response function for all four tasks was a quadratic fit, $F = 219.11$, $p = 0.001$, $\eta_p^2 = 0.99$ for Dense-Sparse, $F = 88.38$, $p = 0.003$, $\eta_p^2 = 0.97$ for Top-Bottom, $F = 113.22$, $p = 0.002$, $\eta_p^2 = 0.97$ for Left-Right, $F = 40.90$, $p = 0.008$, $\eta_p^2 = 0.93$ for Yellow-Cyan. Second, we correlated proportion correct with the distance in bins or levels from the breakpoint of the discrimination, and we found correlations of 0.70, 0.79, 0.76, and 0.80, respectively, in the Sparse-Dense, Top-Bottom, Left-Right, and Yellow-Cyan tasks. All these correlations were significant beyond $p < 0.01$. Both these analyses confirm what the figure suggests—that

the lowest levels of performances occurred for the trials nearest the breakpoint of the discriminations, and that higher performances occurred for trial levels trials farther from the central levels of each task. These changes in performance level are the backdrop against which one may see complementary changes in uncertainty responding.

The analysis of uncertainty responding revealed these complementary changes. Figure 2 (A-D) shows the proportion of uncertainty responses by level for each task. There was a significant effect of level on the proportion of uncertainty responses in all 4 tasks, $F_{39,117} = 4.66$, $p < 0.001$, $\eta_p^2 = 0.61$ for Sparse-Dense, $F_{47,141} = 4.08$, $p < 0.001$, $\eta_p^2 = 0.58$ for Top-Bottom, $F_{47,141} = 4.02$, $p < 0.001$, $\eta_p^2 = 0.57$ for Left-Right, and $F_{37,111} = 4.82$, $p < 0.001$, $\eta_p^2 = 0.62$ for Cyan-Yellow. Figure 2 (A-D) shows that monkeys generally declined easier trials (far from the discrimination breakpoint) less often, and difficult trials (near the discrimination breakpoint) more often.

We conducted three additional analyses to confirm this relationship between level/bin and uncertainty responding. First, a test of within-subject contrast indicated that the best fit to the uncertainty response function for three of the four tasks was a quadratic fit, $F = 13.61$, $p = .035$, $\eta_p^2 = .82$ for Dense-Sparse, $F = 9.07$, $p = .057$, $\eta_p^2 = .75$ for Top-Bottom, $F = 7.92$, $p = .067$, $\eta_p^2 = .73$ for Yellow-Cyan. For the Left-Right task, however, the quadratic fit was best fit to the function was a linear fit $F = 18.01$, $p = .024$, $\eta_p^2 = .86$. Second, we correlated uncertainty responding with the distance in bins or levels from the breakpoint of the discrimination, and we found correlations of $-.83$, $-.86$, $-.77$, and $-.54$, respectively, in the Sparse-Dense, Top-Bottom, Left-Right, Yellow-Cyan tasks. All these correlations were significant beyond $p < 0.01$. Third, we correlated uncertainty responding with proportion correct, and we found correlations of $-.84$, $-.73$, $-.69$, and $-.66$, respectively, in the Sparse-Dense, Top-Bottom, Left-Right, Yellow-Cyan tasks. All these correlations were significant beyond $p < 0.01$. All of these analyses confirm what the figure suggests--that the highest levels of uncertainty responding occurred for the trials that were nearest the discrimination breakpoint and that produced the poorest performance.

Monkey Lou, however, performed differently. Lou never used the uncertainty response. He performed as well, however, in all four primary discrimination tasks as did the other four monkeys. However, his failure to use the uncertainty response led to his spending about 57% of his time in this experiment making errors, serving timeouts, and receiving no rewards. In all, he spent more than 6 hours of his time in the task experiencing timeouts for errors. Thus, the task was sufficiently difficult for him that uncertainty responding would have been adaptive.

Discussion

In many uncertainty-monitoring experiments, animals were highly trained in a single task domain. They could have learned that some stimuli are associated with lean reinforcement. They could have adopted stimulus-specific aversion-avoidance responses facing those stimuli. Thus, we asked whether monkeys could transcend these low-level strategies in uncertainty-monitoring tasks. To do so, we randomly interleaved trials from four uncertainty-monitoring tasks. Training was structured so that the animals never received single-task training on the difficult and presumably uncertain trials in any of the stimulus domains. That is, they never received the focused, single-task trial experience that would have fostered the development of stimulus-based reinforcement histories and response aversions. Under these conditions, four monkeys made adaptive uncertainty responses across four domains simultaneously, suggesting the multi-functionality of the psychological signal that occasions these responses.

The findings are a step toward showing that monkeys have an uncertainty-monitoring capacity that is like that of humans in transcending task-specific cues and extending simultaneously to multiple domains. For one thing, the complexity of the monkeys' situation—with four 40-level discriminations underway concurrently—decreased the likelihood that monkeys could develop four sets of reinforcement histories simultaneously or maintain four different sets of associative/conditioning processes involving the same response icons. For example, successive difficult trials from the Sparse-Dense task would have been separated on average by about 16 highly variable trials, yet the monkeys would have had to associate these to each other, to their common reinforcement history, and to a common response mapping. Given the difficulty of this memory and mapping problem, another plausible explanation might be more parsimonious. That is, animals could have performed all four tasks simultaneously with no memory or mapping complexity, by simply monitoring the difficulty and indeterminacy of trials and responding Uncertain in those instances.

For another thing, animals demonstrated in the present case a generalizability and flexibility to their use of the uncertainty response that had not been previously demonstrated. Their performance approached the multi-functional flexibility with which humans use and report uncertainty. When animals verge on showing a crucial aspect of human metacognition, it becomes necessary to consider the theoretical possibility that they are demonstrating an isomorphic capability.

However, we acknowledge the difficulty of proving that animals have human-like metacognition. In the present case, it is possible, though unlikely and unparsimonious, that animals did accomplish the difficult stimulus-memory and response-mapping problem posed by four interdigitated tasks. This possibility would remain even if animals interdigitated forty tasks, and even if they intermixed perceptual and non-perceptual tasks. Therefore, though these would be interesting extensions of the present research, different scholars could and would still bring different proof standards to them.

Moreover, we accept that there are general cues that cut across tasks and reliably indicate difficult trials, so in an abstract way uncertainty monitoring might be construed to be cue-driven. For example, the perceptual system could send out a general indeterminacy signal when it encounters unreliable perceptual information and/or decisional difficulty. This indeterminacy signal would be common across tasks, and it would occasion adaptive uncertainty responses across several stimulus domains at once, simply and without the need for feats of stimulus memory and associative connection. (Note: there were no specific cues that could be used across all tasks such as density, color, and edge location--none of these could account for the strong performances shown in the four primary discriminations).

In a sense, the goal of our article was to use concurrent performance in four difficult and different tasks to focus attention on the possible existence of these more general signals of indeterminacy. These cues are more general, more derived, more second order, and they have a "monitoring" character that one could label metacognitive.

This description dovetails with Carruthers's (2008) discussion of animal metacognition. He proposed that some species have in place "a mechanism (most likely evolved, but perhaps constructed through some sort of learning) which when confronted with conflicting plans that are too close to one another in strength will refrain from acting on the one that happens to be strongest at that moment, and will initiate alternative information-gathering behavior instead." This "gate-keeper" mechanism (Carruthers, 2008) is more complex and demanding than first-order cognitive processes. It is second order, because it operates on the outputs of lower cognitive processes. It is meta-, because it is responsive to indeterminacy in lower-level cognitive processes. It contains an extra layer of representational complexity as

metacognitive processes must. The gatekeeper meets the definition of controlled cognitive processes (voluntary, deliberate, etc.; see Atkinson and Juola, 1971; Shiffrin and Schneider, 1977). Indeed, a 100-year history in psychophysics suggests that gate-keeper processes near participants' perceptual thresholds are higher-level and metacognitive in psychological character (Boring, 1920; Fernberger, 1914; Watson et al., 1973). The gate-keeper also engenders a qualitative change in behavior and cognition (hesitancy, information seeking, uncertainty responses, etc.). This gate-keeper process is close to what we envision for animal minds, and that process typifies the traditional construct of metacognition.

However, we also propose, as does also Caruthers (2008), that there are characteristics that general indeterminacy signals need not possess. They would not need to be fully conscious. They would not have to involve a self-awareness by the organism that it is in difficulty (e.g., Proust, 2003, 2007). They would not need to be explicitly reportable (except insofar as the adaptive uncertainty response is a report of that signal). Thus, by no means do the present results, or the presence of generalized indeterminacy signals in animal mind, demand that one elevate what animals are doing to the full status of human metacognition.

In fact, it is a central theme of our ongoing research that researchers and theorists not take an all-or-none perspective toward animal metacognition. This perspective could distract researchers from the theoretically rich middle ground wherein one can grant animals the cognitive sophistication of their performance, and acknowledge the monitoring quality of their performance, without elevating what they do too far. In this middle ground lies the whole phylogenetic emergence of human metacognition, and probably also the ontogenetic emergence of metacognition in human development. Therefore, we urge that this field not adopt an all-or-none perspective, because its focal construct—animal metacognition—is not all of none either.

Our research joins other approaches that are addressing concerns about animals' uncertainty-monitoring performances by exploring their sophistication. For example, Washburn et al. (2006) showed that monkeys could generalize their use of the uncertainty response to the first trial of new tasks (see also Kornell et al., 2007). The immediate transfer of the uncertainty response to new tasks helps confirm the flexibility with which uncertainty responses are applied to difficult trials. These transfer results with the uncertainty response complement the present results on multi-tasking.

Call and Carpenter (2001) and Suda-King (2008) tested apes in another complementary approach to ours. They combined a naturalistic uncertainty task (which container has food) with a naturalistic uncertainty response (e.g., seeking information about which food tube was baited). Animals showed adaptive uncertainty behaviors with essentially no training. In that research, the training-specificity problem was addressed by the absence of training, whereas here we addressed it through multi-tasking.

Animals' use of the uncertainty response is also being explored in abstract cognitive domains like memory. Hampton (2001), Kornell et al. (2007), and Smith et al. (1998) showed that animals make adaptive uncertainty responses when they face indeterminate memories. The higher and more sophisticated the cognitive state that the uncertainty response is responsive to, the less tenable are low-level, associative descriptions grounded in stimulus-response connections.

Finally, Kornell et al. (2007) and Shields et al. (2005) let animals make a low- or high-risk betting response after answering each trial. These studies verged on asking animals to rate their confidence about the trial's outcome.

Thus, there is strong promise that this research area is moving beyond the low-level behavioral descriptions that have shadowed it, and there is recent evidence that some species may meet more stringent tests of their uncertainty systems (e.g., by showing the multi-functionality of the uncertainty response as here, the immediacy of transfer in other studies, and so forth). In a sense, researchers in this area are taking aim at animal awareness and consciousness from the ground up.

Of course there is more progress to be made in assessing whether some animals have a metacognitive capacity like that of adult humans. For example, regarding the present research, it might be possible to extend farther the idea of multi-functionality. Human uncertainty is momentary, situational, and infinitely variable—beyond 4 or even 200 tasks. Human uncertainty arises with every forgotten PIN number and every annual tax form. Human uncertainty is also explicit, declarative, and conversational—we share these states and seek clarification from others. One might collect situational uncertainty reports from animals by giving them a portable uncertainty response for use through their day and seeing if they spontaneously report when they do not remember or when they need clarification. Such a response could be intrinsically rewarding to animals, even though it was not extrinsically rewarded. For it could facilitate their interactions with trainers, help them learn new symbols or lexigrams, and let them understand better the course of upcoming activities. If animals made these reports, they would be untrained, un-conditioned, situational and declarative reports about the momentary contents of mind. They could turn out to be a strong indication of declarative consciousness in animals. Though researchers of comparative metacognition have not yet reached empirical endpoints of this kind, the strong empirical base they have laid down now puts those demonstrations potentially within reach.

Acknowledgments

This research was supported by National Institutes of Health Grant HD - 38051 and by Grant BCS-0634662 from the National Science Foundation.

All applicable institutional rules and regulations regarding animal care and use were followed in the care and testing of the monkeys, and the experiment complied with all laws of the United States of America. The authors thank Theodore Evans for his assistance with data collection.

References

- Atkinson, RC.; Juola, JF. Search and decision processes in recognition memory. In: Krantz, DH.; Atkinson, RC.; Luce, RD.; Suppes, P., editors. *Learning, memory, and thinking*. W.H. Freeman; San Francisco CA: 1974. p. 242-293.
- Beran MJ, Smith JD, Redford JS, Washburn DA. Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *J Exp Psychol Anim Behav Proc*. 2006; 32:111–119.
- Boring EG. The control of attitude in psychophysical experiments. *Psychol Rev*. 1920; 27:440–452.
- Call J, Carpenter M. Do apes and children know what they have seen? *Anim Cogn*. 2001; 4:207–220.
- Carruthers P. Meta-cognition in Animals: a skeptical Look. *Mind Lang*. 2008; 23:58–89.
- Fernberger SW. The effect of the attitude of the subject upon the measure of sensitivity. *Am J Psychol*. 1914; 25:538–543.
- Flavell JH. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *Am Psychol*. 1979; 34:906–911.
- Footo A, Crystal J. Metacognition in the rat. *Curr Biol*. 2007; 17:551–555. [PubMed: 17346969]
- Gallup GG. Self-awareness and the emergence of mind in primates. *Am J Primatol*. 1982; 2:237–248.
- Ghetti S, Lyons K, Lazzarin F. The development of metamemory monitoring during retrieval: The case of memory strength and memory absence. *J Exp Child Psychol*. 2008; 99:157–181. [PubMed: 18191139]

- Hampton RR. Rhesus monkeys know when they remember. *Proc Nat Acad Sci*. 2001; 98:5359–5362. [PubMed: 11274360]
- Koriat A. How do we know that we know? The accessibility model of the feeling of knowing. *Psychol Rev*. 1993; 100:609–639. [PubMed: 8255951]
- Koriat, A. Metacognition and consciousness. In: Zelazo, PD.; Moscovitch, M.; Thompson, E., editors. *The Cambridge handbook of consciousness*. Cambridge University Press; Cambridge, UK: 2007. p. 289-325.
- Koriat A, Ma'ayan H, Nussinson R. The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *J Exp Psychol Gen*. 2006; 135:36–69. [PubMed: 16478315]
- Kornell N, Son L, Terrace H. Transfer of metacognitive skills and hint seeking in monkeys. *Psychol Sci*. 2007; 18:64–71. [PubMed: 17362380]
- Kuhn D. Metacognitive development. *Curr Dir Psychol Sci*. 2000; 9:178–181.
- Lockl K, Schneider W. Developmental trends in children's feeling-of-knowing judgements. *Int J Behav Dev*. 2002; 26:327–333.
- Metcalfe, J. Metamemory: Theory and data. In: Tulving, E.; Craik, FIM., editors. *The Oxford handbook of memory*. Oxford University Press; New York: 2000. p. 197-211.
- Nelson, TO., editor. *Metacognition: Core readings*. Allyn and Bacon; Toronto: 1992.
- Nelson TO. Consciousness and metacognition. *Am Psychol*. 1996; 51:102–116.
- Proust J. Does metacognition necessarily involve metarepresentation? *Behav Brain Sci*. 2003; 26:352.
- Proust J. Metacognition and metarepresentation: is a self-directed theory of mind a precondition for metacognition? *Synthese*. 2007; 2:271–295.
- Roebbers CM, von der Linden N, Schneider W, Howie P. Children's metamemorial judgments in an event recall task. *J Exp Child Psychol*. 2007; 97:117–137. [PubMed: 17306823]
- Rumbaugh DM, Richardson WK, Washburn DA, Savage-Rumbaugh ES, Hopkins WD. Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *J Comp Psychol*. 1989; 103:32–38. [PubMed: 2924530]
- Schwartz BL. Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychon Bull Rev*. 1994; 1:357–375. [PubMed: 24203520]
- Shields WE, Smith JD, Guttmanova K, Washburn DA. Confidence judgments by humans and rhesus monkeys. *J Gen Psychol*. 2005; 132:165–186. [PubMed: 15871299]
- Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol Rev*. 1977; 84:127–90.
- Smith JD. The homunculus at home. Commentary on J A Gray, The contents of consciousness: A neuropsychological conjecture. *Behav Brain Sci*. 1995; 18:697–698.
- Smith JD, Beran MJ, Couchman JJ, Coutinho MVC, Boomer J. Animal metacognition: Problems and prospects. *Comp Cognit Behav Revs*. 2008 In press.
- Smith JD, Beran MJ, Coutinho MVC, Couchman JC. The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychon Bull Rev*. 2008; 154:679–691. [PubMed: 18792496]
- Smith JD, Beran MJ, Redford JS, Washburn DA. Dissociating uncertainty states and reinforcement signals in the comparative study of metacognition. *J Exp Psychol Gen*. 2006; 135:282–297. [PubMed: 16719654]
- Smith JD, Shields WE, Allendoerfer KR, Washburn WA. Memory monitoring by animals and humans. *J Exp Psychol Gen*. 1998; 127:227–250. [PubMed: 9742715]
- Smith JD, Shields WE, Washburn DA. The comparative psychology of uncertainty monitoring and metacognition. *Behav Brain Sci*. 2003; 26:317–373. [PubMed: 14968691]
- Suda-King C. Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Anim Cogn*. 2008; 11:21–42. [PubMed: 17437141]
- Sutton JE, Shettleworth SJ. Memory without awareness: Pigeons do not show metamemory in delayed matching-to-sample. *J Exp Psychol Anim Behav Proc*. 2008; 34:266–282.
- Washburn DA, Rumbaugh DM. Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. *Behav Res Method Ins C*. 1992; 24:157–164.

- Washburn DA, Smith JD, Shields WE. Rhesus Monkeys (*Macaca mulatta*) immediately generalize the uncertain response. *J Exp Psychol Anim Behav Proc.* 2006; 32:85–89.
- Watson CS, Kellogg SC, Kawanishi DT, Lucas PA. The *uncertain* response in detection-oriented psychophysics. *J Exp Psychol.* 1973; 99:180–185.

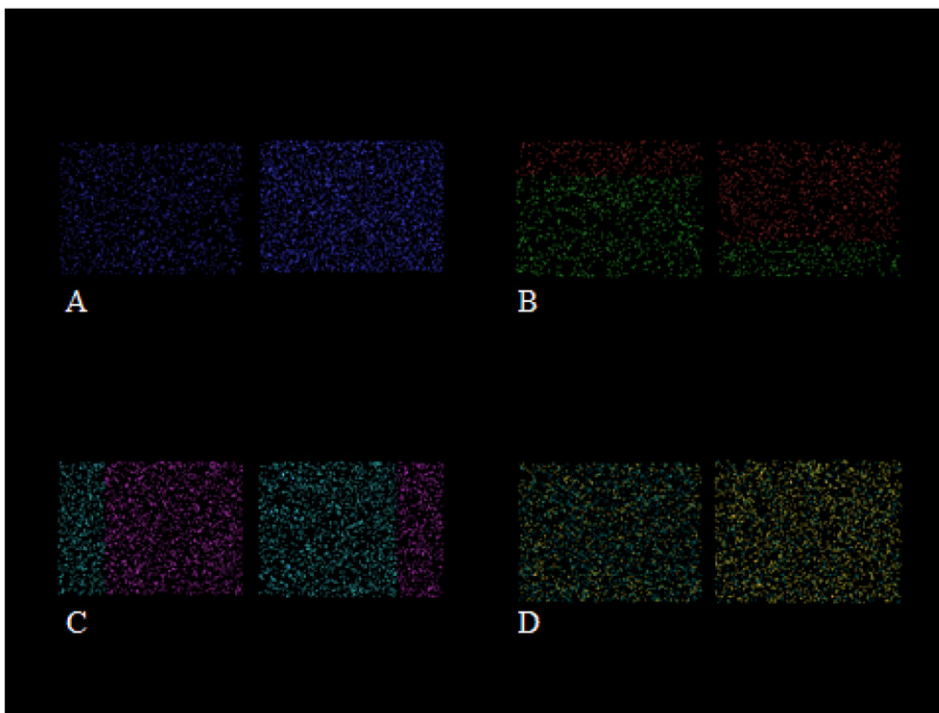


Figure 1.

A. Sample stimuli from the Sparse-Dense task. For Levels 1 to 41, the number of blue pixels was given by the formula $\text{Pixels} = \text{round}(1400 \times 1.018^{\text{Level}})$. Levels 1-20 and 22-41 were designated Sparse and Dense trials, respectively. The discrimination's breakpoint (Level 21) was not presented. Levels 1 and 41 are shown. B. Sample stimuli from the Top-Bottom task. The dividing line between the colors varied from Level 1 (26-75 red-green) to Level 50 (75-26 red-green). Levels 25 (50-51 red-green) and 26 (51-50 red-green) were not presented. Levels 1 and 50 are shown. C. Sample stimuli from the Left-Right task. The dividing line between the colors varied from Level 1 (51-150 cyan-magenta) to Level 100 (150-51 cyan-magenta). Levels 50 (100-101 cyan-magenta) and 51 (101-100 cyan-magenta) were not presented. Levels 1 and 100 are shown. D. Sample stimuli from the Cyan-Yellow task. For Levels 1 to 41, the number of cyan or yellow pixels was given by the formula $\text{Pixels} = \text{round}(700 \times 1.018^{\text{Level}})$. On More Cyan trials, lower yellow levels (Levels 1 to 20) were paired with higher cyan levels (Levels 22-41). On More Yellow trials, the reverse was true. Level 21 was not presented in either color. Difficulty Levels -40 (Cyan 41, Yellow 1) and +40 (Cyan 1, Yellow 41) are shown.

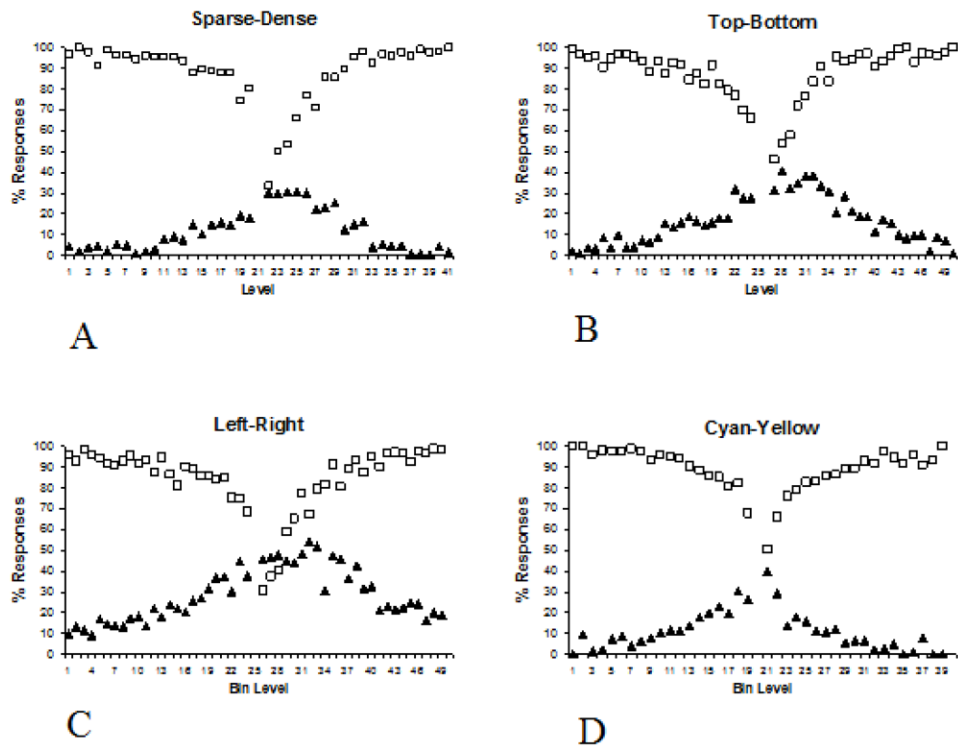


Figure 2. Monkeys' performance in the Sparse-Dense (A), Top-Bottom (B), Left-Right (C), and Cyan-Yellow (D) psychophysical tasks. The horizontal axis indicates the trial or bin level. The percentage of trials receiving the uncertainty response at each level is shown (filled triangles) along with the percentage of correct responses made when animals completed trials with a primary discrimination response (open squares).