Research Article

Transfer of Metacognitive Skills and Hint Seeking in Monkeys

Nate Kornell,[1] **Lisa K. Son,**[2] **and Herbert S. Terrace**[3]

[1]*University of California, Los Angeles;* [2]*Barnard College; and* [3]*Columbia University and New York State Psychiatric Institute*

ABSTRACT—*Metacognition is knowledge that can be expressed as confidence judgments about what one knows (monitoring) and by strategies for learning what one does not know (control). Although there is a substantial literature on cognitive processes in animals, little is known about their metacognitive abilities. Here we show that rhesus macaques, trained previously to make retrospective confidence judgments about their performance on perceptual tasks, transferred that ability immediately to a new perceptual task and to a working memory task. We also show that monkeys can learn to request "hints" when they are given problems that they would otherwise have to solve by trial and error. This study demonstrates, for the first time, that nonhuman primates share with humans the ability to monitor and transfer their metacognitive ability both within and between different cognitive tasks, and to seek new knowledge on a need-to-know basis.*

Metacognition, the ability to monitor and control one's knowledge, is a basic feature of human thought (Flavell & Wellman, 1977; Koriat, 1997; Nelson & Narens, 1990). Metacognitive *monitoring* has been studied in experiments in which subjects are asked to state some fact and are then asked to evaluate the accuracy of their response. For example, after responding to a question such as "What is the capital of Australia?" a subject might be asked, "On a scale from 0 to 100, rate how confident you are that your answer is correct." In experiments on metacognitive *control*, subjects are typically allowed to determine how much time they spend studying each item during a memory task (Dunlosky & Hertzog, 1998; Metcalfe & Kornell, 2005; Son, 2004). The usual finding, an inverse relationship between

Address correspondence to Herbert S. Terrace, Professor of Psychology and Psychiatry, Columbia University, 406 Schermerhorn Hall, 1190 Amsterdam Ave., New York, NY 10027, e-mail: terrace@columbia.edu.

accuracy and study time, shows that subjects use their knowledge of what they do not know to determine which items to study.

Although there is a substantial body of research on animal cognition (Shettleworth, 1998), this literature provides very little evidence of metacognitive monitoring and virtually none of metacognitive control processes. Here we present evidence that monkeys trained to make metacognitive judgments about perceptual problems learned to do so in a way that was not task-specific. They were able to transfer their metacognitive skills immediately to a qualitatively different working memory task. We also show that monkeys can use metacognitive control to improve their performance on a serial memory task by seeking the information they need while learning that task.

A major difficulty in studying animal metacognition is the paucity of paradigms that can rule out the influence of exteroceptive stimuli in purported instances of metacognition. Some recent psychophysical experiments on monkeys (Shields, Smith, & Washburn, 1997) and dolphins (Smith et al., 1995) illustrate this problem. Subjects were required to make one response ("high") on trials on which one stimulus (S1) was presented and another response ("low") when a second stimulus (S2) was presented. Subjects were also given the option of ending each trial when they were "unsure" of the correct response. Because subjects opted out mainly on trials on which the difference between S1 and S2 was small, opting out was interpreted as evidence of the subjects' uncertainty. However, just as S1 and S2 exerted stimulus control over "high" and "low" responses, they could also exert an influence over the opting-out response. As noted by Metcalfe (2003) and Shettleworth and Sutton (2003), when a behavior occurs in the presence of the discriminative stimuli to which the subject has been trained, there is no reason to interpret that behavior as metacognitive unless it can be shown to transfer to a new task.

The problem of control by exteroceptive stimuli can be obviated by requiring subjects to make retrospective judgments of their memories. In one experiment, rhesus macaques were given the opportunity to choose between two types of trials on a

delayed-matching-to-sample task after the sample was presented but before the actual test could be administered (Hampton, 2001). One option was to wait for the test, at which time the subject could earn a desirable reward if it responded correctly, or earn nothing if it made an error. The second option was to end the trial and settle for a less desirable, but certain, reward. Subjects were also given forced-choice trials, on which they could not avoid the test phase. Accuracy was significantly higher for trials on which subjects opted to take the test than for forced-choice trials, a result suggesting that subjects declined to take the test when their memory of the sample was relatively weak.

In a different paradigm, which is also the one we used in the current experiment, monkeys were trained to make retrospective judgments of their accuracy on perceptual tasks (Son & Kornell, 2005). They were differentially rewarded for selecting one of two confidence icons (high or low) after each response on a given perceptual task. Reward was maximized when subjects chose the high-confidence icon following correct responses and the low-confidence icon following errors. Subjects trained on this paradigm learned to respond to the confidence icons appropriately.[1]

Task-specificity is another issue that needs to be addressed in experiments on metacognition in animals. Although judgments of confidence and uncertainty by human subjects have been observed in many contexts (e.g., Schwartz, 2002), it has yet to be shown that purported metacognitive behavior in animals transfers to novel types of problems (Shettleworth & Sutton, 2003). Experiment 1 shows, for the first time, that a monkey can transfer its ability to judge the accuracy of its performance from a perceptual task to a working memory task.

EXPERIMENT 1

Method

Subjects

The subjects in Experiment 1 were 2 male rhesus macaques (*Macaca mulatta*), Lashley and Ebbinghaus. Both were approximately 6 years of age at the start of the experiment and had been trained previously in experiments on simultaneous chaining (Terrace, 2005b) and numerical discrimination (Brannon & Terrace, 1998) and in pilot studies on metacognition (Son & Kornell, 2005). The subjects were housed individually in adjoining cages in a colony of 20 rhesus macaques at the New York State Psychiatric Institute (NYSPI). The colony was maintained in accordance with guidelines issued by the National Institutes of Health and by the Institutional Animal Care and Use Committees at NYSPI and Columbia University.

---

[1]Shields, Smith, Guttmannova, and Washburn (2005) subsequently performed a similar experiment with monkeys.

Apparatus

The apparatus was identical to that used in earlier experiments on monkey cognition (e.g., Subiaul, Cantlon, Holloway, & Terrace, 2004). All training and testing sessions took place in chambers that were housed in sound-attenuated booths. Each chamber was equipped with a touch-sensitive video monitor that was used to present stimuli and to detect subjects' responses.

Pilot Experiment

Experiment 1 is an extension of a pilot experiment in which we determined optimal values of the parameters needed to train reliable metacognitive performance (Son & Kornell, 2005). During the pilot experiment, subjects were trained on two perceptual tasks. The first was a line-discrimination task, in which they were required to select the longest of nine lines; the second was a numerical discrimination task, in which one subject (Lashley) was required to select the stimulus containing the largest number of geometric objects, and the other (Ebbinghaus) was required to select the stimulus containing the fewest objects. Training on the metacognitive paradigm (described in the next section) was added after subjects learned the line-discrimination task. Both subjects learned to respond metacognitively. They were subsequently trained on the numerical discrimination task. After that task was mastered, the metacognitive paradigm was added to this second task without further metacognitive training. Both subjects responded metacognitively immediately after the metacognitive paradigm was added to the numerical discrimination task.

Metacognitive Paradigm

The metacognitive paradigm used in this experiment is similar to the one used in the pilot experiment. As shown in Figure 1, two confidence icons were presented at the end of each trial, immediately after the subject responded on a perceptual task. One icon signified high confidence; the other, low confidence. Reward was contingent on the relation between the accuracy of the subject's response on the perceptual task and the subject's choice of confidence icon. Choosing the high-confidence icon was a "risky" bet: The subject won three tokens if its response on the perceptual task was correct, but lost three tokens if its response was incorrect. Responses to the "safe," low-confidence icon were always rewarded with one token.

Tokens were deposited in a "bank" located in the lower right-hand corner of the video monitor. The bank contained 9 tokens at the start of each session. Tokens were added to and subtracted from the bank following each metacognitive response. Food reward (a 190-mg banana pellet) was provided when 12 or more tokens had accumulated in the bank. The number of tokens was then reset to 9. If the balance dropped to zero, no further tokens could be lost. We note that this appears to be the first experiment with animals in which tokens were subtracted following incorrect responses. In earlier experiments with token economies,
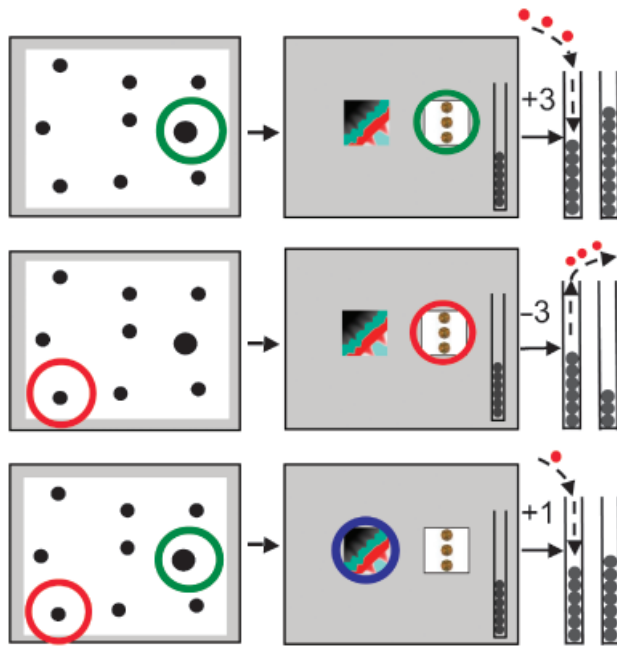
Fig. 1. Example of the task sequence (perceptual task, followed by metacognitive task) in Experiment 1. For the reader's benefit, correct responses are circled in green; incorrect responses, in red. In the trial illustrated here, a response to the largest stimulus was correct on the perceptual task. The high- and low-confidence icons appeared immediately after the subject made its selection on the perceptual task. A response to the high-confidence icon (a column of dots) resulted in a gain of three tokens if it followed a correct response (top), but a loss of three tokens if it followed an error (middle). A response to the low-confidence icon (a multicolored geometric pattern) always resulted in a gain of one token, regardless of whether the response on the perceptual task was correct or incorrect (bottom). The addition or subtraction of tokens that followed each metacognitive response is shown on the right.

incorrect responses did not incur any penalties (Jackson & Hackenberg, 1996; Wolfe, 1936).

Transfer to a New Perceptual Task
In experiments on human metacognition, subjects do not require any pretraining when asked to make metacognitive judgments on a new task. Having established the efficacy of the metacognitive paradigm in our pilot study, we sought to determine whether monkeys could transfer their metacognitive skills to a new perceptual problem and to a qualitatively different memory task. Our measure of transfer was based on subjects' immediate performance during the first four sessions in which the metacognitive paradigm was added to each new task.

The subjects were trained on four tasks, in the following order: line length and then numerical estimation (during the pilot experiment), followed by circle size and then serial working memory (in Experiment 1). In all four cases, the metacognitive task was added after training on the primary task. Experiment 1 began after subjects completed the pilot study. They were first trained on a perceptual task based on circle size; the level of difficulty on this task varied randomly across trials. One subject
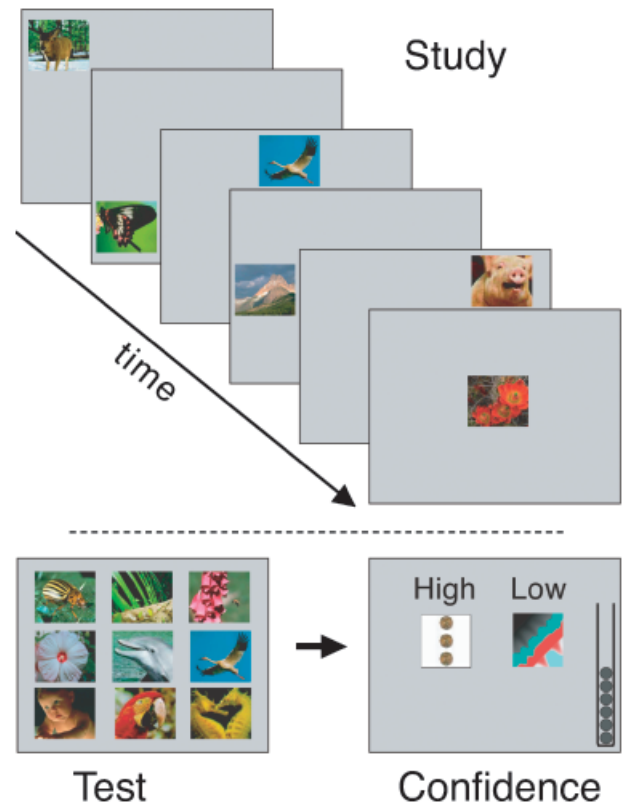

Fig. 2. Trial structure of the serial working memory task in Experiment 1. Six sample photographs were displayed successively, followed by a test during which one of the six samples was presented with eight distractors. The subject's task was to select the sample and then select the appropriate confidence icon.

(Ebbinghaus) was trained to select the circle that had the largest area; the other (Lashley), to select the circle that had the smallest area. Subjects initiated each trial by touching a "start" stimulus on the monitor. Correct responses on the size-discrimination task were followed by the simultaneous delivery of primary reinforcement (two 190-mg banana-flavored Bio-Serv© food pellets) and visual and auditory feedback. Subjects were trained daily in sessions that were 20 min long. The metacognitive paradigm was added after a subject completed at least 65% of the trials in one session correctly.

Transfer to a Serial Working Memory Task
Following metacognitive testing on the size-discrimination problem, subjects were trained on a working memory task that was similar to the tasks used in experiments on serial working memory (SWM) with human subjects (Sternberg, 1969) and monkeys (Wright, Santiago, Sands, Kendrick, & Cook, 1985). The trial structure of the SWM task is shown in Figure 2.[2] A sequence of six trial-unique photographs was presented during

[2]Video clips of Ebbinghaus performing the SWM task can be found on the Web at http://www.columbia.edu/cu/psychology/primatecognitionlab/videos/MetaSPR.mov.

the study portion of each trial. Attention to the samples was maintained by requiring subjects to touch each photograph before the next one was presented. The response to each of the first five samples was followed by a 2-s interstimulus interval. The test portion of each trial began 0.5 s after the subject responded to the sixth sample. Nine photographs, one of which had appeared previously as a sample, were displayed simultaneously during test. The remaining eight photographs were distractors. The subject's task was to select the sample. Correct responses were rewarded with two banana pellets. A response to any of the distractors ended the trial without reward. The samples and distractors were selected at random from a library of more than 2,500 photographs. A given photograph was never repeated during the same session. When subjects satisfied the accuracy criterion of 65% correctly completed trials in one session, the two metacognitive icons were introduced at the end of each trial. The contingencies for responding to the high- and the low-confidence icons were the same as those used in the pilot experiment during the size task: a gain or loss of three tokens on high-risk trials (depending on test accuracy), or a gain of one token on low-risk trials (regardless of test accuracy).

Results

Evidence of metacognition in a nonverbal animal requires a correlation between task accuracy and a metacognitive judgment. The measure we used was the phi correlation coefficient (ϕ) between accuracy and confidence. The value of ϕ is the same as the value of Pearson's r when r is calculated from data arranged in a 2×2 table, in this instance, a table composed of two levels of correctness (correct vs. incorrect) with two levels of risk (high vs. low). The value of ϕ would approach 1.0 to the extent that subjects chose the high-confidence icon after responding correctly and the low-confidence icon after making an error.

Perceptual Task

Both subjects responded metacognitively as soon as the risk icons were introduced to the perceptual task. This means that they chose the high-confidence icon more frequently after correct responses than after errors, and that they chose the low-confidence icon more frequently after errors than after correct responses. For each subject, Figure 3 shows the value of ϕ for the first 4 sessions of responding on the new perceptual task, along with the 2×2 table from which the value of ϕ was derived. In each instance, the value of ϕ was significant beyond the .01 level (as computed by Fisher's exact test).[3] Given that no

---

[3]The values of ϕ we obtained compare favorably with those reported in experiments in which adult human subjects were asked to make metacognitive judgments immediately after learning a set of paired associates (Nelson & Dunlosky, 1991). In those studies, in which metacognition was evaluated by the magnitude of gamma correlations, the values of gamma were typically low, with a mean of .38 (Dunlosky & Nelson, 1992) and range of .09 through .48 (Vesonder & Voss, 1985). The ϕ statistic is a more conservative estimate of metacognitive performance than gamma because gamma can return values that are higher, but never lower, than the value of ϕ (Nelson, 1984).
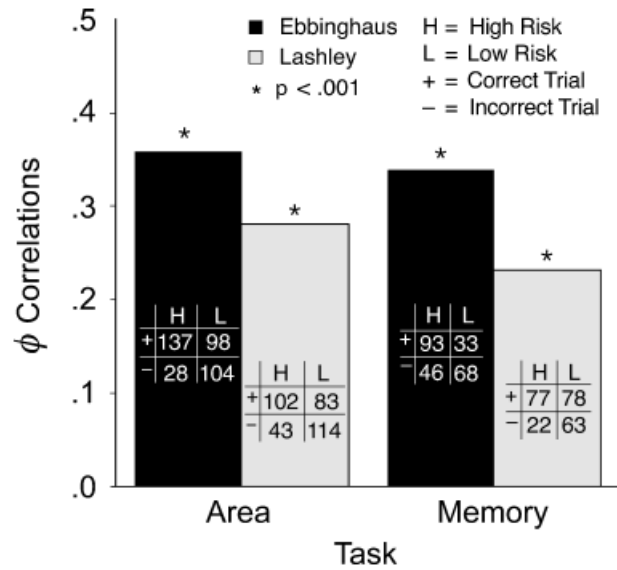

Fig. 3. Results from Experiment 1: ϕ correlations for each subject during the first four sessions of training on the perceptual (circle-size) and serial working memory tasks. The absolute frequencies used for computing ϕ are shown in the 2 (accuracy) × 2 (confidence) table in each bar. Significant values of ϕ ($p < .001$) are marked with an asterisk.

additional metacognitive training was provided during the area task, these data provide clear evidence that an animal can transfer its metacognitive ability from one perceptual task to another.

Memory Task

Metacognitive responding also transferred immediately when the confidence icons were added to the SWM task. As shown in Figure 3, there was no decrement in the value of ϕ, as compared with its value on the area-discrimination task, the most recent perceptual task. For both subjects, the value of ϕ was significant beyond the .001 level. This result is especially striking because of the long delays that occurred between the subjects' responses to sample items and the presentation of the confidence icons. In the SWM task, the median duration of the interval between Ebbinghaus's and Lashley's responses to samples and the appearance of the confidence icons was 22 and 23 s, respectively, for Sample 1 and 4 and 5 s, respectively, for Sample 6.

Serial Position of Probe

On most serial tasks, the difficulty of a list item is correlated with its serial position (primacy and the recency effects). Only a recency effect was obtained in this experiment. Accuracy increased significantly with serial position for both Ebbinghaus ($r = .14, p < .001$) and Lashley ($r = .12, p < .01$). By contrast, the position of the sample had no effect on the value of ϕ. Indeed, the value of ϕ differed significantly from zero at each serial position ($p < .05$), with the sole exception of the value for Ebbinghaus at Serial Position 6. These data show that difficulty,

as defined by an item's serial position, cannot account for the subjects' confidence in the accuracy of their responses.

Influence of Reaction Time

It might be argued that subjects' responses to the confidence icons were based on the difference between their reaction times (RTs) for correct and incorrect responses. In experiments with human subjects (e.g., Benjamin, Bjork, & Schwartz, 1998), RTs are typically shorter for correct responses than for errors. The same was true on the SWM task in Experiment 1: There was a negative correlation between RTs during test and accuracy during the first 10 sessions of metacognitive training for both Ebbinghaus, $r = -.11$, $t(9) = -2.05$, $p = .07$, and Lashley, $r = -.40$, $t(9) = -9.77$, $p < .0001$; these results confirm that RT decreased as accuracy increased.

There is no simple method for controlling for possible influences of RT duration on confidence judgments. It is, however, possible to partial out the contribution of RT to the correlation between accuracy and confidence choice. The resulting values during the first 10 sessions of metacognitive training on the SWM task were significantly greater than zero for both Ebbinghaus, $r = .28$, $t(9) = 5.74$, $p < .001$, and Lashley, $r = .24$, $t(9) = 6.62$, $p < .0001$. It follows that the duration of RTs on correct and incorrect trials was not the sole determinant of the subjects' metacognitive performance.

EXPERIMENT 2

Experiment 1 showed that a monkey can monitor its accuracy on perceptual tasks and transfer that ability to monitoring its memory. In Experiment 2, we asked whether a monkey can acquire a second component of human metacognition, control of one's knowledge. Previous studies (Call & Carpenter, 2001; Hampton, Zivin, & Murray, 2004) showed that apes and monkeys sought information by looking for food when they were uncertain about its location, but that they reached for the food directly if they saw where it was hidden. However, similar behavior occurs naturally during foraging. In certain contexts, animals search for food by default unless they can acquire it directly. It is therefore possible that the animals in these studies simply used a familiar food-seeking strategy in a new context. They did not have to know what they knew to search for food. Instead they simply needed to know where the food was, and if they did not, they searched for it.

The task used in Experiment 2 required subjects to learn novel four-item sequences composed of arbitrary photographs. Subjects were given the opportunity to ask for "hints" as to the next response in a sequence. Without hints, they would have to learn each sequence by trial and error, a skill they had acquired in a previous experiment (Terrace, Son, & Brannon, 2003). Thus, Experiment 2 addressed two questions: Would subjects learn to request hints when starting to learn a new list, and would hint
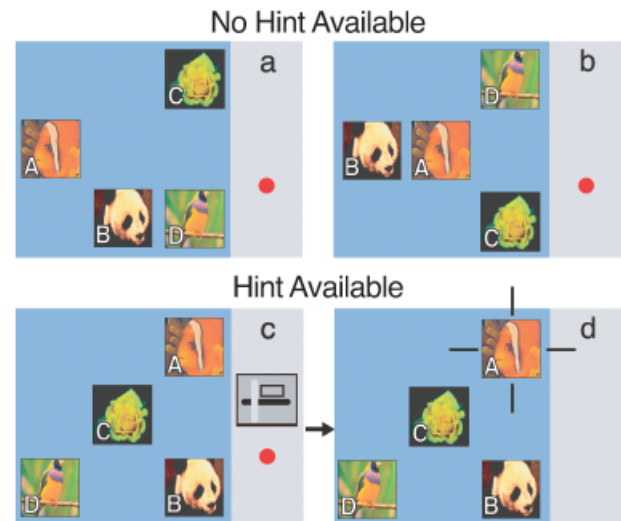

Fig. 4. Examples of displays used in Experiment 2. Subjects were rewarded for responding to all items in a particular order ($A \rightarrow B \rightarrow C \rightarrow D$). As illustrated in (a) and (b), the position of the items varied randomly from trial to trial. On 50% of trials, no hint was available (a and b). On the other 50% of trials, a hint was available, as indicated by an icon that appeared on the right side of the monitor, above a red circle (c and d). Touching the hint icon resulted in the appearance of four blinking lines around the correct item for that position in the sequence. The red circle signaled that the subject could earn an M&M, instead of a less desirable banana pellet, by completing the trial correctly. The first request for a hint eliminated the red circle (d).

seeking decrease as they became more proficient at executing a particular list?

Method

Subjects

The subjects were two 10-year-old male rhesus macaques (Macduff and Oberon). Both monkeys had extensive experience learning three-, four-, and seven-item lists as simultaneous chains (Terrace et al., 2003).

Apparatus

The apparatus was identical to that used in Experiment 1.

Simultaneous-Chaining Paradigm

In the simultaneous-chaining paradigm, all list items are displayed simultaneously throughout each trial, typically on a touch-sensitive video monitor (Figs. 4a and 4b). The subject's task is to respond to the items in a particular order, regardless of their spatial position on the monitor. The physical location of list items is randomized on each trial to prevent subjects from performing the required sequence as a fixed motor pattern. All of the lists used in Experiment 2 were composed of photographs (1.5 in. × 2 in.) of natural or man-made objects (e.g., animals,

people, scenery, flowers, cars, bridges). All the stimuli were novel, and none were repeated on a second list.

Simultaneous chains are typically learned by trial and error from feedback that follows each response (Terrace, 2005a). In Experiment 2, correct responses were followed by brief (0.5-s) visual and auditory feedback; errors were followed by a 4-s time-out during which the screen was darkened. A trial ended either when the subject responded to an item incorrectly or when the subject responded to all of the items in the correct order. Food reward (a 190-mg banana pellet) was provided only after a correct response to the last list item.

Training Simultaneous Chains With a Hint Option

As shown in Figures 4c and 4d, the subjects in Experiment 2 could obtain a hint by responding to an icon to the right of the list items. The hint consisted of four blinking lines that appeared around the item to which the subject should respond next. If the subject completed a trial correctly without requesting any hints, it was rewarded with a highly desirable M&M candy. If the subject requested one or more hints during the course of completing a trial correctly, it received the usual banana pellet as a reward. Subjects could obtain as many as four hints on each trial, one for each list item.

Availability of Hints

For 6 days, subjects were pretrained on a familiar list, with the hint icon present on 50% of the trials.[4] Each subject was then trained on 10 new lists. Training for each list took place during four 60-trial sessions, with the familiar list presented on half the trials and the new list on the other half. During each session, the subject could request one or more hints on half of the trials, at any point in the sequence. Thus, out of 60 trials, hints were available on 15 trials on which a new list was presented and on 15 trials on which the familiar list was presented. No hint was available on the remaining 30 trials: 15 with the new list and 15 with the familiar list. Trial types alternated randomly.

Results

Hint seeking was inversely related to accuracy on each of the 10 novel lists, for both Macduff ($r = -.74$, $p < .0001$) and Oberon ($r = -.87$, $p < .0001$). As Figure 5a shows, requests for hints, when a hint was available, decreased across the 4 sessions of training on each list, from .80 of trials to .53 of trials. At the same time, the proportion of correct responses increased from .11 to .38 when a hint was not available.

The inverse relation between hint seeking and accuracy was also evident in performance on new and familiar lists, for both Macduff ($r = -.74$, $p < .0001$) and Oberon ($r = -.94$, $p < .0001$). As Figure 5b shows, requests for hints when a hint was

---

[4]The percentage of trials completed correctly when the hint was not available during the last session of training on the familiar list was 70% for Macduff and 87% for Oberon.
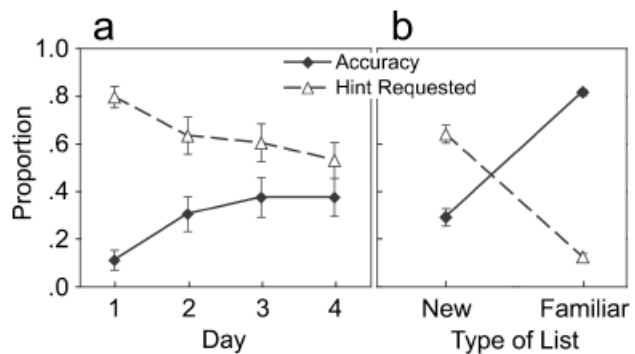

Fig. 5. Proportion of trials completed correctly when no hint was available and proportion of trials on which a hint was requested when the hint was available: (a) across the four sessions of training on new lists and (b) on new lists and on the familiar list.

available were less frequent for the familiar list (.13) than for new lists (.65). In contrast, proportion correct was higher on the familiar list (.83) than on the new lists (.30).

GENERAL DISCUSSION

Because experiments on metacognition rely on verbal paradigms, and because metacognitive abilities have not been observed in children until the age of 4, some researchers have argued that metacognition requires verbal ability (Metcalfe & Shimamura, 1994; Tulving, 1994). Using behavioral measures, Experiment 1 showed that monkeys can learn to make metacognitive judgments that are functionally similar to those obtained from human subjects. After learning to monitor their performance on a series of perceptual tasks, monkeys were able to transfer that ability to a qualitatively different task based on working memory. This shows that a monkey's metacognitive ability is not task-specific. Experiment 1 also provides the first demonstration that an animal can learn to respond metacognitively about the match (or lack thereof) between its memory of a sample and the item it selected during test. Because the monkey's response was based on its memory of stimuli that were no longer present, its choice of confidence icon cannot be attributed to the physical characteristics of a co-occurring exteroceptive stimulus. Experiment 2 provided the first evidence to date that a monkey can apply appropriate control strategies to correct deficiencies in its knowledge. When subjects lacked the information they needed to execute new lists, they acquired that information by asking for hints. Taken together, the results of Experiments 1 and 2 show that a monkey's metacognitive behavior provides analogues of the two basic features of human metacognition: monitoring and controlling knowledge.

Given that monkeys do not have verbal ability, there is no reason to assume that they are conscious of their metacognitive judgments or, for that matter, of the metacognitive states of other individuals (Povinelli & Vonk, 2003). Although humans are conscious of many of their metacognitive processes, they often

make metacognitive judgments before those judgments can be processed consciously (Reder & Schunn, 1996; Son & Metcalfe, 2005). Thus, rhesus macaques appear to share with humans the ability to make unconscious metacognitive judgments and the ability to differentiate between the cognitive states of knowing and not knowing across different tasks. This ability adds to the growing list of cognitive processes that evolved without language, including concept formation (Cook & Wasserman, 2006), discrimination of numerically defined stimuli (Brannon & Terrace, 1998), timing (Church, 2006), and serial learning (Terrace, 2005b). In each of these experiments, it was shown that non-verbal animals can solve problems that require the memory of a previously experienced stimulus that was not present during test (i.e., a representation; Kinsbourne, 2005; Terrace, 2002, 2005a). Thus, contra Descartes, animals are clearly able to think without language. How they do so is currently a mystery whose solution is key to understanding the evolution of cognition.

REFERENCES

Benjamin, A.S., Bjork, R.A., & Schwartz, B.L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55–68.

Brannon, E.M., & Terrace, H.S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*, 746–749.

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

Church, R.M. (2006). Behavioristic, cognitive, biological and quantitative explanations of timing. In E.A. Wasserman & T.R. Zentall (Eds.), *Comparative cognition* (pp. 249–269). New York: Oxford University Press.

Cook, R.G., & Wasserman, E.A. (2006). Relational discrimination learning in pigeons. In E.A. Wasserman & T.R. Zentall (Eds.), *Comparative cognition* (pp. 307–324). New York: Oxford University Press.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D.J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah, NJ: Erlbaum.

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*, 373–380.

Flavell, J., & Wellman, H. (1977). Metamemory. In R. Kail & J. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3–33). Hillsdale, NJ: Erlbaum.

Hampton, R.R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences, USA*, *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*, 239–254.

Jackson, K., & Hackenberg, T.D. (1996). Token reinforcement, choice, and self-control in pigeons. *Journal of the Experimental Analysis of Behavior*, *66*, 29–49.

Kinsbourne, M. (2005). A continuum of self-consciousness that emerges in phylogeny and ontogeny. In H.S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 142–156). New York: Oxford University Press.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349–370.

Metcalfe, J. (2003). Drawing the line on metacognition. *Behavioral and Brain Sciences*, *26*, 350–351.

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*, 463–477.

Metcalfe, J., & Shimamura, A.P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133.

Nelson, T.O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science*, *2*, 267–270.

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.

Povinelli, D.J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, *7*, 157–160.

Reder, L.M., & Schunn, C. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L.M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45–77). Mahwah, NJ: Erlbaum.

Schwartz, B.L. (2002). *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval*. Mahwah, NJ: Erlbaum.

Shettleworth, S.J. (1998). *Cognition, evolution, and behavior*. New York: Oxford University Press.

Shettleworth, S.J., & Sutton, J.E. (2003). Metacognition in animals: It's all in the methods. *Behavioral and Brain Sciences*, *23*, 353–354.

Shields, W.E., Smith, J.D., Guttmannova, K., & Washburn, D.A. (2005). Confidence judgments by humans and rhesus monkeys. *Journal of General Psychology*, *132*, 165–186.

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.

Smith, J.D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncates*). *Journal of Experimental Psychology: General*, *124*, 391–408.

Son, L.K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 601–604.

Son, L.K., & Kornell, N. (2005). Meta-confidence 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: Oxford University Press.

Son, L.K., & Metcalfe, J. (2005). Judgments of learning: Evidence for a two-stage model. *Memory & Cognition, 33,* 1116–1129.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist, 57,* 421–457.

Subiaul, F.X., Cantlon, J.F., Holloway, R., & Terrace, H.S. (2004). Cognitive imitation in rhesus macaques. *Science, 305,* 407–410.

Terrace, H.S. (2002). Serial expertise and the evolution of language. In J.H.A. Wray & F.J. Newmeyer (Eds.), *The transition to language* (pp. 64–90). New York: Oxford University Press.

Terrace, H.S. (2005a). Metacognition and the evolution of language. In H.S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 84–115). New York: Oxford University Press.

Terrace, H.S. (2005b). The simultaneous chain: A new approach to serial learning. *Trends in Cognitive Sciences, 9,* 202–210.

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

Tulving, E. (1994). Preface. In J. Metcalfe & A.P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. vi–viii). Cambridge, MA: MIT Press.

Vesonder, G., & Voss, J. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language, 24,* 363–376.

Wolfe, J.B. (1936). Effectiveness of token-rewards for chimpanzees. *Comparative Psychology Monographs, 12,* 1–72.

Wright, A.A., Santiago, H.C., Sands, S.F., Kendrick, D.F., & Cook, R.G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science, 229,* 287–289.