Boise State University
ScholarWorks

Curriculum, Instruction, and Foundational Studies Faculty Publications and Presentations

Department of Curriculum, Instruction, and Foundational Studies

1-1-2010

Evidence of Metacognitive Control by Humans and Monkeys in a Perceptual Categorization Task

Joshua Redford
Boise State University

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Journal of Experimental Psychology: Learning, Memory, and Cognition*, published by American Psychological Association. Copyright restrictions may apply. DOI: 10.1037/a0017809

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Evidence of Metacognitive Control by Humans and Monkeys in a Perceptual Categorization Task

Joshua S. Redford
Boise State University

Keywords: Metacognition, Rhesus Macaques, Categorization, Dot Distortions

Abstract

Metacognition research has focused on the degree to which nonhuman primates share humans' capacity to monitor their cognitive processes. Convincing evidence now exists that monkeys can engage in metacognitive monitoring. By contrast, few studies have explored metacognitive control in monkeys and the available evidence of metacognitive control supports multiple explanations. The current study addresses this situation by exploring the capacity of human participants and rhesus monkeys (*Macaca mulatta*) to adjust their study behavior in a perceptual categorization task. Humans and monkeys were found to increase their study for high-difficulty categories suggesting that both share the capacity to exert metacognitive control.

Metacognition—defined as thinking about thinking—involves a monitoring and control component (Nelson & Narens, 1990). The monitoring component is responsible for assessing the mind's basic mental processes (Dunlosky & Nelson, 1992; Koriat & Bjork, 2005; Koriat & Ma'ayan, 2005; Serra & Dunlosky, 2005; Thiede, Anderson, & Therriault, 2003). Metacognitive monitoring provides information including how difficult information will be to learn and how well information has been learned. Metacognitive control uses this information to control study. For example, students that quit studying because they decide that the information is well-learned are displaying metacognitive control. In this case, metacognitive monitoring provided information regarding the state of learning (i.e., the information was well-learned) and metacognitive control enacted the appropriate action (i.e., to terminate study). Effective metacognitive monitoring and control have been shown to be critical for learning (Thiede, 1999).

Research has explored whether rhesus macaques (*Macaca mulatta*–hereafter referred to as monkeys) can demonstrate metacognitive monitoring. Smith, Shields, Schull, and Washburn (1997) used a perceptual discrimination task similar to tasks used in psychophysics research. In the original paradigm, monkeys encountered unframed boxes filled with varying pixel densities. They classified the stimulus as either sparsely or densely populated with pixels or they escaped from the trial. The escape response was termed the uncertainty response as monkeys (and humans) used it selectively for the more difficult trials. However, the uncertainty response usage in this original paradigm may have been either conditioned or used as a third category response. That is, monkeys may have used the uncertainty response to escape punishment associated with the difficult trials or because those difficult trials were learned as a third category, not because they were experiencing uncertainty.

Recent research has provided stronger evidence supporting the monitoring ability of monkeys. First, a betting paradigm (Shields, Smith, Guttmannova, & Washburn, 2005; Son & Kornell, 2005) enabled monkeys to indicate their confidence in responses by "betting". High confidence bets resulted in larger rewards and larger penalties for correct and incorrect responses, respectively. Low confidence bets resulted in smaller rewards and smaller penalties for correct and incorrect responses, respectively. The major finding of these studies was that monkeys made bets that were consistent with their accuracy. That is, they made large bets more often when their response was correct and small bets more often when their response was incorrect. Kornell, Son, and Terrace (2007) used this paradigm to demonstrate that monkeys can accurately identify their confidence level across different tasks. Of particular import, these confidence judgments transferred across tasks that were qualitatively different (i.e., from a perceptual task to a working memory task). Second, Smith, Beran, Redford, and Washburn (2006) showed that the uncertainty response is used on the most difficult trials of a task even if monkeys are not given feedback after each trial. Specifically, a monkey used the uncertainty response on the most difficult trials in a sparse-dense task despite accrued feedback

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

being reordered and presented after every four trials. This accurate uncertainty response usage also persisted despite pixel range shifts that modified which boxes were classified as sparse and which boxes were classified as dense. These studies suggest that monkeys can accurately monitor across situations/tasks and without direct feedback.

By contrast, little research exists regarding the capacity of monkeys to exert metacognitive control. Previous research (Call, 2004; Call & Carpenter, 2001; Hampton, Zivin, & Murray, 2004) showed that monkeys sought more information about the location of food when they did not know where the food was hidden. In cases where monkeys were shown the food while it was hidden, they did not seek additional information, but chose to immediately reach for the container with the food. Kornell et al. (2007) comment that any potential metacognitive interpretation of these data is compromised because the behavior of interest (i.e., whether monkeys search before reaching for food) reflects food-seeking strategies that are default behavioral patterns that monkeys typically display when the location of food is unknown. In other words, searching for food may only signify that the monkey wants food and does not have it, not that the monkey knows that he does not know the location of a hidden food. Kornell and his colleagues advanced research on metacognitive control in monkeys in their second experiment. They used a sequence-learning task that the monkeys had been trained on in a previous study (Terrace, Son, & Brannon, 2003). They modified the paradigm in this experiment by providing the monkeys with a hint option to assist in their attempts to recall the correct order of an image sequence. When selected, these hints identified the next image that the monkey should select. Their primary finding was that hint usage was greatest when accuracy was lowest.

Kornell et al. (2007) do offer stronger evidence that monkeys can exert metacognitive control. However, certain limitations still undermine a pure metacognitive interpretation of the data. The most problematic aspect of this study was that hint usage basically guaranteed a (less preferred) reward by making the task trivial. This was also one of the most problematic elements of earlier uncertainty monitoring paradigms (e.g., Smith, Shields, Schull, and Washburn, 1997)—the uncertain response guaranteed some form of reward. Therefore, the monkeys may have been conditioned to seek hints rather than risk penalties with unfamiliar sequences. Hint usage may have decreased as sequences became known because they were paired more often with a better reward, not because monkeys "knew that they knew." This possibility is made more likely because each session forced the monkeys to attempt sequence recall without the hint option for half of the trials. Therefore, unknown sequences were guaranteed to be paired with penalties and learned sequences were guaranteed to be paired with rewards.

To explore metacognitive control in monkeys, I used a dot distortion category-learning task. This task seemed optimal for several reasons. First, categorization tasks often involve a learning and test phase. The learning phase involves the study of category exemplars and the test phase involves the classification of exemplars as either belonging or not belonging to the studied category. This separation of phases permits metacognitive control to be measured independent of test performance. This separation also eliminates the risk of feedback at test from influencing study behavior. To further ensure that feedback did not influence behavior, the difficulty levels were randomly ordered and the test items were randomly ordered. Second, dot distortion categorization has been extensively studied in humans (e.g., Little & Thulborn, 2006; Minda & Smith, 2002; Zaki, Nosofsky, Stanton, & Cohen, 2003) and recently demonstrated in monkeys (Smith, Redford, & Haas, 2008). This similar ability to learn and categorize dot distortions ensures that this task can be administered to both humans and monkeys. Lastly, dot distortion categories can vary in difficulty which, in turn, influences their rate of learning and future test performance (Homa & Cultice, 1984). A common approach to vary the difficulty of category learning is to strengthen or weaken the degree of similarity (family resemblance) across a category's exemplars. Recent evidence has shown that monkeys perceive these changes in similarity (Smith, Redford, Haas, Coutinho, & Couchman, 2008). Therefore, monkeys may be able to use perceived changes in family resemblance to guide how they learn the dot distortion categories. Specifically, if humans and monkeys are exerting metacognitive control, they should choose to view more study trials when they are learning high-difficulty categories (i.e., those with a weak family resemblance across exemplars) relative to low-difficulty categories (i.e., those with a strong family resemblance across exemplars). I first present humans with this categorization task.

Experiment 1

In this experiment, human participants studied categories that varied in family resemblance—the weaker the family resemblance, the more difficult the category was to learn. If the participants exert accurate metacognitive control in this task, then they should choose to view more study trials as the difficulty is increased.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Method

Participants

Fifty students from the University at Buffalo participated in this experiment in partial fulfillment of their introductory psychology course requirements. All participants were treated in accord with APA ethical standards.

Design

This experiment was a single-factor design with family resemblance (strong, intermediate, weak) serving as the within-participant independent variable.

Materials

The dot distortion categories were created with a method described in Smith and Minda (2002). As the distortion level of the exemplars was increased, the family resemblance across dot distortions weakened which, in turn, made categories more difficult to learn (Homa & Cultice, 1984). Specifically, a weaker family resemblance increases category-learning difficulty because a learner has to view more exemplars in order to successfully use category knowledge at test. Figures 1A, 1B, and 1C display exemplars sharing a strong, intermediate, and weak family resemblance, respectively.

Procedure

Participants were seated in one of three different experiment rooms and read instructions. These instructions described the task's requirements and encouraged participants to move onto the test phase when they felt prepared.

Figure 2A shows a screenshot of a study trial. Each study trial displayed a red dot distortion at the top middle of the screen, a cursor in the center of the screen, and a red-filled circle at the bottom middle of the screen. On every trial, the dot distortion was presented alone onscreen for 2.5 s followed by the appearance of the cursor and red circle. Once the cursor and circle appeared, participants had the option of either moving the cursor to the dot distortion or to the circle. A quarter second pause followed all choices. Moving the cursor to the dot distortion gave participants another dot distortion to view. Moving the cursor to the circle transitioned the participants to the test phase

Each study phase involved a new to-be-learned category consisting of either level 3, level 5, or level 7 dot distortions. Distortion level 3, 5, and 7 corresponded to low, intermediate, and high-difficulty conditions, respectively. The study phases were randomized without replacement until participants completed each difficulty level. After all three difficulty levels were finished, the iteration repeated with a new randomly ordered set of three difficulty levels. This randomization-without-replacement process repeated until the end of the experiment. The overall number of studied categories was dependent on how quickly participants went through the study and test phases in the time allotted.

During the test phase, participants decided if new dot distortions belonged to the studied category. Each test phase consisted of the prototype, five level 5 trials, five level 7 trials, and 11 trials where the test item was a level 7 dot distortion of a randomly generated prototype (hereafter, referred to as random dot distortions). These trial types were randomly ordered anew for each test phase. Figure 2B shows a screenshot of a test trial. Each test trial displayed the to-be-categorized dot distortion at the top middle of the screen, a cursor in the center of the screen, and a large "N" at the bottom middle of the screen. To accept a dot distortion as a category member, the participant moved the cursor to the dot distortion at the top middle of the screen (the same cursor movement performed during the study phase to view additional dot distortions). To reject a dot distortion as a category member, the participant moved the cursor to the large "N" at the bottom middle of the screen (the same cursor movement performed during the study phase to enter the test phase). Correct responses received a beep to indicate that the participant was correct and a point was added to the participant's score. Incorrect responses received a 10 s buzzing sound and two points were subtracted from the participant's score. After completion of the final test phase in a cycle, a participant began another cycle with a new category to learn. The experiment was programmed to run for about 45 min.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Results

Study phase difficulty influenced the number of trials that participants chose to view prior to test, $F(2, 98) = 11.19$, $p < .05$, $MSE = 21.58$ (Table 1, Rows 1-3). Follow-up tests found that participants chose to view more trials in the high-difficulty condition, than in the intermediate- or low-difficulty condition. However, the intermediate- and low-difficulty conditions were not different. These data suggest that humans exerted metacognitive control. In line with most of the previous research exploring metacognitive control, participants chose more trials when learning high-difficulty material—in this case, categories with the weakest family resemblance.

As this category learning paradigm provided potential evidence of metacognitive control by human participants, it was presented to monkeys in the next experiment.

Experiment 2

In this experiment, I gave monkeys a modified version of the task used in Experiment 1. Human participants chose to view more study trials during the more difficult study phases as they prepared for a categorization test, thereby providing evidence that humans can employ metacognitive control in this task. This experiment determined if monkeys display the same pattern.

Method

Participants

Three monkeys—Murph (14 years old), Gale (24 years old), and Han (4 years old)—participated in this experiment. The monkeys were housed at the Language Research Center of Georgia State University in Atlanta, GA. They were singly housed in rooms that offered constant visual and auditory access to other monkeys. They also were periodically group-housed with compatible members of the same species in outdoor-indoor housing units. All monkeys were maintained on a healthy diet including fresh fruits, vegetables, and monkey chow each day independent of their computer test schedule. The monkeys were not restricted in food intake for the purposes of testing. Each monkey was tested using the Language Research Center's Computerized Test System (LRC-CTS; described in Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989; Washburn & Rumbaugh, 1992) that consists of a PC computer, a digital joystick, a color monitor, and a pellet dispenser. Monkeys manipulated the joystick through the mesh of their home cages, producing isomorphic movements of an onscreen cursor. Rewarded responses resulted in the delivery of a 94-mg fruit-flavored chow pellet (Bioserve, Frenchtown, NJ) using a Gerbrands 5120 dispenser interfaced to the computer through a relay box and output board (PIO-12 and ERA-01; Keithley Instruments, Cleveland, OH). Each test session was started by the facility's Research Coordinator or a Research Technician. Once initiated, monkey tests were autonomous and required no human monitoring. Murph, Gale, and Han have participated in dozens of studies associated with cognitive psychology and animal learning, including spatial and working memory, numerical cognition, judgment and decision making, psychomotor control, discrimination learning, planning, concept learning, categorization, and metacognition. All experiments were conducted after obtaining approval from the Institutional Animal Care and Use Committee.

Training

A series of training sessions prepared the monkeys for the experiment. The study phase of these training sessions involved 120 mandatory level 7 dot distortions presented for 3.0 s each. After 3.0 s passed, each monkey was permitted to move the cursor up to the dot distortion to view the next dot distortion. They received a food pellet for each study trial selected. After the final study trial, the test phase began. The test phase involved 352 randomly ordered trials and consisted of 16 prototypes, 80 level 5 dot distortions, 80 level 7 dot distortions, and 176 random dot distortions. Like the human participants in Experiment 1, monkeys made category endorsements by moving the cursor up to the displayed dot distortion at the top middle of the screen and rejected dot distortions by moving the cursor to the “N” at the bottom middle of the screen. Each monkey trained until he performed well on the test phase. During their final training session, Murph obtained an overall accuracy of 79.3%, Gale obtained an overall accuracy of 63.6%, and Han obtained an overall accuracy of 74.1%.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

At this point, the monkey transitioned to the experimental program for additional piloting. This pilot program was identical to the program described below except the test phase was not equated for difficulty in the transitional program. After they demonstrated proficiency on this task they transitioned to the experimental program described below.

Design

The design of this category learning task was identical to that of Experiment 1.

Materials

The materials were similar to Experiment 1. The stimuli were generated using the same distortion algorithms used in Experiment 1. The single differences was that the test phase contained 11 level 7 dot distortions and 11 random dot distortions.

Procedure

Four modifications were made to the procedures used in Experiment 1. First, instead of verbal instructions, monkeys learned the task's requirements through experience across multiple sessions. Second, rewarded responses yielded food pellets in addition to a beep sound and errors resulted in a longer buzzing sound (20 s instead of 10 s). Third, the probability of earning pellets during the study phase slowly decreased as additional dot distortions were viewed. For the first trial of each study phase, monkeys were guaranteed a pellet for choosing to view a second dot distortion. Thereafter, for each additional selected dot distortion that produced a pellet, the probability of receiving another pellet decreased by 2% (1% for Han) until pellet earnings were at chance levels (50%). Pellet dispersal remained at chance levels for the rest of the study phase. This declining reward rate encouraged monkeys to view variable numbers of study trials. The fourth and final change involved the composition of the test phase. In Experiment 1, test phases consisted of the prototype, five level 5 trials, five level 7 trials, and 11 random trials. As these test trials were randomized, the test phases could vary in their overall difficulty—an unintentional manipulation. For instance, if the first test trial involved the prototype or a level 5 dot distortion, then performance could improve by comparing the perceptual similarity of this first trial to the subsequent trial(s), eliminating the need for category learning during the study phase. The solution was to modify the randomized test set so that it contained only level 7 dot distortions and random dot distortions. This increased the likelihood that all test phases were of equal difficulty.

Results

Analyses of the monkey study data differed from the analyses of the human participant data. Instead of performing analyses on whole sessions, the monkey's study phases were partitioned into completed sets of the three difficulty levels (low, intermediate, and high). The set-to-set variability provided the error variance to perform tests of significance.

For each monkey data were collected and summarized for each complete set of three phases. Study phase difficulty influenced the number of trials that Murph chose to view prior to test, $F(2, 954) = 151.31$, $p < .05$, $MSE = 97.12$ (Table 1, Rows 4-6). Follow-up tests found that Murph chose to view more high-difficulty study trials than intermediate-difficulty study trials and he chose to view more intermediate-difficulty study trials than low-difficulty study trials. Murph's data pattern suggests that he exerted metacognitive control.

Study phase difficulty influenced the number of trials that Gale chose to view prior to test, $F(2, 528) = 7.27$, $p < .05$, $MSE = 223.91$ (Table 1, Rows 7-9). Follow-up tests found that Gale chose to view more high-difficulty study trials than either intermediate- or low-difficulty study trials, but he did not choose to view more intermediate- than low-difficulty study trials. Gale was as sensitive to difficulty as the humans. Gale and the human participants chose to view more high-difficulty study trials than any others, but neither Gale nor the human participants differentiated the low- and intermediate-difficulty study conditions. Therefore, Gale may have exerted metacognitive control in this task.

Study phase difficulty did not influence the number of trials that Han chose to view prior to test, $F(2, 436) < 1.00$ (Table 1, Rows 10-12). Therefore, Han did not display any pattern consistent with metacognition.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Alternative Hypotheses

Although the data have been presented as indicating metacognitive control, alternative hypotheses exist. The following section considers two promising hypotheses. Both of these hypotheses offer another explanation for why study trial views increase as the difficulty of the study phase is increased.

Test Performance Hypothesis

One plausible explanation for this effect is that more study trials are viewed because the test is less rewarding as the study phase becomes more difficult. An increase in study trial views may indicate a postponement of the inevitable punishment at test. This hypothesis depends on poorer test performance as the difficulty of category learning increases. In other words, the fewer rewards associated with poorer performance is the factor that would make these tests less attractive to take.

Study phase difficulty influenced humans' test performance, $F(2, 98) = 20.72$, $p < .05$, $MSE = 29.33$ (Table 2, Row 1). Participants performed better at test after the distortion level 5 study phase than after the distortion level 7 study phase, $F(1, 49) = 17.57$, $p < .05$, $MSE = 31.90$ and they performed better at test after the distortion level 3 study phase than after the distortion level 5 study phase, $F(1, 49) = 4.72$, $p < .05$, $MSE = 22.60$. Therefore, the test performance hypothesis may account for the human data.

Study phase difficulty did not influence Murph's test performance, $F(2, 942) = 2.17$, $p > .05$ (Table 2, Row 2) or Gale's test performance, $F(2, 506) < 1.00$ (Table 2, Row3). Murph and Gale performed equally well after all study phase conditions. As the tests appear to be equally rewarding across difficulty levels for Murph and Gale, it is unlikely that test performance is influencing the number of study trial views. Han's test performance data is also presented for the interested reader (Table 2, Row 4).

Only humans displayed poorer test performance as study phase difficulty was increased. Therefore, the test performance hypothesis can be discounted as it is improbable that only monkeys display metacognitive abilities in this task.

Novelty Preference Hypothesis

A second hypothesis is that study trial views increase because the weaker family resemblance across exemplars makes those stimuli more interesting to view. Because participants see greater changes occur from one trial to the next, they are inclined to view more of those types of trials. Indirect evidence supports this hypothesis as humans do prefer novel visual stimuli starting in infancy (e.g., Thompson, Petrill, DeFries, Plomin, & Fulker, 1994). However, this hypothesis makes another prediction: If novelty is driving the effect, then study trial views should decline over time as the novelty of these trials dissipates.

To address the novelty preference hypothesis, I reexamined the human participants' data using split-group analyses. To do these analyses, the number of viewed study trials was recorded for each completed set of three study phases (low-, intermediate-, and high-difficulty). The last set of study phases was omitted if participants completed fewer than the three study phases. These data were split into two equal halves. In cases where participants completed an odd number of study sets, the middle set was also excluded from the split-group analyses. For example, if a participant completed five study-test cycles, the third cycle was excluded from the split-group analyses. Hereafter, the first half will be referred to as the Start Phase and the second half will be referred to as End Phase. When discussed together, the factor will be labeled *time-phase*.

For humans, split-group analyses did not show a main effect of time-phase, $F(1, 294) < 1.00$, but did show a main effect of difficulty level, $F(2, 294) = 5.16$, $p < .05$, $MSE = 151.44$ (Table 3, Rows 1-3). No interaction between time-phase and difficulty level emerged, $F(2, 294) < 1.00$. Novelty is unlikely to account for the influence of difficulty level on study trial views as study trial views remained constant over time.
Indirect evidence also supports the novelty preference hypothesis as an explanation for the monkey data because primates exhibit a novelty preference beginning in infancy (Gunderson & Sackett, 1984). Moreover, novelty preference has been reported specifically in rhesus monkeys (Golub & Germann, 1998).

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Split-group analyses showed a main effect of both time-phase, $F(1, 1428) = 17.89$, $p < .05$, $MSE = 151.44$, and difficulty level, $F(2, 1428) = 97.04$, $p < .05$, $MSE = 151.44$ on Murph's study trial views (Table 3, Rows 4-6). No interaction between time-phase and level emerged, $F(2, 1428) = 1.26$, $p > .05$, indicating that both time-phase and level separately influenced the number of trials that Murph viewed. The absence of an interaction is due to the similar increase in viewed study trials observed at each difficulty level from Start Phase to End Phase. The increase in study trial views directly invalidates the novelty preference hypothesis as dot distortions can not become more novel over time.

Split-group analyses showed a main effect of both time-phase, $F(1, 786) = 18.97$, $p < .05$, $MSE = 303.40$, and difficulty level, $F(2, 786) = 5.35$, $p < .05$, $MSE = 303.40$ on Gale's study trial views (Table 3, Rows 7-9). An interaction between time-phase and level also emerged, $F(2, 786) = 4.33$, $p < .05$, $MSE = 303.40$. Declines occurred in the low-difficulty condition, $F(1, 262) = 7.68$, $p < .05$, $MSE = 159.14$, and the high-difficulty condition, $F(1, 262) = 12.99$, $p < .05$, $MSE = 538.56$, but not in the intermediate difficulty condition, $F(1, 262) < 1.00$. These declines offer some support for the novelty preference hypothesis. However, the reason why novelty decreases over time in the study conditions that offered low and high novelty, but not for the study condition that offered a moderate level of novelty remains unclear. Han's split-group data is presented for the interested reader (Table 3, Rows 10-12).

General Discussion

The present research updates the literature with a demonstration of metacognitive control by monkeys without relying on natural behavioral patterns and without the risk of the data being conditioned via stimulus-response pairings. Critically, humans and monkeys both increased their study trial views when the study materials were (perceived to be) more difficult (i.e., the category exemplars shared a weaker family resemblance). Alternative hypotheses were explored but found lacking. The test difficulty hypothesis was only supported by the human data from Experiment 1, making it implausible. Murph's split-group data from Experiment 2 directly invalidated the novelty preference hypothesis.

Previous research has demonstrated that monkeys are sensitive to their level of knowing (Smith, et al., 2006; Son & Kornell, 2005). This research demonstrates that monkeys are capable of greater metacognitive sophistication—they can intentionally act upon their environment based on the difficulty of the to-be-learned materials. At this point, it remains unknown whether this behavior demonstrates actual preparation for the impending test. Future research will need to address that issue. However, this research does provide evidence that monkeys can exert metacognitive control – one of the most sophisticated cognitive processes displayed by humans.

Acknowledgements

Many thanks to David Smith, David Washburn, Michael Beran, Gail Mauner, Jim Sawusch, Mark Kristal, Janet Metcalfe, and Keith Thiede for their guidance and support of this research. The research reported here was supported by Grant HD-38051 from NICHD and by Grant BCS-0634662 from NSF. The opinions expressed are those of the author and do not represent the views of either funding body.

References

Call, J. (2004). Inferences About the Location of Food in the Great Apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, and Pongo pygmaeus). *Journal of Comparative Psychology, 118*(2), 232-241.

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition, 3*(4), 207-220.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20*(4), 374-380.

Golub, M. S., & Germann, S. L. (1998). Perinatal bupivacaine and infant behavior in rhesus monkeys. *Neurotoxicology and Teratology, 20*(1), 29-41.

Gunderson, V. M., & Sackett, G. P. (1984). Development of pattern recognition in infant pigtailed macaques (Macaca nemestrina). *Developmental Psychology, 20*(3), 418-426.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

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*(4), 239-246.

Homa, D., & Cultice, J. C. (1984). Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 83-94.

Koriat, A., & Bjork, R. A. (2005). Illusions of Competence in Monitoring One's Knowledge During Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 187-194.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language, 52*(4), 478-492.

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of Metacognitive Skills and Hint Seeking in Monkeys. *Psychological Science, 18*(1), 64-71.

Little, D. M., & Thulborn, K. R. (2006). Prototype-distortion category learning: A two-phase learning process across a distributed network. *Brain and Cognition, 60*(3), 233-243.

Minda, J. P., & Smith, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(2), 275-292.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125-141). New York: Academic Press.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (Macaca mulatta), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology, 103*(1), 32-38.

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*(6), 1258-1266.

Shields, W. E., Smith, J. D., Guttmannova, K., & Washburn, D. A. (2005). Confidence Judgments by Humans and Rhesus Monkeys. *Journal of General Psychology, 132*(2), 165-186.

Smith, J. D., Beran, M. J., Redford, J. S., & Washburn, D. A. (2006). Dissociating Uncertainty Responses and Reinforcement Signals in the Comparative Study of Uncertainty Monitoring. *Journal of Experimental Psychology: General, 135*(2), 282-297.

Smith, J. D., & Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(4), 800-811.

Smith, J. D., Redford, J. S., & Haas, S. M. (2008). Prototype abstraction by monkeys (Macaca mulatta). *Journal of Experimental Psychology: General, 137*(2), 390-401.

Smith, J. D., Redford, J. S., Haas, S. M., Coutinho, M. V. C., & Couchman, J. J. (2008). The comparative psychology of same-different judgments by humans (Homo sapiens) and monkeys (Macaca mulatta). *Journal of Experimental Psychology: Animal Behavior Processes, 34*(3), 361-374.

Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in humans and animals. *Cognition, 62*(1), 75-97.

Son, L. K., & Kornell, N. (2005). Metaconfidence Judgments in Rhesus Macaques: Explicit Versus Implicit Mechanisms. In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness.* (pp. 296-320). New York, NY, US: Oxford University Press.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science, 14*(1), 66-73.

Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review, 6*(4), 662-667.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*(1), 66-73.

Thompson, L. A., Petrill, S. A., DeFries, J. C., Plomin, R., & Fulker, D. W. (1994). Longitudinal predictions of school-age cognitive abilities from infant novelty preference *Nature and nurture during middle childhood.* (pp. 77-85). Malden, MA, US: Blackwell Publishing.

Washburn, D. A., & Rumbaugh, D. M. (1992). Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. *Behavior Research Methods, Instruments & Computers, 24*(2), 157-164.

Zaki, S. R., Nosofsky, R. M., Stanton, R. D., & Cohen, A. L. (2003). Prototype and Exemplar Accounts of Category Learning and Attentional Allocation: A Reassessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1160-1173.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Table 1

Mean Study Trial and Category Data by Experiment

Group	Experiment Condition	Study Trial Mean	Viewed Category Mean (per session)
Humans	1 L3	12.35 (1.41)	5.8 (0.21)
	1 L5	13.53 (1.53)	5.6 (0.20)
	1 L7	16.61 (1.27)	5.7 (0.19)
Murph	2 L3	14.64 (0.44)	10.91 (0.91)
	2 L5	18.24 (0.53)	11.00 (0.93)
	2 L7	25.53 (0.70)	11.05 (0.95)
Gale	2 L3	21.68 (0.78)	8.65 (1.12)
	2 L5	22.60 (0.89)	8.06 (1.11)
	2 L7	26.35 (1.46)	8.47 (1.11)
Han	2 L3	6.26 (0.46)	7.43 (0.95)
	2 L5	5.78 (0.35)	7.50 (0.94)
	2 L7	5.77 (0.38)	7.69 (0.93)

standard errors of the means presented in parentheses

Note. L3 = Level 3, L5 = Level 5, L7 = Level 7

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Table 2

Test Performance by Experiment

Experiment	Group	Study Condition (percent correct)		
		L3	L5	L7
1	Humans	80.7 (0.89)	78.6 (0.88)	73.9 (1.29)
2	Murph	77.2 (0.52)	77.6 (0.54)	78.7 (0.54)
2	Gale	63.9 (0.84)	65.1 (0.81)	65.1 (0.82)
2	Han	57.2 (1.07)	58.8 (0.83)	58.8 (0.84)

standard errors of the means presented in parentheses

Note. L3 = Level 3, L5 = Level 5, L7 = Level 7

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

Table 3

Mean Viewed Study Trials of Start Phase and End Phase by Experiment

Group	Experiment Condition	Start Phase	End Phase
Humans	1 L3	12.56 (1.27)	12.08 (1.28)
	1 L5	13.76 (1.61)	13.09 (1.27)
	1 L7	16.66 (1.39)	16.81 (1.71)
Murph	2 L3	13.49 (0.59)	15.80 (0.66)
	2 L5	16.15 (0.67)	20.33 (0.79)
	2 L7	24.64 (0.92)	26.41 (1.05)
Gale	2 L3	23.80 (1.21)	19.49 (0.98)
	2 L5	23.37 (1.41)	21.80 (1.11)
	2 L7	31.48 (2.58)	21.18 (1.23)
Han	2 L3	4.99 (0.39)	7.50 (0.82)
	2 L5	5.06 (0.34)	6.48 (0.61)
	2 L7	5.07 (0.39)	6.39 (0.66)

standard errors of the means presented in parentheses

Note. L3 = Level 3, L5 = Level 5, L7 = Level 7

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at Journal of Experimental Psychology: Learning, Memory and Cognition, published by American Psychological Association. Copyright restrictions may apply. doi: 10.1037/a0017809

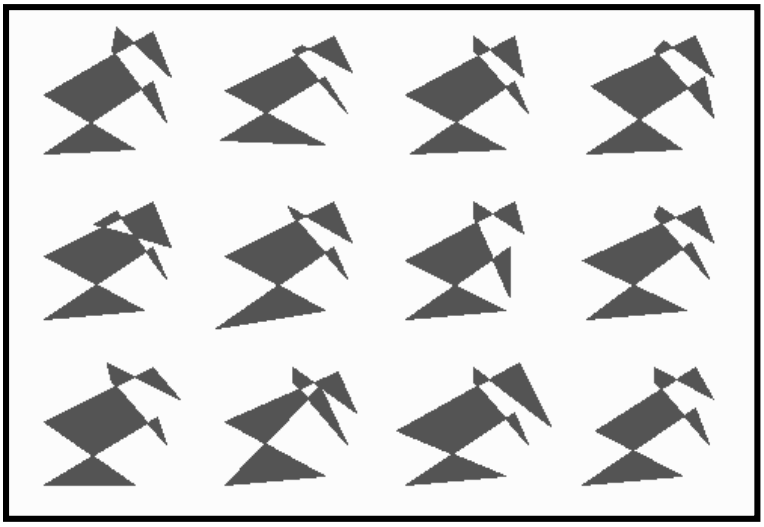
Figure Captions

Figure 1. A. A collection of exemplars that share a strong family resemblance. *B.* A collection of exemplars that share an intermediate family resemblance. *C.* A collection of exemplars that share a weak family resemblance.

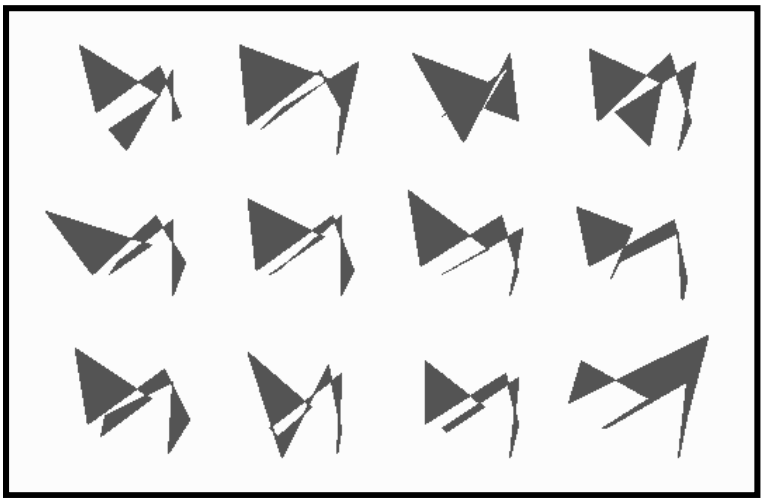
Figure 2. A. A screenshot taken from the study phase of Experiment 1 with a red dot distortion at the top middle of the screen, the cursor in the center of the screen, and a red-filled circle at the bottom middle of the screen. *B.* A screenshot taken from the test phase of Experiment 1 with a to-be-categorized dot distortion at the top middle of the screen, the cursor in the center of the screen, and a large "N" at the bottom middle of the screen.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at
Journal of Experimental Psychology: Learning, Memory and Cogni
Copyright restrictions may apply. (

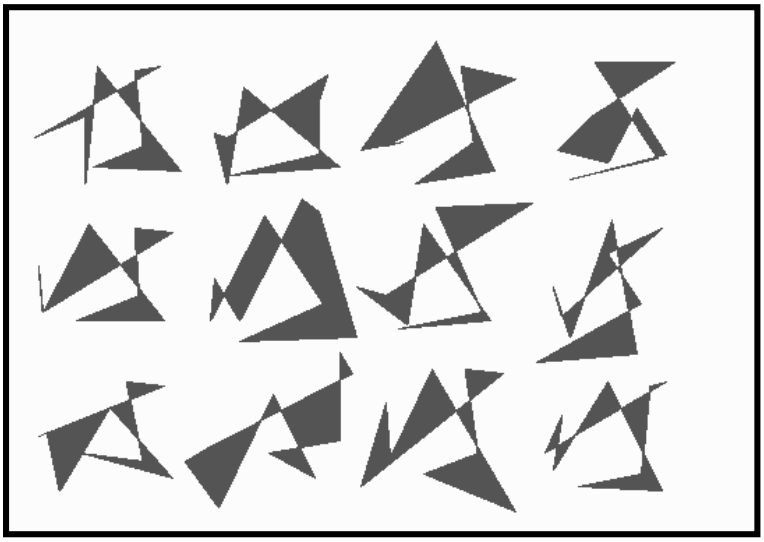
A.



B.

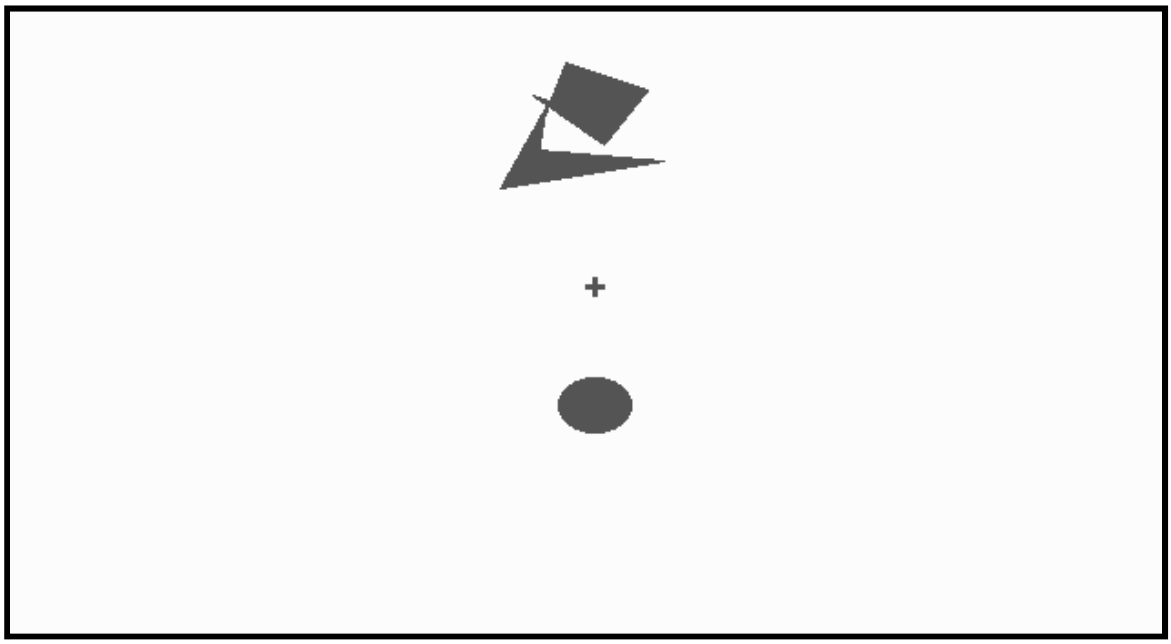
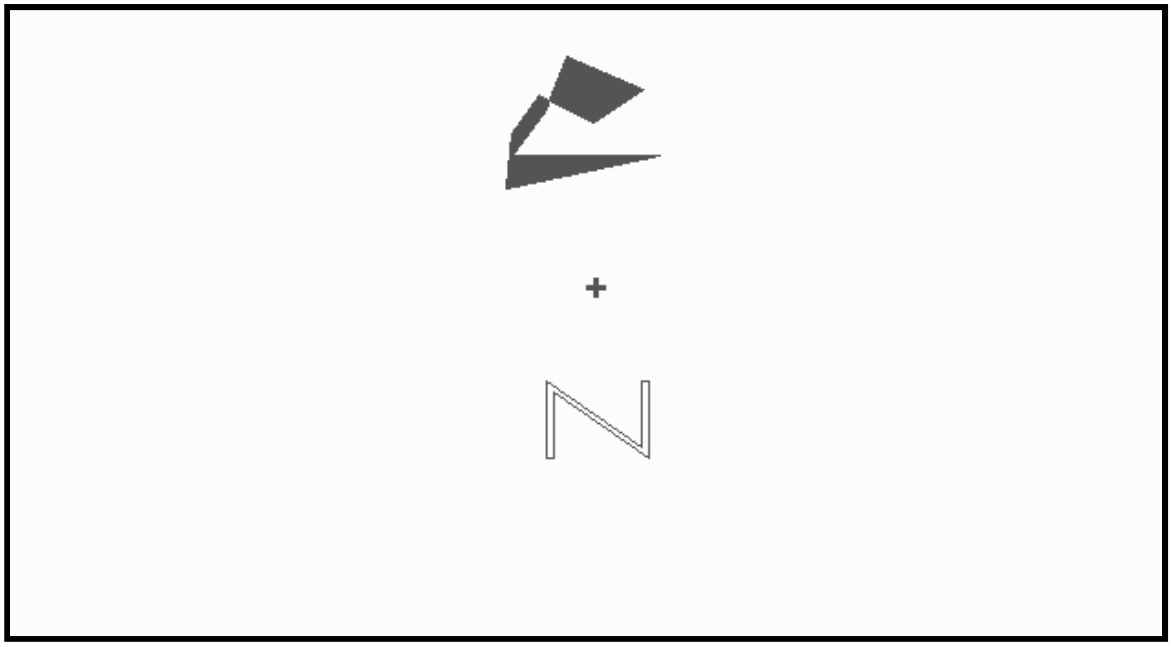


C.



This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at
Journal of Experimental Psychology: Learning, Memory and Cog
Copyright restrictions may apply

A.



B.