CrossMark

ORIGINAL PAPER

Rats know when they remember: transfer of metacognitive responding across odor-based delayed match-to-sample tests

Victoria L. Templer[1] · Keith A. Lee[1] · Aidan J. Preston[1]

Abstract Metamemory entails cognitively assessing the strength of one's memories. We tested the ability of nine Long-Evans rats to distinguish between remembering and forgetting by presenting a decline option that allowed a four-choice odor-based delayed match to sample (DMTS) tests to be by-passed. Rats performed significantly better on tests they chose to take than on tests they were forced to take, indicating metacognitive responding. However, rather than control by internal mnemonic cues, one alternative explanation is that decline use is based on external test-specific cues that become associated with increased rewards overtime. To examine this possibility, we tested rats on three generalization tests in which external contingencies were inconsistent and therefore could not serve as discriminative cues. Rats transferred adaptive use of the decline response in tests that eliminated memory by presenting no sample, increased memory by presenting multiple samples, and both weakened and strengthened memory by varying the retention interval. Further, subjects chose to take or decline the test before encountering the memory test, providing evidence that rats based their metacognitive responding on internal cues rather than external ones. To our knowledge, this is the first robust evidence for metamemory in rats using the DMTS decline-test paradigm in which several possible sources of external stimulus control can be ruled out.

✉ Victoria L. Templer
  vtempler@providence.edu

[1] Department of Psychology, Providence College,
    1 Cunningham Square, Providence, RI 02918, USA

Springer

Introduction

Metamemory is the ability to monitor the strength of one's memories, an ability routinely demonstrated by humans (e.g. Sutton and Shettleworth 2008). For example, members of a trivia team with competing answers to the same question may gauge their memory strengths before submitting a final answer. In doing this, they have made real-time assessments of their level of knowledge concerning a specific subject. Such metacognitive abilities rely on a central executive that controls and optimizes one's cognitive processes (e.g. Benjamin et al. 1998; Dunlosky and Bjork 2008; Flavell 1979; Nelson 1992; Schwartz 1994) and is thought to be one of humans' most cognitively complex abilities (Metcalfe and Kober 2005).

While humans are able to demonstrate metacognitive ability by verbal communication, the extent to which non-human animals possess metamemory remains a debated topic (Carruthers 2008; Jozefowiez et al. 2009; Smith et al. 2014). The initial nonverbal hurdle has been largely overcome with clever designs suitable for nonverbal species that rely on operational definitions of metacognition based on observable behavior. Many species have demonstrated metacognitive responding (for reviews see: Crystal and Foote 2009; Hampton 2009; Kornell 2009; Smith 2009; Roberts et al. 2012, Smith et al. 2014), but the current challenge lies in teasing apart potential sources of stimulus control that govern putative metacognitive responding.

The earliest group of researchers to assess metacognitive abilities in nonhumans gave subjects the option to decline

tests. Presumably, if subjective knowledge states can be monitored, tests should be declined when confidence is low. Rhesus monkeys, bottlenose dolphins, and humans show metacognitive responding by successfully declining difficult visual or auditory discrimination tests (e.g. recognizing auditory stimulus as high- and low-pitched; Shields et al. 1997) when the stimuli were near the discriminatory threshold. Results of similar tasks demonstrate that, in addition to monkeys, birds, and rats selectively avoid difficult psychological discriminations (e.g. monkeys: Brown et al. 2017; rats: Foote and Crystal 2007; pigeons and bantams: Nakamura et al. 2011).

These response patterns may be a result of monitoring one's own knowledge state, which is an internal and private source of stimulus control. However, they also may be explainable by reliance on external cues that are publicly available (Hampton 2009). Because introspective cues are internal and private, they can only be inferred by first excluding public ones. One possible external cue is the fact that difficult trials in discrimination tasks may be declined simply because intermediate stimuli, like pixel densities that are not unmistakably dense or sparse, become associated with the decline response over time. Stimulus magnitude, image similarity, and other features inherent to the test stimuli itself (Sutton and Shettleworth 2008; Kornell et al. 2007) are examples of *environmental cue associations* (Hampton 2009) that may serve as a discriminative stimulus to decline certain trials which include no internal component. Associative simulation models (Jozefowiez et al. 2009; Le Pelley 2012; Smith et al. 2008) have also been found to predict use of the decline response on psychophysical discriminations such as those in Smith et al. (1997) and Foote and Crystal (2007). The best way to rule out control of metacognitive responding (i.e. declining difficult tests) by external cues, like environmental cue associations, is to provide generalization tests across which external test-specific features are not consistent (Brown et al. 2017; Hampton 2009). In addition, use of metamemory paradigms in which relative memory strength serves as the natural continuum of discriminanda is advantageous (Brown et al. 2017; Hampton 2001; Smith et al.1998; Templer and Hampton 2012). This is because metamemory tests offer fewer ostensible stimulus features and potential cues in the test itself compared to psychophysical tests. Instead, memory strength is responsible for creating variation in performance.

Generalization of metacognitive responding across varied experimental conditions is one of the main reasons evidence for metamemory in rhesus monkeys has been so strong (Smith et al. 2014). The other reason is because of positive results when prospective metamemory tests have been used (Basile et al. 2015; Hampton 2001). In a prospective metamemory test the secondary metamemory options are present before the primary memory test (Goto and Watanabe 2012; Hampton 2009; Shettleworth 2010, p. 249). In contrast, in a concurrent metamemory test the primary memory test and secondary metamemory options are present at the same time. In Hampton's (2001) study monkeys chose to take or decline the test before engaging with the primary Delayed Match to Sample (DMTS) memory test. After training on the DMTS test, delays between study and test were titrated to an intermediate level that was above chance but below ceiling. Monkeys were then presented with a metacognitive response at the end of the delay that allowed them to accept the memory test, and receive a preferred reward if correct, or decline the memory test which resulted in a guaranteed, but less preferred reward. Monkeys were more accurate on trials they chose to take than they were on other trials that were presented in which there was no option to decline the memory test. A performance advantage on chosen trials as compared to forced trials results from adaptively declining difficult tests and indicates metacognitive responding. This is because forced trials are a combination of accuracy on tests animals would have declined had they had the option and trials they would have chosen to take. Monkeys also transferred adaptive decline use to conditions in which memory was manipulated by eliminating memory of the sample (by presenting no sample) or increasing memory difficulty (by increasing the delay).

The prospective, rather than concurrent, metamemory method is advantageous because it rules out potential sources of external stimulus control like *behavioral cue associations* and *response competition* (Hampton 2009). Behavioral cue associations include behavioral tendencies a subject might engage in when the animal doesn't know how to respond, like taking longer to respond, vacillating, or hesitating in any way (Goto and Watanabe 2012; Hampton 2009; Terrace and Son 2009). Longer response latencies on such trials, when the correct answer is not immediately evident, could therefore serve as a publicly available cue to decline when the metacognitive response is concurrent with the memory test, especially since response latencies in such MTS-based tasks are longer for incorrect than correct tests (e.g. Hampton and Hampstead 2006). Response competition refers to the fact that, when the primary memory test is in direct competition with the secondary metacognitive response, subjects might learn to associate absence of an inclination to select a choice response with selection of the decline response. In contrast, when the test is less difficult animals might simply have a strong tendency to select the correct choice response without making the internally-guided decision to take the test because they know they remember. Hampton's (2001) and Basile et al. (2015) results are especially strong because of the use of prospective tests where control by

behavioral cue associations and response competition are impossible.

Metamemory and metacognitive tests with multiple generalization procedures (Basile et al. 2015; Brown et al. 2017; Hampton 2001; Templer and Hampton 2012) combined with results from other information-seeking paradigms, and retrospective confidence judgment paradigms, have led several students of comparative metacognition to contend that evidence for introspective metacognition in Old World monkeys (Kornell et al. 2007; Rosati and Santos 2016; Smith et al. 1998) and apes (Beran et al. 2013; Call 2010; Call and Carpenter 2001; Marsh and MacDonald 2012; Suda-King 2008; Suda-King et al. 2013) is strong and functionally similar to that of humans (Fujita 2009; Roberts et al. 2009; Smith et al. 2014). Such compelling results, including those of Hampton (2001), and Smith et al. (2003, 2006) have since been shown to be impossible to predict from modeling using only associative processes (Smith et al. 2014).

Evidence for metacognition in other species like pigeons (Castro and Wasserman, 2013; Inman and Shettleworth 1999; Iwasaki et al. 2013; Roberts et al. 2009, 2012; Sutton and Shettleworth 2008), New World monkeys (Basile et al. 2008; Beran et al. 2016a, b; Paukner et al. 2006; Vining and Marsh 2015) and rats (Foote and Crystal 2007, 2012) has been less robust, making it unclear whether metamemory is phylogenetically limited to Old World monkeys and apes, or whether is it a more ancient cognitive ability. Demonstrations of some aspects of metacognitive behavior in pigeons (Adams and Santi 2011; Zentall and Stagner 2010) and capuchin monkeys (Beran et al. 2016) have been more convincing, but certainly not as substantial as those in macaque monkeys (Beran et al. 2016a, b). In the few studies that exist, evidence for metacognition in rats is also mixed (Foote and Crystal 2012; Kirk et al. 2014; Yuki and Okanoya 2017). The lack of robust results in rats are somewhat surprising given that researchers have found many similarities between human and rat explicit memory in recent years (e.g., Crystal 2013; Crystal et al. 2013; Eacott and Norman 2004; Eichenbaum 2005; Fortin et al. 2004; Panoz-Brown et al. 2016). One reason that the metacognitive results in rats are equivocal is due to the methods employed and fact that few attempts have been made to systematically rule out external sources of stimulus control.

Sprague–Dawley rats showed improved performance on classification tests of short versus long sound durations when they chose to take the test than on trials they were forced to complete. Rats also successfully declined difficult intermediate durations (Foote and Crystal 2007). In a later study, Foote and Crystal expanded upon this paradigm by also including trials in which subjects were forced to repeat the sample stimulus or could choose to repeat the stimulus.

Three of seven rats were more accurate at making the duration discriminations when they were forced to repeat it than when they chose to repeat it (Foote and Crystal 2012). While such responding could be explained by control by internal cues, rats were not more likely to repeat the stimulus on difficult trials, which would be a clear prediction of internal stimulus control. Further, with a temporal discrimination task, cueing by environmental associations is difficult to rule out because animals might simply have learned that middle stimuli (i.e. intermediate durations) signal decline use and this hypothesis was not tested with generalization tests (Hampton 2009).

In a separate study, Long-Evans rats obtained information that signaled correct response locations by pressing a lever in a T-maze or a Radial Arm Maze, when it worked to their advantage in gaining food (Kirk et al. 2014). However, when the food reward that resulted from obtaining more information via a lever press was eliminated, the behavior diminished in some cases. Because rats did not complete generalization tests like those used in other information-seeking paradigms (Basile et al. 2015) it is possible that external sources of stimulus control could have controlled information-seeking behavior. In a very recent study, rats did generalize adaptive use of the decline response on a delayed-match-to-position test when the sample was omitted, but this metacognitive behavior was restricted to a six-choice test and was not found in a two-choice test (Yuki and Okanoya 2017). This study made a significant step in the right direction, but positive results were obtained by only two rats in one generalization test.

Given the lack of consensus on the extent to which rats show performance consistent with metacognitive responding by internal sources of stimulus control, in Experiment 1 we employed a Delayed Matching to Sample (DMTS) test with Long-Evans rats modeled after the influential paradigm used in monkeys by Hampton (2001) and later adapted and expanded with a novel generalization test by Templer and Hampton (2012). Rats were tested in a four-choice odor based paradigm in which they dug in sand-filled cups, either to receive a preferred reward or to decline the test and obtain a less preferred, but guaranteed reward. To make the memory test reasonably difficult so that rats would experience both remembering and forgetting, the same four odors appeared on every trial, increasing memory interference. Subjects that have have access to internal memory cues should choose to take tests in which they remember the sampled odor and decline tests in which they forget, which would result in a performance advantage on tests in which the decline option is available as compared to tests on which it is not available. Experiments 2–4 presented three generalization tests in which memory strength was directly manipulated to control for possible cueing by environmental cue associations.

General methods

Subjects

Subjects were nine male Long-Evans rats (*Rattus norvegicus*) that arrived from Charles River Laboratories on post-natal day (PND) 21. The original testing group began with ten rats, but one was eliminated from testing after repeated failure to pass food preference testing. Rats were housed individually in enriched environments throughout development and into adulthood when they were tested on experiments described here, which occurred between PND 329-458. Subjects were housed in one colony room, maintained at 21.6 degrees C and kept on a 12:12 reversed light:dark cycle, with light onset at 8 PM. Cages were large ($17 \times 12.75 \times 24$ in.; Ferret Nation) with multiple platforms and solid flooring, enclosed by wire meshing. Rats had constant access to enrichment in their cages, which included additional platforms, running wheels, plastic toys/enclosures, an open-topped plastic shoebox cage ($17 \times 8 \times 8$ in.) and wooden chew toys. Corncob bedding (Enrich-o'cobs) was provided at the bottom of all enclosures, including the shoebox cage that was used to transport the rats into the testing room. The large size of the cages and surrounding metal bars provided ample opportunities for climbing and exercise.

Rat chow (LabDiet) was provided ad libitum until rats reached adulthood (approximately PND 60), after which point weighed rations of chow were provided daily after testing was complete. Rats were weighed weekly and chow rations were adjusted accordingly to maintain their body weights at approximately 90% of their estimated free-feeding weight. Water was available ad libitum, except for when the rats were being tested. All procedures described in the current study were approved by the Institutional Animal Care and Use Committee (IACUC) of Providence College.

Subjects had previous experience on unrelated learning and memory tasks. With the goal of piloting the methods for the experiments presented here, four of the subjects included in this study had participated in a previous DMTS metacognitive task. Seven of the subjects had also received limited training on the odor-based MTS procedure.

Apparatus

All training and testing sessions occurred inside a social interaction chamber (Noldus; 30×45 in., Fig. 1) divided into three separate rooms. The center room was divided into the sample (10×15 in.) and delay (10×15 in.)

Fig. 1 Photograph of the testing apparatus, which was divided into the *sample chamber*, *delay chamber, and testing chamber* (on the *right*). The testing chamber was divided into two sides: (1) the decline option area, which was bumpy in texture and covered in red contact paper, and contained the decline cup and (2) the MTS task area, which was corrugated in texture, covered with black contact paper, and contained the four choice cups

chambers by a cardboard insert that could be easily lifted by pulling the insert up along metal tracks secured to the walls divided the rooms. A small rectangular hole connected the delay chamber and the large room to the right: the testing chamber (10×45 in.), where the choice and decline cups were located. A removable opaque Plexiglas door could easily be lifted to allow the rat to travel from the delay chamber to the testing chamber.

The sample, choice, and decline cups were ceramic (2.75 in. in diameter and 1.5 in. tall) and filled half-way with play sand. The sample and choice cups were filled with play sand and the decline cup was filled with orange play sand that was slightly finer. To further differentiate the test and the decline option, two different floor textures (Fig. 1) were created. Where the choice cups were located, the floor was covered in black contact paper with elevated stripes created by layering duct tape underneath, while the floor where the decline option was located was covered in red contact paper with sand granules underneath. The four odors that were used throughout all experiments were: cinnamon, thyme, paprika, and coffee. Scented sand was created by mixing approximately three cups of sand with a small pinch of each scented material. All choice cups contained tiny crumbled pieces of cereal (that could not be eaten) to control for odor of the correct choice cup that contained a food reward. The scented sand was replaced each week.

General procedure

Rats were tested between 8:30 a.m. and 5 p.m. 6 days a week. Subjects were transported from their home cages to the testing room and given at least 30 min to acclimate before being tested under red-light conditions. Rats were fed food rations after testing each day.

Rats were placed in the sample room and a right-angled flat-spatula holding the sample cup was lowered to the floor so that the subject could dig to the bottom of the sand-scented cup to obtain a half piece of cereal. A food reward was provided to ensure that subjects sufficiently sampled the odor. After the food reward was retrieved, the door that connected the sample chamber to the delay chamber was opened and the rat remained in the delay chamber for the duration of the retention interval. After the retention interval, a small opaque door connecting the delay chamber and the testing chamber was opened so that the rat could enter the testing chamber. The testing chamber contained the choice cups on the right side of the room and decline cup on the left side of the room. On trials in which the decline response was not available, there was no decline cup present. On trials in which the decline response was available, a quarter piece of cereal was buried in the unscented decline cup. Each of the four choice cups was filled with a scented sand (i.e. cinnamon, coffee, paprika, thyme). In one of the cups, the scent matched that of the sample cup (e.g. coffee), and this cup also contained the buried preferred food reward: a whole piece of cereal. Remaining choice cups were odor controlled with crumbled cereal mixed into the sand, so that all four cups smelled equally of the food reward. Rats indicated choices by digging and were prevented from attempting to dig in more than one cup. After digging in any one of the cups, and consuming the food reward if correct, the rat was immediately removed from the chamber to prevent multiple choices. Rats waited in their home cage adjacent to the testing chamber during the 4-min inter-trial interval (ITI) before being moved back into the sample chamber to begin the next trial.

The sample odor was semi-randomized and each of the odors appeared as the sample three times within a block of 12 trials. The locations of the incorrect and correct choice cups were also counterbalanced and randomized such that each odor was presented at each location an equal number of times.

Data analysis and behavior scoring

During each trial, the first cup in which the rat started digging was considered his choice. Rats were immediately removed from the chamber after digging in the first cup to prevent them from selecting another choice cup or the decline cup. Proportions were arcsine transformed before statistical analyses to better approximate the normality assumption underlying parametric statistics (Keppel and Wickens 2004, p. 155).

Training procedures

Preference testing and preference checks

A hierarchy of food rewards was established so that the larger reward (whole piece of cereal) could be used as the "preferred" option in the correct choice cup and the smaller reward (1/4 piece of cereal) could be used as the "less preferred" in the decline cup. Depending on the individual rat's preference, the cereal reward was either a FrootLoop (Kellogg's), Capn' Crunch (Quaker Oats), or dehydrated marshmallow (Medley Hills Farm). The two rewards were placed in two corners of the chamber and the rat was placed on the opposite side and then allowed to eat one of the rewards. Once the rat chose the larger reward four out of five trials for two consecutive sessions, MTS training began. As mentioned above, one rat did not establish a preference and was therefore not included in the following experiments. This testing began on PND 288 and lasted until PND 298.

Preference checks were performed each week throughout testing to ensure that preferred food rewards did not change. Preference checks operated in the same way as preference testing, but were comprised of two trials that occurred after normal test sessions were conducted. All animals maintained their original preferences throughout testing.

Training to dig in sand-filled cups

Rats were trained to dig in sand-filled cups to obtain a food reward at the bottom. First, these sand filled cups were placed in home cages overnight. After two nights of successful reward retrieval, rats were required to dig for food rewards in the sample chamber of the apparatus twice before MTS training began.

Match to sample (mts) training

Although seven rats had already undergone MTS training several months earlier, all were retrained to ensure that their behavior was still under the control of the task contingencies. Trials occurred as explained above in general procedure. The delay (retention interval) between study and test was as short as possible: as long as it took to open the door to the testing chamber, which was less than a few seconds. Rats received MTS trials until the sample odor was accurately matched to the choice odor for at least three out of six trials for two consecutive six-trial sessions.

Delay titration in delayed match to sample (DMTS)

This phase was identical to MTS training except that the delays between study and test were increased for each rat in order to ensure that performance was above chance (25%) but below ceiling. Delays were therefore titrated until performance stabilized between 40 and 70% correct. Beginning with the effective delay of 0 s from MTS training, the delay was increased to 30 s if performance exceeded 70% for two consecutive sessions of six trials. Subsequent delay increases were calculated by multiplying 1.5 times the current delay. If performance fell below 40% on two consecutive sessions, the delay was decreased back to its previous retention interval. When the delay had not been changed for two sessions, this delay was set as the *criterion delay*, specific to each rat (Table 1).

Decline use training

While delay titration was in progress, subjects were also trained to select the decline response option so that they could learn that this cup always held a less preferred, but guaranteed food reward. After each six-trial DMTS session, rats were given two forced decline trials. After

sampling the odor and waiting for the currently titrated delay, the rat entered the testing chamber and encountered only the unscented decline cup (no choice cups were present). The decline cup contained the less preferred, yet guaranteed reward, a quarter of a cereal. After the rat consumed the cereal, he was removed from the chamber.

Experiment 1

Experiment 1 was the first situation in which rats had the opportunity to demonstrate metacognitive responding. Subjects were given the option to decline the DMTS test on two-thirds of the trials, but in the remaining third of trials, the decline option was not available. If rats have metamemory, then when given the option, they should decline tests when memory for the sample is weak, and take tests when memory is comparatively strong and they are more likely to answer the memory test correctly. Such adaptive use of the decline response would cause significantly higher performance on trials in which the decline option is available compared to trials on which it is not. This is because performance on forced tests includes trials that rats might have declined if they had the option. If rats

Table 1 Criterion delays and decline rates in Experiments 1–4 and accuracies in Experiments 1 and 4

Stage of testing	Specifics	R1	R2	R3	R4	R5	R6	R7	R8	R9	Mean
Experiment 1	Criterion delay (s)	0	60[a]	0	0	60	90[a]	30	90	0	
	Forced accuracy	.55	.58	.43	.6	.45	.65	.5	.5	.5	.52
	Chosen accuracy	.51	.58	.63	.7	.51	.69	.78	.53	.6	.61
	Decline rate	.31	.2	.11	.13	.21	.16	.25	.08	.3	.2
Experiment 2	Criterion delay (s)	0	60	0	0	60	90	30	90	0	
	Decline rate no sample trials	.3	.2	.33	.5	.25	.33	.15	.38	.5	.32
	Decline rate normal trials	.13	.18	.06	.2	.2	.13	0	.29	.2	.15
Experiment 3	Criterion delay (s)-Phase 1	20	80	20	20	80	110	50	110	20	
	Criterion delay (s)-Phase 2	120	160	90	100	120	150	90	150	60	
	Decline rate double sample trials	.05	.13	0	.3	.07	.05	.03	.3	0	.1
	Decline rate normal trials	.21	.33	.05	.5	.13	.24	.11	.62	.1	.26
Experiment 4	Criterion delay (s)	120	160	90	100	120	150	90	150	60	
	Forced accuracy (0 s)	.69	.56	.83	.88	.5	.88	.65	.73	.6	.70
	Forced accuracy criterion delay	.43	.5	.64	.5	.43	.21	.57	.43	.29	.44
	Forced accuracy long delay (4 min)	.45	.63	.33	.41	.33	.45	.38	.33	.36	.41
	Chosen accuracy (0 s)	.72	.97	.71	.97	.77	.82	.84	.78	.88	.83
	Chosen accuracy (criterion delay)	.71	.97	.61	.75	.83	.75	.76	.74	.67	.75
	Chosen accuracy long delay (4 min)	.8	.67	.83	1	.69	.88	.81	.17	.5	.71
	Decline rate short delay (0 s)	0	.06	.07	0	.19	.07	.03	.09	0	.06
	Decline rate criterion delay	.11	.23	.2	.6	.13	.19	.12	.49	.2	.24
	Decline rate long delay (4 min)	.78	.72	.67	.9	.42	.69	.51	.64	.5	.64

[a] Criterion delay was readjusted during Experiment 1

do not have metamemory, decline use should not track internal memory strength and should instead be chosen at random, causing performance on chosen and forced trials to be equivalent.

Methods

Subjects and apparatus

All subjects and apparatus were the same as those used in training.

Procedure

The procedure was identical to that used in training except that four *forced* trials in which the decline cup was not available were randomly intermixed with eight *choice* trials in which the decline cup was available. Rats were tested on one 12-trial session a day and received a total of ten sessions.

To ensure that each rat's performance remained within in the range of 40–70%, accuracy over three consecutive sessions (12 forced trials) was calculated using a sliding window after each session was completed. If a subject's performance fell outside of the range for three consecutive sessions, the rat re-entered the delay titration phase (consisting of 6 forced trials) to gauge performance. Delay titration occurred as described above and continued until performance stabilized within the desired range for one 12-trial session. At this point the rat re-entered testing in Experiment 1, beginning at session one with the adjusted delay. The performance of two rats exceeded 70% on forced trials, which lead to an increase in their respective delays (Table 1).

Results and discussion

Performance was significantly higher on chosen trials than on forced trials (Fig. 2; two-tailed, paired-sample t test: $t_8 = 3.20$, $P = .013$), indicating that rats adaptively used the decline option on chosen trials when memory was weak (average decline use is displayed in Table 1). There was no main effect of odor on forced trials (RMANOVA, $F_{3,35} = 1.09$, $P = .368$), indicating that odors were equally memorable to subjects. Correspondingly, no odors were declined more than others (Cinnamon: M = .31, Coffee: M = .22; Paprika: .24; Thyme: .20; RMANOVA, $F_{3,35} = .87$, $P = .469$), indicating that the effect was not driven simply by a tendency to decline one odor over another.

The results of Experiment 1 are consistent with the hypothesis that rats possess the ability to monitor their memories. Results mirror findings by Foote and Crystal
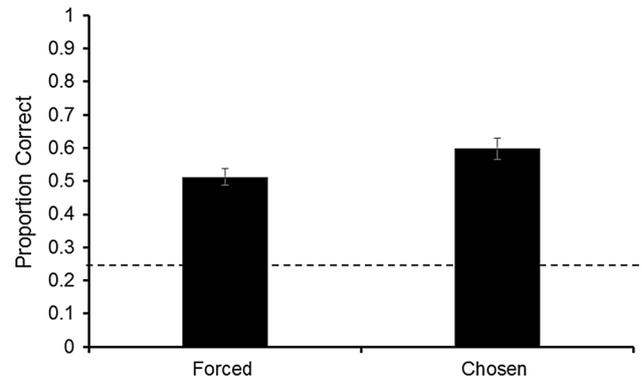

Fig. 2 Proportion correct (±SEM) on *forced* and *chosen* trials in Experiment 1. The *dotted line* represents chance performance

(2007) that rats perform better on chosen tests as compared to forced tests, but this is the first demonstration of a chosen performance advantage on a memory-based test in rats. Because there are fewer ostensible environmental stimuli that could have cued metacognitive responding here than in the psychophysical arena (i.e. intermediate duration signals decline use), it is more likely that subjects here were responding to the detection of internal memory strengths. However, while variation in performance is driven by memory strength rather than externally presented stimuli, there may be environmental cues that subjects are attending to of which we are not aware. The following experiments test for generalization of appropriate decline use across multiple contexts in which no association with external stimuli was consistent. By manipulating memory in several ways, we test for the possibility that behavior is controlled by external stimuli or if is in fact driven by an internal assessment of memory strength.

Experiment 2

In Experiment 1 rats used the decline response adaptively, which lead to higher accuracy on chosen trials than forced trials. In order to control for external cues that may be correlated with the probability of reinforcement, Experiment 2 utilized "no-sample" trials, in which no sample was presented prior to the memory test. Here, memory strength is brought under direct experimental control because rats have no memory of a sample odor on these trials. If adaptive use of the decline response in Experiment 1 was controlled by an environmental cue association, rats should decline no-sample trials *less* frequently than normal trials because no relevant cue of the sample is available as a discriminative cue. Alternatively, if adaptive use of the decline response was controlled by an internal assessment of memory strength, rats should decline no-sample trials

significantly *more* than normal trials because there are no external stimulus features of the sample that could be associated with reinforcement, as no sample is presented.

Methods

Subjects and apparatus

The subjects and apparatus were the same as those used in Experiment 1.

Procedure

The procedure was similar to Experiment 1 in all but two ways; forced choice trials were not included, and no-sample trials were semi-randomly intermixed among normal choice trials. Each twelve-trial session therefore consisted of eight normal sample trials and four no-sample trials. During no-sample trials, a rat was placed in the sample chamber and after approximately 1 min (the time it took to normally present a sample), he entered the delay chamber to wait for his criterion delay. A random choice cup was baited so that rats had a 25% chance of obtaining a reward if they did not decline. Subjects completed one session a day and received a total of ten sessions.

Results and discussion

Rats declined no-sample trials significantly more often than the normal sample trials (Fig. 3, two-tailed, paired-sample *t* test; $t_8 = 5.15$, $P = .001$), suggesting that use of the decline response is modulated by identifying an absence of memory, rather than by external cues that signal trial difficulty.
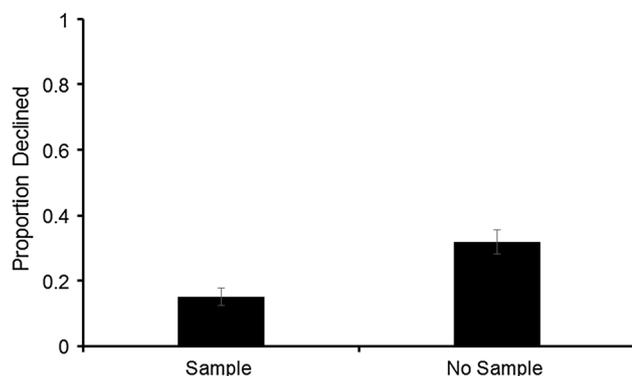
Fig. 3 Proportion (±SEM) of choice trials on which rats declined the memory test when a sample was presented (*sample*) and when no sample was presented (*no-sample*) in Experiment 2

While subjects declined no-sample trials more frequently than they did normal trials, average proportion declined on no-sample trials was .32, which is lower than would be expected. Presumably, if memory is truly being monitored and no sample was presented, subjects should *almost always* decline no-sample trials, which would result in decline rates close to 1.0, as seen in monkeys in an analogous experimental manipulation (Templer and Hampton 2012). However, we rewarded animals randomly on no-sample tests so that the test phase of no-sample trials was identical to normal sample trials, while monkeys never received a reward on no-sample trials (Templer and Hampton 2012). It is possible that rats declined these trials a third of the time, rather than 100% of the time, because they were attempting to remember the last presented sample and use that information to make their choice at test. In the current design, memory interference between trials is very high as the same four odors are present during each choice phase of every trial, increasing the possibility that rats were trying to remember the previously presented odor from the last normal trial in which a sample odor was experienced. To determine if rats were erring based on interference from the previous trial, we examined which odor was selected on trials that were not declined. During no-sample trials, if rats accepted the test, they were significantly more likely to choose the scent from the previous trial than chance (one sample *t* test; $t_8 = 2.85$, $P = .020$), suggesting that behavior was likely under the control of the last presented odor on no-sample trials.

A related hypothesis worth considering is that subjects declined no-sample tests on the basis of weak signal strength of the last presented sample. In other words, the rat could be declining no-sample tests with no regard to the current no-sample test itself. However, if this was the case, such responding would necessitate metacognitive responding based on assessment of memory strength for previously presented trials. While this plausible adopted tactic by rats was unforeseen by the experimenters, there is no way to determine if decline use was under the control of the last presented sample or the omitted sample on the current trial. However, the strategy of remembering previously presented samples is consistent with the fact that subjects did not decline all or close to all no-sample tests.

In addition to the 25% chance of receiving a reward on a no-sample trial and the confirmed proactive interference, a general tendency towards risk taking behavior may have also contributed to rats' tendency to take no-sample tests. Previous studies have indicated that if rats are satiated in terms of caloric intake, they may be more likely to take risks for possible but not guaranteed food rewards; when presented with choices similar to those presented here, rats are risk-prone (Mazur 1988). Given that rats were fed full rations of food each day to maintain healthy weight, it is

plausible that choosing to take the test, and having a 25% chance of receiving the quarter piece of cereal, was well worth the risk in most cases.

While we considered plausible reasons why rats did not decline the majority of no-sample trials, our primary metamemory hypothesis that no-sample trials should be declined significantly more than normal sample trials was strongly supported. Decline use tracked experimental manipulation of memory strength, reinforcing the hypothesis that rats chose to decline tests in which they knew they did not remember.

Experiment 3

In Experiment 2 we aimed to eliminate memory by not providing a sample, simulating trials on which subjects happened to forget the sample. Due to proactive inference from the previous trial's sample on no-sample trials, it is likely that rats were attempting to remember a sample, and therefore were not fully aware of a true absence of memory. In Experiment 3, we therefore brought memory under direct experimental control in the opposite direction: increasing memory by providing two sample-study periods rather than one. If use of the decline response is controlled by an internal awareness of memory strength, rats should transfer adaptive use of the decline response to this new condition. Double-sample trials should be declined significantly *less* than normal sample trials as the second sample should increase memory (Roberts 1972). However, if adaptive use of the decline response is controlled by some external cue, such as experiencing an unusual situation of receiving two samples, rats should decline double-sample trials significantly *more* than normal trials. Importantly, the hypothesis-driven predictions are in direct opposition to each other. Metacognitive responding by an assessment of memory strength predicts less decline use on double-sample trials as compared to normal sample tests. Responding based on environmental cue associations of probe tests predicts more decline use on double-sample trials as compared to normal sample tests. These bidirectional predictions will provide more convincing evidence of what decline use is tracking: either an internal memory-based cue or an external cue.

Methods

Subjects and apparatus

The subjects and apparatus were the same as those used in Experiments 1 and 2.

Procedure

By the start of Experiment 3, rats had been tested on the MTS odor task for four months. Because we suspected performance may have increased due to experience, all rats underwent a performance check of ten forced trials with a slightly longer delay than their criterion delay (criterion delay plus 20 s). The results of the performance check revealed that all accuracies were within the range of 40–70% with the new delay, so testing for Experiment 3 began using this new delay.

Details of these trials were similar to Experiment 2 with two exceptions. Instead of no-sample trials, subjects received double-sample trials. Eight trials per session were normal sample trials, while the remaining four trials were double-sample trials. Additionally, to avoid satiation and hold reward value constant across sample periods, each presentation of the sample on double-sample trials contained a quarter piece of cereal rather than the half piece on normal trials. During the double-sample trials, the rat received the first sample scent and consumed the sample cup reward (quarter piece of cereal). Immediately after the rat had finished eating, a second sample of the same scent and a quarter piece of cereal was presented to the rat. The rest of the trial progressed as usual. Rats received one twelve-trial session a day and were tested for ten sessions.

Halfway through testing in Experiment 3, when five sessions had been completed, we noticed that decline use on normal sample trials had increased compared to Experiment 2. Therefore, we completed a second performance check halfway through Experiment 3 and confirmed that performance on forced trials had increased. Delays were subsequently re-titrated for each rat such that performance again fell between 40 and 70% correct. Rats were then tested on the remaining five sessions (Phase 2) with increased delays from the first half of Experiment 3 (considered Phase 1; see Table 1).

Results and discussion

Subjects declined normal sample trials significantly more than they declined double-sample trials in both Phase 1 with the criterion delay + 20 s (two-tailed, paired-sample t test; $t_8 = -6.71$, $P = .000$) and Phase 2 with the further increased criterion delay (see Table 1; two-tailed, paired-sample t-test; $t_8 = -6.65$, $P = .000$). Due to the analogous results across phases, we collapsed the data into a single set, which again indicated normal sample trials were declined significantly more than double-sample trials (Fig. 4; two-tailed, paired-sample t test; $t_8 = 6.71$, $P = .000$). That rats did not decline double-sample trials significantly more than normal sample trials, even though they could have appeared unique and unusual, indicates that metacognitive responding was not dependent upon
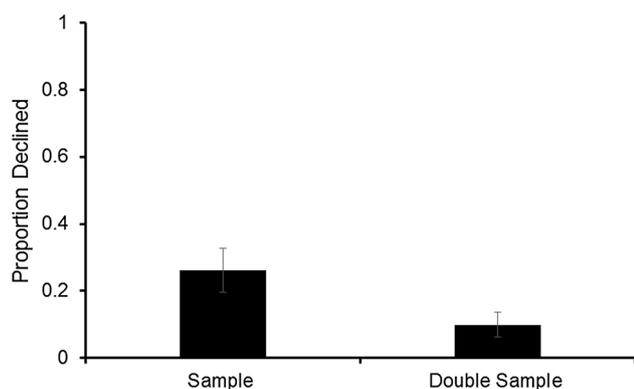
Fig. 4 Proportion (±SEM) of choice trials on which rats declined the memory test when a sample was presented (*sample*) and when two samples were presented (*double-sample*) in Experiment 3

external features unique to the probe trials. Instead, subjects were more likely to choose to take tests after a sample was presented twice and memory was increased. This provides additional evidence that decline use is likely based on an internal assessment of memory strength.

In addition to initial metacognitive evidence presented in Experiment 1, rats transferred adaptive use of the decline response in both Experiments 2 and 3, providing growing refutations of potential sources of behavioral cue associations. Use of the decline response seems to be flexible enough to respond to bidirectional experimental manipulations of memory, as seen in the previous two experiments. To test bidirectional manipulations of memory in a single experiment, we attempted to both increase and decrease memory in Experiment 4.

Experiment 4

Memory in Experiment 4 was manipulated subtly by changing the delay between sample and test. Rats were tested on trials with their normal criterion delay, with no delay, and with a long delay of 4 min. To be sure that decreasing and increasing the delay altered memory strength, we included both forced trials and chosen trials at each of the three delays. If rats have metamemory they should decline fewer trials when memory is strong at the short delay and decline many trials when memory is weak at the long delay. Rats adaptively using the decline response should also display a performance advantage on chosen tests as compared to forced tests, mirroring the original finding in Experiment 1.

Methods

Subjects and apparatus

The subjects and apparatus were the same as those used in the first three experiments.

Procedure

Procedures followed those described in general procedure. Rats received one 18-trial session per day for a total of ten sessions. Within each session, there were six trials at each delay: the short delay (0 s), the adjusted criterion delay (used in Experiment 3, Phase 2), and the long delay (4 min). Forced and choice trials were also semi-randomly intermixed (constrained so that no more than two identical trial types occurred in a row) during each session so that there was a total of four choice trials and two forced trials per delay over the course of the experiment.

Results and discussion

Subjects were more accurate on short delay trials than they were on long delay trials overall, and were more accurate on chosen trials than on forced trials (Fig. 5; 2×3 factorial ANOVA: main effect of delay: $F_{2,48} = 7.81$, $P = .001$; main effect of trial type [forced vs. chosen]: $F_{1,48} = 32.51$, $P = .000$; delay $\times$ trial type interaction: $F_{2,48} = 1.54$, $P = .225$). Rats were most likely to decline trials when memory was weak at long delays and least likely when memory was strong at short delays (one-way ANOVA: $F_{2,24} = 43.82$, $P = .000$). When considering decline rates and accuracy on forced tests as dependent measures, decline use tracked accuracy (2×3 factorial ANOVA: main effect of delay: $F_{3,48} = 40.22$, $P = .000$; main effect of trial type [forced vs. decline]: $F_{1,48} = 51.08$, $P = .000$; delay $\times$ trial type interaction: $F_{1,48} = 11.33$, $P = .002$), indicating rats used the decline response adaptively and according to memory strength. Examining
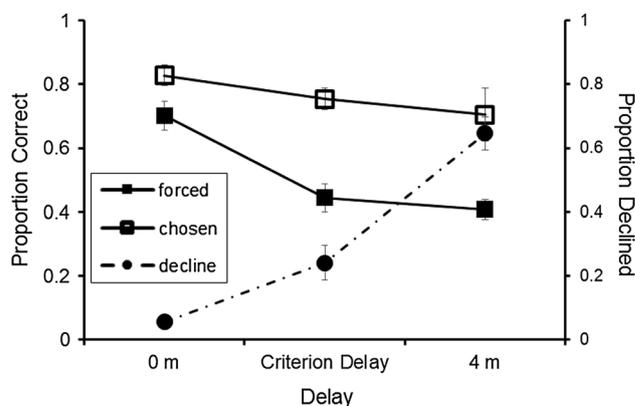

Fig. 5 Proportion correct (±SEM) on *forced* and *chosen* trials after a short delay (0 min), intermediate delay (individually titrated *criterion delays*, which included 60, 90, 100, 120, 150, and 160 s), and long delay (4 min), and proportion of trials declined in Experiment 4. The proportion of correct trials on forced tests (*filled squares*), chosen tests (*open squares*) correspond to the primary y-axis and the proportion of trials declined (*filled circles*) correspond to the secondary y-axis

Fig. 5, it is unknown why performance on forced and chosen criterion trials did not differ dramatically from performance at the long delay (proportion correct respectively, forced: .44 vs. .40, $P = .53$; chosen: .75 vs. .70; $P = .79$), but decline use did increase dramatically (decline criterion delay: .24 vs. decline 4 m delay: .64; $P = .000$).

In Experiment 4 memory was manipulated in a graded manner by varying delay times that corresponded to varying levels of difficulty, as confirmed by performance on forced trials. Memory on long delay trials here theoretically mirrors the results found in Experiment 2 in which memory was decreased by presenting no sample; while memory on short delay trials here theoretically mirrors the results found in Experiment 3 in which memory was increased by presenting two samples. As in Experiment 2 and 3, use of the decline response tracked trial difficulty as subjects were less likely to decline trials with shorter delays (.057) and more likely to decline trials with longer delays (.64). Importantly, decline use doubled on long delay trials as compared to no-sample trials in Experiment 2 (.32), suggesting that subjects could detect a weak memory state and that high interference is likely to account for relatively low decline rates on no-sample trials. The results of this experiment further support the hypothesis that internal memory monitoring is the cause of the metacognitive responding as opposed to external cue associations.

General discussion

In each of the four experiments presented here, results are consistent with the hypothesis that rats declined tests in which memory was weak and accepted tests in which memory was strong. In Experiment 1, performance on chosen trials exceeded that of performance on forced trials, suggesting that rats adaptively used the decline option when memory was weak. Three additional experiments examined transfer of decline use in generalization tests across which external stimuli were varied and not associated with memory status. In Experiment 2, rats declined no-sample trials more frequently than normal sample trials, indicating that memory state was weak much like in occasional normal sample trials when animals happened to forget the sample odor. However, a potential issue with this manipulation is that if subjects noticed something odd about these probe trials, such as the absence of the sample cup, they might have declined based on that novel feature rather than the absence of memory. We therefore increased memory in Experiment 3 by presenting two samples, thus predicting bidirectional external versus internal cueing hypotheses. If rats selected the decline response whenever

something odd or unexpected occurred, subjects should have declined these double-sample trials significantly more than normal sample trials. Instead, rats declined double-sample trials significantly less, thereby directly refuting this external cue hypothesis and lending support to the hypothesis that decline use was based on memory strength. Finally, in Experiment 4, both forced and choice trials were reintroduced as in Experiment 1 with the added manipulation of both increasing and decreasing the retention interval. Rats rarely declined trials after short delays and frequently declined trials after long delays. Subjects also performed significantly more accurately on chosen than forced tests, further reinforcing the conclusion that rats used the decline response adaptively and that its use is flexible enough to vary according to different memory strengths trial-to-trial. Successful transfer of the decline response across these test situations reveals that use of the decline response is most likely controlled by internal memory state rather than external cues like environmental cue associations.

A prospective metacognitive choice

Normally, when the decline response and memory test are present at the same time, the test is considered a concurrent metacognitive task (e.g. Basile et al. 2008; Hampton et al. 2004; Smith et al. 1997; Templer and Hampton 2012). However, due to the fact that the metacognitive choice (turn to the left or right) occurred before, and was not the same response as the primary test responses, the present task can more aptly be described as a pseudo-prospective. If the rat chose to take the test, he would then travel to the right side of the testing room and systematically start sampling odors at one end of the line of choice cups and work his way through the options until arriving at his selection, which was indicated by digging. Locations of odors were always counterbalanced and randomized, so for rats to encounter the test, they needed to physically sample the choice odors. Although rats were physically able to change their take-the-test or decline-the-test decision, as they could travel to the decline side of the room after they had inhaled scents from the choice cups, this almost never occurred. Instead, 99.95% of the time rats made a decision to take or decline the test immediately upon entering the testing chamber: if they turned right to take the test they stayed right and chose one of the four choice cups; if they turned left they stayed right and dug in the decline cup.

Evaluation of internal and external sources of stimulus control

Using a broad definition of metacognition in which control for metacognitive responding can occur by both internal or

external sources of stimulus control (Hampton 2009), we are able to consider several explanations for adaptive use of the decline response within a unified framework. In animals, metacognitive responding by internal cues can only be inferred by excluding possible sources of external stimulus control (Hampton 2009). Here we thoroughly consider the extent to which publicly available cues may have been available to subjects in this paradigm.

The environmental cue association hypothesis posits that animals respond to external test-specific cues to guide decline use (Hampton 2009). This hypothesis is extremely unlikely to account for the results of the present study because rats made a pseudo-prospective choice and generalized adaptive decline use across several tests. In Experiment 1, rats appropriately used the decline response, demonstrated by increased performance on chosen compared to forced tests, mirroring what was found in an analogous visual task in monkeys (Hampton 2001; Templer and Hampton 2012). Rats were no more accurate after sampling any of the four odors and decline use did not vary as a function of odor. Though we can safely eliminate the environmental cue of individual odors cueing decline use, there may have been other environmental cue associations inherent to the testing stimuli of which we were unaware. In Experiment 2, we eliminated the sample, and found that rats declined these no-sample trials more than when the sample was presented, suggesting that features of the sample itself could not have served as a discriminative cue for decline use. Subjects did just the opposite in Experiment 3 when the sample odor was presented twice to increase memory of it, further reinforcing evidence that decline use was not controlled by stimulus-specific cues. When retention intervals were both increased and decreased, subjects used the decline response adaptively and according to memory strength. This strong generalization of the decline response across tests in which environmental cues were not constant suggests that rats based decline use on an internal memory cue.

According to the behavioral cue association hypothesis, subjects may use self-generated behaviors elicited by the primary memory test as cues to select the decline response. The most common possibilities of such publicly observable cues an animal could emit are hesitation or vacillation (Goto and Watanabe 2012; Hampton 2009; Terrace and Son 2009). If the test is truly prospective, behavioral cue associations are not possible (Hampton 2009). Here, the design is reminiscent of a concurrent task, but also resembles a prospective one. Because the choice and decline cups were in the same room, and subjects may have been able to smell the same four odors in the center of the room on every trial, it is possible that some external behavior was generated *before* the subject turned left or right to either the decline or choice cups. This could have

resulted in the behavioral cue of relatively long versus short response latencies to turn left or right. Unfortunately, with the manual design we employed, we were unable to collect response latency data, which limits the certainty with which we can rule out the potential behavioral cue of response latencies cueing decline use.

The response competition hypothesis posits that animals seeing (in the case of a visual task) or smelling (for our purposes) a correct answer overpowers other responses and the absence of this propensity leads to decline use (Hampton 2009). This hypothesis was originally developed for tasks in which the metacognitive choice and selection of a primary test response are the same (e.g. Smith et al. 1998; Templer and Hampton 2012), and are thus are in direct competition with each other. In the present task, the metacognitive choice (turn left or right) is not the same response as selection of the choice or decline cups. While this pseudo-prospective design may limit such external cueing to some extent, it is possible that response competition could have cued decline use. Though the same odors were present on every trial, and in order to determine which cup contained which odors rats had to physically sample each odor before choosing, it is possible rats could have detected the sample odor at the metacognitive decision point before they turned left or right. This external odor cue could have therefore activated a relatively salient sample representation and corresponding strong inclination to move towards the choice cups. Thus, response competition could result in a metacognitive behavior without complete control by an internal memory cue. Future studies should develop designs to improve upon the prospective nature of this paradigm by having the decline and choice cups in different rooms so that the response competition hypothesis can be tested more directly.

The final hypothesis to consider is rote response (Basile et al. 2015), which posits that, over many training trials, animals learn which response leads to maximized reward. In designing our study, we aimed to decrease the likelihood of this influence by presenting animals with the minimum number of trials required to sufficiently analyze the data. It is unlikely that associative learning guided adaptive decline use as rats shifted behavior to adjust for new circumstances immediately and in several conditions. In Experiment 2 the rote response hypothesis would posit that overtime animals learn that no sample tests should be declined to maximize reward. However, unlike other tests with omitted samples (e.g. Templer and Hampton 2012 and Yuki and Okanoya 2017), here we chose to randomly reward one of the choice cups so that animals could not learn that choosing to take the test on no-sample trials never results in a reward, thereby eliminating the possibility of rote response in Experiment 2. In Experiment 3, we directly tested this hypothesis by pitting rote response against cueing by

internal memory strength. If rats had learned to simply decline all non-typical sample (probe) tests, then they should have declined double-sample trials significantly more than normal sample trials. Subjects declined these probe tests less frequently, which is instead consistent with responding based on increased memory strength. In Experiment 4, short and long delay trials were randomly intermixed among criterion delay trials, making it difficult to learn that long delays resulted in incorrect responses and no rewards and short delay resulted in correct responses and rewards in just 16 forced trials at each delay.

That subjects quickly and adaptively transferred decline use across several generalization tests use allows us to safely eliminate the environmental cue association hypothesis as a possible explanation for the results of Experiments 1–4. We cannot rule out behavioral cue associations and response competition with certainty. However, based on the fact that rats made a prospective decision to take or decline the test, these hypotheses are less likely to apply to this data. Though rote response learning is a potential threat to many comparative cognition studies, it is unlikely that this can account for the aforementioned behaviors because animals experienced so few trials on several different generalization tests. We conclude that subjects most likely based their decision to take or decline the memory test on an internal memory cue.

Internal mnemonic cues: executive or associative control

A recurring debate in the field of animal metacognition is whether metacognitive behavior can be explained entirely by associative principles of learning and conditioning without the need to appeal to "more cognitive" mechanisms (Jozefowiez et al. 2009; Le Pelley 2012; Smith et al. 2008). In response to this issue, some have taken the approach of obfuscating the reinforcement contingencies (Beran et al. 2006, 2014; Smith et al. 2006), thus attempting to make "low-level" accounts based on reinforcement less likely. These debates, and the experimental designs developed to circumvent them, often assume that "high-level" accounts are more interesting than "low-level" accounts. We see no need for such stratification, as there is little utility in differentiating between the worthiness of hypotheses. Any given behavior is likely supported by multiple mechanisms, and identifying the cues that support metacognitive decision making is important regardless of the answer (Basile et al. 2015; Basile and Hampton 2014; Hampton 2009). We therefore do not deny that pairing weak memory states with a particular response and strong memory states with another response follows well-recognized associative principles, but this does not negate the inference that that memory state is monitorable

(Basile et al. 2015). Furthermore, we endorse the argument by Basile et al. (2015) that for many memory-monitoring paradigms "current associationist accounts are objectively indistinguishable from accounts of introspective memory monitoring, and may actually be in agreement" (p. 16; see also Basile and Hampton 2014).

The goal of the experiments presented here was to minimize external sources of stimulus control by using multiple generalization tests so that an internal source of stimulus control, if present, could be inferred. We have shown when memory state was experimentally manipulated, rats used that variation as a discriminative cue to respond adaptively, indicating that memory state was monitored. While we have argued that rats know when they remember, and have identified the monitored cue as an internal memory state, it is still unclear how much their decision represents a controlled "executive" process or a more automatic rule. We look forward to future studies that are designed to determine the degree to which associative or executive control support metacognitive behavior.

Broader implications and conclusions

Here we present a novel and valid test for investigating the presence of metamemory in rodents, which was modeled after a visual-based task in monkeys by Hampton (2001). Many methodological details were refined from previous iterations piloted to optimize this design which should help others aiming to replicate and expand upon this paradigm. For example, an interesting challenge we faced was the fact that several rats could remember odors for over 20 min, even with a minimal inter-trial interval, when odors were drawn from a pool of 24. We therefore found that for experimenters to be able to test nine rats in a day's time, a set-size of four was optimal. High interference allowed us to more quickly titrate memory delays to an optimal level so that rats experienced both remembering and forgetting the sample. We suspect that using olfaction made to-be-remembered samples particularly salient and facilitated learning, especially in such an ecologically relevant task where subjects had to forage for food rewards. In future studies it would be interesting to examine metacognitive ability as a function of working memory (Smith et al. 2010) by increasing and decreasing the set-size as researchers have done in separate memory paradigms (Basile and Hampton 2010).

One of the major controversies in comparative cognition is the degree to which nonhuman animals have declarative or explicit memory. We have examined this important question by operationalizing the monitorability feature of explicit (i.e. declarative) memory (Hampton 2003a, b). Researchers have made significant progress in other aspects of explicit memory in rats, including its source memory

component (Crystal et al. 2013), episodic memory (Babb and Crystal 2005, 2006; Eacott and Norman 2004; Eichenbaum 2005; Panoz-Brown et al. 2016), memory for the future (Crystal 2013), and recollective processes (Eacott et al. 2005; Fortin et al. 2004). While putative metacognitive behavior has been reported in other studies (Foote and Crystal 2007, 2012; Kirk et al. 2014; Yuki and Okanoya 2017) the former two studies are especially subject to non-introspective accounts, and no previous studies have been able to rule out several possible sources of external stimulus control. Lack of evidence that clearly supports metacognitive or non-metacognitive accounts had led students of comparative metacognition to question whether metacognitive interpretations in rats are even appropriate (Foote and Crystal 2012). Until now, robust evidence for metamemory in the rat had yet to be presented.

Evidence presented here suggests that rats respond adaptively to variation in memory strength caused by several experimental manipulations of memory. We argue that this behavior is very unlikely to be explainable by external sources of stimulus control, and is best explained by an internal source of stimulus control: memory state. These results, suggesting rats know when they remember, add to existing neurobiological evidence that the explicit memory system is largely conserved across humans and rats (Manns and Eichenbaum 2006) and sets the stage for future cognitive and neurobiological studies of explicit memory. Together these studies will be shed light on the evolution of metamemory and will help provide an essential translational animal model for investigations of the treatment and prevention of disorders of memory.

References

Adams A, Santi A (2011) Pigeons exhibit higher accuracy for chosen memory tests than for forced memory tests in duration matching-to sample. Learn Behav 39:1–11

Babb SJ, Crystal JD (2005) Discrimination of what, when, and where: implications for episodic like memory in rats. Learn Motiv 26:177–189

Babb SJ, Crystal JD (2006) Episodic-like memory in the rat. Curr Biol 16(13):1317–1321. doi:10.1016/j.cub.2006.05.025

Basile BM, Hampton RR (2010) Rhesus monkeys (Macaca mulatta) show robust primacy and recency in memory for lists from small, but not large, image sets. Behav Process 83(2):183–190. doi:10.1016/j.beproc.2009.12.013

Basile BM, Hampton RR (2014) Metacognition as discrimination: commentary on Smith et al. Comp Psych 128:135–137. doi:10.1037/a0034412

Basile BM, Hampton RR, Suomi S, Murray EA (2008) An assessment of memory awareness in tufted capuchin monkeys (Cebus apella). Anim Cogn. doi:10.1007/s10071-008-01801

Basile BM, Schroeder GR, Brown EK, Templer VL, Hampton RR (2015) Evaluation of seven hypotheses for metamemory performance in rhesus monkeys. J Exp Psychol Gen 144(1):85–102. doi:10.1037/xge0000031

Benjamin AS, Bjork RA, Schwartz BL (1998) The mismeasure of memory: when retrieval fluency is misleading as a metacognitive index. J Exp Psychol Gen 127:55–68. doi:10.1037/00963445.127.1.55

Beran MJ, Smith JD, Redford JS, Washburn DA (2006) Rhesus macaques (Macaca mulatta) monitor uncertainty during numerosity judgments. J Exp Anim Behav Process 32:111–119. doi:10.1037/0097-7403.32.2.111

Beran MJ, Smith JD, Perdue BM (2013) Language-trained chimpanzees (Pan troglodytes) name what they have seen but look first at what they have not seen. Psychol Sci 24:660–666. doi:10.1177/0956797612458936

Beran MJ, Perdue BM, Smith JD (2014) What are my chances? closing the gap in uncertainty monitoring between rhesus monkeys (Macaca mulatta) and capuchin monkeys (Cebus apella). J Exp Anim Learn Cogn 40:303–316. doi:10.1037/xan0000020

Beran MJ, Menzel CR, Parrish AE, Perdue BM, Sayers K, Smith JD, Washburn DA (2016a) Primate cognition: attention, episodic memory, prospective memory, self-control, and metacognition as examples of cognitive control in nonhuman primates. Wiley Interdiscip Rev Cogn Sci 7(5):294–316. doi:10.1002/wcs.1397

Beran MJ, Perdue BM, Church BA, Smith JD (2016b) Capuchin monkeys (Cebus apella) modulate their use of an uncertainty response depending on risk. J Exp: Anim Learn Cogn 42(1):32–43. doi:10.1037/xan0000080

Brown EK, Templer VL, Hampton RR (2017) An assessment of domain-general metacognitive responding in rhesus monkeys. Behav Process 35:132–144. doi:10.1016/j.beproc.2016.12.004

Call J (2010) Do apes know that they can be wrong?. Anim Cogn 13:689–700. doi:10.1007/s10071010-0317-x

Call J, Carpenter M (2001) Do apes and children know what they have seen? Anim Cogn 3(207):220. doi:10.1007/s100710100078

Carruthers P (2008) Meta-cognition in animals: a skeptical look. Mind Lang 23(1):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

Crystal JD (2013) Remembering the past and planning for the future in rats. Behav Process 93:39–49. doi:10.1016/j.beproc.2012.11.014

Crystal JD, Foote AL (2009) Metacognition in animals: trends and challenges. Comp Cogn Behav Rev 4:54–55

Crystal JD, Alford WT, Zhou WY, Hohmann AG (2013) Source memory in the rat. Curr Biol 23(5):387–391. doi:10.1016/j.cub.2013.01.023

Dunlosky J, Bjork RA (eds) (2008) Handbook of memory and metamemory. Psychology Press, New York

Eacott MJ, Norman G (2004) Integrated memory for object, place, and context in rats: a possible model of episodic-like memory?

J Neurosci 24(8):1948–1953. doi:10.1523/jneurosci.297503.2004

Eacott MJ, Easton A, Zinkivskay A (2005) Recollection in an episodic-like memory task in the rat. Learn Mem 12(3):221–223. doi:10.1101/lm.92505

Eichenbaum H (2005) The hippocampus and episodic memory: cognitive and neural mechanisms. Neuropsychopharmacol 30:S37–S37

Flavell JH (1979) Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. Am Psychol 34(906):911. doi:10.1037/0003-066X.34.10.906

Foote AL, Crystal JD (2007) Metacognition in the rat. Curr Biol 17(6):551–555

Foote AL, Crystal JD (2012) "Play it Again": a new method for testing metacognition in animals. Anim Cogn 15(2):187–199. doi:10.1007/s10071-011-0445-y

Fortin NJ, Wright SP, Eichenbaum H (2004) Recollection-like memory retrieval in rats is dependent on the hippocampus. Nature 431(7005):188–191

Fujita K (2009) Metamemory in tufted capuchin monkeys (Cebus apella). Anim Cogn 12(575):585. doi:10.1007/s10071-009-0217-0

Goto K, Watanabe S (2012) Large-billed crows (Corvus macrorhynchos) have retrospective but not prospective metamemory. Anim Cogn 15(1):27–35. doi:10.1007/s10071-011-0428-z

Hampton RR (2001) Rhesus monkeys know when they remember. Proc Nat Acad Sci 98(9):5359–5362

Hampton RR (2003a) Rhesus monkeys know when they remember. Proc Natl Acad Sci USA 98(3):5359–5362

Hampton RR (2003b) Metacognition as evidence for explicit representation in nonhumans. Behav Brain Sci 26(3):346–347

Hampton RR (2009) Multiple demonstrations of metacognition in nonhumans: converging evidence or multiple mechanisms? Comp Cogn Behav Rev 4:17–28

Hampton RR, Hampstead BM (2006) Spontaneous behavior of rhesus monkeys (Macaca mulatta) during memory tests suggests memory awareness. Behav Process 72:184–189 PMID: 16460887

Hampton RR, Zivin A, Murray EA (2004) Rhesus monkeys (Macaca mulatta) discriminate between knowing and not knowing and collect information as needed before acting. Anim Cogn 7:239–254

Inman A, Shettleworth SJ (1999) Detecting metamemory in nonverbal subjects: a test with pigeons. J Exp Psychol Anim B 25:389–395. doi:10.1037/0097-7403.25.3.389

Iwasaki S, Watanabe S, Fujita K (2013) Do pigeons (Columba livia) seek information when they have insufficient knowledge? Anim Cogn 16:211–221. doi:10.1007/s10071-012-0566-y

Jozefowiez KJ, Staddon JER, Cerutti DT (2009) Metacognition in animals: how do we know that they know? Comp Cogn Behav Rev 4:29–39

Keppel G, Wickens TD (2004) Design and analysis, a researchers handbook, 4th edn. Pearson, Upper Saddle Rober

Kirk CR, McMillan N, Roberts WA (2014) Rats respond for information: metacognition in a rodent? J Exp Psychol Anim B 40(2):249–259. doi:10.1037/xan0000018

Kornell N (2009) Metacognition in humans and animals. Curr Dir Psychol Sci 18:11–15

Kornell N, Son LK, Terrace HS (2007) Transfer of metacognitive skills and hint seeking in monkeys. Psychol Sci 18(1):64–71

Le Pelley ME (2012) Metacognitive monkeys or associative animals? simple reinforcement learning explains uncertainty in nonhuman animals. J Exp Psychol Learn Mem Cogn 38:686–708

Manns JR, Eichenbaum H (2006) Evolution of declarative memory. Hippocampus 16(9):795–808

Marsh HL, MacDonald SE (2012) Information seeking by orangutans: a generalized search strategy? Anim Cogn 15:293–304. doi:10.1007/s10071-011-0453-y

Mazur JE (1988) Choice between small certain and large uncertain reinforcers. Anim Learn Behav 16:199–205

Metcalfe J, Kober H (2005) Self-reflective consciousness and the projectable self. In: Terrace HS, Metcalfe J (eds) The missing link in cognition: origins of self-reflective consciousness. Oxford University Press, New York, NY, pp 57–83

Nakamura N, Watanabe S, Betsuyaku T, Fujita K (2011) Do birds (pigeons and bantams) know how confident they are of their perceptual decisions? Anim Cogn 14(1):83–93. doi:10.1007/s10071-010-0345-6

Nelson TO (ed) (1992) Metacognition: core readings. Allyn & Bacon, Toronto

Panoz-Brown D, Corbin HE, Dalecki SJ, Gentry M, Brotheridge S, Sluka CM, Crystal JD (2016) Rats remember items in context using episodic memory. Curr Biol 26(20):2821–2826. doi:10.1016/j.cub.2016.08.023

Paukner A, Anderson JR, Fujita K (2006) Redundant food searches by capuchin monkeys (Cebus apelia): a failure of metacognition? Anim Cogn 9(2):110–117. doi:10.1007/s10071-005-0007-2

Roberts WA (1972) Short-term memory in pigeons-effects of repetition and spacing. J Exp Psychol Gen 94(1):74

Roberts WA, Feeney MC, McMillan N, MacPherson K, Musolino E, Petter M (2009) Do pigeons (Columba livia) study for a test? J Exp Psychol Anim Behav Process 35:129–142

Roberts WA, McMillan N, Musolino E, Cole M (2012) Information seeking in animals: metacognition? Comp Cogn Behav Rev 7:85–109. doi:10.3819/ccbr.2012.70005

Rosati AG, Santos LR (2016) Spontaneous metacognition in rhesus monkeys. Psychol Sci 27(9):1181–1191. doi:10.1177/0956797616653737

Schwartz BL (1994) Sources of information in metamemory: judgments of learning and feelings of knowing. Psychon B Rev 1:357–375. doi:10.3758/BF03213977

Shettleworth SJ (2010) Cognition, evolution, and behavior, 2nd edn. Oxford University Press, New York

Shields WE, Smith JD, Washburn DA (1997) Uncertain responses by humans and Rhesus monkeys (Macaca mulatta) in a psychophysical same-different task. J Exp Psychol Gen 126(2):147–164

Smith JD (2009) The study of animal metacognition. Trends Cogn Sci 13:389–396

Smith JD, Shields WE, Schull J, Washburn DA (1997) The uncertain response in humans and animals. Cognition 62(1):75–97. doi:10.1016/s0010-0277(96)00726-3

Smith JD, Shields WE, Washburn DA (1998) Memory monitoring by animals and humans. J Exp Psychol Gen 127(3):227–250. doi:10.1037//00963445.127.3.227

Smith JD, Shields WE, Washburn DA (2003) The comparative psychology of uncertainty monitoring and metacognition. Behav Brain Sci 26:317–374

Smith JD, Beran MJ, Redford JS, Washburn DA (2006) Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring. J Exp Psychol Gen 135(2):282–297

Smith JD, Beran MJ, Couchman JJ, Coutinho MVC (2008) The comparative study of metacognition: sharper paradigms, safer inferences. Psychon B Rev 15(4):679–691. doi:10.3758/pbr.15.4.679

Smith JD, Redford JS, Beran MJ, Washburn DA (2010) Rhesus monkeys (Macaca mulatta) adaptively monitor uncertainty while multi-tasking. Anim Cogn 13(1):93–101. doi:10.1007/s10071-009-0249-5

Smith JD, Couchman JJ, Beran MJ (2014) Animal metacognition: a tale of two comparative psychologies. J Comp Psychol 128(2):115–131. doi:10.1037/a0033105

Suda-King C (2008) Do orangutans *(Pongo pygmaeus)* know when they do not remember? Anim Cogn 11(1):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(1):65–84. doi:10.1007/s10071-012-0551-5

Sutton JE, Shettleworth SJ (2008) Memory without awareness: pigeons do not show metamemory in delayed matching to sample. J Exp Psychol Anim B 34(2):266–282. doi:10.1037/00977403.34.2.266

Templer VL, Hampton RR (2012) Rhesus monkeys *(Macaca mulatta)* show robust evidence for memory awareness across multiple generalization tests. Anim Cogn 15(3):409–419. doi:10.1007/S10071-011-0468-4

Terrace HS, Son LK (2009) Comparative metacognition. Curr Opin Neurobiol 19(1):67–74. doi:10.1016/j.conb.2009.06.004

Vining AQ, Marsh HL (2015) Information seeking in capuchins *(Cebus apella)*: a rudimentary form of metacognition? Anim Cogn 18(3):667–681. doi:10.1007/s10071-015-0835-7

Yuki S, Okanoya K (2017) Rats show adaptive choice in a metacognitive task with high uncertainty. J Exp Psychol Anim Learn Cogn 43(1):109–118

Zentall TR, Stagner JP (2010) Pigeons prefer conditional stimuli over their absence: A comment on Roberts et al. J Exp Psychol Anim B 36:506–509. doi:10.1037/a0020202