NIHMS PDF Receipt

Journal title:	Animal cognition
Manuscript title:	Dissociation of memory signals for metamemory in rhesus monkeys (Macaca mulatta)
Submitter:	Springer Publishing
NIHMSID:	1521709

Manuscript Files

Type	Label	File Name	Upload Date
Manuscript		10071_2019_1246_ReferencePDF.pdf	2019-02-14 12:34:40

Running Head: Monkeys monitor multiple signals

Dissociation of memory signals for metamemory in rhesus monkeys (*Macaca mulatta*)

Emily Kathryn Brown[1], Benjamin M. Basile[2], Victoria L. Templer[3], and Robert R. Hampton[1]

[1] Department of Psychology and Yerkes National Primate Research Center, Emory University

[2] Laboratory of Neuropsychology, NIMH, NIH

[3] Department of Psychology, Providence College

Conflict of Interest: All authors declare no conflict of interest.

Acknowledgements: We thank Steven L. Sherrin, Jessica A. Joiner, and Tara A. Dove-VanWormer for assistance with testing animals.

Correspondence concerning this article should be addressed to Emily Brown, Department of Psychology, 36 Eagle Row, Atlanta, GA, 30322. Phone: 404.727.9619. Email: emily.brown@emory.edu

Keywords: metacognition, information-seeking, monitoring, working memory, familiarity

Abstract

Some nonhuman species demonstrate metamemory, the ability to monitor and control memory. Here, we identify memory signals that control metamemory judgments in rhesus monkeys by directly comparing performance in two metamemory paradigms while holding the availability of one memory signal constant and manipulating another. Monkeys performed a four-choice match-to-sample memory task. In Experiment 1, monkeys could decline memory tests on some trials for a small, guaranteed reward. In Experiment 2, monkeys could re-view the sample on some trials. In both experiments, monkeys improved accuracy by selectively declining tests or re-viewing samples when memory was poor. To assess the degree to which different memory signals made independent contributions to the metamemory judgement, we made the *decline-test* or *review-sample* response available either prospectively, before the test, or concurrently with test stimuli. Prospective metamemory judgements are likely controlled by the current contents of working memory, whereas concurrent metamemory judgements may also be controlled by additional relative familiarity signals evoked by the sight of the test stimuli. In both paradigms, metacognitive responding enhanced accuracy more on concurrent than on prospective tests, suggesting additive contributions of working memory and stimulus-evoked familiarity. Consistent with the hypothesis that working memory and stimulus-evoked familiarity both control metamemory judgments when available, metacognitive choice latencies were longer in the concurrent condition, when both were available. Together, these data demonstrate that multiple memory signals can additively control metacognitive judgements in monkeys and provide a framework for mapping the interaction of explicit memory signals in primate memory.

Introduction

Performance on memory tasks is supported by a variety of memory systems, each characterized by distinct functional properties and underlying neural substrates (Sherry and Schacter 1987; Squire and Zola-Morgan 1991; Yonelinas 2002). For example, a contestant on a trivia show may answer based on a combination of retrieval of the correct answer and a vague sense of which choice feels most familiar. The memory systems giving rise to these signals differ in flexibility, robustness against interference, and the conditions under which they are available for introspective monitoring. Monitorable memory signals have gained particular attention because they afford subjects the opportunity to control their cognition, such as by seeking additional information or altering response strategy. The way in which subjects monitor and control different memory signals can be assessed using metamemory paradigms.

Metamemory is the ability to monitor memory processes and adapt behavior or cognition in accord with monitored memory signals. Metamemory is exemplified by the game show “Who wants to be a millionaire?” wherein contestants must answer trivia questions and make judgments about whether they have answered correctly. Contestants are given the opportunity to reconsider their answers after reflection, or to collect additional information by contacting a friend. The metamemory judgments contestants make can be controlled by whether a memory was successfully retrieved, has been successfully held in working memory, or evokes strong relative familiarity compared to other test options (Flavell 1979; Kornell 2013; Nelson 1996). Just as memory performance often reflects a combination of memory signals, so too do metamemory judgements.

A substantial body of work with nonhuman primate species indicates that they can monitor memory, as indicated by their ability to selectively decline difficult tests, re-view previously studied but forgotten information, seek information when ignorant, or adaptively wager rewards based on recent test choices (Basile et al. 2015; Brown et al. 2017; Hampton 2001; Kornell et al. 2007; Templer and Hampton 2012; Washburn et al. 2010). Only recently has experimental work been directed at determining which specific memory systems are subject to memory monitoring and how different memory signals interact to control metamemory judgments in nonhumans (e.g., Coutinho et al. 2015; Smith et al. 2013; Takagi and Fujita 2018).

Because memory systems provide different types of information, and multiple systems contribute to performance on memory tasks, metamemory is likely controlled by a combination of memory signals. The way that these signals are weighted in metamemory decisions in monkeys is unknown. One possiblity is that in the presence of multiple memory signals, metacognitive choice may be guided by only one, such as the strongest signal. A second possibility is that multiple memory signals contribute additively to metamemory judgments, such that congruent positive signals strengthen the likelihood of a high-confidence metamemory judgment.

To evaluate the degree to which monkey metamemory is controlled by multiple memory signals, we compared metamemory judgments under conditions that always allowed for the use of one memory signal but systematically manipulated the availability of a second memory signal. In nonhuman primates, there is strong evidence that recognition memory performance is supported by at least two types of memory: working memory and relative familiarity (Basile and Hampton 2013; Wittig et al. 2016; Wittig and Richmond 2014). Working memory is a limited-capacity system that allows the active, relatively short-term maintenance and manipulation of information (Baddeley 2000; Baddeley 2003). Familiarity is a passive signal evoked by the re-presentation of a stimulus that has been seen previously (Yonelinas 2002; Yonelinas et al. 2010). In matching to sample memory tests, the image seen as a sample on the current trial can potentially be held in working memory during the delay. The sample image is also expected to be relatively more familiar than the distractors presented with it at test, because the sample is the image that has been seen most recently.

The degree to which working memory and familiarity are available for metamemory judgments can be experimentally manipulated. Prospective metamemory judgments are made before the test options have been seen and thus favor monitoring of working memory because information about the relative familiarity of the test options is not yet available. Concurrent metamemory judgments are made in the presence of the memory test. Monkeys can still monitor working memory in concurrent metamemory judgments, and additional information about the relative familiarity of the sample and distractors is also available.

Here, we assessed both prospective and concurrent metamemory judgments. In Experiment 1 we used a decline-test paradigm, in which monkeys chose to avoid some tests for a small, but guaranteed, reward (Fujita 2009; Hampton 2001; Suda-King 2008; Suda-King et al. 2013; Templer and Hampton 2012; Templer et al. 2017; Washburn et al. 2010). In Experiment 2, we used an information-seeking paradigm, in which monkeys chose to re-view the sample on some trials (Basile et al. 2009; Basile et al. 2015; Beran and Smith 2011; Call and Carpenter 2001; Castro and Wasserman 2013;

Iwasaki et al. 2013; Kirk et al. 2014; Kornell et al. 2007; Marsh 2014; Marsh and MacDonald 2012; McMahon et al. 2010; Vining and Marsh 2015; Watanabe and Clayton 2016). Using these two paradigms to assess metacognition, presented both prospectively and concurrently, provides a powerful test of the generalizability of our findings.

To compare the contributions of working memory and familiarity in both decline-test and information-seeking paradigms, we required a common metric. In prior work with the decline-test paradigm, monkeys showed a performance advantage on chosen tests compared to performance on forced tests when there was no option to decline the memory test (Brown et al. 2017; Hampton 2001; Templer and Hampton 2012). When monkeys monitored memory to choose when to use the *decline-test* response, they disproportionately chose to take trials when memory was strong and selectively avoided tests on which memory was weak, resulting in higher accuracy on tests they chose to take. Forced trials included both trials on which memory was strong, which would have been chosen had that option been available, and trials on which memory was weak, which may have been declined, had the option been available. This accuracy benefit can be applied equivalently to both decline-test and information-seeking paradigms. On choice trials in the information-seeking paradigm, subjects can either choose to take the test immediately, if memory is strong, or first re-view the sample, if memory is weak. Forced trials, when the option to re-view the sample is omitted, will include some tests that would have been taken immediately and some tests when subjects would have first re-viewed the sample, had the option been available. Thus, adaptive use of the *re-view sample* response in the information-seeking paradigm should result in a performance advantage on chosen trials over forced trials. If both working memory and familiarity signals are monitorable, we should see the performance advantage for chosen over forced trials in both prospective and concurrent choice conditions of the decline-test and information-seeking paradigms.

If both working memory and familiarity contribute to metamemory performance, then the benefit resulting from use of both the *decline-test* and *review-sample* options will be larger in the concurrent than the prospective condition. This is because both working memory and the familiarity evoked by the test stimuli may jointly guide concurrent metacognitive choices, but only working memory can guide prospective choices.

EXPERIMENT 1 – DECLINE-TEST PARADIGM

Methods

Subjects

Subjects were 8 pair-housed male rhesus macaque monkeys (*Macaca mulatta*), with a mean age of 7 years at the beginning of these studies. Three subjects had previous experience with a manual metacognition task (Templer and Hampton 2012). All subjects had prior training with the concurrent metamemory version of the information-seeking paradigm (Basile et al. 2015). All subjects also had prior training and generalization tests with the decline-test paradigm across a variety of perceptual discriminations, as well as memory tests conducted across a range of retention intervals (Brown et al. 2017).

Apparatus

We tested monkeys in their home cages, using portable touch-screen computer rigs consisting of a laptop computer (Dell, Round Rock, TX) with generic speakers, a 15” color LCD touchscreen (ELO, Menlo Park, CA), and two automated food dispensers (Med Associates Inc., St. Albans, VT) that dispensed into food cups beneath the screen. Food reinforcement consisted of 94 or 97 mg nutritionally complete primate pellets (Bio-Serv, Frenchtown, NJ and Purina TestDiet, Richmond, IN). Calories from pellets earned during testing were subtracted from monkeys’ daily primate biscuit chow rations, such that monkeys consumed the same number of calories daily, regardless of testing performance. Daily calorie budgets were established by veterinary staff based on weight trajectories and clinical assessments. Monkeys had *ad libitum* access to water. We presented stimuli and collected responses using programs written in Presentation (Neurobehavioral Systems, Albany, CA).

Procedure

Monkey housing and testing conditions

During testing, paired monkeys were separated by dividers that allowed visual and physical contact through large slots, but prevented access to adjacent testing equipment. Monkeys had *ad libitum* access to their testing rigs up to seven hours per day.

Monkeys completed trials of a four-choice delayed-matching-to-sample (DMTS) task with a metacognitive choice phase (Figure 1). Four clipart images were used across all sessions, such that every image was seen at test on every trial. All responses required two touches (FR2) to prevent recording undirected contacts with the touchscreen as responses. To start a trial, monkeys touched a green ready square at the bottom center of the screen. A sample image then appeared in the center of the screen. Touches to the image resulted in a blank screen for a retention interval ranging from 4 to 28 seconds, depending on the individual monkey.

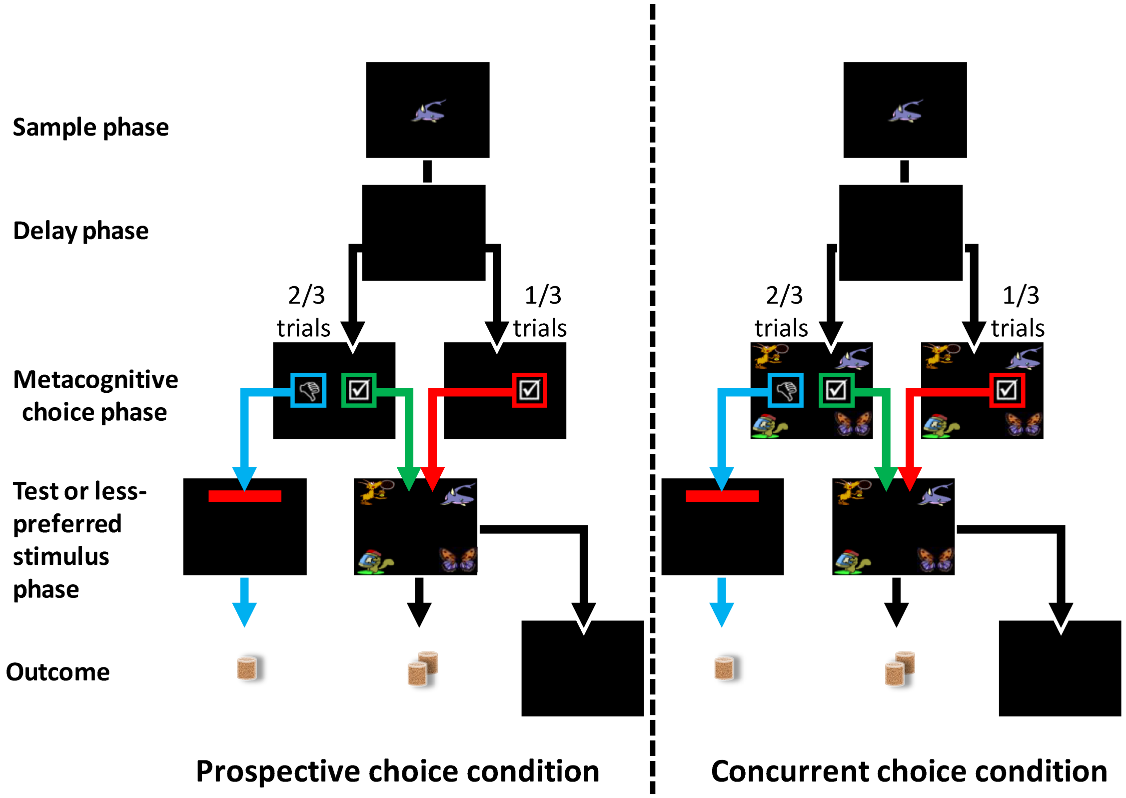


Fig. 1 Trial progression in the decline-test task. Monkeys touched a green ready square to initiate each trial (not shown). A sample clipart image then appeared on

screen. Monkeys touched the sample image to advance the trial and initiate the delay interval. At the end of the delay, metacognitive choice images appeared. On *prospective choice* sessions (left), metacognitive choice stimuli appeared before presentation of the test images. On *concurrent choice* sessions (right), metacognitive choice images appeared at the same time as the test stimuli. In concurrent choices, test stimuli were unresponsive to touches until after the metacognitive choice was made. On 2/3 of trials, the *accept-test* and *decline-test* choice stimuli appear together. On 1/3 of trials, the *decline-test* choice did not appear. Selection of the *accept-test* stimulus extinguished choice stimuli and activated test stimuli. Correct choices resulted in food reinforcement of two pellets; incorrect choices resulted in a black time out screen. Selection of the *decline-test* response caused the *guaranteed small reward* stimulus screen to appear. Touches to this stimulus resulted in guaranteed food reinforcement of one pellet.

Metacognitive choice stimuli appeared after the delay, which allowed monkeys to take the DMTS test for a large reward if correct or avoid the test for a small but guaranteed reward. The metacognitive choice phase consisted of two black and white clipart choice stimuli, which could appear concurrently, at the same time as the test stimuli, or prospectively, before the test stimuli (Figure 1). The *accept-test* stimulus, a check-marked square, was vertically centered on the right side of the screen. Touches to the *accept-test* stimulus extinguished metacognitive choice stimuli and made the test stimuli responsive to touch. Selection of the target image seen at study resulted in a distinctive auditory signal and two food pellets. Selection of a distracter resulted in auditory feedback and black screen for a brief timeout period. The *decline-test* stimulus, a thumbs-down, was vertically centered on the left side of the screen. Selection of the *decline-test* stimulus resulted in the immediate presentation of a red bar at the top center of the screen. Touches to this *guaranteed small reward* stimulus resulted in a distinctive auditory signal and one food pellet.

To ensure that monkeys declined some, but not all trials, we titrated the number of touches required to obtain the guaranteed small reward after each session. The number of touches was increased or decreased by two if the overall decline rate was greater than 70% or less than 30%, respectively, with the minimum possible touches being two. Our titration of the number of responses required for the guaranteed small reward took place between sessions and thus affected the overall rate of use of the decline-test response within a session but did not differ on a trial-by-trial basis. Thus, titrating this response kept behavior in a range that permitted detection of differences in the use of the *decline-test* response, but could not create such differences.

On 2/3 of trials, monkeys were presented with both metacognitive choice stimuli. On the other 1/3 of trials, only the *accept-test* stimulus was presented, forcing subjects to take the test. Each session consisted of 120 trials, with trial types pseudorandomly intermixed, such that each session contained 80 choice trials and 40 forced trials.

Prospective choice sessions, in which the metacognitive choice stimuli were presented before the test options, were alternated with concurrent choice sessions, in which the metacognitive choice stimuli were presented at the same time as the test options. Prospective and concurrent choice trials were not intermixed within a single session. Monkeys completed 10 sessions of each trial type.

Data analysis

All proportions were arcsine transformed before statistical analysis to better approximate the normality assumption underlying parametric statistics (Keppel and Wickens 2004, p. 155). Geisser–Greenhouse correction was used, and appropriately adjusted degrees of freedom reported, whenever the sphericity assumption was violated (Keppel and Wickens 2004, p. 378).

For all experiments, we assessed accuracy by calculating the proportion correct on forced trials and on trials that monkeys chose to take, without declining or re-viewing the sample, when they had the option. We assessed the interaction between trial type (forced, chosen) and timing of the metacognitive judgement (prospective, concurrent) using a repeated measures ANOVA. We used follow-up planned paired t-tests to compare accuracy on forced and chosen trials.

Results and discussion

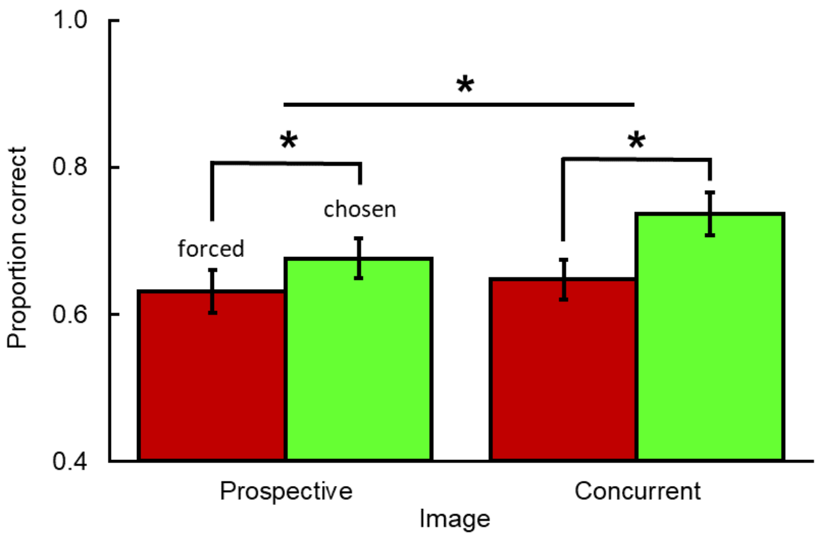


Fig. 2 Monkeys were more accurate on chosen tests than forced tests and this benefit was larger when the metamemory judgment was made concurrently rather than prospectively. Displayed are mean group accuracy (± SEM) as a function of whether the monkeys were forced to take the test (dark red) or chose to take the test (light green) and whether the metamemory judgment was made prospective to the test (left) or concurrently with the test (right). * = $p < .05$ for ANOVA interaction and follow-up t-tests that compared forced and chosen performance in the prospective and concurrent conditions, respectively.

Monkeys improved accuracy when the *decline-test* option was available, and did so to a greater extent in concurrent compared to prospective judgements. Monkeys were more accurate on chosen tests than they were on forced tests (Figure 2; main effect of forced or chosen: $F_{(1,7)} = 39.915$, $P < .001$, partial $\eta^2 = .851$). This benefit was significant on both concurrent and prospective tests when each comparison was analyzed separately (prospective: $t_{(7)}=-4.468$, $P=.003$, d=1.580; concurrent: $t_{(7)}=-6.458$, $P<.001$, d= 2.283). Further, there was a significant interaction between trial type (forced or chosen) and the time of the metacognitive choice (prospective or concurrent), such that the benefit for chosen test accuracy was greater in the concurrent than the prospective condition ($F_{(1,7)} = 17.025$, $P = .004$, partial $\eta^2 = .709$). The difference in forced test accuracy across conditions was not significant ($t_{(7)}=.868$, $P=.414$). The greater benefit of choosing to take the test on concurrent choices than prospective choices is consistent with the hypothesis that additional information controlled metamemory judgments in the concurrent condition, and that this information was provided by the sight of the test items. It is likely that prospective metamemory judgments are controlled by monitoring of working memory, whereas the concurrent metamemory judgements are controlled both by monitoring working memory and familiarity evoked by presentation of the test images.

EXPERIMENT 2 – INFORMATION-SEEKING PARADIGM

Test accuracy on chosen and forced trials in Experiment 1 provided evidence that more or better information is available for metacognitive judgments made concurrently with memory tests than is available when judgments are made prospectively, before presentation of the tests. This benefit is consistent with the hypothesis that multiple memory signals control metamemory judgments in a manner that is independent and additive. In Experiment 2, we used a second established metamemory paradigm, information-seeking, to further test whether multiple memory systems control metamemory judgments and to evaluate the generalizability of our finding. In Experiment 2, adaptive metacognitive responding would result in the choice to take tests immediately when memory is relatively strong, and re-view the sample prior to taking the test on trials when memory is relatively weak. We hypothesized that if the *decline-test* and *review-sample* responses are controlled by similar underlying mechanisms, monkeys will show a benefit on trials they choose to take immediately over forced trials. Additionally, if working memory and familiarity signals contribute to performance on these tasks, then we expect that the accuracy benefit of choosing which tests to complete immediately will be larger on concurrent choices, in which familiarity signals resulting from the appearance of test stimuli can additionally guide metacognitive choices, as compared with prospective choices, which are completed before familiarity signals are available.

Methods

Subjects and apparatus

Experiment 2 used all monkeys from Experiment 1 and three additional monkeys (male, full group mean age = 6) that had undergone the training described in Basile et al., (2015). Monkeys were housed in the same conditions, and tested on the same apparatus, as described in Experiment 1.

Procedure

Specifics of the information-seeking paradigm have been published in detail (Basile et al., 2015). Briefly, monkeys studied either a spatial location that could occupy any one of the four corners of the screen (Basile et al., 2015, Experiments 1 and 5) or a color photograph (Basile et al., 2015, Experiment 6; see also Figure 3), and had the option to either proceed directly to the memory test or re-view the sample. In the image condition, the same four photographs were used across all sessions, such that every image was seen at test on every trial. In spatial tests, the same four locations were similarly used on all trials. We included the spatial condition because the “tubes task” that this information-seeking paradigm is based on was originally a spatial task and so including the spatial condition fulfills the secondary objective of this study: to evaluate the degree to which the different common metamemory paradigms produce similar results when compared directly. As in Experiment 1, we tested monkeys with a concurrent metacognitive choice, in which the test options were visible while choosing whether to re-view the sample, and with a prospective metacognitive choice, in which the test options were not presented until after the monkey chose whether to re-view the sample (Figure 3).

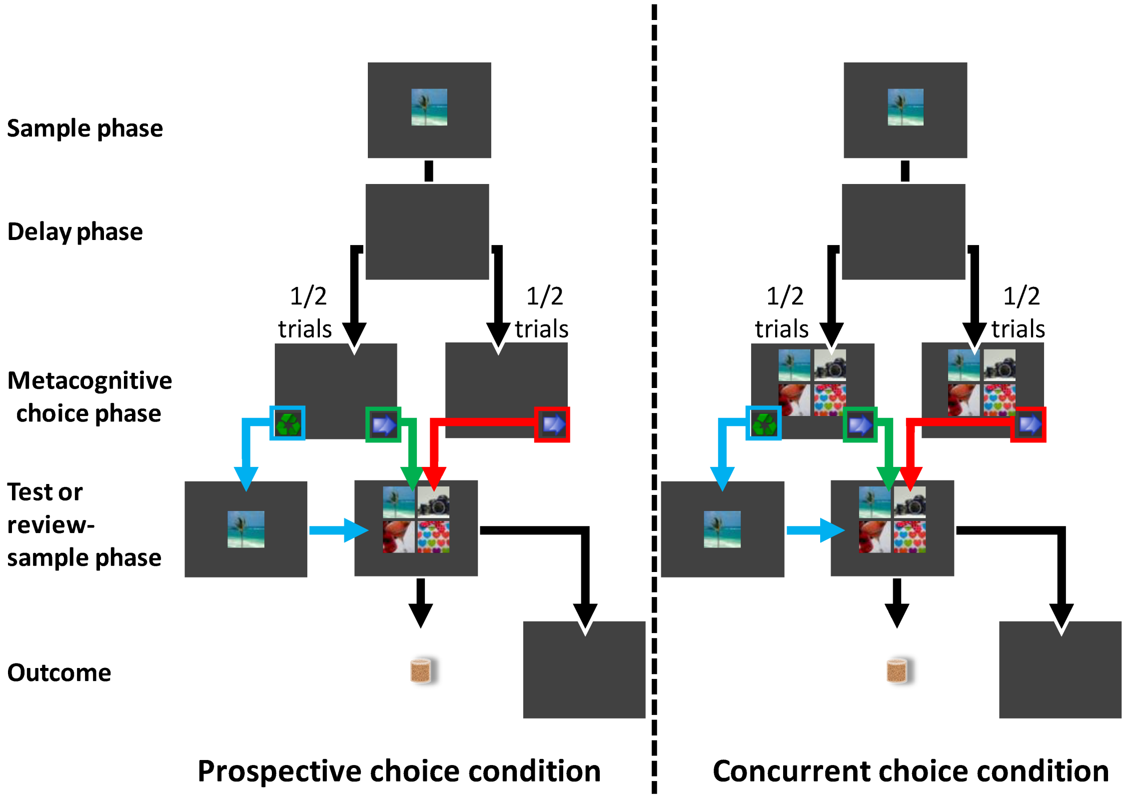


Fig. 3 Progression of trials of the information-seeking task. Monkeys touched the green ready square to initiate trials (not shown). A sample (image condition pictured) then appeared on screen for 200ms and then disappeared. At the end of the programmed delay, metacognitive choice images appeared. On *prospective choice* sessions (left), the metacognitive choice stimuli appeared before the test options. On *concurrent choice* sessions (right), the metacognitive choice images appeared concurrently with test stimuli. On 1/2 of trials, the *accept-test* and *review-sample* choice stimuli appear together. On 1/2 of trials, the *review-sample* response did not appear. Choice of the *accept-test* stimulus caused the metacognitive choice stimuli to disappear and the test stimuli to become responsive to touch. Tests resulted in food reinforcement of one pellet (correct) or a blank time out screen (incorrect). Selection of the *review-sample* response caused the sample phase to be shown again. After monkeys

had the opportunity to review the sample, the test was presented, with the outcome contingencies described above. The location memory version was identical except that all stimuli were identical red dots and the monkeys had to remember the screen location of the sample dot.

To ensure that accuracy and metacognitive bias were stable and at appropriate levels to detect metacognition, we titrated the retention interval and the number of touches required to select each monkeys' preferred metacognitive option. Sessions were 80 trials, half choice trials (Figure 3, left) and half forced test trials (Figure 3, right), intermixed and pseudorandomized such that no trial type appeared more than four times in a row. The retention interval was increased by four seconds or decreased by two seconds at the end of each session if accuracy on forced-test trials was above 67.5% or below 57.5%, respectively. Requiring accuracy to be midway between ceiling and chance ensured that monkeys were performing the task correctly and that we could detect any accuracy benefit of chosen trials relative to forced trials. As done previously (Basile et al. 2015), we also required monkeys to use both the *review-sample* and *accept-test* options regularly, with titration as described previously (Basile et al. 2015). Briefly, the number of touches required to select the preferred metacognitive option was increased or decreased by two following every session in which the monkey chose that option on greater than 75% or fewer than 25%, respectively, of all choice trials. For each of the four tasks, when accuracy and metacognitive bias were within those parameters for two consecutive sessions, we considered those stable data to use for analysis. Monkeys were tested until at least 100 chosen trials were available for analysis from stable performance sessions for each task. The total number of trials completed before reaching this criterion depended on how frequently each monkey chose to take tests. Monkeys progressed through the tasks in the following order: spatial prospective, spatial concurrent, image prospective, and image concurrent.

Results and discussion

Final titrated retention intervals ranged from 2 to 32 seconds. All monkeys preferred the *review-sample* over the *accept-test* metacognitive option. Consequently, the *review-sample* metacognitive option required between 6 and 52 touches to select, depending on the strength of the monkey's preference.

As a result of titrating performance, accuracy did not vary as a function of sample type (Figure 4; spatial or image; $F_{(1,10)} = 3.04$, $P = .11$). Additionally, sample type did not interact with any other factor (all Ps > .098). Monkeys were more accurate on chosen tests than they were on forced tests regardless of sample type ($F_{(1, 10)}= 14.63$, $P=.003$, partial $\eta^2 = .594$). Further, as with the decline-test paradigm, there was a significant interaction between trial type (forced or chosen) and the time of the metacognitive choice (concurrent or prospective), such that the benefit to chosen test accuracy was greater in the concurrent than the prospective condition ($F_{(1,10)} = 7.11$, $P = .024$, partial $\eta^2 = .416$). The difference in monkeys' forced test accuracy across all conditions was not significant ($F_{(3,30)} = 1.83$, $P = .16$). This reproduces the main finding from Experiment 1 using *review-sample* in the place of the *decline-test* response. The larger memory benefit of choosing to take the test on concurrent choices than prospective choices with both paradigms provides converging evidence that information from multiple memory systems act additively to control metamemory choices. Because the most obvious difference between the concurrent and prospective choices is the presence of the test options, the improved accuracy likely results from the additional information available from comparing the relative familiarity of the test stimuli, which is available only on concurrent tests.

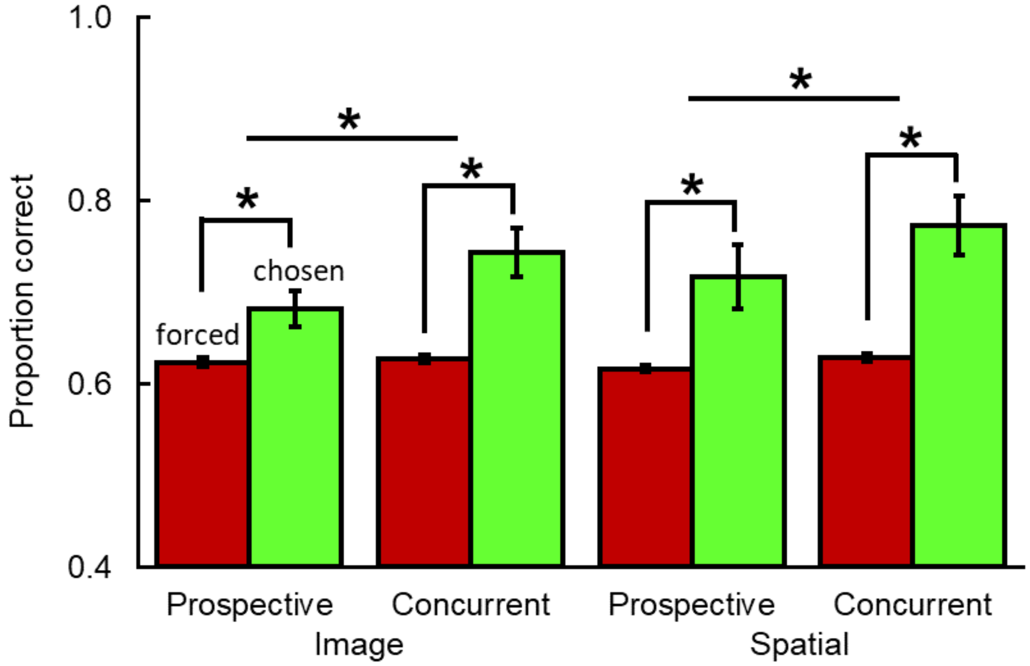


Fig. 4 Monkeys were more accurate on chosen tests than forced tests and this benefit was larger for concurrent than prospective tests. Mean group accuracy (± SEM) as a function of sample type (spatial or image), timing of the metacognitive choice (prospective or concurrent), and whether the monkeys chose to take the test without reviewing the answer or were forced to take the test.

Analysis of Latency Data

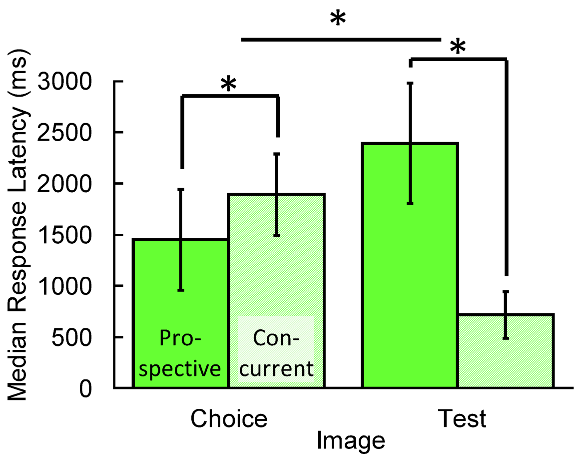
We hypothesized that the results we obtained in Experiments 1 and 2 were due to monitoring working memory in the both conditions, supplemented by additional stimulus-evoked familiarity in the concurrent condition. The analyses of accuracy in the decline-test and information-seeking paradigms were consistent with the hypothesis that metacognitive decisions are additively controlled by multiple memory signals. To further evaluate whether an additional familiarity signal contributed to concurrent metacognitive choices, but did not contribute to prospective judgements, we conducted additional analyses of decision latency. If monkeys do indeed evaluate the additional information from the familiarity evoked by the sight of the test items in concurrent tests, then this should be evident as an increased decision time during the metacognitive choice epoch on concurrent judgments as compared to prospective judgments. Although familiarity is a passive automatic process, monkeys would need the additional time to scan the available test responses and compare the relative familiarity signals. In contrast, if both prospective and concurrent metamemory judgments are controlled solely by working memory, monkeys should be equally quick during the metacognitive choice epoch in both concurrent and prospective tests because working memory is equally available in both conditions.

Data Analysis

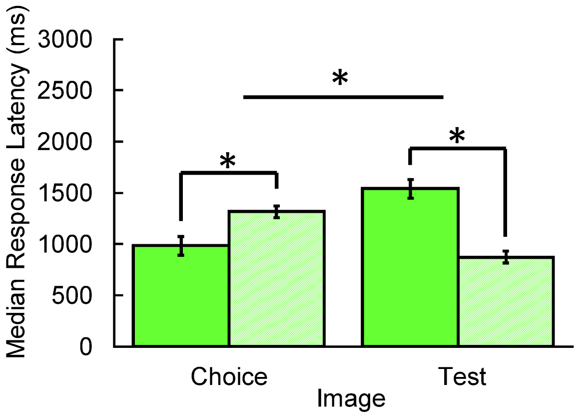
We evaluated the median latency to complete the metacognitive choice and the latency to complete the memory test for all chosen trials, regardless of whether the monkeys selected the correct response at test. Latencies were calculated based on the first touch in the information-seeking paradigm and the required number of touches to make the primary metacognitive choice in the decline-test paradigm was the same across trials. Therefore, titration of metacognitive responding could not have differentially affected latency in the concurrent condition.

Latency Results and Discussion

A) Decline-test, image



B) Information-seeking, image



C) Information-seeking, spatial

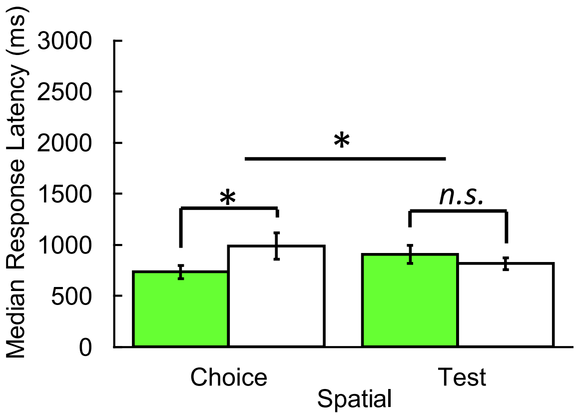


Fig. 5 Monkeys allocated more time to the metacognitive choice in concurrent judgments than in prospective judgments. Panel A depicts decline-test paradigm with image stimuli, panel B depicts information-seeking paradigm with image stimuli, panel C depicts information-seeking paradigm with spatial stimuli. Median response latency in milliseconds (±SEM) as a function of metacognitive choice placement and latency epoch. Metacognitive choice was prospective (solid) or concurrent (striped) with the presentation of the test. Latency epoch was divided into time spent making the metacognitive choice (left) and the test choice (right).

In both paradigms, processing time differed by epoch, as indicated by significant interactions of metacognitive choice placement (prospective or concurrent) with trial epoch (metacognitive choice or memory choice; Figure 5; decline-test paradigm: $F_{(1,7)}$ = 113.227, $P < .001$, partial $\eta^2 = .942$; information-seeking, images: $F_{(1,10)}$ = 99.007, $P < .001$, partial $\eta^2 = .908$; information-seeking, spatial: $F_{(1,10)}$ = 5.696, $P = .038$, partial $\eta^2 = .363$). When making the metacognitive choice to *decline-test* or *review-sample*, monkeys devoted significantly more time to selecting a metacognitive response in the concurrent condition compared with the prospective condition (Figure 5; decline-test paradigm: $t_{(7)} = 10.382$, $P < .001$, d = 3.671; information-seeking, images: $t_{(10)} = 6.899$, $P < .001$, d = 2.080; information-seeking, spatial: $t_{(10)} = 2.425$, $P = .036$, d = .73). When making a memory choice at test, monkeys showed the opposite pattern for both image-memory tests, devoting significantly more time in the prospective condition compared with the concurrent condition (Figure 5; decline-test paradigm: $t_{(7)} = -9.357$, $P < .001$, d = 3.308; information-seeking, images: $t_{(10)} = 7.924$, $P < .001$, d = 2.389), but this difference was not significant for the spatial memory tests of the information-seeking paradigm ($t_{(10)} = 1.793$, $P = .103$, d = .54). Longer time spent on the metacognitive choice epoch is consistent with our hypothesis that information from both working memory and stimulus-evoked familiarity additively control concurrent metacognitive judgments. This is because it should take longer to evaluate two sources of metamemory control than it does to evaluate just a single source of control.

General Discussion

Across the decline-test and information-seeking metamemory paradigms, monkeys were more accurate on trials they chose to take than those they were forced to take, replicating the basic metamemory findings associated with these paradigms. The accuracy benefit on chosen trials was consistently greater, across paradigms, when monkeys made metacognitive judgments in the concurrent condition than it was in the prospective condition. Monkeys also took longer to make concurrent metamemory judgments, when more information was available, than they did to make prospective metamemory judgments. In the prospective condition, the monkeys must make decisions based on the contents of working memory alone; however, in the concurrent condition, monkeys must take additional time to scan the test responses, compare relative familiarity signals, and conceivably check whether the most familiar item is consistent with the item held in working memory. These findings support the hypothesis that there is more mnemonic information available to cue metacognitive judgments in the concurrent condition than the prospective condition.

The most plausible memory signals controlling monkeys' metacognitive choices in this study are working memory and stimulus-evoked familiarity. In both the concurrent and prospective conditions, working memory for the sample is potentially available through the retention interval, and is a signal likely to control metamemory judgments. It is likely that monkeys actively kept the sample image in working memory because we used task parameters (e.g., small image sets and relatively short retention intervals) that have been shown in previous research to promote active working memory (Basile and Hampton 2013). In addition, we manipulated the availability of stimulus-evoked familiarity by manipulating the timing of the metacognitive choice. In the prospective condition, the metacognitive choice took place prior to the appearance of the test. Requiring monkeys to make the metacognitive choice prior to the appearance of the test should encourage reliance primarily on working memory because the test stimuli were not present to evoke familiarity. In the concurrent condition, the metacognitive choice appeared simultaneously with the test, such that monkeys could base metacognitive choices on the relative familiarity of the sample and distractors, in addition to monitoring working memory. Familiarity is most often characterized as a signal automatically evoked by the sight of previously-seen stimuli (Jacoby 1991). Thus, it is reasonable that in the concurrent condition, heightened familiarity for the recentlyviewed sample, taken additively with working memory strength, would increase the accuracy of metamemory judgments. We do note that the familiarity of the most recently seen sample image is probably only slightly greater than the familiarity of the distractor images because all of the images have been seen recently in preceding trials. The extent to which familiarity would control test choice, as well as metacognitive choice, would presumably be much greater if memoranda were trial unique images. It is also possible that other memory signals, as well as other non-mnemonic cues contribute to metamemory performance. Identifying these signals and how they interact to support behavior will be an interesting problem for future research.

The longer metacognitive decision times in concurrent tests is consistent with the hypothesis that when multiple memory signals are present, they control metamemory judgements additively. Monkeys spent more time making the metacognitive choices and used that option to greater benefit when the relative familiarity of the test stimuli was available for evaluation compared to when only working memory was available. This increased decision time likely results from additional evaluation of the relative familiarity signals evoked by the sight of the test items. These signals might reinforce or countermand the contents of the monkey’s working memory. A slightly different, but not incompatible explanation is that monkeys already have a planned test response in mind, and allocation of response time represents a visual search for that planned test response. This visual search would take place prior to the metacognitive choice in the concurrent test and after the metacognitive choice in the prospective choice. An alternative explanation for the effect of concurrent vs. prospective condition on accuracy is that the choice latency in the prospective condition imposes a longer retention interval than monkeys experienced in the concurrent condition. Because prospective choice latencies were so brief relative to the total retention interval, this explanation is unlikely to be correct.

Although monkeys showed a similar pattern of cognitive processing time in the spatial test, the longer metacognitive choice epoch latency for concurrent trials was greatly attenuated and the longer test epoch latency for prospective trials was absent. One explanation for the difference between image and spatial tests is that monkeys devote less processing time to search for the remembered location in spatial tests compared to tests with images. Although the red dots that mark potential response sites are absent in the prospective test, the remembered screen location is likely encoded relative to landmarks on the screen that are always present because the screen itself is present. This means that monkeys can make a decision about where to respond without scanning the whole screen. In contrast, on tests with images, the location of the correct response must be identified before a selection can be made. Although it is less clear what familiarity means in the case of spatial memory, because space is essentially always present in these tests, the intriguing similarities in performance across paradigm and stimulus type warrant further study. The addition of eye-tracking data to this task would provide more evidence to use in identifying the specific search strategy that monkeys employ. For example, monkeys might engage in an exhaustive search prior to metacognitive choice in the concurrent condition with images, but saccade immediately to the intended response in tests of spatial memory.

As in all studies of metacognition, monkeys did not make perfect metamemory judgments (Basile et al. 2015; Brown et al. 2017; Hampton 2001; Templer and Hampton

2012). Across conditions, monkeys never approached perfect accuracy on trials that they chose to take. Although it is tempting to characterize this as unusually poor performance relative to what we feel we might do as humans, humans have not been tested under these conditions on these paradigms. Even in humans, cognition is not always accurate and metacognition is also subject to errors (Maniscalco and Lau 2012; Nelson 1996). It is likely that the cues controlling metamemory responding are subtle and noisy in both species. Metacognitive sensitivity may represent a continuum across taxa, with some species more attuned to the often-subtle cues that control metacognitive judgments. It is likely that metacognitive responding is less precise in macaques than in humans, and that metacognitive signals are less robust in monkeys than in humans. However, humans usually have years of explicit metacognitive training in school and in other settings, so direct comparisons are problematic. There is some evidence that monkey improve metacognitive responding over the course of multiple generalizations (e.g., Brown et al. 2017). It may be worthwhile to explore the degree to which training on metacognition tasks enhances metacognitive sensitivity. We cannot speak to whether monkeys, like humans, have subjective experiences of certainty or uncertainty when making metamemory responses. However, we have here manipulated the information available to control metamemory decisions in our best attempt to understand which memory signals are accessible to monitoring.

Strictly associative accounts have been proposed to explain the results obtained from some nonhuman metacognition paradigms (Carruthers 2008; Jozefowiez et al. 2009; Le Pelley 2012). In response to these criticisms, some researchers have made attempts to obscure the relation between the metacognitive response and primary reinforcement (e.g., Couchman et al. 2010; Smith et al. 2006). Here, we fully acknowledge that the monkeys likely pair specific responses to specific mental states via well-understood associative mechanisms. Manipulating the available information and measuring the change in responding, which is presumably driven by maximization of reinforcement, is what allows us to infer changes in mental state. Thus, to the degree that associative accounts posit that different mental states are the discriminative cues controlling behavior, we agree.

The increased benefit to metamemory judgments under the concurrent condition over the prospective condition is likely the result of additive information from multiple memory signals rather than a shift to basing judgments on different systems in the different conditions. Because the same four stimuli were seen on every trial, all stimuli likely evoked high familiarity at tests. Though it appears that the relative familiarity of the sample still provided a useful memory signal for making accurate metamemory judgments, this signal is likely weak and noisy. The addition of information from familiarity to the information from working memory would provide a modest but reliable benefit to accuracy, as obtained here. It has sometimes been suggested that monkeys use the additional information provided in concurrent test conditions to make metamemory judgments (e.g., Hampton 2009); however, this is the first study to use a direct comparison between prospective and concurrent judgments to provide strong evidence that this is the case. Future studies might utilize more direct manipulation of working memory and familiarity signals, for example, through the manipulation of image set size. A more graded accuracy difference obtained across a range of familiarity strengths would support the hypothesis that the additive effects of multiple memory signals control metamemory judgments.

In humans, a distinction is made between the monitoring and control aspects of metamemory. For example, a student can monitor their memory to report on whether they remember the answer to a test question, and they can control their memory by studying information that they do not remember. Because nonhumans cannot provide verbal response, they necessarily “self-report” the status of memory by engaging in control, re-viewing answers or avoiding tests. The paradigms currently used to test metamemory in nonhumans blur the monitoring-control distinction, though the separability of monitoring and control in nonhumans would be an interesting topic of future research.

In conclusion, monkeys show similar patterns of accuracy and latency across decline-test and information-seeking paradigms, two metamemory tasks commonly used with nonhumans. A similar pattern of performance in both paradigms provides converging evidence that multiple memory signals can additively control metacognitive judgements in monkeys and provides a framework for mapping the interaction of explicit memory signals in primate memory.

Compliance with ethical standards

Ethical approval:

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Conflict of interest:

All authors declare no conflict of interest.

Funding:

This work was supported by the National Science Foundation (grants IOS-1146316; BCS-0745573; BCS-1632477) and the National Institutes of Health (grants RO1MH082819; T32HD071845). This project was supported in part by ORIP/OD P51OD011132.

References

Baddeley A (2000) The episodic buffer: a new component of working memory? Trends Cogn Sci 4:417-423

Baddeley A (2003) Working memory: looking back and looking forward Nature reviews neuroscience 4:829

Basile BM, Hampton RR (2013) Dissociation of active working memory and passive recognition in rhesus monkeys Cognition 126:391-396

Basile BM, Hampton RR, Suomi SJ, Murray EA (2009) An assessment of memory awareness in tufted capuchin monkeys (Cebus apella) Anim Cogn 12:169-180

Basile BM, Schroeder GR, Brown EK, Templer VL, Hampton RR (2015) Evaluation of seven hypotheses for metamemory performance in rhesus monkeys Journal of Experimental Psychology: General 144:85

Beran MJ, Smith JD (2011) Information seeking by rhesus monkeys (Macaca mulatta) and capuchin monkeys (Cebus apella) Cognition 120:90-105 doi:10.1016/j.cognition.2011.02.016

Brown EK, Templer VL, Hampton RR (2017) An assessment of domain-general metacognitive responding in rhesus monkeys Behavioural processes 135:132-144

Call J, Carpenter M (2001) Do apes and children know what they have seen? Anim Cogn 4:207-220

Carruthers P (2008) Meta- cognition in animals: a skeptical look Mind & Language 23:58-89

Castro L, Wasserman EA (2013) Information-seeking behavior: exploring metacognitive control in pigeons Anim Cogn 16:241-254 doi:10.1007/s10071-012-0569-8

Couchman JJ, Coutinho MV, Beran MJ, Smith JD (2010) Beyond stimulus cues and reinforcement signals: a new approach to animal metacognition J Comp Psychol 124:356

Coutinho MV, Redford JS, Church BA, Zakrzewski AC, Couchman JJ, Smith JD (2015) The interplay between uncertainty monitoring and working memory: Can metacognition become automatic? Mem Cogn 43:990-1006

Flavell JH (1979) Meta-cognition and cognitive monitoring - new area of cognitive-developmental inquiry American Psychologist 34:906-911

Fujita K (2009) Metamemory in tufted capuchin monkeys (Cebus apella) Anim Cogn 12:575

Hampton RR (2001) Rhesus monkeys know when they remember Proceedings of the National Academy of Sciences of the United States of America 98:5359-5362 doi:10.1073/pnas.071600998

Hampton RR (2009) Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? Comparative cognition & behavior reviews 4:17

Iwasaki S, Watanabe S, Fujita K (2013) Do pigeons (Columba livia) seek information when they have insufficient knowledge? Anim Cogn 16:211-221

Jozefowiez J, Staddon J, Cerutti D (2009) Metacognition in animals: how do we know that they know? Comparative Cognition & Behavior Reviews 4

Keppel G, Wickens TD (2004) Design and Analysis: A Researcher's Handbook. Prentice Hall, Kirk CR, McMillan N, Roberts WA (2014) Rats respond for information: Metacognition in a rodent? Journal of Experimental Psychology: Animal Learning and Cognition 40:249

Kornell N (2013) Where is the "meta" in animal metacognition? J Comp Psychol

Kornell N, Son LK, Terrace HS (2007) Transfer of metacognitive skills and hint seeking in monkeys Psychol Sci 18:64-71 doi:10.1111/j.1467-9280.2007.01850.x

Le Pelley M (2012) Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals Journal of Experimental Psychology: Learning, Memory, and Cognition 38:686

Maniscalco B, Lau H (2012) A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings Conscious Cogn 21:422-430

Marsh HL (2014) Metacognitive-like information seeking in lion-tailed macaques: a generalized search response after all? Anim Cogn 17:1313-1328

Marsh HL, MacDonald SE (2012) Information seeking by orangutans: a generalized search strategy? Anim Cogn 15:293-304

McMahon S, Macpherson K, Roberts WA (2010) Dogs choose a human informant: Metacognition in canines Behav Processes 85:293-298 doi:10.1016/j.beproc.2010.07.014

Nelson TO (1996) Consciousness and metacognition Am Psychol 51:102-116

Sherry DF, Schacter DL (1987) The Evolution of Multiple Memory-Systems Psychological Review 94:439-454

Smith JD, Beran MJ, Redford JS, Washburn DA (2006) Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring Journal of Experimental Psychology: General 135:282

Smith JD, Coutinho MV, Church BA, Beran MJ (2013) Executive-attentional uncertainty responses by rhesus macaques (Macaca mulatta) Journal of Experimental Psychology: General 142:458

Squire LR, Zola-Morgan S (1991) The medial temporal lobe memory system Science 253:1380-1386

Suda-King C (2008) Do orangutans (Pongo pygmaeus) know when they do not remember? Anim Cogn 11:21-42 doi:10.1007/s10071-007-0082-7

Suda-King C, Bania AE, Stromberg EE, Subiaul F (2013) Gorillas' use of the escape response in object choice memory tests Anim Cogn 16:65-84 doi:10.1007/s10071-012-0551-5

Takagi S, Fujita K (2018) Do capuchin monkeys (Sapajus apella) know the contents of memory traces?: A study of metamemory for compound stimuli J Comp Psychol 132:88

Templer VL, Hampton RR (2012) Rhesus monkeys (Macaca mulatta) show robust evidence for memory awareness across multiple generalization tests Anim Cogn 15:409-419 doi:10.1007/s10071-011-0468-4

Templer VL, Lee KA, Preston AJ (2017) Rats know when they remember: transfer of metacognitive responding across odor-based delayed match-to-sample tests Anim Cogn 20:891-906 doi:10.1007/s10071-017-1109-3

Vining AQ, Marsh HL (2015) Information seeking in capuchins (Cebus apella): A rudimentary form of metacognition? Anim Cogn 18:667-681

Washburn DA, Gulledge JP, Beran MJ, Smith JD (2010) With his memory magnetically erased, a monkey knows he is uncertain Biol Lett 6:160-162 doi:10.1098/rsbl.2009.0737

Watanabe A, Clayton NS (2016) Hint-seeking behaviour of western scrub-jays in a metacognition task Anim Cogn 19:53-64

Wittig JH, Morgan B, Masseau E, Richmond BJ (2016) Humans and monkeys use different strategies to solve the same short-term memory tasks Learning & Memory 23:644-647

Wittig JH, Richmond BJ (2014) Monkeys rely on recency of stimulus repetition when solving short-term memory tasks Learning & Memory 21:325-333

Yonelinas AP (2002) The nature of recollection and familiarity: A review of 30 years of research J Mem Lang 46:441-517

Yonelinas AP, Aly M, Wang WC, Koen JD (2010) Recollection and familiarity: Examining controversial assumptions and new directions Hippocampus 20:1178-1194