Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Victoria Lynn Templer Date

Memory Monitoring in Rhesus Monkeys
(*Macaca mulatta*)

By

Victoria L. Templer
B.A. Franklin & Marshall College, 2007

Psychology

_____
Robert Hampton, Ph.D.
Advisor

_____
Jocelyn Bachevalier, Ph.D.
Committee Member

_____
Stella Lourenco, Ph.D.
Committee Member

_____
Joseph Manns, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Memory Monitoring in Rhesus Monkeys
(*Macaca mulatta*)

By

Victoria L. Templer
B.A. Franklin & Marshall College, 2007

Advisor: Robert Hampton, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2009

Abstract

Memory Monitoring in Rhesus Monkeys *(Macaca mulatta)*
By Victoria L. Templer

Taxonomies of human memory distinguish between memories that are accessible to monitoring (explicit) and memories that are inaccessible (implicit). Because nonhumans cannot provide the verbal reports that typify accessible memories in humans, it is difficult to directly assess the accessibility of memory nonhumans. However, recently developed metamemory paradigms provide objective behavioral measures of the accessibility of memory in nonhumans. Animals that can monitor their memory should perform better on memory tests if given the choice to decline trials when memory is weak, compared to when there is no option to decline tests. There is growing evidence that rhesus monkeys selectively decline tests when memory is weak (Hampton, 2001; Smith et al., 2003), but such evidence is limited. The goal of the present study was to test the robustness of monkeys' ability to monitor their own memory using a new foraging-like spatial test. Six rhesus monkeys performed a four choice delayed matching to location task. On each trial monkeys had to remember the location in which a preferred food was hidden in order to collect the reward after a delay. A decline-test response was concurrently available on some trials. Selection of the decline-test response allowed monkeys to avoid the memory test and resulted in a guaranteed but less preferred reward. Monkeys performed significantly better on trials with the decline option available than on trials without it, indicating that when possible, they appropriately declined tests when memory was weak. Monkeys transferred appropriate use of the decline-test response under three conditions that assessed generalization: two tests weakened memory and one test enhanced memory. These results provide converging evidence that rhesus monkeys are able to monitor their memory. This objective behavioral demonstration of memory monitoring in monkeys validates this animal model for the study of the neurobiology of accessible memory in nonhumans.

Memory Monitoring in Rhesus Monkeys
(*Macaca mulatta*)

By

Victoria L. Templer
B.A. Franklin & Marshall College, 2007

Advisor: Robert Hampton, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2009

Table of Contents

LIST OF TABLES AND FIGURES

Memory Monitoring in Rhesus Monkeys *(Macaca mulatta)*

A variety of evidence supports neuroanatomical and functional distinctions between types of memory. Perhaps the most striking evidence comes from dissociations resulting from brain damage, by which it has been repeatedly shown that damage to one neural structure impairs performance on particular memory tests while leaving other memory functions intact (Baddeley 2003; Jonides 2008). Ecological and evolutionary reasoning corroborates the presence of memory systems. Processing, storing, and retrieving one type of information is often functionally incompatible with the processing of other types of information. When functional incompatibilities affect fitness, different selection pressures may lead to divergent memory systems as the properties of memory evolve to match properties of the environment (Shettleworth 1998). Diverse cognitive demands therefore select for distinct memory mechanisms or systems (Sherry & Schacter 1987; Sherry 2006). If we want to understand the evolution of human memory, it is critical to do comparative research, particularly with primates. However, the extent to which major taxonomies of human memory map onto nonhuman animal memory is unknown. Most modern taxonomies of human memory make a fundamental distinction between memories about which we are aware (declarative or explicit memories) and memories that are unconscious (non-declarative or implicit; Clark, Manns, & Squire 2002; Squire & Zola-Morgan, 1991, Figure 1). This distinction has been difficult to

operationalize in studies of nonhuman animals

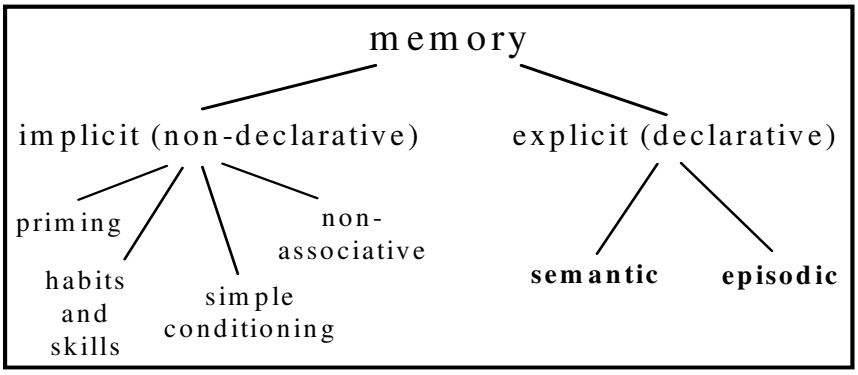

Figure 1. Human memory taxonomies distinguish between explicit (declarative) memories and implicit (non-declarative) memories. (Squire & Zola-Morgan, 1991).

It has been argued that a large "cognitive gap" exists between human and nonhuman memory (e.g. Suddendorf & Corballis 2007) at least in part due to the uniquely human ability to consciously access and monitor the contents of memory. Because much of the evidence for conscious access to memory in humans comes from verbal reports not available from nonhumans, it is difficult to determine the size and nature of this gap, if it exists at all. While it is not feasible to measure the subjective conscious properties of animal memories, memory monitoring can be objectively operationalized in nonhumans and used to make the distinction between accessible and inaccessible memory. In memory monitoring paradigms animals are given a choice between taking memory tests and declining the tests. Accessible memory is inferred when monkeys selectively decline trials on which memory is weak. Though researchers often discriminate between memory systems in humans on the basis of consciousness, (Reder Park, & Kieffaber 2009) this term is unnecessarily ambiguous, especially when applied to nonhuman animals. I will therefore use the term memory accessibility to address the distinction between implicit and explicit memory. Memory monitoring is

therefore presently given a narrower operational definition by asking if monkeys can access a memory trace and guide a behavioral decision accordingly.

Metamemory is the ability to monitor and appropriately control one's cognitive activities on the basis of a particular accessible memory. A student studying for an exam is able to assess his memory strength for the material being tested and study accordingly. For example, because Stan knows that he knows all the countries of Africa, he does not waste time studying country names; instead he allocates his time more wisely to study capital cities, which he knows he does not know well. Whether the motivation is to save time, limit the number of working memory stores being utilized, or maximize one's score on a test, humans frequently use metamemory adaptively.

Recent evidence suggests that nonhuman animals can also monitor their own memory. Traditionally in animal cognition studies, animals are tested in forced choice situations. While these paradigms measure memory well, they allow no opportunity for subjects to demonstrate memory monitoring. The study of nonhuman animal memory monitoring is made possible by the addition of a decline-test option. Selection of the decline-test response ends a test, such that the animal is not forced to complete the test. Typically, correct choice responses (on accepted test trials) result in a highly preferred reward; incorrect responses result in no reward; and the decline-test response results in a less preferred, but guaranteed reward. Therefore, if the subject does not remember the correct answer, it is optimal to choose the decline-test response rather than making an incorrect test choice. Animals able to access and monitor their memory states should be more accurate when given the opportunity to decline tests than they are when forced to take tests.

Smith et al. (1998) conducted the first study of memory monitoring in rhesus macaques. The monkeys were tested on a serial probe recognition (SPR) task (Sands & Wright 1980). During the study phase, monkeys were shown a list of images presented one at a time. In the test phase subjects judged whether a target image (the probe) had been on the list or not by selecting the target or another distinct image, which indicated that the target image was not from the list. SPR performance is typically lowest when the probe presented is from the middle of the study list and highest when the probe is either first or last in the list (primacy and recency effects, respectively; Sands & Wright 1980). In this study, on some conditions, monkeys were given the option of using a decline response, which ended the trial. Declining the test guaranteed a forced-win trial in which only the correct answer appeared on the screen and resulted in a reward when selected. Use of the decline test option is advantageous for an animal if it does not know whether the image was from the list or not because it guarantees an easy-win trial, but after a short delay. Smith et al.'s subjects showed the expected primacy and recency effects. They also chose the decline test response most often when the probe stimulus was from the middle of the list, and memory was presumably weak. This correlation between primary performance in the memory task and use of the decline response is a major source of evidence used to infer metamemory. Importantly, performance on trials on which the decline response was present was significantly better than on forced trials when the decline response was not available. Superior performance on chosen trials indicates that monkeys monitored their memory and took tests only when memory was comparatively good.

Building on the approach used by Inman & Shettleworth (1999) to test for metamemory in pigeons, Hampton (2001) tested rhesus monkeys in a delayed matching to sample (DMTS) task. Subjects studied a sample stimulus presented on a touch sensitive monitor and were required to select an identical matching stimulus from a set of four choice stimuli at test. The delay between study and test was increased such that, on some trials, the monkeys were likely to forget the image before the test was presented. In two-thirds of trials subjects could decline the test by touching a particular icon on the screen, which resulted in a guaranteed but less preferred reward, or choose to take the test and potentially earn a preferred reward, by touching the alternate icon which initiated the choice phase of the trial. In the remaining one-third of trials, monkeys were forced to take the test and not given an opportunity to decline the test. Given these contingencies of reinforcement, monkeys with metamemory they should decline trials when they do not remember the matching stimuli, resulting in superior performance on chosen compared to forced trials. The monkeys did indeed perform significantly better on chosen than forced trials.

Novel probe tests were presented to assess the mechanism by which the monkeys decided to take or decline the tests. It is possible that the monkeys were cued to decline tests by certain external events, such as noises, grooming, changes in motivation, etc. that occurred during the delay and interfered with memory. Subjects could learn through trial and error that such external events predict low reward frequencies. This could therefore reliably cause subjects to decline tests when memory was weak, leading to the false conclusion that the monkeys' behavior was a result of monitoring and assessing one's own memory state, rather than an external cue (Hampton, 2001). Subjects were presented

with random no sample probe trials to control for such external cueing. These unpredictable trials were identical to normal trials, except that no sample image was presented at study. The proportion of declined tests was significantly higher on no sample probe trials as compared to normal trials. Therefore the absence of memory, rather than other external cues, guides the decision to take or decline a memory test. Delays were both increased and decreased to further rule out confounding environmental or behavioral cues that could account for the subject's performance. At long delays, when subjects were more likely to forget the sample stimulus, subjects declined tests more frequently; on trials with short delays, when memory was strong, the monkeys rarely declined tests. This study showed that, like humans, monkeys can appropriately monitor their own memory states.

This ability to adaptively monitor one's cognitive activity supported previous studies on information seeking in children, orangutans, chimpanzees (Call & Carpenter 2001) and rhesus macaques (Hampton, Zivin, & Murray, 2004). Animals were presented with 2 to 4 tubes arranged in parallel. On some trials the experimenter hid a food reward in one of the tubes in full view of the subject. In other trials, the subjects could not see the baiting take place. In these cases, subjects could learn the location of the food by bending down and looking though the tube. Both apes and rhesus monkeys looked through tubes more frequently when ignorant (when the baiting of a food reward was not visible) than when the baiting was shown, indicating that the subjects knew when they remembered. Call (2005) repeated this study with bonobos, gorillas, chimpanzees, and orangutans and found similar results. This foraging hint-seeking paradigm was also used to test capuchin monkeys' metacognitive abilities, but the results were not as robust (Basile & Hampton,

2008). Equivocal results were also obtained when capuchin monkeys were tested on a DMTS metamemory task (Fujita, 2009). Though it is therefore unclear whether new world monkeys (e.g. capuchins) can monitor their own cognitive states, it is likely that metamemory abilities are widely distributed in our closest relatives: apes and old world monkeys (e.g. rhesus monkeys).

The ability to monitor one's memory state may be advantageous for a monkey or ape in the wild. Instead of using trial and error to recover critical resources such as food, assessing a memory store would allow primates to travel to particularly promising locations. Such an accurate selection of bountiful foraging sites rather than returning to a location that has already been depleted would save both time and energy (Suda-King 2008). Accordingly, the experiments in this thesis will evaluate the extent to which monkeys use metamemory to appropriately control behaviors in novel forage-like decline-response test.

Converging approaches

Triangulation, the use of converging methods, (Heyes, 1993; Shettleworth 1998) is critical for studying memory monitoring in nonhuman primates. Tests have different parameters which control for, or are subject to particular behavioral and environmental cues dependent upon certain task unique characteristics. In the face of changing context-specific variables, consistent positive results for metamemory make the argument for metamemory very compelling. Testing if a nonverbal species is aware of the contents of his memory is challenging endeavor because it is difficult to discriminate between subjective and objective sources of behavioral control. The current research therefore

capitalizes on the importance of triangulation so that alternate tenable explanations, other than memory monitoring, can be eliminated. With the goal of testing the robustness of the presence of metamemory in monkeys, converging operations are implemented in the present study by: 1) using a decline-test, delayed response memory test 2) testing subject on 3 generalization tasks, 3) testing subjects in a new testing setting, and 4) implementing an approach that requires few training trials which may therefore reduce possible associative solutions. In sequence, these implementations will be briefly described before introducing the methods of this research.

The following experiments evaluate if monkeys can appropriately monitor the presence or absence of memory in a spatial delayed response (DR) matching to location task. In the DR task, monkeys are required to maintain a representation of an absent stimulus over the retention interval. Critical for these studies, the delay intervals used ensured that subjects experienced some trials in which they forgot and some trials in which remembered the location of the hidden food. The present DR task therefore created an opportunity for monkeys to discriminate between forgetting and remembering such that subjects could learn to decline memory tests when memory was poor. If the monkeys can monitor their memories, they should perform better on memory tests when given the choice to decline trials in which their memory is weak (choice trials) than when there is no option to decline (forced trials).

Memory manipulations in Hampton's 2001 study included no sample presentations and changes to the delay. Although these manipulations alter memory performance, it is unclear when during a trial (e.g. study, delay, test) memory processes (e.g. encoding, retention, retrieval, respectively) are affected. These variations in probe

trials will be evaluated in generalization tests in this study. On short delay trials memory is enhanced by simply shortening the delay, but because these are probe trials, the subject is unaware that the studied information needs to be remembered for a short period of time. The encoding of the sample is therefore similar across sessions with mixed delays. In these tests, a memory trace is most noticeably manipulated during the retention interval or retrieval. In no sample probe trials, the subject has not encoded any sample information, and is likely surprised to later be presented with a test after the delay period. In this case, the unique aspect of these probe trials might occur during the study phase (lack thereof) or during test. While it is unclear whether altering the probe trial occurs at study, delay, or test; the present study presents a new manipulation of memory in which memory was notably enhanced during the study phase of a trial. Time-specific manipulations of memory that target certain processes, i.e. encoding, has never been evaluated in task of metamemory before.

Finally, while many training sessions were required for monkeys to show evidence of metamemory in previous studies, the present study kept training to a minimum to decrease the opportunity for subjects to learn task-specific associations not related to memory monitoring. In addition to testing subjects on several novel probe tests, initial sessions on new tests were analyzed separately whenever possible to evaluate transfer performance and therefore examine the possibility of mere association learning.

General Methods

Subjects

Six male rhesus macaques (*Macaca mulatta*; mean age = 6.2 years) were tested in this study. Monkeys were pair-housed with food and water available *ad libitum*. Each monkey had access to his cage-mate at all times except during testing when the two monkeys were separated by a cage divider (20-90 mins/day).

Apparatus

A Plexiglas manual test tray (19.8" in length x 12" in width) containing five food wells was used (Figure 3). Four identical black PVC pipe end-caps (3.2" in diameter x 2.5" in height) were used to conceal the food rewards. A fifth smaller well located approximately half way between the front two cups and the back two cups, was concealed by a flat circular green Plexiglas sliding panel. A small green rectangular lip was attached to the circular panel on the side closest to the subject such that the lid could be easily pushed to the side to reveal the food reward in the well. The tray was placed on a short table such that the height of the table matched the height of a chow hole on the front of the monkey cages, allowing the monkeys to readily reach out to the food cups. During the test phase of a trial, the table was slid in front of the cage such that the

monkeys could reach the cups.

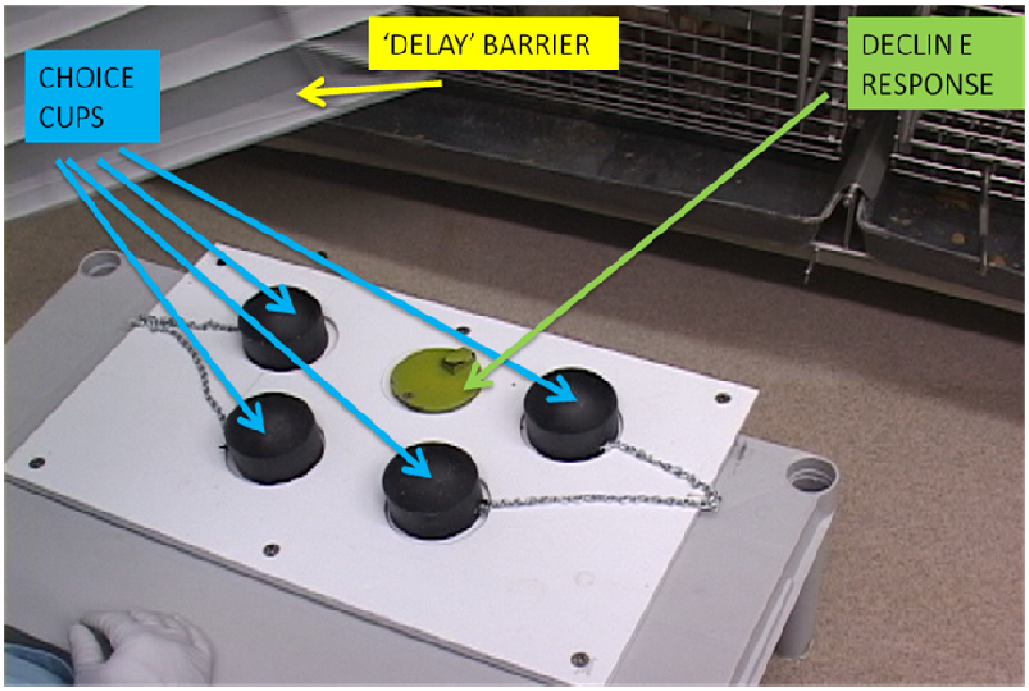

Figure 2. Photograph of the testing tray. The tray could be slid forward to the monkey's cage during the choice phase of trials. The experimenter sat where white glove can be seen at the bottom of the picture. See Procedure for further description of how the individual pieces were used during testing.

A grey plastic 'delay' barrier (18'' long, 15'' width, and 1.5'' in depth; Figure 2) was be used to completely block the monkey's view of the cups.

Rewards

Two raisins were always used as the food rewards hidden under one of the choice cups. On choice trials, when a distinct decline response sliding panel was present, one raisin was always hidden under this sliding panel (see Figure 2; *Apparatus section*) but the monkeys never saw this baiting occur. The baiting of the decline response was always done in between trials when the delay barrier was raised. The location containing the

raisins was randomized across trials with the constraint that no location was baited consecutively for more than four trials.

General Procedure

Monkeys were tested in their home cages. Prior to testing one paired-housed monkey, separation panels were placed between the pair of connected cages to confine each monkey to one cage. These opaque plastic separation panels had slits in them such that physical but not visual access to a subject's cage-mate was blocked. During the study phase of each trial, monkeys were presented with the testing tray, and in full view of the subject, the experimenter lifted one of the black food well covers, placed two raisins in the well, and replaced the cover on the food well. After the expiration of a delay interval, the tray was immediately moved up to the edge of the monkeys' cage so that the cups could be reached.

If the monkey chose the baited cup he retrieved two raisins; if not the tray was withdrawn so that the monkey could not overturn additional cups. The visual barrier was then placed in between the experimenter and subject, such that the subject could not see the cups rearranged back to starting position or any remaining food rewards removed from under the cups. This was necessary because without the barrier raised, the subject could have seen the re-baiting of the decline response on trials preceding trials in which the decline response selected. Due to proactive interference, watching the baiting of decline response and the retrieval of raisins under the unselected choice cup could mistakenly cue the subject to pick these locations.

Because the experimenter needed to ensure that the monkey was attending to the baiting (i.e. was facing forward), there was no specific inter-trial interval (ITI)

instantiated. However, the ITI was as short as possible: a trial began once the previous trial ended, the tray was reset for a new trial, and the monkey was attending. A darkened face shield, similar to a welders mask, was worn throughout testing in order to eliminate experimenter facial and gaze cues.

Data analysis & behavior scoring

On each trial the cup the monkey first touched was recorded. Before all statistical analyses were performed all proportions were arcsine transformed to better approximate the normality assumption underlying parametric statistics (Keppel and Wickens 2004, p.155).

Familiarization Training

Phase 1: selecting one baited cup out of two identical cups

In order to gradually expose the monkeys to the conditions used in the main task, subjects were first oriented to the apparatus and were to choose the baited cup out of two identical cups under which a reward was placed by the experimenter in full view of the subject. Only the first two cups on the tray (closest to the subject) were used in this phase; the back two cups (further from the subject) were removed from the tray (Figure 2). The shortest possible delay, 3 seconds, was used.

In familiarization phase 1 up to two correction trials were run in which the correct choice cup was re-baited. If the subject's cup selection was still incorrect after two correction trials, a forced trial was given in which only the baited cup was presented and the remaining un-baited cups were removed from the tray. Though these second and third

attempts to retrieve the reward were recorded, only the first response was used to calculate proportion correct. Sessions consisted of 10 trials. An 80% accuracy rate was required on two consecutive sessions in order for a subject to progress to familiarization phase 2.

Phase 2: visual barrier introduction

Phase 2 was identical to Phase 1 except that a visual barrier was placed between the tray and the subject during the delay immediately after the subject watched the experimenter bait the cup. The barrier was removed once the delay passed and the tray was slid directly in front of the subject. The purpose of introducing the barrier was that the subjects had to learn to solve the task by holding the baited location in memory, even when it was out of sight. Correction trials were given, as indicated in phase 1. Sessions consisted of 10 trials. An 80% criterion level was required for a subject to progress to familiarization phase 3.

Phase 3: transition to 4 cups

Phase 3 was identical to phase 2 expect that 4 cups were used rather than 2. The purpose of this phase of familiarization training was to gradually adapt the subjects to the final testing conditions in which the subjects choose the baited cup from among 4 cups. Using four cups rather than two increased the range of possible accuracies between chance (25%) and perfect accuracy (100%). In addition to the front two wells and cups, the back two wells (located further from the subject) were used (see Figure 2). Because the back two wells are closer together and the front two are further apart the spacing of

the cups forms a radial (semi-circle) configuration such that the ease in reaching each cup is intended to be approximately the same (see Appendix C for observed cup location preferences). Sessions consisted of 10 trials. Correction trials were given as in previous phases. A subject was moved onto familiarization phase 4 once accuracy was 70% or above on two consecutive sessions.

Phase 4: delay titration

The general procedure of phase 4 was identical to phase 3 except that delays between study and test were increased. The goal of the titration procedure was to find a delay at which performance was between 40 and 70% correct. Such intermediate performance ensures that monkeys had experiences both remembering and forgetting the correct hiding location. The delays, which were incrementally increased based on meeting criterion, were: 3 (as in all previous phases), 6, 12, 24, and 48 seconds (see Table 1). Before the session began a countdown timer was set to the given delay of the particular session. The timer was started right after the cup was baited, and stopped at the end of the delay interval at which time the visual barrier was removed and the tray was immediately slid in front of the subject. If a subject's accuracy was at or below 40% on two consecutive sessions, his delay was decreased back to the lower delay in the sequence (e.g. moved from 24 seconds back to 12 seconds). Likewise, if a subject's accuracy was 70% or above on two consecutive sessions at one delay, his delay was increased on the next session.

If the subject failed to reach this passing criterion level after 10 sessions, he remained at this delay for another 5 sessions. If the subject remained unable to reach

criterion on the additional 5 sessions, additional "tutor" training sessions were administered. This occurred with one monkey, M2, who was given additional training sessions where the number of choice cups was incrementally increased from 1 to 2, to 3, and 4. The subject was moved onto sessions with 1 added cup after 70% accuracy was achieved on 2 consecutive sessions with a fixed number of cups (see Table 1; *Note*).

If the subject did not reach this fail criterion or the pass criterion (which increased the delay) after ten sessions, this particular unchanging delay was his determined "criterion delay" (CD). These subject specific delays ensured the subjects' accuracies were between 40% and 70%. Correction trials were given as in previous phases. All subjects progressed onto phase 5 after 10 sessions (10 trials/sessions) were received. Accuracies for each subject are indicated in Table 1.

Table 1. Number of sessions required to reach criterion for subjects during each phase of familiarization training. Darkened cells indicate the stabilized criterion delay; 10 sessions without reaching the pass or fail criterion resulting in no change in delay in phase 4 of familiarization. The two even darker cells under M1 and M4 indicate that these monkeys' established criterion delays during training changed in Experiment 1 (from the initial criterion delay to the final criterion delay). The average number of sessions required to reach criterion for each delay are indicated on the far right. The averages for 10 sessions are not included in these means because after 10 sessions at one delay, the subjects did not progress to a new delay. Phase 4 accuracy is indicated for each subject. These percentages measure accuracy on the 10 sessions of phase 4 training at the corresponding delay. The average phase 4 accuracy across 6 monkeys is shown. Below this, phase 5 accuracies are indicated for each subject. Please note that the difference between phase 4 and phase 5 is the introduction of the decline response in phase 5. The average accuracy on phase 5 is also shown. Delays used in training and testing sessions (Experiments 1-4) according to trial type are displayed for each subject. *s*= seconds of the presented delay.

Stage of training/testing	Delay	M1	M2	M3	M4	M5	M6	Mean
Phase 1	3 s	5	13	12	12	3	4	8.17
Phase 2	3 s	2	3	2	3	2	2	2.33
Phase 3	3 s	2	17*	3	3	3	2	2.6
Phase 4	6 s	10	10	10	10	3	2	2.5
	12 s					2	6	4
	24 s					9	2	5.5
	48 s					10	10	

Phase 5	Initial criterion delay	6 s	6 s	6 s	6 s	48 s	48 s	
Phase 4 accuracy	no decline available	54%	58%	51%	53%	53%	61%	55%
Phase 5 accuracy	decline available	77%	78%	58%	76%	59%	54%	67%
Experiment 1	Final criterion delay	12 s	6 s	6 s	12 s	48 s	48 s	
Experiment 2	No delay	-	-	-	-	-	-	
	Final criterion delay	12 s	6 s	6 s	12 s	48 s	48 s	
Experiment 3	Short delay	12 s	6 s	6 s	12 s	48 s	48 s	
	Final criterion delay	12 s	6 s	6 s	12 s	48 s	48 s	
	Long delay	96 s	96 s	96 s	96 s	96 s	96 s	
Experiment 4	Increased delay	24 s	18 s	12 s	36 s	144s	96 s	
	Double sample	24 s	18 s	12 s	36 s	144s	96 s	

Note. * M2 reached 10 sessions at a 3 second delay without meeting the pass criteria (70% accuracy on two consecutive sessions) and was given 5 additional sessions. After still not reaching criteria to progress to a 6 second delay, he was given 5 additional training sessions with intermixed numbers of cups (2, 3, and 4). He then began sessions with only 2 cups (2 sessions required to reach criteria), 3 cups (4 sessions required), and 4 cups (2 sessions required). M2 then received 2 more normal phase 4 sessions, making the total number of phase 4 sessions (not including the additional training sessions) at 3 seconds equal to 17 sessions.

Phase 5: introduction of the decline response

The testing procedures used in phase 5 were the same as phase 4 except that the sliding decline response panel well was also present during study and test phases of each trial (see apparatus section above; Figure 2). The purpose of familiarization phase 5 was to give monkeys the opportunity to benefit from memory monitoring by learning that correct responses are always rewarded with 2 raisins, errors with none, and the decline response with 1. In addition to the four choice cups, monkeys were presented with a fifth option, a decline-test response, which ended the memory test (no choice cup was to be chosen). One raisin was always hidden under the decline response panel. To ensure that monkeys knew that a single raisin was available under the decline response panel, monkeys were given 4 trials a day in which only the decline response was available (no

choice cups) before free trials were introduced in which a decline response and a four choice cups were presented.

Rather than only a 25% chance of choosing the correct choice cup and getting the preferred reward of 2 raisins (an average of 0.5 raisins per trial), declining a memory test when the correct answer is not known would be the optimal choice as it results in one guaranteed raisin. However, to make such appropriate decisions, the monkeys must know on which trials they do not remember. If the monkeys can adaptively monitor their memory, performance should be better on freely chosen tests used here in phase 5 than on the forced tests used in phase 4. Subjects did in fact perform better on tests in which the decline was available, then they it was not (Table 1). No further analysis or discussion will be discussed here as these performance levels are simply observational results of training when the decline response was first introduced. Further analyses of familiarization results, however, can be found in Appendix A.

All six monkeys were tested in 20 sessions at the established CDs (see Table 1) determined, on which performance ranged from 40-70% correct. No correction trials were used in phase 5. Sessions consisted of 10 trials. After all sessions were completed, subjects progressed onto Experiment 1. Accuracies on forced and chosen trials on phase 5 are shown in Table 1.

Experiment 1

The purpose of Experiment 1 was to provide a direct comparison between forced and freely chosen trials. Free and forced trials were randomly intermixed in Experiment 1.

Methods

The procedures used in Experiment 1 were identical to familiarization phase 5 except that free and forced trials were randomly intermixed in each session of testing. One-third of trials presented to subjects were forced tests, in which no decline response was available. In two-thirds of the trials monkeys either chose to take the test by selecting one of the choice cups or declined the test by sliding the decline response panel to the side to receive the less preferred reward. To ensure that performance remained between 40 and 70%, if a subject's accuracy on forced trials averaged across 5 sessions was 70% or above, the monkey's delay was doubled. The performance of two monkeys met this criterion. Because both M1's and M4's performance at 6 seconds was above 70%, their new criterion delay became 12 seconds (see Table 1; *darkened cells*). Accuracies on forced trials were monitored for the remainder of this experiment, but no further adjustments to delay periods were needed. Each monkey received 15 sessions of 18 trials each, at their final CD; 12 of the trials in each session were choice; 6 were forced. Cup location was randomized and counterbalanced as done in familiarization training, as was forced and choice trials.

Results & Discussion

Accuracy was higher on chosen tests than on forced tests (80%, SEM= 3.81, and 61%, SEM=6.28, respectively; t_5 =-7.76, $p< .001$; Figure 3). These results indicate that when given the option of not taking a test, monkeys appropriately monitored their memory and selectively declined tests in which memory was poor. The first session subjects were tested on at their final criterion delay was analyzed separately to evaluate if

there was immediate transfer to intermixed trials. In this first session accuracy was again significantly higher on chosen tests than forced tests (82%, SEM=6.46, and 67%, SEM=7.45, respectively; t_5 =-2.83, p= .037). This reveals that when first introduced to intermixed trials of forced and choice tests, monkeys immediately showed evidence for memory monitoring. It is therefore unlikely that learning task contingencies can account for the ability of monkeys to appropriately use the decline response.

It is possible that factors other than memory monitoring could have guided subjects' decisions to take or decline the test. Subjects seemed more likely to select the correct cup on trials in which the raisins were placed in one of the front two cups (closest to subject; see Figure 2) as compared to trials in which one of the back two cups (further from the subject) was baited. This raises the concern that subjects were basing their decision to take tests on the location of the baited cup. To address this concern, performance on front cup trials was compared to performance on back cup trials. Performance on front cup trials was better than performance on back cup trials (t_5=3.87, p=.012; paired-sample two-tailed t-test). This indicates that it was in fact easier for subjects to answer correctly when the baited cup was in closer proximity to them. Subjects accordingly declined back cup trials significantly more than front cup trials (t_5=-7.185, p=.001; paired-sample, two-tailed t-test).

The fact that monkeys declined tests more often when they were more likely answer incorrectly is consistent with the possibility that monkeys' decisions to take and decline tests is based on their internal memory state. However, it is also possible that it was the location of the food, rather than the quality of memory that guided use of the decline-test response. Performance on forced trials vs. chosen trials on the front two cups

was therefore analyzed separately, as was performance on the back two cups. There was no significant difference between performance on forced trials as compared to chosen trials for both front and back cups (paired-sample, two-tailed t-tests; front cup trials: t_5=.603, p=.573; back cup trials: t_5=-1.44, p=.211). Thus, while monkeys showed apparent memory monitoring in that they performed significantly better on chosen than forced tests as a whole (when all cup locations were averaged together), this result does not hold up when cup locations were grouped into back and front cups and analyzed separately. Additional analyses of performance for each cup location are located in Appendix C. Because memory was confounded with cup location in this experiment additional experiments were conducted in which cup location was not confounded with differences in memory. A crucial purpose of the three following experiments is to rule out the possibility that the decision to take tests was based cup preference or other external cues, rather than memory monitoring itself.

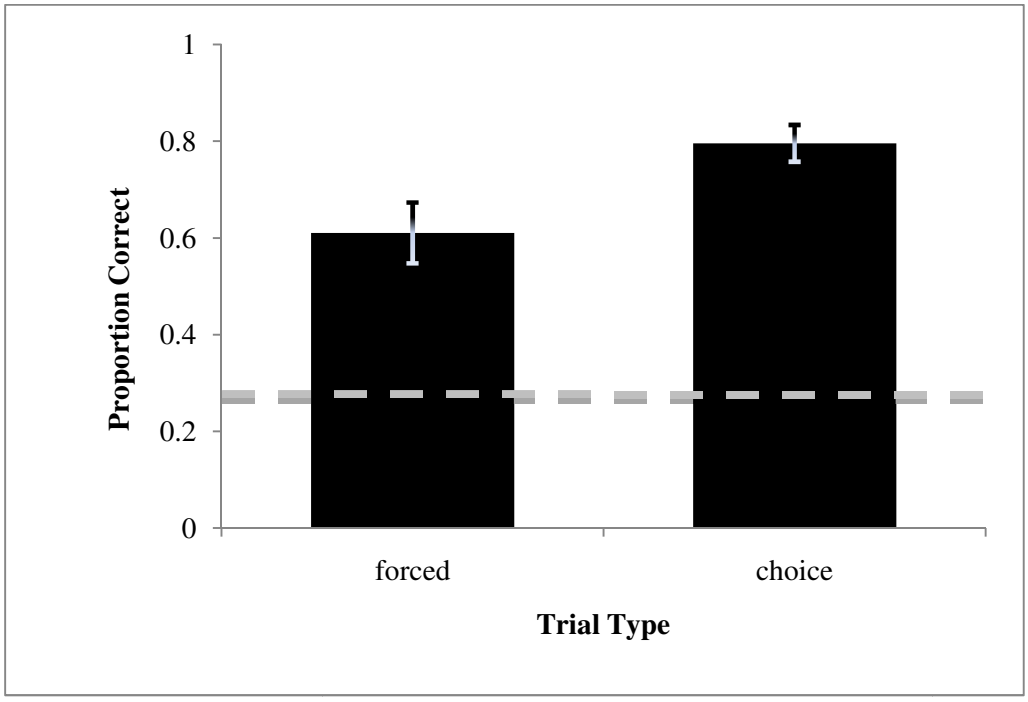

Figure 3. Experiment 1. Average accuracy on freely chosen and forced tests. Forced and choice tests were intermixed within sessions. Monkeys were tested on 15 sesssions. Chance, 25%, is indicated by the dashed line. Error bars are standard errors of the means.

Experiment 2

While the discrepancy between forced and choice trials found in Experiment 1 presents possible evidence for metacognition and could be due to monitoring memory strength, it is possible that other factors guided the subject's decision to take or not take certain memory tests. Most evidently, cup preference analyses revealed that no differences between forced and chosen trials were observed when front cup trials and back cup trials were analyzed separately. In addition, other cueing, which is not directed by memory monitoring, could have also contributed to this performance. Such cueing includes behaviors such as noises, changes in motivation, grooming which could result in forgetting (Hampton 2001). These environmental or behavioral stimuli could become associated with particular task contingencies, misleading the conclusion that monkeys are judging their own memory states.

To control for the possibility that monkeys might simply be responding to these event cues (including which cup is baited during the study phase of a trial), novel probe trials were presented. Normal choice trials were randomly intermixed with novel trials in which no sample was presented. Unpredictable, no sample probes in which no choice cup was baited allowed the absence of memory to be experimentally controlled. If the absence of memory does in fact cause the monkeys to decline tests, they should decline tests much more frequently when no sample to remember was presented. Conversely, if these distracting events location are what causes subjects to forget and reliably decline tests, then monkeys should decline roughly the same proportion of normal tests as probe

tests. Similarly, if monkeys are only guiding their decision to decline a test based on if one of the back two cups is baited, they should decline no sample tests as frequently as they decline normal tests.

Methods

The subjects, apparatus, and procedure were identical to previous experiments and training. Monkeys were tested at their CDs (see Table 1). Probe trials were executed exactly as normal trials except that no cup was baited. The delay period between trials was the same for both types of trials. This ensured that the monkeys could not simply associate an external cue, time, with whether they should decline the test or not. In between trials the cups were moved back to starting positions and left over raisins that had not been retrieved were removed with the delay barrier in front of the tray such that the monkey could not see any resetting of the tray. For no sample probes the timer was started right after this short resetting period and the delay barrier remained raised. After the delay, the tray was revealed and pushed within reaching distance of the monkey just as in the normal test phase of a trial. Monkeys received 5 sessions. Six no sample probe trials were randomly intermixed with 12 normal trials to produce sessions of 18 trials.

Results & Discussion

Monkeys declined tests on no sample probe trials more than on normal trials, 96% and 31% of the time, respectively (t_5=-16.14, p <.001; Figure 4). Because the experimental manipulation of memory produced this significant difference, these results strongly support the conclusion that the monkeys' ability to selectively decline tasks is

based on detecting an absence of memory rather than cueing by external events occurring during the delay (bouts of grooming, noises, etc.) or cup location.

Though not many sessions were run, it is possible that over time the monkeys learned a distinctive feature of the probe trials and based their decision on this association (Hampton 2001). The first session for all six monkeys was analyzed separately to rule out learning. In the first session with this novel task monkeys declined probe tests significantly more than normal trials (paired two-tailed t-test: t_5= -9.37, p<.001; Figure 4; *grey bars*). These results indicate that while familiarization with no sample probe trials may have increased over five sessions, simply learning to decline any probe test did not underlie performance. Association learning can therefore not account for monkeys' decisions to decline tasks. In addition, two monkeys, M3 and M6, never took a memory test on no sample probe trials, and thus never experienced the negative consequences of choosing the choice cup on probe trials and only having a 25% of getting a reward. The other four monkeys had minimal experiences with such low reward consequences (M4, M1 and M5 declined probe tests 97% of the time; M2 declined 83% of the time).

By experimentally manipulating an absence of memory no sample probe trials simulate those cases on normal trials when monkeys forget the correct reward location. The subjects were able to generalize this adaptive capability of monitoring their own memory to a novel task. These results provide convincing evidence that monkeys' decision to decline memory tests is based on the ability to monitor one's own memory strength or lack thereof.

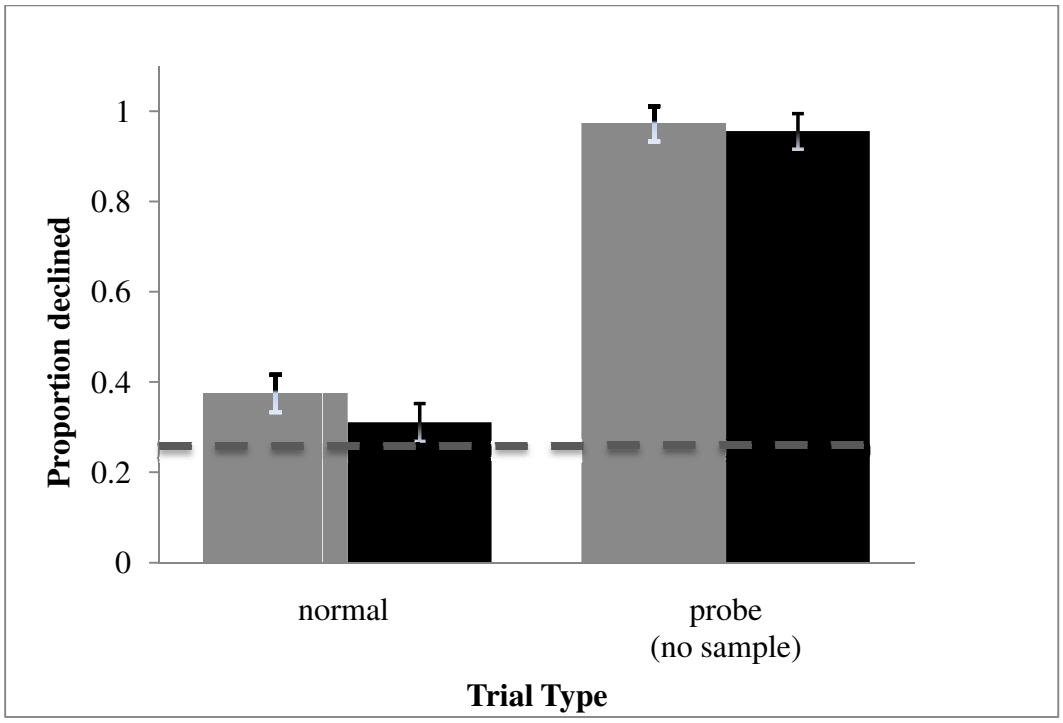

Figure 4. Experiment 2. Probability of declining tests on normal trials and on trials in which no choice cup was baited (probe; no sample). All trials in Experiment 2 in the above graph, were tested at each subjects' criterion delays. Grey bars represent the first session monkeys received. Black bars indicate the average performance across all 5 sessions. Error bars are standard errors of the mean. Chance, 25%, is indicated by the dashed line.

Experiment 3

Memory was experimentally manipulated in Experiment 2 by eliminating any to-be-remembered stimulus by presenting no sample probe trials. In Experiment 3, memory was also experimentally manipulated but rather than eliminating a memory trace by presenting no sample, the likelihood of remembering the sample was manipulated. The delay over which monkeys had to remember the baited cup was manipulated in the present experiment; monkeys were tested at both long and short delays. If monkeys are able to monitor their memory they should decline tests following long delay intervals

more often than those following short intervals because memory should be weaker after long than short delays.

Methods

All procedures were identical to previous experiments. Short and long delays (3 and 96 seconds, respectively) were randomly intermixed with the established CDs for each monkey, which fell between these extremes (see Table 1). Each 18 trial session contained 6 trials at each of the three delays, 4 of which were chosen, and 2 of which were forced. In addition to the randomization and counterbalancing of forced and choice trials, short, medium, and long delays were also randomized. Monkeys received 20 sessions.

Results & Discussion

Memory strength at each delay is reflected in performance on forced trials. The average accuracy on forced trials for all six monkeys was 92% on short delays, 72% at criterion delay, and 29% at long delays (see Figure 5; *filled boxes*). A within subjects one-factor ANOVA was conducted to test if accuracy declined significantly on forced trials. As delay increased, performance significantly decreased ($F=_{2, 10}=64.15$, $p<.001$). To determine whether monkeys declined long delay tests more than short delay tests a single-factor within subjects ANOVA was conducted. Monkeys were most likely to decline trials at long delays, less likely at intermediate delays, and least likely at short delays ($F_{2, 10}=95.01$, $p< .001$; bonferroni test corrected for multiple pair-wise comparisons and showed significant comparisons between all delays, $p< .005$; Figure 5;

filled circles). At 3 second delays monkeys declined tests less than 1% of the time, while at a 96 second delay period they declined tests 91% of the time. Combined, these two results signify the main objective of this study: they indicate that monkeys' decisions to take tests are guided by their ability to assess their memory strength.

A within subjects' two-factor (delay x trial type) analysis of variance (ANOVA) was conducted to evaluate the effect of delay on performance on both forced and chosen tests. Because the sphericity assumption was violated in this ANOVA, the Geisser-Greenhouse correction was used, which normalized the distribution (Keppel and Wickens, 2004, p.378). As delay increased, accuracy fell: performance on forced trials decreased with delay and differed significantly from performance on chosen trials (delay $F_{2, 10}$=26.17, p=.003; trial type: $F_{1, 5}$=17.24, p=.009; delay x trial type: $F_{2, 8}$=0.58, p=.582; Figure 5).

Though at a group level there was a main effect of both trial type and delay, no interaction was found. However, it is not clear that an interaction would necessarily be expected if animals have metamemory. Because previous studies showed individual significant differences (Hampton 2001; Suda-King 2008), exploratory analyses for individual monkeys were performed. No difference in performance is expected when monkeys frequently remembered; this corresponds to the small gap between the forced and chosen lines at short and medium delays (Figure 5; *squares*). When accuracy is close to ceiling at the short and medium delays there is not room for significant improvement on choice trials as compared to forced. However, at long delays, when accuracy is lower, some argue that a significant difference between forced and choice trials should emerge (Hampton 2001). In other words, of those very few tests that monkeys chose to take at

high delays, they should be much more likely to be correct on these trials than on forced trials at this delay. Using a two-way contingency table analysis, forced and chosen tests at the high delay were found to be significantly related in 2 monkeys, Pearson X^2 ($p<.05$, see Table 2). M1, M4 and M6 were also more accurate on freely chosen tests at the long delay but these differences were not statistically significant (Table 2). Because subjects rarely chose to take the test at long delay tests (Table 2 indicates how many test each subject choose to take), it is not particularly surprising that 4 monkeys accuracies on long delays did not reveal a significant effect between chosen and forced tests. Yet, this pattern found in which two monkeys performed significantly better on chosen tests than forced tests at a high delays is interesting. It indicates that monkeys are able to spontaneously generalize the ability to selectively decline tests when they don't remember. Even when harder novel tasks with long delays are presented, monkeys choose to take tests when they were sure, or knew, that they remembered.

These results indicate that monkeys choose to take tests when they remember and decline tests when they don't. Experiment 2 showed that subjects based their decision to take memory tests on the absence or presence of a memory, the study phase. Rather than manipulating memory to an extreme with no study phase (no visibly baited cup), the present experiment manipulated memory strength. Monkeys were therefore basing their decision to take tests on the particular degree of memory strength resulting from an unpredictable delay period. Like Experiment 2, it is not possible to explain these results by cup preference biases as explained above. This experiment further supports the conclusion that monkeys can monitor their memories.

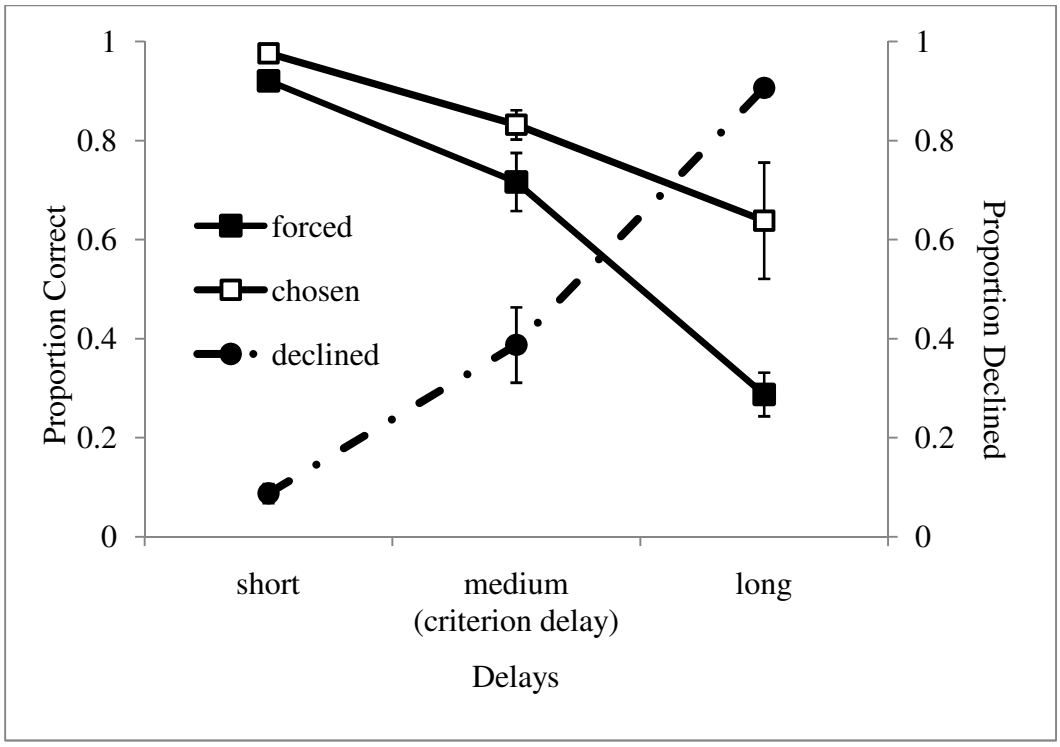

Figure 5. Experiment 3. Accuracy on forced and chosen trials after short (3 s), medium (6, 12, 24, 48 s depending on the monkey's criterion delay), and long delays (96 s). The proportion of correct trials for forced (open squares) and chosen (closed squares) corresponds with the far left y-axis. Proportion of trials declined (closed circles) corresponds to the far right y-axis. Error bars represent standard errors of the means.

Table 2. Chi square values in the first two rows from Experiment 3 were obtained from contingency table analyses on forced vs. chosen trials at the high delay (96 s). *F* and *p* values in respective rows are from the Pearson chi square test. Two-tailed test were used to determine if these probabilities in Experiment 3 were significant. The number of memory tests each subject chose to take across 20 sessions (2 forced high delay trials per session; total=40 trials) is shown in the next row. Accuracies at the doubled criterion delay (and tripled criterion delay for M2, M4, and M5) are shown that were established before Experiment 4 began. These delays are indicated in Table 1. Chi square values are shown in the next two rows and are from analysis of contingency tables classifying the number of correct and incorrect trials on single and double sample trials. Percentage of tests declined on single sample testes and double sample tests is shown in the next row (total=80 trials for single sample; total= 40 trials for double sample). Chi square values in the last two rows were obtained from contingency tables analyses of the number of incorrect and correct trials on single vs. double sample trials. Two-tailed tests were used to determine if these probabilities in Experiment 4 were significant.

Monkey		M1	M2	M3	M4	M5	M6	Mean
Experiment 3	Forced vs. chosen chi square, *df*=1	0.20	7.38	0.017	2.02	11.48	2.63	
	Chi square probability	0.654	0.007	0.895	0.156	0.001	0.105	
	Number of chosen tests at high delay	10	1	8	8	9	9	7.5
Experiment 4	Accuracy at increased delay	56%	44%	56%	61%	61%	44%	54%
	Single vs. double sample chi square, *df*=1	6.652	4.195	3.851	2.967	2.143	0.534	
	Chi square probability	0.01	0.041	0.05	0.085	0.143	0.465	
	Percentage of single sample tests declined	64%	58%	44%	53%	50%	66%	56%
	Percentage of double sample tests declined	40%	23%	38%	8%	48%	40%	33%
	Declined single vs. double sample chi square, *df*=1	7.42	17.3	2.12	21.98	0.02	6.79	
	Chi square probability	0	0	0.07	0	0.45	0	

Experiment 4

In Experiment 2 no sample probe trials served as equivalents to tests when monkeys forgot the hiding location on normal trials (criterion delays ensured 40-70% accuracies) as did long delay probes in Experiment 3 when monkeys rarely remembered. These previous experiments lead to the same prediction and result: more declined tests due to a decrease or absence of memory. As a result, it is possible that monkeys use the decline response more whenever something unusual happens in a trial, rather than because of the effect these manipulations have on memory. Therefore, a possible concern is that that in Experiments 2 and 3 monkeys simply chose to decline tests when novel probe trials were presented as subjects are not used to these types of trials. Perhaps they became confused and by default, frequently declined probe tests. Experiment 4 aimed to test this hypothesis by increasing memory in an atypical way.

To address this concern that monkeys might decline any unusual trial, the present experiment tested the reverse pattern by increasing memory. Experiment 4 investigated if monkeys base their decisions to take memory tests on the absence of memory rather than the simple detection of an unusual or novel trial. Double sample probe trials were determined to be an appropriate means to increase memory. In accordance with the trace strength and decay theory, a sample presented multiple times should increase memory, as it did in pigeons (Grant 1976; Roberts 1972). If the correct choice cup is double-baited the probability of retrieving the target memory should increase. The goal of Experiment 4 was to increase monkeys' memories in a novel way, and if memory was increased, evaluate if memory tests are declined less frequently on these increased-memory trials. If it is memory rather than unusual events that directs subjects when and when not to take tests, then the proportion of declined tests should decrease on double sample tests. If, however, any unusual event guides the decision to take tests, the number of declined tests should increase on double sample tests.

In previous experiments it is not clear exactly when the novel manipulation took place: either during the study or test phase of a trial. When no cup was baited (no sample trials) in Experiment 2, the manipulation may have occurred during the study phase when the barrier simply remained raised and no study period was presented. At the same time, however, the novel aspect of the trial could have been during the test phase of the trial when the subject may have been surprised to be presented with a test (because no study phase was presented). In Experiment 3 the unusual event of the probe trial was simply the short and long delays presented, such that memory was either manipulated 1) during the retention interval or 2) at test when the subject may have been surprised to be presented

with a test after only 3 seconds or after waiting as long as 96 seconds. The present experiment therefore created a situation where this timing of the unique trial manipulation was more definite: it clearly occurred during the study phase. Double baiting therefore must affect the encoding process of memory.

Methods

Apparatus, subjects, and general procedure were identical to previous experiments. Double sample probe trials were presented on one-third of the trials to increase the subjects' memory of the baited cup. On these probe trials the correct choice cup was slowly lifted and the two preferred raisins were placed in the corresponding well in the tray. The cup was then placed over the raisins as in previous experiments. The cup was then lifted for a second time such that the monkey could clearly see and be reminded of which cup hid the reward. The timer was started after this second presentation and the delay barrier was introduced as usual.

For this experiment it was important that baseline memory was low enough to provide an opportunity for memory to improve on double sample trials. Because performance at the criterion delay had likely increased from when it was established in Experiment 1, delays were doubled. Monkeys were tested on normal forced trials at the doubled delay in one session (see Table 1). If their performance fell in between 40 and 70%, this delay was used in the current experiment. M1, M3, and M6 obeyed this criterion. Because M2, M4, and M5s' performances were above 70% correct, their CDs were tripled. These final delays are indicated in Table 1 and performance at these delays is displayed in Table 2. These delays were used on both double sample and normal trials.

Monkeys received 10 sessions; each session consisted of 18 trials, 6 of which were double sample trials, and 12 of which were single sample (normal trials). Sample type (single vs. double) and trial type (forced vs. choice) were randomized and counterbalanced.

Results & Discussion

To test whether accuracy was higher on double sample trials on both forced and chosen tests a two factor within subjects ANOVA (trial type x sample type) was conducted. There was a significant effect of trial type (forced vs. chosen), no significant effect of sample type (normal vs. double sample baiting), and no significant interaction, (trial type: $F_{1, 5}$=11.03, p=.021; sample type: $F_{1, 5}$=4.52, p=.658; trial type x sample type: $F_{1, 5}$=0.221, p=.658; Figure 6). The significant effect of trial type replicates the main finding of Experiments 1-3 that accuracy increases when monkeys are given the opportunity to decline tests in which they do not remember. Though there was no main effect of sample type and no interaction, a significant effect is not in fact predicted, particularly on chosen trials. However, what are critical to compare is 1) forced test accuracies on single versus double sample trials and, 2) the amount of tests monkeys declined on single vs. double sample trials.

A paired one-tailed t-test was used to compare accuracy on forced choice double and single sample trials. Accuracy on double sample trials was significantly higher than on single sample trials (t_5=2.167, p=.041). Monkeys also declined tests significantly more frequently on normal trials than on double sample trials (paired, one-tailed t-test: t_5=3.221, p=.0115). This latter result is particularly important because it shows that

monkeys do not simply decline tests based on the fact that the test was novel or odd. If this were the case, the subjects would have declined the double sample tests more than the single sample tests, basing their decision on unique features of the task rather than memory. The opposite was found: monkeys were more likely to decline memory tests on normal trials. Subjects were able to appropriately use the decline response on novel tests. It was the purpose of this experiment to create unordinary probe trials which increased memory because it reveals that monkeys are not basing their decision to decline tests simply on some distinguishing feature of probe trials.

Inspection of data from individual monkeys revealed large differences in response to the double sample procedure. While 5 of 6 monkeys were numerically more accurate on double than single sample trials, one monkey showed the opposite pattern (Table 2). The performance of individual monkeys was therefore analyzed to determine which monkeys significantly benefited from the double sample presentations. Chi square analyses were performed to evaluate which monkeys performed better when the correct choice cup was baited two times. Four of the monkeys, M1, M2, M3, and M4 performed significantly better with double sample baiting (Pearson X^2, two-tailed, $p < .05$, Table 2). Unlike Experiment 3 where it was fairly transparent that short delays would increase memory and long delays would decrease memory, it was not known how the manipulation in Experiment 4 (double sample baiting) would affect accuracy. It is also unknown exactly why 4 monkeys significantly benefited from this novel sample type presentation and 2 did not. Possible factors could include differences in motivation, attention, study methods, etc. Nonetheless, of these 4 monkeys who did significantly benefit from double sample baiting, 3 of them, M1, M2, and M4, declined single sample

tests significantly more than double sample tests (Pearson X^2, $p < .05$, Table 2). This supports the conclusion that at least some monkeys base their decisions to take tests on their assessment of their own memory strength.

Though it is difficult to say how one would expect performance on chosen tests with each sample type would interact with an increase in performance (forced tests) with double sample baiting, the critical questions of this experiment were answered. Because monkeys declined tests more on single sample tests, when they were less likely to remember, it is clear that they did not simply decline new or odd tests on the basis of such a non-cognitive, external cue. Rather, they used their memory and ability to monitor their own cognitive state to decide whether or not they should take a memory test. This experiment provides additional evidence that monkeys know when they remember.

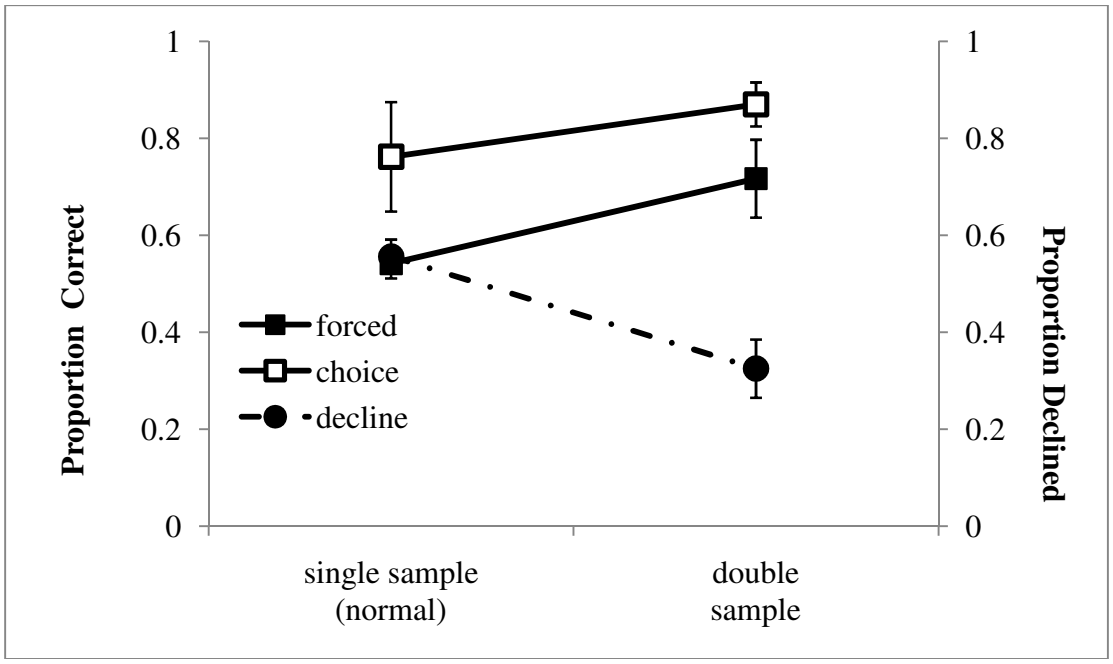

Figure 6. Experiment 4. Average accuracy on forced and chosen trials for single sample (normal) and double sample (probe) trials are displayed with filled and open squares, respectively. These accuracies correspond to the proportion of correct tests on the far left y-axis. Filled circles represent the proportion of trials declined for single sample and

double sample trials and correspond to the far right y-axis. Error bars are standard errors of the means.

General Discussion

Metacognition, having knowledge of one's thoughts (Nelson 1996), has been thought to be a uniquely human faculty (Carruthers 2008). However, recent research indicates that other animals can distinguish between when they know and when they do not know (Kornel et al 2007; Smith 2003; Hampton 2001; Call & Carpenter; 2001). Reinforcing these recent findings, in the present study monkeys avoided tests when their memories were weak and choose to take tests when their memories were strong. The battery of different probe trial types subjects received demonstrates that monkeys can apply metamemory judgments adaptively in novel situations. Such generalization indicates that the monkeys' metamemory judgments are not associatively tied to specific objective stimuli present in particular testing circumstances, but are rather based on subjective assessment of the quality of memory. In Experiment 1, monkeys performed significantly better on chosen than forced tests, indicating that when given the opportunity to decline tests, monkeys appropriately monitoring their memory and selectively declined tests in which memory was poor. Experiment 2 controlled for the experimental confound of cup location guided decisions to take or decline tests as well as other environmental cueing. Subjects declined tests significantly more when no sample was presented, indicating monkeys are able to detect when they have no memory trace of a baited cup and decline tests accordingly. Experiment 3 showed that monkeys frequently decline tests when memory is weak after long delays and rarely decline tests when memory is strong after short delays. In Experiment 4 a novel generalization test was used

in which memory was increased by repeating the sample phase twice in probe trials. Monkeys declined fewer tests when their memory was increased in this way than they did on normal, single sample, trials. Rather than declining the novel probe trials more frequently than normal trials, which monkeys would presumably do if they were simply learning an association to decline new or different trials, subjects applied their ability to appropriately use the decline response to these tests. Together, the four experiments presented here provide strong evidence that monkeys directly assess memory strength, rather than using behavioral or environmental cues that correlate with memory strength, to guide their decision to take or decline tests (Hampton, 2009). Because monkeys transferred flawlessly to all novel probe trials tested, this spatial memory task has proved to be a practical, foraging-like test paradigm for assessing metamemory that rhesus monkeys readily learn.

The findings reported here and elsewhere indicate that some memories in monkeys is accessible to monitoring suggests that human memory taxonomies that distinguish between implicit and explicit memory can be applied to nonhumans. Old world monkeys (rhesus monkeys) and humans shared a common ancestor approximately 25 million years ago (Gibbs et al 2007). The present study, along with previous research (Smith et al 1998; Hampton 2001; Hampton, Zivin, & Murray 2004; Kornell et al) suggests that the capacity for memory monitoring may have originally evolved in this common ancestor before old world monkeys diverged from apes and humans. While it is possible that metacognition evolved even earlier before the split between new world monkeys (i.e. capuchins) and old world monkeys, strong evidence of metacognition in new world monkeys is currently lacking (Basile & Hampton 2008).

Monkeys vs. orangutans in a decline response paradigm

In addition to testing use of the decline-test option in an ecologically relevant way never before used with monkeys, this study also provided a more direct comparison between monkeys and apes as monkeys were tested using an experimental design similar to that used with apes. Interestingly, however, orangutans *(Pongo pygmaeus)* tested in a similar paradigm failed to show evidence for metamemory (Suda-King 2008). The lack of evidence for metamemory in orangutans was surprising because there was existing evidence of metamemory in apes (Call & Carpenter 2001). Because apes are most closely related to humans there is no reason to assume that monkeys evolved a capacity for memory awareness and apes did not.

Because this discrepancy between rhesus monkeys and orangutans (Suda-King 2008) could be attributable to differences in motivation, attention, metacognitive abilities, or modifications in the test design, apes need to be further tested. Yet, because evaluating the possibility that methodological contingencies may help account for this apparent discrepancy, the methods of Suda-King's study will be briefly analyzed. The methods used in Suda-King (2008) differed from this study in a few possibly relevant ways: 1) no delay was interposed between study and test, 2) two choice cups were used rather than four, and 3) the choice cup locations were moved after the subject chose to take the test but before the subject chose a particular choice cup. The apes watched the baiting of a preferred reward in one of the two identical choice cups and either chose to take the test by pointing to the choice cups or declined the test by selecting the decline-response cup which contained a less preferred reward. If the choice cups were chosen, the two cups

which were arranged in a straight line in front and back of each other were then separated and slid to distinct locations on the tray such that the apes had to visually track a moving hiding location. This extra work may have caused the subjects to perform poorly on the task due to mistakes made in visually tracking the moving cups rather than a lack of metamemory. Importantly, this made the task difficult after the subject had already made a decision of whether to take the test or not. Any impact moving the cups had on the ability of the ape to remember the location of the food could not possibly be reflected in the subjects' decision to take or decline the test.

While the apes may have found it difficult to visually track the moving cups and this in turn may have decreased their overall performance on the task, it is also possible that the task was not hard enough. Using only two cups as hiding locations and instantiating no delay might have made the task too easy such that the apes may not have found the decline response option particularly advantageous. Because performance on this task depends on many cognitive abilities other than metacognition, negative results may not be attributable to metacognition per se, but could instead be a failure of any of these needed skills. Such abilities might include object permanence, spatial learning, and some sort of reward probability analysis. This raises an important aspect of this study: monkeys' success in the present experiments should be attributed not only to metamemory, but also to other cognitive skills listed above.

Neuroanatomy of metamemory

The study of nonhuman primate memory accessibility may provide insight into the evolutionary foundations of memory awareness and possible precursors to

consciousness. This behavioral evidence of memory monitoring in monkeys presents a good animal model for neurobiological explorations of accessible memory such that direct comparisons can be made between nonhuman and human cognitive processing. Though at present little is known about the neural correlates of metamemory in nonhuman primates, it has been shown that the human metacognitive control depends on midfrontal regions of the brain (Fernandez-Duque et al 2000). Because metacognition involves the cognitive control of cognition, Shimamura (2000) explored in greater detail the relationship between executive control and metacognition. While Shimura concurs that the midfrontal brain regions are involved in metacognitive regulation, he argues that metacognition is also comprised of other complex memory processes associated with the frontal lobes, such as working memory and executive control.

Based on the structural similarity of rhesus macaque and human brains, it is likely that the midfrontal regions were recruited when the subjects in the present study made metacognitive decisions. However, to be more exact, future neuroanatomical studies of metamemory in monkey models are needed. Not only is it hard to infer neural correlates of monkey cognition from human data, but much of the work in human metacognition has not focused exclusively on memory monitoring per say. It would be useful to identify the specific structures involved in metamemory because of implications it has for improving treatments for patients with deficits in higher cognitive functioning, such as memory loss and cognitive control. For example, memory loss from injury, stroke, or Alzheimer's disease can starkly weaken the ability to plan for the future and make appropriate decisions. Because of the lack of ability to reflect upon one's owns cognitive states (and infer about other's mental states and attitudes), metacognition is also impaired

in autism. Deficits in metacognitive abilities also exist in behavioral disorders, such as ADHD, in which attention and impulse control are aberrant. The present behavioral study importantly adds to the growing evidence that rhesus macaques are an ideal model for studying the neural basis of memories that are accessible to monitoring.

References

Baddeley, A. (2003). Working memory: Looking back and looking forward. [Review]. *Nature Reviews Neuroscience, 4*(10), 829-839.

Call, J. (2001). Object permanence in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes), and children (Homo sapiens). *J. Comp. Psychol.* ***115***(2).

Call, J. (2005). The self and other: a missing link in comparative social cognition. In: Terrace HS, Metcalfe K (eds) The missing link in cognition: origins of self-reflective consciousness. Oxford University Press, New York: 321-152.

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

Carruthers, P. (2008). Meta-cognition in animals: A skeptical look. [Article]. *Mind & Language, 23*(1), 58-89.

Clark, R. E., Manns, J. R., & Squire, L. R. (2002). Classical conditioning, awareness, and brain systems. [Review]. *Trends in Cognitive Sciences, 6*(12), 524-531.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. [Review]. *Consciousness and Cognition, 9*(2), 288-307.

Gibbs, R. A., Rogers, J., Katze, M. G., Bumgarner, R., Weinstock, G. M., Mardis, E. R., et al.(2007). Evolutionary and biomedical insights from the rhesus macaque genome. [Article]. *Science, 316*(5822), 222-234.

Grant, D.S. (1976). Effect of sample presentation time on long-delay matching in pigeons. *Learning & Motivation, 7*:580-590.

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences of the United States of America* ***98***: 5359-5362.

Hampton, R. R. (2006). Memory awareness in rhesus monkeys. *Diversity of Cognition.* K. Fujita and I. Shoji. Kyoto, Kyoto University Press**:** 282-299.

Hampton, R.R. (2009). Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? Comparative Cognition & Behavior Reviews, 4, 17-28.

Hampton, R. R., A. Zivin, et al. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition* ***7***: 239-254.

Heyes, C. M. (1993). Anecdotes, training, trapping and triangulating: do animals attribute mental states? *Anim. Behav., 46*(1), 177-188.

Inman, A. and S. J. Shettleworth (1999). Detecting metamemory in nonverbal subjects: A test with pigeons. *Journal of Experimental Psychology-Animal Behavior Processes* ***25***(3): 389-395.

Jonides, J., Lewis R.L., Nee, D.E., Lustig, C.A., Berman, M.G., Moore, K.S. (2008). The Mind and Brain of Short-Term Memory. *Annual Review of Psychology, 59*, 193-224.

Kornell, N. (2009). Metacognition in Humans and Animals. *Current Directions in Psychological Science* ***18***(1): 11-15.

Kornell, N., L. K. Son, et al. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science* ***18****(1)*: 64-71.

Nelson, T. O. and L. Narens (1996). Why investigate metacognition? Metacognition. J. Metcalfe and A. P. Shimamura. Cambridge, MA, The MIT Press.

Parker, A. (1998). Primate cognitive neuroscience: What are the useful questions? *Behavioral and Brain Sciences, 21*, 128.

Paukner, A., J. Anderson, et al. (2006). Redundant food searches by capuchin monkeys (*Cebus apella*): a failure of metacognition? *Animal Cognition* ***9***(2): 110-117.

Purdy, J. E., & Domjan, M. (1998). Tactics in theory of mind research. *Behavioral and Brain Sciences, 21*(1), 129-130.

Reder, L. M., Park, H., & Kieffaber, P. D. (2009). Memory Systems Do Not Divide on Consciousness: Reinterpreting Memory in Terms of Activation and Binding. [Article]. *Psychological Bulletin, 135*(1), 23-49.

Roberts, W. A. (1972). Short-term memory in pigeon- effect of repetition and spacing. *Journal of Experimental Psychology, 94*(1), 74-&.

Sherry, D. F. and D. L. Schacter (1987). The Evolution of Multiple Memory-Systems. Psychological Review **94**(4): 439-454.

Sherry, D. F. (2006). Neuroecology. [Review]. *Annual Review of Psychology, 57*, 167-197.

Shettleworth, S. J. (1998). Cognition, Evolution, and Behavior. New York Oxford University Press.

Shields, W. E., J. D. Smith, et al. (2005). Confidence judgments by humans and rhesus monkeys. *Journal of General Psychology* ***132***(2): 165-186.

Shields, W. E., J. D. Smith, et al. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology-General* ***126****(2)*: 147-164.

Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. [Editorial Material]. *Consciousness and Cognition, 9*(2), 313-323.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26*, 317-374.

Smith, J. D., J. S. Redford, et al. (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., J. Schull, et al. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*)." Journal *of Experimental Psychology-General* **124**(4): 391-408.

Smith, J. D., W. E. Shields, et al. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology-General* ***127***(3): 227-250.

Son, L. K. and N. Kornell (2005). Metaconfidence judgments in rhesus macaques: Explicit versus Implicit mechanisms. The missing link in cognition: Origins of self-reflective consciousness. H. S. Terrace and J. Metcalfe. Oxford, Oxford University Press.

Son, L. K., B. L. Schwartz, et al. (2003). Implicit metacognition, explicit uncertainty, and the monitoring/control distinction in animal metacognition. *Behavioral and Brain Sciences* ***26****(3):* 355-+.

Squire, L. R. and S. Zola-Morgan (1991). The Medial Temporal-Lobe Memory System. *Science* ***253***(5026): 1380-1386.

Suda-King, C. (2008). Do orangutans (Pongo pygmaeus) know when they do not remember? *Animal Cognition* ***11****(1)*: 21-42.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? [Review]. *Behavioral and Brain Sciences, 30*(3), 299-+.

Tomasello, M. and J. Call (1997). Primate cognition. New York, Oxford University Press.

Trepel, C., Fox, C. R., & Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. [Review]. *Cognitive Brain Research, 23*(1), 34-50.

Washburn, D. A., J. D. Smith, et al. (2006). Rhesus monkeys (Macaca mulatta) immediately generalize the uncertain response. *Journal of Experimental Psychology-Animal Behavior Processes **32**(2)*: 185-189.

Weiskrantz, L. (2001). Commentary responses and conscious awareness in humans: The implications for awareness in non-human animals. *Animal Welfare, 10*, S41-S46.

Sands, S. F., & Wright, A. A. (1980). Serial probe recognition performance by a rhesus monkey and a human with 10-item and 20-item lists. [Article]. *Journal of Experimental Psychology-Animal Behavior Processes, 6*(4), 386-396.

Appendix A

Familiarization training phase 4 & 5: additional analyses

The mean accuracy in phase 5 when the decline response was available is depicted in Figure A1. Performance increased significantly from 55%, SEM= 1.53 in phase 4, to 67%, SEM=4.53 in phase 5 when the decline-test response was introduced (paired, one-tailed t-test, t_5=-2.43, p=.0295). On average, the monkeys used the decline response on 19% of the trials, SEM=4.9%. Accuracy may have increased between phase 4 and phase 5 because monkeys used the decline response adaptively, declining tests when they did not know the reward location. However, it is also possible that with additional experience on the task, learning alone could account for this increase in performance. Because a direct comparison between forced and chosen trials was needed, free and forced trials were intermixed in Experiment 1.

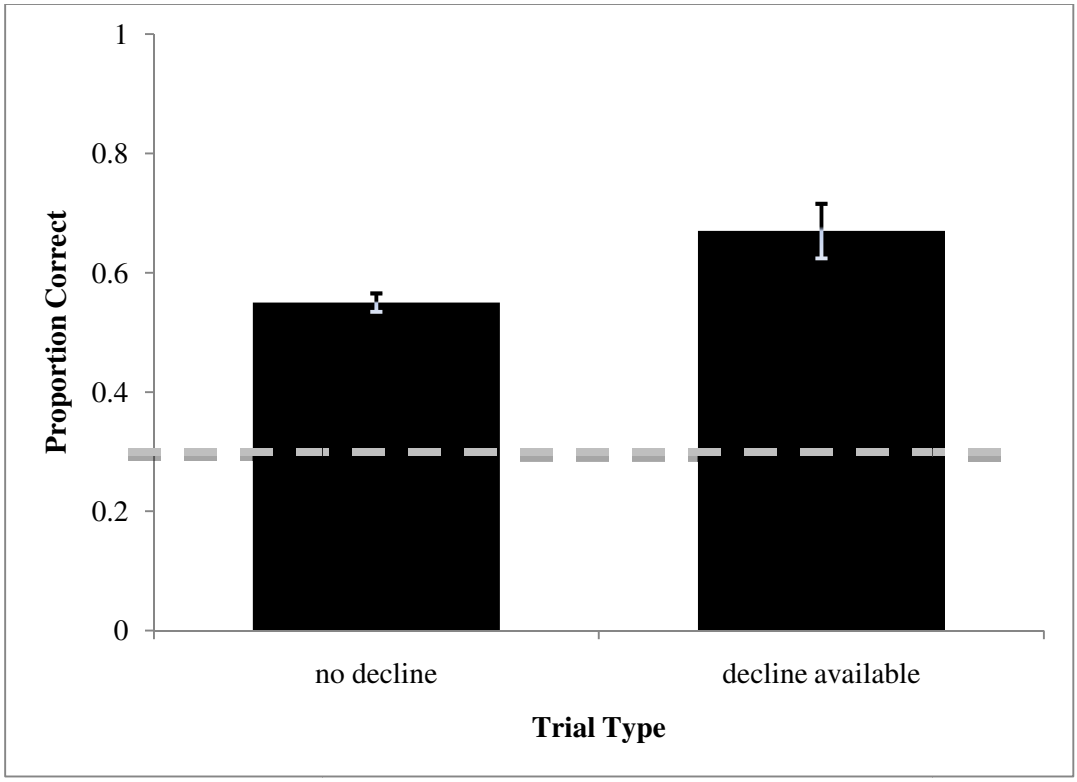

Figure A1. Average accuracy on forced (no decline, Phase 4) and freely chosen (decline available, phase 5) trials in familiarization training. The far left bar represents familiarization training phase 4 at the fixed criterion delay (10 sessions) when no decline response was available. The second bar represents familiarization phase 5 when the decline response was introduced (20 sessions). Chance, 25%, is depicted by the dashed line.

Appendix B

Individual Analyses

Each subjects' criterion delays (CDs) were titrated in familiarization training such that accuracy on forced trials was between 40-70% correct. Using individually titrated delays for each monkey ensured that all subjects experienced trials in which they forgot and trials in which they remembered. It also assured that task difficulty was about the same for each subject so that monkeys' performances could be analyzed as a group. A potential concern, however, is that there might be differences in the use of the decline response and thus metamemory performance depending on a subject's particular CD. Specifically, because the range of CDs used is so large (6- 48 seconds), perhaps it is the monkeys with longer CD's (who are particularly high performers on the task) that account for the positive metamemory results in Experiments 1-4. To determine whether the monkeys with longer lasting memory were also better in using the decline test response I report the criterion delay for each monkey and the difference in accuracy on chosen and forced trials in Table B1. In order to see if there was a correlation between longer CDs and better metacognitive performance (i.e. discrepancy between forced vs. chosen accuracy in Experiment 1, a regression analyses was performed. No significant correlation between CD and difference in accuracy between forced and chosen trials was found (p=.399, R^2=.227; Figure B1). Thus it does not appear that differences in memory duration predict differences in metamemory performance.

Table B1. Final criterion delay used for each monkey is indicated in the first row. Subjects are grouped from left to right according to CD. Accuracies and percentages of tests declined on each type of trial presented in Experiments 1-4 are shown. Mean accuracies and percentages of tests declined are indicated in the far right column. CDs and particularly relevant performance comparisons within each experiment are indicated with incremental shades of gray: values for subjects with 6 second delays (M2 and M3) are indicated with the lightest shade of gray; values for subjects with 12 second delays (M1 and M4) are indicated with the middle shade of gray; values for subjects with 48 second delays are indicated with the darkest shade of gray.

Monkey		M2	M3	M1	M4	M5	M6	Mean
	Criterion delay	**6 s**	**6 s**	**12 s**	**12 s**	**48 s**	**48 s**	
Experiment 1	Forced trial accuracy	64%	59%	74%	81%	46%	42%	61%
	Chosen trial accuracy	83%	84%	85%	89%	69%	66%	80%
	Difference between forced and chosen trial accuracy	18%	25%	11%	8%	24%	24%	19%
Experiment 2	Normal trials declined	27%	43%	37%	22%	20%	38%	31%
	No sample probe trials declined	83%	100%	97%	97%	97%	100%	96%
	Difference between normal and no sample trials declined	56%	57%	60%	75%	77%	62%	65%
Experiment 3	Forced short delay accuracy	93%	88%	90%	95%	98%	90%	92%
	Chosen short delay accuracy	97%	100%	97%	99%	99%	94%	98%
	Short delay trials declined	11%	1%	8%	6%	14%	13%	9%
	Forced medium delay accuracy	63%	83%	80%	85%	73%	48%	72%
	Chosen medium delay accuracy	71%	92%	89%	84%	80%	82%	83%
	Medium delay trials declined	30%	24%	29%	29%	49%	73%	39%
	Forced long delay accuracy	10%	40%	33%	25%	38%	28%	29%
	Chosen long delay accuracy	100%	38%	40%	50%	100%	56%	64%
	Long delay trials declined	99%	90%	88%	90%	89%	89%	91%
	Difference between short and long delay trials declined	88%	89%	80%	84%	75%	76%	82%
Experiment 4	Forced single sample accuracy	55%	63%	45%	58%	60%	45%	54%
	Chosen single sample accuracy	85%	100%	86%	79%	85%	22%	76%
	Single sample trials declined	58%	44%	64%	53%	50%	66%	56%

Forced double sample probe accuracy	90%	85%	80%	80%	40%	55%	72%
Chosen double sample probe accuracy	94%	100%	96%	73%	76%	83%	87%
Double sample trials declined	23%	38%	40%	8%	48%	40%	33%
Difference between single and double trials declined	35%	6%	24%	45%	3%	26%	23%

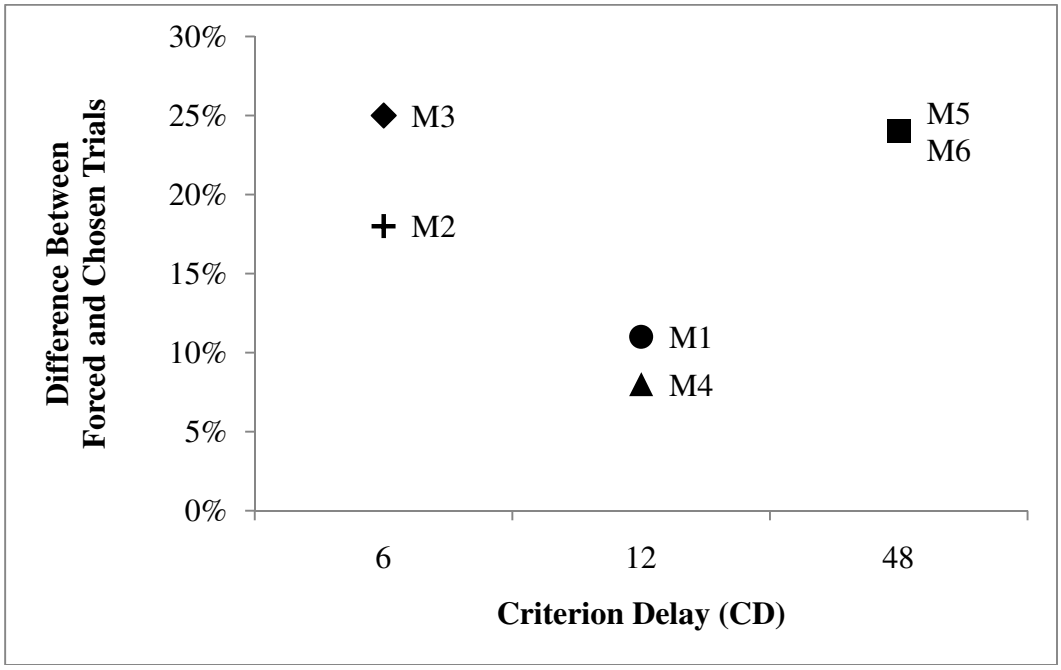

Figure B1. Difference in accuracies on chosen minus forced trials in Experiment 1. The criterion delay each monkey was tested with is indicated on the x-axis. Each monkey's difference in accuracy is plotted and labeled. Only one symbol can be seen at the 48 second delay because the performances' of monkeys M5 and M6 produced the same difference between chosen and forced scores.

Appendix C

Cup preference analyses from Experiment 1

During testing, subjects appeared more likely to find the correct cup location when one of the front cups was baited as compared to one of the front two cups. This could potentially be problematic because subjects could have used the location of baiting, rather than the quality of memory, to cue use of the decline response. In order to determine whether the decline response was used more frequently following baiting of

particular cups in Experiment 1, the four cup locations were divided into front (cups 1&2) and back (cups 3&4). As reported in Experiment 1, subjects both 1) performed better on front cup trials as compared to back cups trials and 2) declined back cup trials significantly more than front cup trials (paired sample t-tests; forced front vs. back cups trials: t_5=3.87, p=.012; declined front cup trials vs. back cup trials: t_5=-7.185, p=.001). There was also no significant difference between performance on forced trials as compared to chosen trials for both front and back cups (paired-sample, two-tailed t-tests; front cup trials: t_5=.603, p=.573; back cup trials: t_5=-1.44, p=.211).

While analyzing front vs. back cups are crucial analyses, one also might wonder how subjects performed at one individual cup location and how these accuracies might compare to performance an another specific cup location. The following results and Tables C1 and C2 therefore analyze and compare performance at each cup.

Dividing the cup locations in back and front cups in the results of Experiment 1 was a legitimate grouping of cup location because monkeys performed similarly when cup 1 and cup 2 were baited (front cup trials) and when 3 and cup 4 were baited (back cup trials; Figure C1).No significant differences were found between the performances on: 1) cup 1 and cup 2 forced trials, and 2) cup 3 and cup 4 forced trials ($p > .05$; Table C1.) Conversely, when comparing all other cup locations, there was a significant difference in performance on forced trials ($p < .05$; except for cup 1 vs. 3; Table C1). In a paired-sample two-tailed t-test, performance on front cup trials significantly differed from performance on back cup trials (t_5=3.87, p=.012). This indicates that it was in fact easier for subjects to answer correctly when the baited cup was in closer proximity to them. This reinforces the possibility that subjects simply declined back cup trials more

frequently than front cup trials. Table C2 shows differences between forced and chosen trials at each cup location. The table shows that only one cup location, cup 1, revealed a significant difference between forced and chosen trials ($p<.05$).

The proportion of declined trials was analyzed for each cup combination in the last section of Table C1. Not surprisingly, there was no significant difference between cup 1 vs. 2 and cup 3 vs. 4, but all other cup combinations did reveal significant differences ($p< .05$). These results reveal that monkeys were less likely to decline cups 1 and 2 (front cups) and were more likely to decline cups 3 and 4 (back cups). Monkeys are biased to decline back cups more than front cups. This does not disprove the idea that the monkeys may also being using memory monitoring to guide their decisions. Much like subjects are more accurate when tested at shorter delays, and worse when tested at longer delays, monkeys are also more accurate when a cup closer to them is baited than when a cup further from them is baited. Monkeys therefore declined tests in accordance with test difficultly. While this modulation of declining trials could be a result of metamemory, it is also possible that monkeys' ability to identify cup locations could account for the difference in performance on forced vs. chosen trials in Experiment 1. This concern is one of the main motives for testing subjects on further controls in the experiments following Experiment 1. There is no way that performance levels at each cup location and resulting decline biases could account for the results found in Experiments 2-4.

Table C1. Comparisons of performance on individual cups. See Figure C1 for location of cups (see also Figure 2; *Apparatus section*). Data were derived from all trials subjects received in Experiment 1. In two-tailed, paired sample t-tests, the difference between performance on two cups were compared (all combinations were analyzed, without correction for multiple comparisons.) In the right section of the table the difference between performances on forced vs. chosen trials for each cup location were compared in

two-tailed, paired sample t-tests. Rows represent one cup comparison such that significant (and non-significant) accuracies on forced trials can be compared to the corresponding declined trials on the right. Values shaded gray indicate compared cups that share the same front (1 vs. 2) or back location (3 vs. 4). Non-shaded values indicate that the cups compared were front vs. back locations.

Trial type	cups compared	t	p	Trial type	cups compared	t	p
forced trials, *df=5*				declined trials, *df=5*			
	1 vs. 2	2.256	0.074		1 vs. 2	0.271	0.797
	3 vs. 4	0.907	0.406		3 vs. 4	0.888	0.415
	1 vs. 3	2.23	0.076		1 vs. 3	9.688	0.00
	1 vs. 4	3.579	0.016		1 vs. 4	5.876	0.002
	2 vs. 3	3.594	0.016		2 vs. 3	5.788	0.005
	2 vs. 4	5.012	0.004		2 vs.4	4.377	0.007

Table C2. Individual cup analysis in Experiment 1. The difference in performance at each individual cup on forced and chosen tests is shown. In two-tailed, paired sample t-tests, the difference between performance on forced and chosen trials at each cup location was analyzed.

Trial type	cup analyzed	t	p
forced vs. choice, *df=5*			
	1	-2.941	0.032
	2	-2.162	0.083
	3	-1.668	0.156
	4	-0.356	0.737

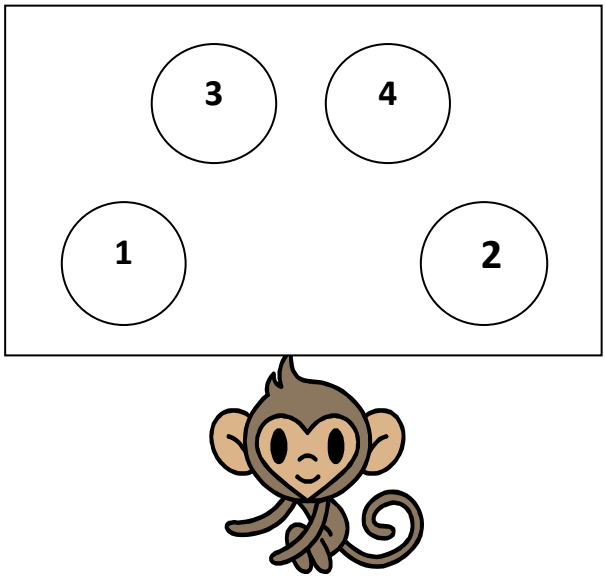

Figure C1. Schematic depiction of the testing tray with each cup location represented by a number label. Location of subject is also shown. Front cups (1&2) are represented in the front row (located closest to the subject). The back cups (3&4) are shown are shown in the back row, which are located slightly further from the subject. See also Figure 2; *Apparatus section*.