



Published in final edited form as:

*Behaviour*. 2015 ; 152(6): 727–756. doi:10.1163/1568539X-00003251.

## The misbehaviour of a metacognitive monkey

Ken Sayers<sup>1,†</sup>, Theodore A. Evans<sup>1</sup>, Emilie Menzel<sup>1</sup>, J. David Smith<sup>2</sup>, and Michael J. Beran<sup>1</sup>

<sup>1</sup>Language Research Center, Georgia State University, 3401 Panthersville Rd., Decatur GA 30034, USA

<sup>2</sup>Department of Psychology and Center for Cognitive Science, University at Buffalo, the State University of New York, 346 Park Hall, SUNY Buffalo, Buffalo, NY 14260, USA

### Summary

Metacognition, the monitoring of one's own mental states, is a fundamental aspect of human intellect. Despite tests in nonhuman animals suggestive of uncertainty monitoring, some authors interpret these results solely in terms of primitive psychological mechanisms and reinforcement regimes, where "reinforcement" is invariably considered to be the delivery and consumption of earned food rewards. Surprisingly, few studies have detailed the trial-by-trial behaviour of animals engaged in such tasks. Here we report ethology-based observations on a rhesus monkey completing sparse-dense discrimination problems, and given the option of escaping trials (i.e., responding "uncertain") at its own choosing. Uncertainty responses were generally made on trials of high objective difficulty, and were characterized by long latencies before beginning visible trials, long times taken for response, and, even after controlling for difficulty, high degrees of wavering during response. Incorrect responses were also common in trials of high objective difficulty, but were characterized by low degrees of wavering. This speaks to the likely adaptive nature of "hesitation," and is inconsistent with models which argue or predict implicit, inflexible information-seeking or "alternative option" behaviours whenever challenging problems present themselves. Confounding models which suggest that nonhuman behaviour in metacognition tasks is driven solely by food delivery/consumption, the monkey was also observed allowing pellets to accumulate and consuming them during and after trials of all response/outcome categories (i.e., whether correct, incorrect, or escaped). This study thus bolsters previous findings that rhesus monkey behaviour in metacognition tasks is in some respects disassociated from mere food delivery/consumption, or even the avoidance of punishment. These and other observations fit well with the evolutionary status and natural proclivities of rhesus monkeys, but weaken arguments that responses in such tests are solely associated with associative mechanisms, and instead suggest more derived and controlled cognitive processing. The latter interpretation appears particularly parsimonious given the neurological adaptations of primates, as well as their highly flexible social and ecological behaviour.

<sup>†</sup>Corresponding author lrckas@gsu.edu.

## Keywords

metacognition; associative learning; self-reward behaviour; Old World monkeys; *Macaca mulatta*; evolution; ecology; Morgan's canon

---

“It would seem that the Lord was simply unaware of drive-reduction learning theory when he created, or permitted the gradual evolution of, the rhesus monkey.”

Harry F. Harlow and Clara Mears (1979, p. 93)

## 1. Introduction

In the classic but controversial 1961 paper “The misbehavior of organisms,” the psychologists-cum-animal trainers Keller and Marian Breland famously catalogued a colorful list of failures which occurred during operant-based attempts to mold the behaviour of an equally colorful list of animals (Breland & Breland, 1961). Raccoons being trained to drop coins in a metal container for a food reward instead rubbed them together, a common behaviour during their natural crayfish foraging. Suids being trained to deposit coins in an appropriately-named “piggy bank” would, over time, drop and “root” at them incessantly—again mirroring feral feeding behaviour—resulting in extraordinarily slow transport rates. The Brelands, rebelling against their former mentor B.F. Skinner, commented sardonically that “perhaps the white rat cannot reveal everything there is to know about behavior” (p. 681) and, perhaps more important, suggested that learning theorists take into account the natural history of species when designing or interpreting experiments. “After 14 years of continuous conditioning and observation of thousands of animals,” they wrote, “it is our reluctant conclusion that the behavior of any species cannot be adequately understood, predicted, or controlled without knowledge of its instinctive patterns, evolutionary history, and ecological niche” (p. 684).

While the Brelands' paper is known best today for its treatment of instinct, the present article focuses more closely on its arguments relating to evolution and ecology more broadly. Specifically, we closely examined the behaviour that one particular rhesus macaque (*Macaca mulatta*) exhibited while performing an uncertainty monitoring task, and explored whether primitive, associative mechanisms—versus more derived, cognitive-oriented ones, which in a phylogenetic sense first appeared later in evolutionary time—best explained the results. Our examination utilized detailed observations of one individual, in tandem with earlier work, to comment on larger, more general, principles (e.g., Premack, 1959). After all, it is individuals, and not averaged groups, that actually “behave,” as has long been noted in behavioral ecology (e.g., Emlen, 1966; Stephens & Krebs, 1986; Sayers et al., 2010). This was attempted with direct reference to the evolutionary status and natural proclivities of the rhesus monkey—factors that we argue have sometimes been neglected by researchers in previous attempts to explain experimental results pertaining to metacognition and other psychological phenomena (see also Hoffman & Schwartz, 2014).

Definitions of metacognition abound (Beran et al., 2012), but at their most basic level involve the monitoring of one's own mental states (“cognition about cognitive phenomena,” Flavell, 1979:906) and include sensations of knowing, not knowing, and doubt. Such

abilities have been documented in children as young as 3.5 years old and indeed are considered a hallmark of normal human cognition (Balcomb & Gerken, 2008). The phylogenetic history of metacognition, however, is debated; while some authors consider it uniquely-human (Carruthers, 2008), others have ushered evidence for metacognition in a range of nonhuman animals (reviewed in Smith, 2009). Part of the disagreement is semantic and/or philosophical and involves consideration of differing levels of metacognition, such as implicit, unconscious metacognition, which is more easily attributed to nonhuman animals, and explicit, conscious metacognition, which is comparatively difficult (e.g., Frith, 2012; Metcalfe & Son, 2012). This paper deals primarily with metacognition in the larger sense, without regard to the difficult, and often untestable, issues concerning “levels of consciousness” (see Karin-D’Arcy, 2005; Couchman et al., 2012). With respect to what follows, however, it should be noted that implicit metacognition is often considered to occur quickly and involuntarily, with explicit metacognition being more slow and methodical (Frith, 2012).

One traditional procedure for investigating putative metacognition in nonverbal animals involves discriminations where the trial-specific difficulty is systematically varied, but the subject is given the ability to “opt-out” of problems at its own choosing (Smith et al., 1995). For example, an animal might discriminate whether a visual pattern on a computer screen is “sparse” or “dense” based on the number of lit pixels comprising it—with correct responses being rewarded with food and incorrect responses resulting in a time out—or utilize a third option that escapes the current trial and moves the subject onto a new one. In a metacognitive animal, the use of this latter “uncertainty response” or UR (Smith et al., 1995) should be selective and associated with the most difficult trials. Using this methodological approach, comparative psychologists have asked whether other species have a functional analog to human metacognition. Pigeons (Sole et al., 2003; but see Adams & Santi, 2011) and capuchin monkeys (Beran et al., 2009; Beran et al., 2014) have to date exhibited little to mixed success at using URs adaptively, while rhesus macaques (Smith, 2009) and great apes (Call & Carpenter, 2001; Suda-King, 2008; Beran et al., 2013; Suda-King et al., 2013) have performed better on these or related tasks.

In the rhesus macaque, the species in which the most systematic work has been conducted, evidence includes myriad extensions of the classic problem, including abstract same-different discriminations (Shields et al., 1997) and situations where the stimuli to be discriminated are not present at the time of decision-making (Hampton, 2001), as well as under conditions of delayed feedback, where rewards and penalties are administered only after subjects complete blocks of trials (Smith et al., 2006; Couchman et al., 2010). Macaques immediately generalize use of the UR to novel problems (Washburn et al., 2006), will bet high on trials they eventually get correct when compared to those they eventually bungle (Kornell et al., 2007), and will selectively increase use of the UR when brain activity is disrupted by transcranial magnetic stimulation (Washburn et al., 2010). To some, such results, which reflect success in numerous tasks from numerous laboratories (i.e., conceptual replication), possess all the hallmarks of mental monitoring and indeed seem indistinguishable from humans’ performance in analogous tasks (reviewed in Couchman et al., 2012).

The consensus, however, is not universal. An alternative viewpoint holds true to a particular reading (see Discussion) of Morgan's famous canon, which holds that "in no case may we interpret an action as the outcome of the exercise of a higher psychological faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale" (Morgan, 1894:53, originally in italics). In relation to potential metacognition in nonhumans, results in the aforementioned tests are linked to primitive, associative mechanisms of reward, punishment, and stimulus-response connections, as opposed to more derived cognitive processes. Indeed, a possible "low-level" basis for URs has been the dominant theoretical concern regarding animal metacognition research (e.g., Smith et al., 2006; Smith et al., 2008; Crystal & Foote, 2009; Hampton, 2009; Smith et al., 2009; Le Pelley, 2012; Smith et al., 2012; Smith et al., 2014). The essence of the associative argument is that animals might learn that some stimuli (in difficult problems) produce comparatively rare food rewards but frequent timeouts during which such rewards are unavailable. These timeout intervals are aversive, as are the stimuli that are associated with them. Therefore, to minimize these intervals, and to maximize the rate at which food rewards are delivered for consumption, animals selectively make URs to wave off difficult trials through associative learning, without necessarily monitoring aspects of their mental state.

The general procedure that has been followed by investigators holding this view involves the development of reinforcement models that capture some elements of the animals' overall patterns of response in the aforementioned tests (Carruthers, 2008; Jozefowicz et al., 2009a; Le Pelley, 2012). Behaviour is invariably considered to be driven solely by the maximization (delivery/consumption) of food rewards, and the avoidance of punishment timeouts (cf. Premack, 1959). If such reinforcement rates adequately explain the patterns of response, it is argued, there is little need to posit higher-order processing such as metacognition. Note that, notwithstanding these models, various psychological mechanisms (including some cognitive ones, as in foraging by modern humans, e.g., Hawkes et al., 1982) could potentially be utilized to achieve reward maximization (Smith et al., 2012); a particular interpretation of Morgan's canon is required to reach the "low-level mechanism" conclusion (Carruthers, 2008; Jozefowicz et al., 2009a; Le Pelley, 2012).

This general associative description for nonhuman metacognition studies has been prominent, although it has itself been subject to criticism. For example, it has been argued that it erroneously predicts that pigeons and capuchins (the former a poster animal for associative learning) should utilize URs as readily as Old World monkeys and apes (Smith et al., 2014). Also, there remain other assumptions of the "low-level mechanism" hypothesis that remain to be adequately tested. Explicitly or implicitly, this viewpoint (see Carruthers, 2008; Jozefowicz et al., 2009a; Jozefowicz et al., 2009b; Le Pelley, 2012) assumes that the URs are just like any other response type; in other words, after controlling for factors such as difficulty, there should be nothing behaviourally to distinguish the UR from correct or incorrect responses. So, while particularly challenging problems may be associated with behavioural signs of hesitation or anxiety, these are by-products of task difficulty and should, notwithstanding idiosyncratic differences in temperament, occur automatically whenever these challenging problems appear (Carruthers, 2008). The associative description also assumes that: 1) responses are linked only to reinforcement history, 2) reinforcement is determined from rules set up by the researcher, in addition to the responses of the subject, 3)

reinforcement is equivalent to the delivery/consumption of a food reward, and 4) responses generally reflect past reward history and trial-by-trial stimuli (local optima) as opposed to a more long-term strategy (global optima) (see also Staddon, 1983).

Our purpose in the present article is twofold. First, we believe that interpretations of the mental processes underlying nonhuman animal performance in uncertainty monitoring (and other psychological) studies should be grounded by a careful ethological consideration of how animals actually behave during these tasks. It has been observed, for example, that a bottlenosed dolphin will exhibit ancillary behaviours such as slowing and wavering when utilizing a UR (Smith et al., 1995), and that a rhesus monkey is more likely to aggressively “strike” a computer touchscreen when selecting an answer that proves to be incorrect than when making a correct response (Hampton & Hampstead, 2006). We would note more broadly that, beyond these examples, the detailed behaviour (e.g., movement patterns) of animals in these tests has only rarely been described, though this could have important implications for the hypotheses utilized to explain the more general results in such studies. As noted above, the associative models developed to date do not predict that any special behaviours, after controlling for task difficulty, will be associated with URs (e.g., Carruthers, 2008; Le Pelley, 2012). A metacognitive interpretation, in contrast, would predict that the use of URs would be associated with flexible behavioural signals associated with high attentional demands and controlled processing, such as long latencies for responses, prolonged trial times, and signs of hesitation or wavering (Shiffrin & Schneider, 1977; Smith et al., 1995; Frith, 2012). Thus, the first goal of this study was to quantify whether and/or at what rate such behaviours were present when a monkey performed an uncertainty-monitoring task, whether they were specifically linked with URs, and, as a corollary, to what degree they were simply associated with the difficulty of problems.

Second, we believe that—lacking this detailed description—the very general associative models that have previously been offered for the results in uncertainty-monitoring tasks potentially threaten to misunderstand the microclimate of rewards and punishments that animals actually experience. As noted above, associative models for putative metacognition assume that the delivery and consumption of food rewards solely constitute “reinforcement” (e.g., Jozefowicz et al., 2009a). Previous work suggests this tenet is problematic, particularly when applied to primates, which possess a nervous system of unparalleled complexity, and whose behaviour, whether in laboratory or field, is markedly flexible and oftentimes inexplicable in terms of simple food-motivated (or, more broadly, drive-motivated) associative explanations (Premack, 1959; Premack, 1962; Harlow & Mears, 1979; Washburn & Rumbaugh, 1992; Rumbaugh & Washburn, 2003; see Discussion). A detailed investigation of the feeding behaviour of a monkey engaged in a metacognition task is therefore appropriate, and relevant to the debates raised above because, as noted, opinions differ as to what degree the patterns of responses in such studies are linked with the food motivator. Previous work has shown that monkeys will complete blocks of trials with food rewards and timeout punishments delayed until after they are all completed (Smith et al., 2006; Couchman et al., 2010), in a sense dissociating response and one aspect of reinforcement. The current paper extends this focus by looking at when pellets earned after completion of successful trials are actually consumed.

Thus, the second goal of this study was to quantify the timing of pellet (the food reward) consumption in relation to task attendance and various response/outcome categories (correct, incorrect, UR), as this is an aspect of the issue that has not been closely investigated. Pellet acquisition and consumption are here considered linked aspects of reward (see Premack, 1959), although we consider consumption of greater hedonic value than delivery (the aforementioned associative models are silent on such issues) and focus on this variable accordingly. It is thus assumed that a bird in the hand (or, more properly, a pellet in the mouth) is better for a monkey than seeing one in the bush (or, more properly, seeing a pellet in the collection tray). While not dismissing the potential value of pellet delivery (as opposed to consumption) as having potential reinforcing properties, the aforementioned studies have already established that monkeys will succeed on metacognition tasks even when food delivery is delayed (Smith et al., 2006; Couchman et al., 2010). The purpose here is to look at the spontaneous consumption of food earned in a more traditional procedure where correct responses result in the immediate delivery of a pellet; in other words, to examine the related idea of “consumption as reinforcer.”

The main questions in relation to this can be succinctly stated as follows: does the temporal consumption of food reward follow or deviate from the contingency put in place by the experimenter, and how does this relate to associative interpretations of possible metacognitive performance? Are there discrete, positively-reinforcing arrivals of food that is consumed? Are there aversive timeout intervals empty of food consumption? The associative descriptions noted previously (e.g., Jozefowicz et al., 2009a) assume all of these. It may be that when the microclimate is fully described, such approaches will need to be recast.

Accordingly, this article examines the trial-by-trial behaviour of a rhesus macaque engaged in a sparse-dense uncertainty-monitoring task, with particular emphasis on the actions exhibited during responses and the temporal acquisition and consumption of rewards. Rhesus macaques possess a number of evolutionary features which make them especially interesting for such investigations. Macaques are catarrhine primates (as are, of course, humans) with sizable, complex brains (Rumbaugh et al., 1996), whose ancestors diverged from the branch leading to *Homo sapiens* approximately 23 million years ago (Glazko & Nei, 2003). They are thus expected to share many plesiomorphic and derived characters with us, including certain cognitive abilities.

Other anatomical and behavioural features suggest that specifying “reinforcement” may be difficult when macaques are the subjects (Harlow & Mears, 1979), as they can engage in short-term food hoarding or behavioural strategies to reduce the negative influence of punishment. This is not to say that associative learning is unimportant to them, but only that concepts such as “reinforcement” may be difficult to operationally define. Like many members of this group, macaques possess hands with exceptional grasping abilities (Roy et al., 2000), including a relatively high opposability index (Napier & Napier, 1967), and thus can hold food items manually while being engaged in other tasks. In addition, macaques, like all cercopithecine monkeys, possess large cheek pouches utilized for the temporary storage of food (Hill, 1974). This adaptation likely relates, in addition to food processing considerations, to the competitive demands of social foraging (Murray, 1975; Lambert,

2005) and thus relates directly to the ability to delay the ingestion of food—note that ingestion relates to swallowing, not necessarily placement in the mouth—if it would otherwise reduce foraging efficiency (cf. Evans & Beran, 2007). In addition, macaques, as in most or all primates (including humans), exhibit marked behavioral flexibility and novelty-seeking behaviors (Butler, 1953; Menzel & Menzel, 1979), suggesting that associative arguments concentrating on food alone as “reinforcement” may be missing valuable psychological phenomena relevant to the interpretation of behaviour. The rhesus macaque is thus an ideal subject to probe for possible behavioural indicators of metacognition, as well as details regarding the complexities of what constitutes “reward,” “reinforcement,” or “punishment.”

## 2. Materials and methods

### 2.1 Subject

The subject was Murph, a 19 year-old male rhesus macaque (*Macaca mulatta*) at the Language Research Center, Georgia State University. He had previously participated in a wide range of psychological tests, including metacognition studies (Beran et al., 2006; Smith et al., 2006; Couchman et al., 2010; Smith et al., 2010; Beran & Smith, 2011; Smith et al., 2013; Beran et al., 2014; Zakrzewski et al., in press).

### 2.2 The task

Murph engaged in a psychophysical sparse-dense discrimination task. The monkey was not food deprived for testing, water was continuously available during testing, and engagement was voluntary. The task was presented on a computer screen, and responses were made with a joystick, held by the monkey, which controlled the cursor. On each trial, he saw a 185 × 185-pixel box in the screen’s top center. The box was filled with a varying number of randomly placed lit white pixels on a black background. Sixty stimulus levels could be presented (Levels 1–60). Each level’s pixel count was given by the formula: pixels = round (base pixels × 1.018<sup>Level</sup>). The base pixels were 400. This formula gave the continuum a logarithmic character where density steps were defined as a constant percentage increase in pixels and not as a constant absolute increase in pixels. The stimulus continuum was divided into sparse and dense regions. Across a session, approximately half of the trials were sparse (Levels 1–30, 407–682 pixels) and dense (Levels 31–60, 694–1163 pixels). Approximately 60 percent of the presented trials were restricted to only the central third of the range (Levels 21–40) whereas the other 40 percent of the trials were sampled from the full continuum (Levels 1–60).

On each trial, Murph brought the cursor into contact with either an S or D icon on the screen to classify the boxes as sparse or dense, respectively. Correct responses led to automated delivery of a single food pellet to a reward cup which Murph could manually access and a 1-second inter-trial interval before presentation of the next trial. Incorrect responses led to a 30-second timeout period during which the screen remained blank. Cursor contact with a “?” icon on the screen (also present on all trials) cleared the screen of that trial, gave Murph no auditory or visual feedback, no food reward, and no timeout penalty. This was the uncertainty response (UR) and after a 1-second interval the next trial was presented. There

was no guarantee that the next trial would be easier (or harder) than the previous trial, as this was randomly determined according to the constraints outlined above. No computer sounds were associated with any response type regardless of outcome.

For each trial, the computer automatically recorded the pixilation of the test box, its true qualitative density region (sparse-dense), the time at which Murph first deflected the joystick, Murph's selection of icon (S, D, or ?), and outcome (correct, incorrect, or UR). The session was video recorded, with one camera focused on Murph in profile and the other on the cup in which pellets were dispensed.

This paper involves a detailed behavioural consideration of one session which occurred on 2 February 2012. Murph was presented the computerized task during a 342 minute test session, during which he could choose when to work or rest. This session was not selected from a larger databank of similar recorded sessions, but instead was conducted and recorded specifically for the assessments outlined below.

### 2.3 Behavioral coding

**Reliability**—For reliability on all behavioral measures outlined below, a subset of > 400 trials, chosen pseudo-randomly from sets of 40 continuous trials, were scored independently by a separate observer. In addition, for the same subset of trials, the variable “wavering” was also scored independently by a third observer naïve to the focus of the study.

**Behavioral correlates of the UR**—The time between Murph first orienting his head to a visible trial to when he deflected the joystick (time before deflection), and the time taken to move the cursor from starting location to the chosen icon (time for response) were recorded for all trials with a split-time stopwatch. In addition, the directness or non-directness of cursor path to icon (“wavering”) was scored, for each trial, using a 1 to 4 scale. A wavering score of “1” reflected approximately direct movement from starting point to chosen icon; a score of “2” reflected movement to one icon with a discernable pause en route, but with no appreciable change in direction; “3” reflected movements towards 2 different icons (e.g., towards S and then a change in direction toward D); and “4” reflected multiple movements between two icons (e.g., towards S, then D, and back towards S) or movements towards 3 icons (e.g., towards S, then ?, and then D).

**Acquisition and consumption of pellets**—The video of the session was also scored on a trial-by-trial basis for the following discrete variables related to pellet acquisition: the total number of pellets accumulated in the reward cup when each trial became visible, the number of pellets placed in the mouth with a given trial visible (before joystick deflection), the number of pellets placed in the mouth during response, and the number of pellets placed in the mouth during the blank screen after a given trial. Note that the above variables include pellets earned from previous trials and/or the current trial, and that the duration of the blank screen was 1 s after correct responses and URs and 30 s after incorrect responses. In addition, instances where masticated food was deposited onto or licked off the cage (“cage-licking”) for each trial was scored as 0 (not observed) or 1 (observed).



## 2.4 Data analysis

All statistical tests described below were two-tailed with significance set at  $p < 0.05$ .

**Reliability**—Reliability for coding was computed, for discrete variables, as proportion agreement for all trials (and, where applicable, trial components) within each data category. For the “wavering” variable, as noted above, there were two reliability coders (and, thus, two reliability estimates). For time variables, reliability was reported as a Pearson’s  $r$  correlation coefficient.

**Behavioural correlates of the UR**—The objective difficulty of each trial was calculated as follows: objective difficulty (untransformed) =  $|30.5 - \text{level}|$ . This value was then transformed to give higher values for more difficult trials: objective difficulty (transformed) =  $30.5 - \text{objective difficulty (untransformed)}$ . All calculations and references to “objective difficulty” below are based on these transformed values. Thus, the trials ranged in difficulty from 1 (easiest) to 30 (most difficult). Pearson’s  $r$  correlation coefficients were utilized to assess relationships between trial number and responding correctly (coded 0/1) for those trials that were discriminated as sparse or dense, trial number and use of the UR (also coded 0/1), objective difficulty and percentage correct, and also between objective difficulty and percentage use of the UR. A Kruskal-Wallis statistic was used to compare the mean objective difficulty for various response types/outcomes (correct, incorrect, UR), and Mann-Whitney tests were utilized to compare the means of objective difficulty between pairs of these response types.

Descriptive statistics were calculated for time before deflection, time for response, and wavering. For these variables, a Kruskal-Wallis statistic was utilized to test the equality of the population means for the categories for correct, incorrect, and uncertain responses. Mann-Whitney tests were used to assess comparisons between pairs of these outcome categories. In order to control for objective difficulty, linear regressions were also constructed with dependent variable (time before deflection, time for response, wavering) predicted by objective difficulty and any one of the three response/outcome types (correct/other, incorrect/other, or uncertain response/other; all coded as 1/0). To test for possible differences in behavior across the test session, Pearson’s correlations between trial number and time before deflection, time taken for response, and wavering, were also calculated.

**Acquisition and consumption of pellets**—Descriptive and/or frequency statistics also were calculated for the following: total number of pellets earned over all trials, total number of pellets placed in the mouth over all trials, pellets accumulated in reward cup at time  $x$ , and the number of pellets placed in the mouth for each trial and trial component (trial visible to subject, subject working on trial, blank screen after trial). In addition, the frequency of cage-licking (i.e., using the tongue to deposit or remove masticated food to and from the cage wire) was also computed for trials in each of the three response/outcome categories (correct, incorrect, UR).

### 3. Results

#### 3.1 Reliability

There was close agreement (designated  $a$ ) on inter-observer reliability for the coded discrete behavioral variables. These include: the total number of pellets accumulated at the beginning of each trial ( $n = 1728$ ,  $a = 0.99$ ), the number of pellets placed in the mouth with a given trial visible ( $n = 447$ ,  $a = 0.96$ ), the number of pellets placed in the mouth during each response ( $n = 447$ ,  $a = 0.96$ ), the number of pellets placed in the mouth during each blank screen ( $n = 444$ ,  $a = 0.995$ ), cage-licking ( $n = 440$ ,  $a = 0.94$ ), and wavering (Observer 2,  $n = 438$ ,  $a = 0.91$ ; Observer 3,  $n = 437$ ,  $a = 0.87$ ). In addition, the two continuous time variables were strongly correlated between independent observers (latency to response, Pearson's  $r = 0.86$ ,  $n = 439$ ,  $p < 0.001$ ; response time,  $r = 0.86$ ,  $n = 439$ ,  $p < 0.001$ ).

#### 3.2 Behavioral correlates of the UR

Murph completed 1728 trials in 342 minutes of testing (5.05 trials/min), with this time frame including several lengthy ( $> 5$  min) rest periods. The pattern of the monkey's responses paralleled those of previous metacognition studies (Smith, 2009). Trials at low or high density levels (those of low objective difficulty) were generally categorized correctly by Murph as sparse or dense, respectively, while those at middle levels (i.e., those of high objective difficulty) were characterized by increased use of the UR (Figure 1) and decreased levels of classification as dense or sparse. With respect to trial responses/outcomes, Murph responded correctly 957 times, incorrectly 285 times, and used the UR 486 times (Table 1). For those trials which Murph categorized as sparse or dense, there was a significant positive correlation between trial number and responding correctly; i.e., he was more accurate later in the test session ( $n = 1242$ ,  $r = 0.23$ ,  $p < 0.001$ ). There was also a significant positive correlation between trial number and use of the UR; i.e., Murph escaped trials more often later in the test session ( $n = 1728$ ,  $r = 0.44$ ,  $p < 0.001$ ).

There was a significant negative correlation between objective trial difficulty and percentage correct when the dense or sparse icons were selected ( $r = -0.92$ ,  $n = 30$ ,  $p < .001$ ). Thus, as objective difficulty increased, performance decreased. There was also a significant positive correlation between objective trial difficulty and percentage choice of the uncertainty response ( $r = 0.94$ ,  $n = 30$ ,  $p < .001$ ). Thus, in accordance with predictions, as objective trial difficulty increased, the percentage of trials on which Murph selected the uncertainty response increased. Looked at in another way, mean objective difficulty differed among trial outcomes (Kruskal-Wallis, Chi-square = 215.92,  $df = 2$ ,  $p < 0.001$ ) with low objective difficulty associated with correct responses, and relatively high objective difficulty associated with incorrect responses and URs (Figure 2). Incorrect responses and URs did not differ significantly with respect to this measure (Mann-Whitney  $U = 68016.00$ ,  $N1 = 285$ ,  $N2 = 486$ ,  $p = 0.68$ ).

The time between Murph orienting towards a visible trial and initial joystick deflection differed by trial outcome (Kruskal-Wallis, Chi-square = 62.61,  $df = 2$ ,  $p < 0.001$ ) with short latencies on correct trials and longer latencies on URs and incorrect responses (Figure 3). Again, incorrect responses and URs did not significantly differ in this measure (Mann-

Whitney  $U = 64818.00$ ,  $N1 = 284$ ,  $N2 = 485$ ,  $p = 0.17$ ). Thus, before joystick movement, Murph generally viewed trials he eventually escaped or got incorrect for longer periods and, as noted above, these were associated with higher objective difficulty than “correct” trials. When objective difficulty was controlled in three independent regression models (one for each response/outcome type), use of the uncertainty response ( $B = 0.13$ ,  $t = 3.64$ ,  $p < 0.001$ ) and responding incorrectly ( $B = 0.10$ ,  $t = 2.23$ ,  $p = 0.026$ ) remained significant positive predictors of time before deflection, while responding correctly was a significant negative predictor ( $B = -0.18$ ,  $t = -5.20$ ,  $p < 0.001$ ). There was a significant negative correlation between trial number and time before deflection; i.e., as the session progressed, Murph tended to take less time to initiate responses ( $n = 1726$ ,  $r = -0.07$ ,  $p = 0.007$ ).

The time taken for the actual response—from movement of cursor to contact with icon—also differed between outcome types (Kruskal-Wallis, Chi-square = 165.49,  $df = 2$ ,  $p < 0.001$ ), with the longest mean response durations associated with URs (Figure 4). In this case, URs were associated with significantly longer response times than correct or incorrect trials (Mann-Whitney  $U$ , all  $p$ 's  $< 0.001$ ), while correct and incorrect responses did not differ in this measure (Mann-Whitney  $U = 132473.50$ ,  $N1 = 957$ ,  $N2 = 284$ ,  $p = 0.52$ ). In this case, however, when objective difficulty was controlled in independent regression models, none of the three response types was a significant predictor of time taken for response (correct,  $B = -0.02$ ,  $t = -0.51$ ,  $p = 0.61$ ; incorrect,  $B = -0.03$ ,  $t = -0.72$ ,  $p = 0.47$ ; UR,  $B = 0.04$ ,  $t = 1.15$ ,  $p = 0.25$ ). There was, in addition, a significant positive correlation between trial number and time taken for response ( $n = 1726$ ,  $r = 0.10$ ,  $p < 0.001$ ).

High degrees of “wavering” accompanied trials where the UR was utilized (Figure 5). Wavering, as defined above, is related to the directness of travel to the chosen icon, with higher scores related to pausing during response and moving back and forth between icons. Wavering scores differed between outcome types (Kruskal-Wallis, Chi-square = 16.26,  $df = 2$ ,  $p < 0.001$ ) with the highest mean scores associated with URs, which differed from correct and incorrect responses (Mann-Whitney, all  $p$ 's  $< 0.01$ ). Interestingly, correct responses did not differ significantly in wavering scores from incorrect responses (Mann-Whitney  $U = 130636.00$ ,  $N1 = 956$ ,  $N2 = 284$ ,  $p = 0.15$ ) and correct responses actually were associated with higher absolute mean values with regards to this measure (Figure 5). This result also held when objective difficulty was controlled in independent regression models; the use of the uncertain response was a positive predictor of wavering ( $B = 0.11$ ,  $t = 2.99$ ,  $p = 0.003$ ) while responding incorrect was a significant negative predictor ( $B = -0.11$ ,  $t = -2.58$ ,  $p = 0.01$ ). Correctly responding, in contrast, was not a significant predictor of wavering ( $B = -0.026$ ,  $t = -0.76$ ,  $p = 0.45$ ). In addition, there was a significant positive correlation between trial number and wavering; this is likely related to the above finding that the UR was employed significantly more often later in the test session ( $n = 1725$ ,  $r = 0.15$ ,  $p < 0.001$ ).

These results paint a unique pattern of behavior that is associated solely with use of the UR. These trials are notable for their comparatively long latencies before joystick deflection and prolonged durations for the responses themselves, but, as noted above, these are not necessarily differentiated from other response types on trials of similar objective difficulty. The most apparent contrast lies in wavering; trials where the UR was utilized, regardless of difficulty, were characterized by high occurrences of pausing during response and

movement back and forth between icons. Such behavior has long been considered a potential indicator of derived, cognitive processing (Tolman, 1927; Frith, 2012), although they also have been viewed more conservatively as manifestations of struggle or confusion which occur near perceptual thresholds (see Tolman, 1938; Smith et al., 1995; Couchman et al., 2012) and an implicit cue signaling a challenging problem, which encourages the seeking of additional information or alternative solutions (Carruthers, 2008). The results given here contrast with the latter interpretation, as incorrect trials, which almost invariably occurred on trials of high objective difficulty, were characterized by the lowest degrees of wavering. This result also hints at the adaptive nature of hesitancy behaviors (see Discussion).

### 3.3 Acquisition and consumption of pellets

Murph's acquisition and consumption of pellets did not follow the temporal patterning assumed in reinforcement models attempting to explain putative metacognitive performance. He earned 957 pellets during the session (Table 1 – "Correct Responses") of which all but one (which was dropped) were placed in the mouth and/or cheek pouches (hereafter, "mouth"). The number of pellets accumulated in the reward cup at the beginning of individual trials during the session ranged from 0 to 5. Pellets were placed in the mouth during and after trials of all response/outcome categories. During incorrect trials, for example, 91 pellets were placed into his mouth with the trial stimuli visible (0.32 pellets/trial), 87 while making incorrect responses (0.31 pellets/trial), and 42 during the 285 "punishment" periods (0.15 pellets/trial). In one striking (though non-exceptional) example, the monkey accumulated 5 pellets and then placed them in his mouth during a highly rewarding timeout after an incorrect response (Figure 6). The timeout (and the response that led to that timeout) thus delivered, in this instance, more reinforcement than the five previous correct trials, if one considers pellets entering the mouth as the most likely candidate for the reinforcing aspect of this task's design. Murph was also observed to masticate pellets, deposit them onto the cage wire, and consume them many seconds or minutes later ("cage-licking"). Cage-licking, like the placement of pellets in the mouth, was again observed during all outcome categories, including incorrect responses, where it occurred during 192/285 (0.67) of trials, sometimes before and almost always during the "punishment" period. Thus, Murph received some variety of reward (whole pellet or masticated pellet pieces from cage-licking) in the majority of the "punishment periods" following incorrect trials (Table 1).

This does not mean, of course, that there was no association between correct responses and ingesting pellets. However, as Murph moved on to new trials rapidly, and generally before dealing with any pellets earned, any such associations would almost invariably need to be from earlier trials. This was often, but with great variance, the immediately preceding trial; the mean number of pellets placed in mouth in the subsequent trial following a correct response was 0.92 (+/- s.d. 0.75). Looking at this in more detail, in the trials following a correct response, the average number of pellets placed in mouth with the next trial visible was 0.39 (+/- 0.57), while working on the next trial was 0.47 (+/- 0.75), and during the blank screen after the next trial 0.07 (+/- 0.39).

Taken together, these data indicate that: 1) Murph was generally working on at least the subsequent trial before an earned pellet was placed in the mouth, 2) there was considerable variation as to when he placed earned pellets in his mouth, and 3) his patterns of behavior (e.g., delayed consumption of pellets, cage-licking) reduced the aversive nature of intended time out punishments, at least if those are viewed as being aversive because they are not associated with pellet consumption. These results contrast with associative models aiming to explain potential metacognitive performance in nonhumans, as “reinforcement” (as assumed in these models, but not necessarily accepted by the authors of the present paper) was only very loosely tied with trial outcomes.

#### 4. Discussion

Murph, the rhesus monkey in this study, “misbehaved” in two senses. The first was that he showed cautious, deliberate action associated specifically with use of the UR and not during other trials of comparable difficulty. This pattern of behavior is suggestive of controlled cognitive processing (see also Smith et al., 1995), is not predicted by general associative models involving primitive psychological mechanisms, and indeed parallels some aspects of explicit metacognition in humans (Frith, 2012). Trials where the UR was employed were characterized by relatively long latencies before beginning visible trials, prolonged times actually working on trials, and, even after controlling for difficulty, high degrees of wavering during the responses. This parallels some aspects of human performance in metacognition tasks, where individual confidence ratings negatively correlate with response time variables (humans, in addition, can estimate another individual’s confidence through such behavioral cues, Patel et al., 2012).

In contrast, the exceptionally low levels of wavering associated with incorrect responses—which were essentially identical to UR trials with respect to their objective difficulty—strongly suggest that hesitation and movements back and forth between icons represented more than a struggle at perceptual threshold. The results in this respect are again not congruent with published associative models which attempt to explain nonhuman (and potentially human, Kornell, 2013) performance in uncertainty monitoring tasks (Jozefowicz et al., 2009a; Le Pelley, 2012). In addition, they point to the essentially adaptive nature of “hesitancy” behavior in the computer task at hand, and which could easily be extended to myriad situations in nature. “Looking before you leap” (Tolman, 1938, p. 27) can be useful, whether practiced at a computer monitor or a break in the canopy of an Indian forest. Working out the genetic and neurological underpinning(s), and particularly the selective pressures that would favor an executive monitoring system for such situations, should represent an important future goal for elucidating the phylogenetic history of metacognition and its precursors. Early studies, for example, suggested that wavering behavior could be specifically interrupted by certain types of cerebral damage (Tolman, 1938).

Is it possible, however, that wavering or similar actions, such as anxiety-related behaviors, could be viewed as merely a “cue” signaling that a particular course of action should be carried out (Hampton, 2009)? Carruthers (2008), for example, has suggested that some animals may possess a “gate-keeping” mechanism that initiates hesitation in problems of great difficulty, and encourages information-seeking and/or alternative solutions (such as

use of the UR). It is argued that the gate-keeper, in nonhuman forms, works implicitly and to some degree automatically, within an associative framework. The appeal of this hypothesis is that it clearly addresses the question of function; with relation to the current and previous studies, however, these same functional concerns present great difficulties. In order for such an inflexible system to be adaptive, behaviours such as wavering should consistently appear when highly challenging problems present themselves. Murph, however, exhibited both the highest (UR) and lowest (incorrect responses) degrees of wavering when presented with such difficult trials. Work with other monkeys and apes, in addition, counters some aspects of an implicit gate-keeper that is driven by hesitation or anxiety; for example, information-seeking behaviour in metacognition tasks is not random, as such a model would suggest, but discriminates between the types of information that the subject does or does not possess (Call, 2012). This research speaks to great complexity in behaviour, as opposed to inflexibility, and is congruent with the nervous system morphology of primates, as well as their particularly nuanced social behavior and marked ecological generalism (Sayers, 2013).

Associative explanations for nonhuman metacognition studies, in addition, rely in part on the seemingly reasonable assumption that reinforcement is meted out according to the dictates of the experimenter. An animal, it is envisioned, engages a task; if successful, the subject receives “reward” at a designated time—and if unsuccessful the subject receives the proscribed punishment—before moving on to another trial. Even in situations of delayed reinforcement (Lattal, 2010) the experimental design is elegant and regimented. The sticking point is that while nonhuman animals generally behave adaptively—even when presented tasks more appropriate for featherless bipeds than the particular subject in question—they are all too often inelegant and unregimented while doing so.

This brings us to the second sense in which Murph “misbehaved,” by consuming pellets in a fashion that conflicted with the reinforcement contingencies outlined in associative models of metacognition. Responses were made in quick succession, with the exception of some uses of the UR, and rarely were rewards consumed immediately after completion of a correct trial and before beginning the next trial. Part of this was undoubtedly driven by the task’s design, as only 1 second separated correct responses and URs from the presentation of the next trial. Even beyond this, however, the monkey showed great variance as to when he actually attended to and consumed pellets. A good visual analogy for this is a human eating popcorn while watching a movie (or, more accurately, playing a video game).

Indeed, pellets were frequently allowed to accumulate, and were eaten while working on—or being punished for—incorrect trials. The monkey stored moistened pellets on the cage wire for later consumption, which, again, frequently occurred during punishments. The data in this study weaken the varied arguments (Jozefowicz et al., 2009a; Jozefowicz et al., 2009b; Le Pelley, 2012) that use of the UR is simply an associative mechanism serving to avoid punishment, as Murph frequently utilized punishment periods to consume his earnings. Like the Brelands’ raccoons and pigs, which behaved like raccoons and pigs, respectively, Murph was simply acting like a rhesus macaque. From a highly social, manually manipulative, cheek-pouched species—where foods are often transported before they are swallowed (Murray, 1975; Lambert, 2005)—he was consuming pellets in a fashion that both reduced the intensity of punishment and hastened his abilities to complete trials

rapidly. The strategy was global and not local, much in line with the long-term rate maximization strategy from optimal foraging theory (Stephens & Krebs, 1986; cf. Sayers & Menzel, 2012).

But while rhesus macaques, like any other animal, have their natural inclinations, this concerns more than instinct or fixed action patterns. It is about the genetically-based flexibility of behavior that is found in all primates; a flexibility that allows rapid and efficient mastering of tasks that the animals have never encountered before (Kummer, 1971). To describe the behaviour of rhesus macaques using only the psychological mechanisms which apply to pigeons is likely to misunderstand the rhesus macaque and to misunderstand evolution.

The main point is not, of course, that the intended reward regime does not in some way guide behaviour. It is rather that the results of this study, like the delayed feedback procedures described earlier (Smith et al., 2006), again suggest a further one-step removal of “reinforcement” as defined in current associative models (i.e., food delivery and consumption) and putative metacognitive performance. Indeed, even describing what constitutes reinforcement is exceptionally difficult in such a study (Jensen, 1963; Martin, 1979) and the associative models described previously, with a focus on food delivery/consumption alone, appear particularly inadequate when applied to primates. Under some circumstances, for example, capuchin monkeys will manipulate freely available apparatuses at higher rates than they ingest freely available food (Premack, 1959). Rhesus monkeys, in addition, will continue to perform computer tasks when free food is available, will show similar performance levels in such circumstances, and will overwhelmingly prefer to work to earn food over 30-minute blocks than to receive free food with the tasks unavailable (Washburn & Rumbaugh, 1992). What constitutes “reinforcement” is not static, can change over time (e.g., ingestion could reinforce task activity, or task activity could reinforce ingestion), and is likely heavily dependent on the state of the animal (Premack, 1959; Premack, 1962). In this regard, it is important to note that in none of the rhesus monkey metacognitive studies performed to date have the animals been food-deprived for testing, a situation (near-satiated monkeys) in which food could potentially lose some of its salience.

The mere ability to play the computer game, the satisfaction of getting a problem right, the delivery of accessible pellets, the manual holding of pellets, the placing of pellets in the mouth or cheek pouches, the depositing of moistened pellets onto the cage, the licking of moistened pellets off the cage, the actual swallowing of said food, and myriad other factors could all serve to influence behaviour in the current study. The task’s microclimate of reinforcement and food suggests that using a simple associative model to explain the results of uncertainty monitoring studies is insufficient because rewards and punishments are so diffuse, blurred, and overlapping. Given all the variables which can guide action, as noted above, it would be difficult to envision an operant approach that could fully circumvent this problem. Such factors likely underpin the difficulties associative models have faced in accounting for large-scale patterns of response in nonhuman metacognitive studies (Smith et al., 2014), as well as the more specific behaviors reported presently.

Such a conclusion meshes well with observations on primates made in the classical days of learning theory. Harlow and Mears (Harlow & Mears, 1979) noted that their rhesus macaques learned more quickly when they were fed before testing, as opposed to when food-deprived. Many monkeys began sessions with engorged cheek pouches, and would add to the store after correct responses and swallow food at varying times whether trials were completed correctly or not. “It is obvious that under these conditions the monkey cannot learn,” Harlow wrote, “but I developed an understandable skepticism of this hypothesis when the monkeys stubbornly persisted in learning, learning rapidly, and learning problems of great complexity” (p. 93). This is what primates, whether human or otherwise, have evolved to do (Kummer, 1971; Count, 1973).

Strict adherence to a purportedly “skeptical,” introductory textbook version of Morgan’s canon—an insistence on labyrinthine models of “lower” processes even when more straightforward models incorporating “higher” ones are available—is likely a disservice to Morgan, to comparative psychology, and, most strikingly, to evolutionary biology. As in papers describing associative models for certain nonhuman metacognition studies (e.g., Carruthers, 2008; Jozefowicz et al., 2009a), this conflates Morgan’s canon with Occam’s razor, and ignores the aspects of phylogeny and evolutionary history that are so vital to interpreting an animal’s behavior. Morgan amended his canon (see Karin-D’Arcy, 2005) to discourage the misuse which occurred then and which persists over a century later; namely, “that the canon by no means excludes the interpretation of a particular activity in terms of higher processes, if we already have independent evidence of the occurrence of these higher processes in the animal under observation” (Morgan, 1903:59). For Old World monkeys and apes, such independent evidence, as noted above, is becoming quite plentiful in the case of metacognition.

In evolutionary biology—which long ago dispensed with terms such as “lower” and “higher” in relation to structures or processes—Morgan’s canon can today be more profitably utilized by substituting the concepts of primitive and derived characters (Sober, 2005). In what respect do rhesus macaques possess derived characters relevant to metacognition? As related above, macaques are primates, share numerous gross and micro-anatomical brain characters with humans (Barbas, 2000; Raghanti et al., 2009), and mirror human performance in numerous comparative tests of metacognition. Parsimony, in such cases, would tend to support the hypothesis of similar underlying mechanisms (Smith, 2007). It would be difficult to convince any primate ecologist who has observed a monkey hesitating before jumping across a gap in the canopy that these organisms are incapable of sensing uncertainty (see Tolman, 1927; Tolman, 1938; Carruthers, 2008; Smith, 2009). While such subjective judgments alone merely provide a means for generating hypotheses, uncertainty monitoring abilities would have clear selective value, and would be expected to evolve given the requisite genetic variation. While humans clearly have a plethora of unique, derived traits (Sayers & Lovejoy, 2008; Sayers et al., 2012), metacognition as a controlled, and perhaps executive, function is looking increasingly unlikely to be one of them, even if our systems differ from those of other animals in terms of certain conscious and self-reflective aspects.



The procedures utilized in this study were approved by the Institutional Animal Care and Use Committee of Georgia State University.

## Acknowledgments

We would like to thank Dr. Charles Menzel, Dr. Bonnie Purdue, Dr. Megan Hoffman, and two anonymous reviewers for detailed critiques of an earlier version of this paper, and Kathryn Anderson for reliability coding. The preparation of the manuscript was supported by NICHD Grant 1R01HD061455 and NSF Grant BCS-0956993.

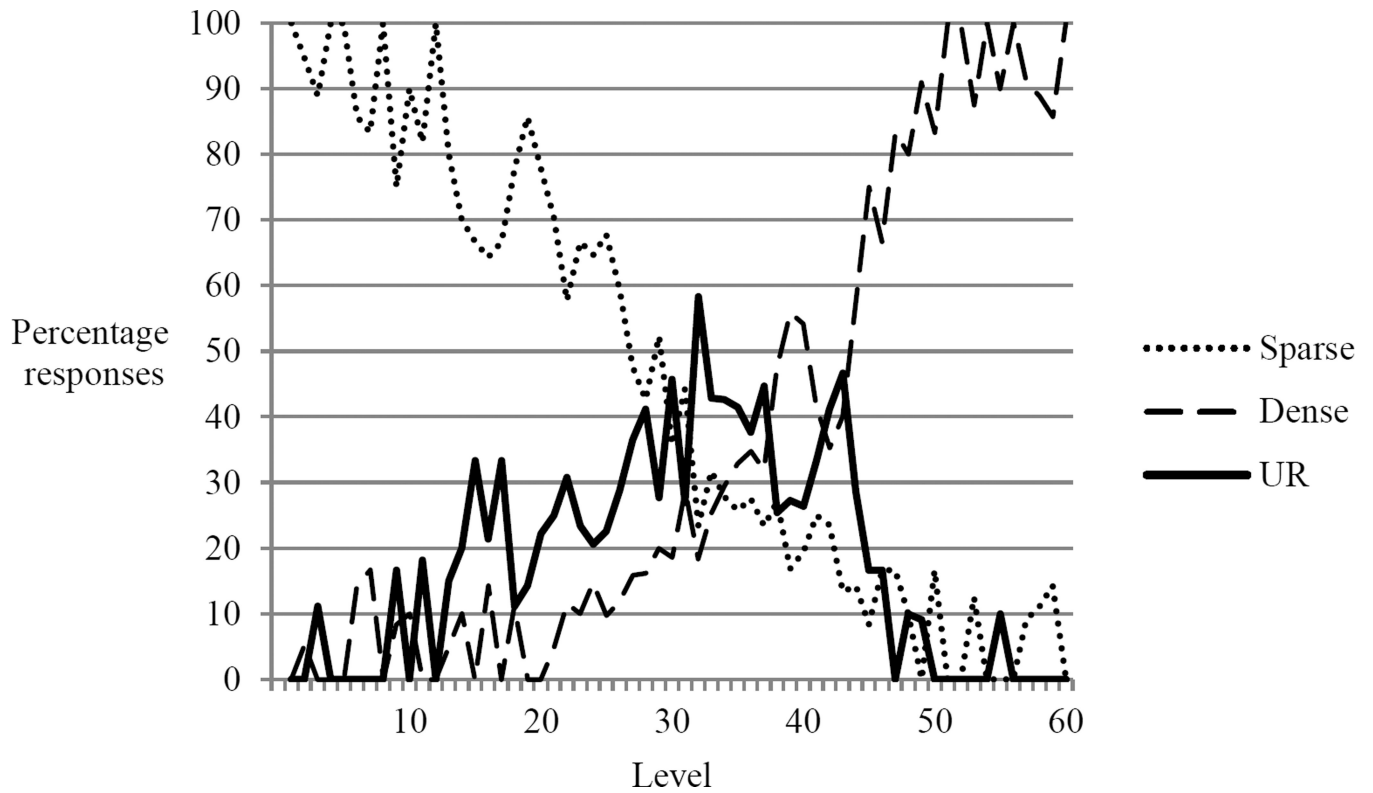
## References

- Adams A, Santi A. Pigeons exhibit higher accuracy for chosen memory tests than for forced memory tests in duration matching-to-sample. *Learn. Behav.* 2011; 39:1–11. [PubMed: 21264568]
- Balcomb FK, Gerken LA. Three year old children can access their own memory to guide responses on a visual matching task. *Dev. Sci.* 2008; 11:750–760. [PubMed: 18801131]
- Barbas H. Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Res. Bull.* 2000; 52:319–330. [PubMed: 10922509]
- Beran, MJ.; Brandl, JL.; Perner, J.; Proust, J. On the nature, evolution, development, and epistemology of metacognition: introductory thoughts. In: Beran, MJ.; Brandl, JL.; Perner, J.; Proust, J., editors. *Foundations of metacognition*. Oxford: Oxford University Press; 2012. p. 1-18.
- Beran MJ, Perdue BM, Smith JD. What are my chances? Closing the gap in uncertainty monitoring between rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Learning and Cognition.* 2014; 40:303–316. [PubMed: 25368870]
- Beran MJ, Smith JD. Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition.* 2011; 120:90–105. [PubMed: 21459372]
- Beran MJ, Smith JD, Coutinho MVC, Couchman JJ, Boomer J. The psychological organization of “uncertainty” responses and “middle” responses: A dissociation in capuchin monkeys (*Cebus apella*). *J. Exp. Psychol. Anim. Behav. Process.* 2009; 35:371–381. [PubMed: 19594282]
- Beran MJ, Smith JD, Perdue BM. Language-trained chimpanzees name what they have seen, but look first at what they have not seen. *Psychol. Sci.* 2013; 24:660–666. [PubMed: 23508741]
- Beran MJ, Smith JD, Redford JS, Washburn DA. Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *J. Exp. Psychol. Anim. Behav. Process.* 2006; 32:111–119. [PubMed: 16634654]
- Breland K, Breland M. The misbehavior of organisms. *Am. Psychol.* 1961; 16:681–684.
- Butler R. Discrimination learning by rhesus monkeys to visual-exploration motivation. *J. Comp. Physiol. Psychol.* 1953; 46:95–98. [PubMed: 13044866]
- Call, J. Seeking information in non-human animals: weaving a metacognitive web. In: Beran, MJ.; Brandl, JL.; Perner, J.; Proust, J., editors. *Foundations of Metacognition*. Oxford: Oxford University Press; 2012. p. 62-75.
- Call J, Carpenter M. Do apes and children know what they have seen? *Anim. Cogn.* 2001; 4:207–220.
- Carruthers P. Meta-cognition in animals: a skeptical look. *Mind Lang.* 2008; 23:58–89.
- Couchman, JJ.; Beran, MJ.; Coutinho, MVC.; Boomer, J.; Smith, JD. Evidence for animal metaminds. In: Beran, MJ.; Brandl, JL.; Perner, J.; Proust, J., editors. *Foundations of metacognition*. Oxford: Oxford University Press; 2012. p. 21-35.
- Couchman JJ, Coutinho MVC, Beran MJ, Smith JD. Beyond stimulus cues and reinforcement signals: a new approach to animal metacognition. *J. Comp. Psychol.* 2010; 124:356. [PubMed: 20836592]
- Count, EW. On the idea of protoculture. In: Menzel, EW., Jr, editor. *Precultural Primate behavior*. Basel: S. Karger; 1973. p. 1-25.
- Crystal JD, Foote AL. Metacognition in animals. *Comp. Cogn. Behav. Rev.* 2009; 4:1–16. [PubMed: 26516391]
- Emlen JM. The role of time and energy in food preference. *Am. Nat.* 1966; 100:611–617.

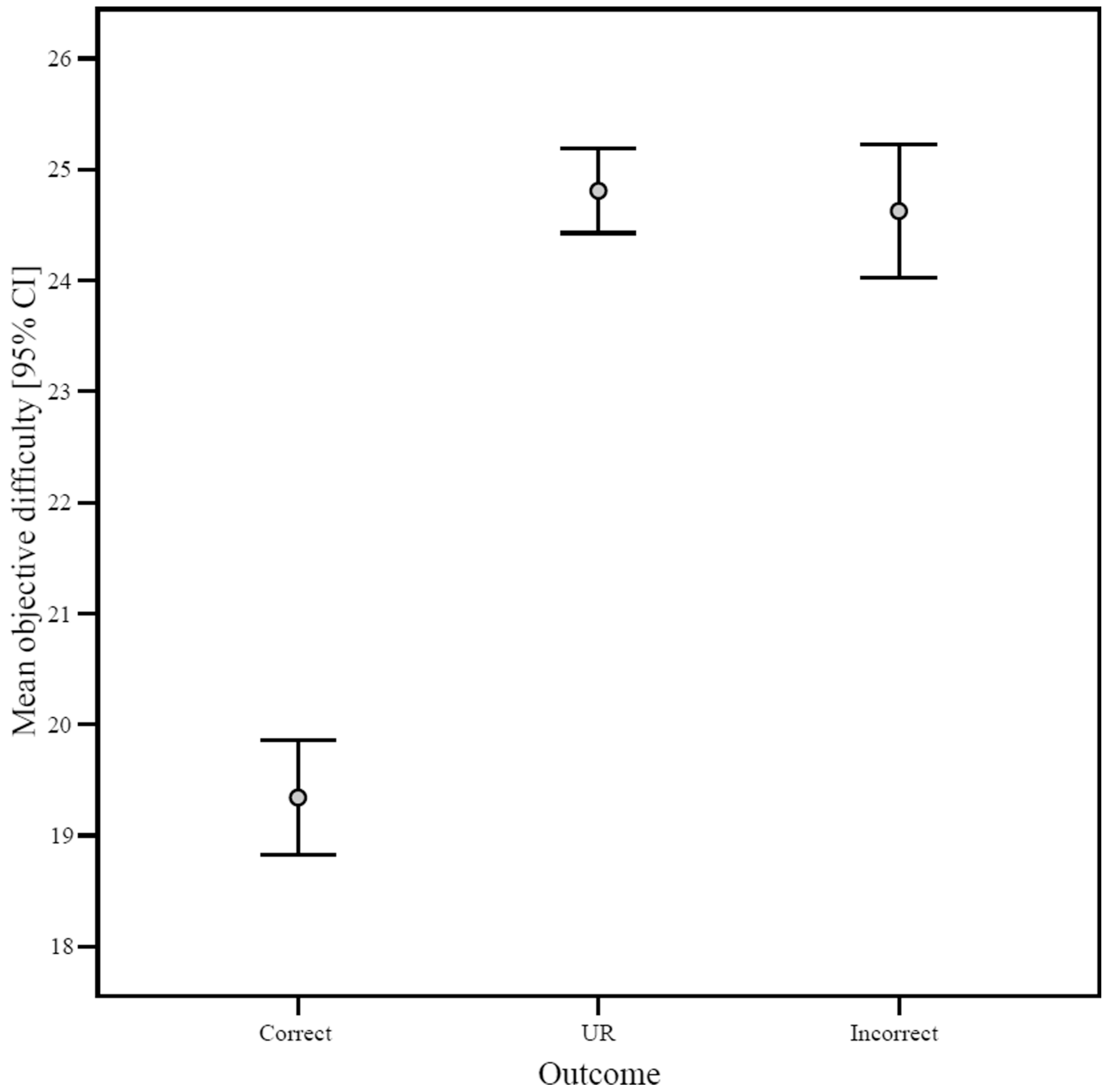
- Evans TA, Beran MJ. Delay of gratification and delay maintenance by rhesus macaques (*Macaca mulatta*). *J. Gen. Psychol.* 2007; 134:199–216. [PubMed: 17503695]
- Flavell JH. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *Am. Psychol.* 1979; 34:906.
- Frith CD. The role of metacognition in human social interactions. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 2012; 367:2213–2223. [PubMed: 22734064]
- Glazko GV, Nei M. Estimation of divergence times for major lineages of primate species. *Mol. Biol. Evol.* 2003; 20:424–434. [PubMed: 12644563]
- Hampton RR. Rhesus monkeys know when they remember. *Proc. Natl. Acad. Sci. U.S.A.* 2001; 98:5359–5362. [PubMed: 11274360]
- Hampton RR. Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms? *Comp. Cogn. Behav. Rev.* 2009; 4:17–28. [PubMed: 20046911]
- Hampton RR, Hampstead BM. Spontaneous behavior of a rhesus monkey (*Macaca mulatta*) during memory tests suggests memory awareness. *Behavioural Processes.* 2006; 72:184–189. [PubMed: 16460887]
- Harlow, HF.; Mears, C. *The human model: Primate perspectives.* Washington, D.C: V.H. Winston and Sons; 1979.
- Hawkes K, Hill K, O'Connell JF. Why Hunters Gather: Optimal Foraging and the Ache of Eastern Paraguay. *American Ethnologist.* 1982; 9:379–398.
- Hill, WCO. *Primates Comparative Anatomy and Taxonomy VII. Cynopithecinae: Cercocebus Macaca Cynopithecus.* Edinburgh: Edinburgh University Press; 1974.
- Hoffman M, Schwartz B. Metacognition does not imply self-reflection, but it does imply function. *J. Comp. Psychol.* 2014; 128:150–151. [PubMed: 24866005]
- Jensen GD. Preference for bar pressing over "freeloading" as a function of number of rewarded presses. *J. Exp. Psychol.* 1963; 65:451–454. [PubMed: 13957621]
- Jozefowicz J, Staddon J, Cerutti D. Metacognition in animals: How do we know that they know. *Comp. Cogn. Behav. Rev.* 2009a; 4:29–39.
- Jozefowicz J, Staddon J, Cerutti D. Reinforcement and metacognition. *Comp. Cogn. Behav. Rev.* 2009b; 4:58–60.
- Karin-D'Arcy MR. The modern role of Morgan's canon in comparative psychology. *Int. J. Comp. Psychol.* 2005; 18:179–201.
- Kornell N. Where is the "meta" in animal metacognition? *J. Comp. Psychol.* 2013; 128:143–149. [PubMed: 23875919]
- Kornell N, Son LK, Terrace HS. Transfer of metacognitive skills and hint seeking in monkeys. *Psychol. Sci.* 2007; 18:64–71. [PubMed: 17362380]
- Kummer, H. *Primate Societies: Group Techniques of Ecological Adaptation.* Arlington Heights: Harlan Davidson, Inc; 1971.
- Lambert JE. Competition, predation, and the evolutionary significance of the cercopithecine cheek pouch: the case of *Cercopithecus* and *Lophocebus*. *Am. J. Phys. Anthropol.* 2005; 126:183–192. [PubMed: 15386232]
- Lattal KA. Delayed reinforcement of operant behavior. *J. Exp. Anal. Behav.* 2010; 93:129–139. [PubMed: 20676272]
- Le Pelley M. Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals. *J. Exp. Psychol. Learn. Mem. Cogn.* 2012; 38:686. [PubMed: 22250905]
- Martin J. Laboratory studies of self-reinforcement (SR) phenomena. *J. Gen. Psychol.* 1979; 101:103–149. [PubMed: 114603]
- Menzel EW Jr, Menzel CR. Cognitive, developmental and social aspects of responsiveness to novel objects in a family group of marmosets (*Saguinus fuscicollis*). *Behaviour.* 1979; 70:251–279.
- Metcalfe, J.; Son, LK. Anoetic, noetic, and auto-noetic metacognition. In: Beran, MJ.; Brandl, JL.; Perner, J.; Proust, J., editors. *Foundations of metacognition.* Oxford: Oxford University Press; 2012. p. 289-301.
- Morgan, CL. *An introduction to comparative psychology.* London: Walter Scott; 1894.

- Morgan, CL. An introduction to comparative psychology. 2nd edition. London: Walter Scott; 1903.
- Murray, P. The role of cheek pouches in cercopithecine monkey adaptive strategy. In: Tuttle, RH., editor. Primate functional morphology and evolution. The Hague: Mouton Publishers; 1975. p. 151-194.
- Napier, JR.; Napier, PH. A handbook of living primates. London: Academic Press; 1967.
- Patel D, Fleming S, Kilner J. Inferring subjective states through the observation of actions. *Proceedings of the Royal Society B: Biological Sciences*. 2012; 279:4853–4860. [PubMed: 23034708]
- Premack D. Toward empirical behavior laws: I. Positive reinforcement. *Psychol. Rev.* 1959; 66:219–233. [PubMed: 13668003]
- Premack D. Reversibility of the reinforcement relation. *Science*. 1962; 136:255–257. [PubMed: 14488597]
- Raghanti MA, Spocter MA, Stimpson CD, Erwin JM, Bonar CJ, Allman JM, Hof PR, Sherwood CC. Species-specific distributions of tyrosine hydroxylase-immunoreactive neurons in the prefrontal cortex of anthropoid primates. *Neuroscience*. 2009; 158:1551–1559. [PubMed: 19041377]
- Roy AC, Paulignan Y, Farne A, Jouffrais C, Boussaoud D. Hand kinematics during reaching and grasping in the macaque monkey. *Behav. Brain Res.* 2000; 117:75–82. [PubMed: 11099760]
- Rumbaugh DM, Savage-Rumbaugh ES, Washburn DA. Toward a new outlook on primate learning and behavior: complex learning and emergent processes in comparative perspective. *Jpn. Psychol. Res.* 1996; 38:113–125. [PubMed: 11541528]
- Rumbaugh, DM.; Washburn, DA. *Intelligence of Apes and Other Rational Beings*. New Haven: Yale University Press; 2003.
- Sayers K. On folivory, competition, and intelligence: generalisms, overgeneralizations, and models of primate evolution. *Primates*. 2013; 54:111–124. [PubMed: 23263563]
- Sayers K, Lovejoy CO. The chimpanzee has no clothes: a critical examination of *Pan troglodytes* in models of human evolution (with comments and reply). *Curr. Anthropol.* 2008; 49:87–114.
- Sayers K, Menzel CR. Memory and foraging theory: chimpanzee utilization of optimality heuristics in the rank-order recovery of hidden foods. *Anim. Behav.* 2012; 84:795–803. [PubMed: 23226837]
- Sayers K, Norconk MA, Conklin-Brittain NL. Optimal foraging on the roof of the world: Himalayan langurs and the classical prey model. *Am. J. Phys. Anthropol.* 2010; 141:337–357. [PubMed: 19844998]
- Sayers K, Raghanti MA, Lovejoy CO. Human evolution and the chimpanzee referential doctrine. *Annu. Rev. Anthropol.* 2012; 41:119–138.
- Shields WE, Smith JD, Washburn DA. Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *J. Exp. Psychol. Gen.* 1997; 126:147. [PubMed: 9163934]
- Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 1977; 84:127–190.
- Smith, JD. Parsimony in comparative studies of cognition. In: Washburn, DA., editor. *Primate Perspectives in Behavior and Cognition*. Washington, D.C: American Psychological Association; 2007. p. 63-79.
- Smith JD. The study of animal metacognition. *Trends Cogn. Sci.* 2009; 13:389–396. [PubMed: 19726218]
- Smith JD, Beran MJ, Couchman JJ, Coutinho MVC. The comparative study of metacognition: sharper paradigms, safer inferences. *Psychon. Bull. Rev.* 2008; 15:679–691. [PubMed: 18792496]
- Smith JD, Beran MJ, Couchman JJ, Coutinho MVC, Boomer JB. Animal metacognition: Problems and prospects. *Comp. Cogn. Behav. Rev.* 2009; 4:40–53.
- Smith JD, Beran MJ, Redford JS, Washburn DA. Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring. *J. Exp. Psychol. Gen.* 2006; 135:282. [PubMed: 16719654]
- Smith JD, Couchman JJ, Beran MJ. The highs and lows of theoretical interpretation in animal-metacognition research. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 2012; 367:1297–1309. [PubMed: 22492748]

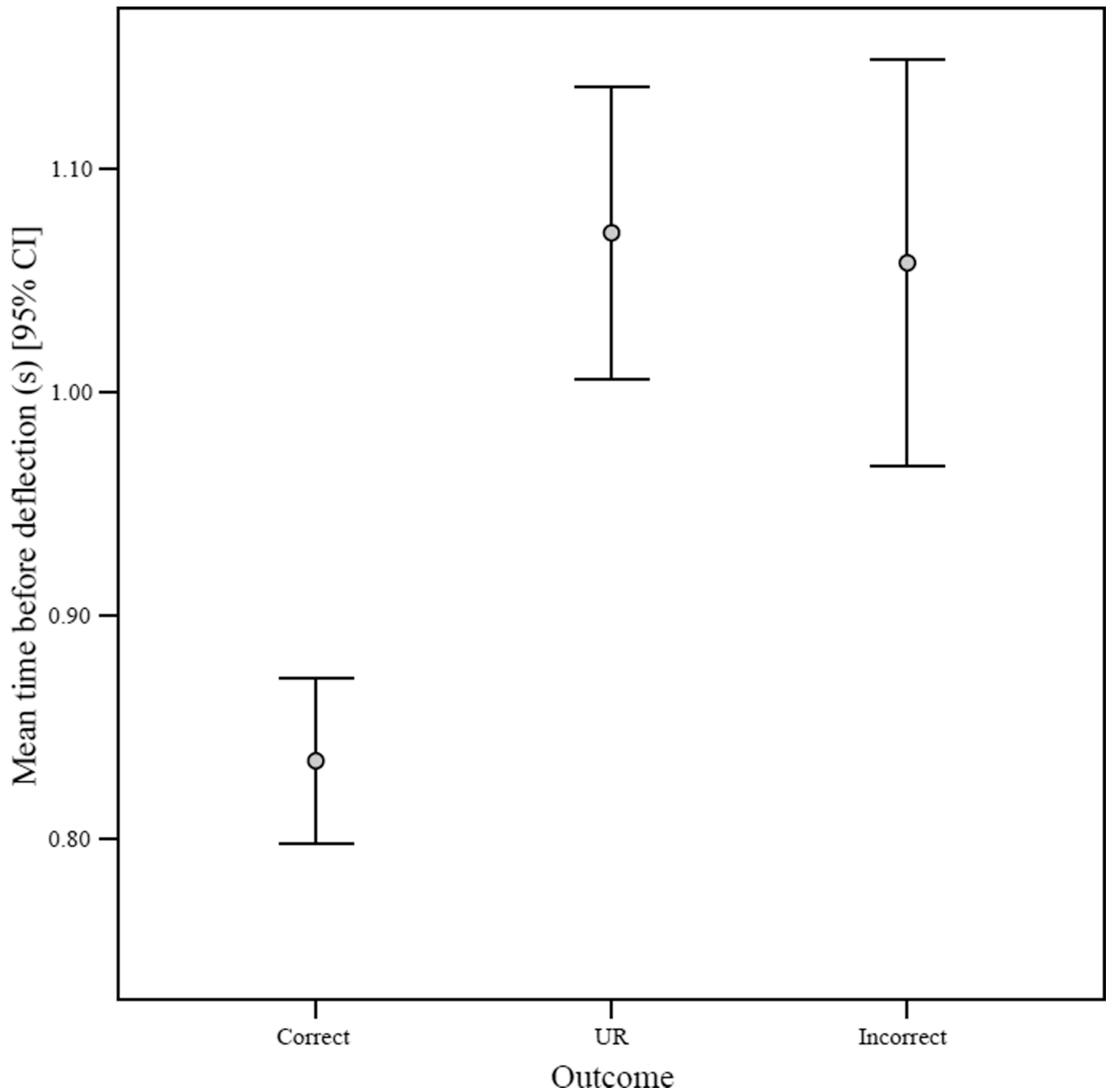
- Smith JD, Couchman JJ, Beran MJ. Animal metacognition: a tale of two comparative psychologies. *J. Comp. Psychol.* 2014; 128:115–131. [PubMed: 23957740]
- Smith JD, Coutinho MVC, Church BA, Beran MJ. Executive-attentional uncertainty responses by rhesus macaques (*Macaca mulatta*). *J. Exp. Psychol. Gen.* 2013; 142:458–475. [PubMed: 22889164]
- Smith JD, Redford JS, Beran MJ, Washburn DA. Rhesus monkeys (*Macaca mulatta*) adaptively monitor uncertainty while multi-tasking. *Anim. Cogn.* 2010; 13:93–101. [PubMed: 19526256]
- Smith JD, Schull J, Strote J, McGee K, Egnor R, Erb L. The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *J. Exp. Psychol. Gen.* 1995; 124:391. [PubMed: 8530911]
- Sober, E. Comparative psychology meets evolutionary biology.. jThinking with Animals. In: Daston, L.; Hitman, G., editors. New York: Columbia University Press; 2005. p. 85-99.
- Sole LM, Shettleworth SJ, Bennett PJ. Uncertainty in pigeons. *Psychon. Bull. Rev.* 2003; 10:738–745. [PubMed: 14620372]
- Staddon, JER. Adaptive behavior and learning. Cambridge: Cambridge University Press; 1983.
- Stephens, DW.; Krebs, JR. Foraging Theory. Princeton: Princeton University Press; 1986.
- Suda-King C. Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Anim. Cogn.* 2008; 11:21–42. [PubMed: 17437141]
- Suda-King C, Bania AE, Stromberg EE, Subiaul F. Gorillas' use of the escape response in object choice memory tests. *Anim. Cogn.* 2013; 16:65–84. [PubMed: 22923213]
- Tolman E. A behaviorist's definition of consciousness. *Psychol. Rev.* 1927; 34:433.
- Tolman E. The determiners of behavior at a choice point. *Psychol. Rev.* 1938; 45:1–41.
- Washburn DA, Gulledge JP, Beran MJ, Smith JD. With his memory magnetically erased, a monkey knows he is uncertain. *Biol. Lett.* 2010; 6:160–162. [PubMed: 19864272]
- Washburn DA, Rumbaugh DM. Investigations of rhesus monkey video-task performance: evidence for enrichment. *Contemporary Topics in Laboratory Animal Science.* 1992; 31:6–10. [PubMed: 11538195]
- Washburn DA, Smith JD, Shields WE. Rhesus monkeys (*Macaca mulatta*) immediately generalize the uncertain response. *J. Exp. Psychol. Anim. Behav. Process.* 2006; 32:185–189. [PubMed: 16634662]
- Zakrzewski AC, Perdue BM, Beran MJ, Church BA, Smith JD. Cashing out: The decisional flexibility of uncertainty responses in rhesus macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Learning and Cognition.* in press.



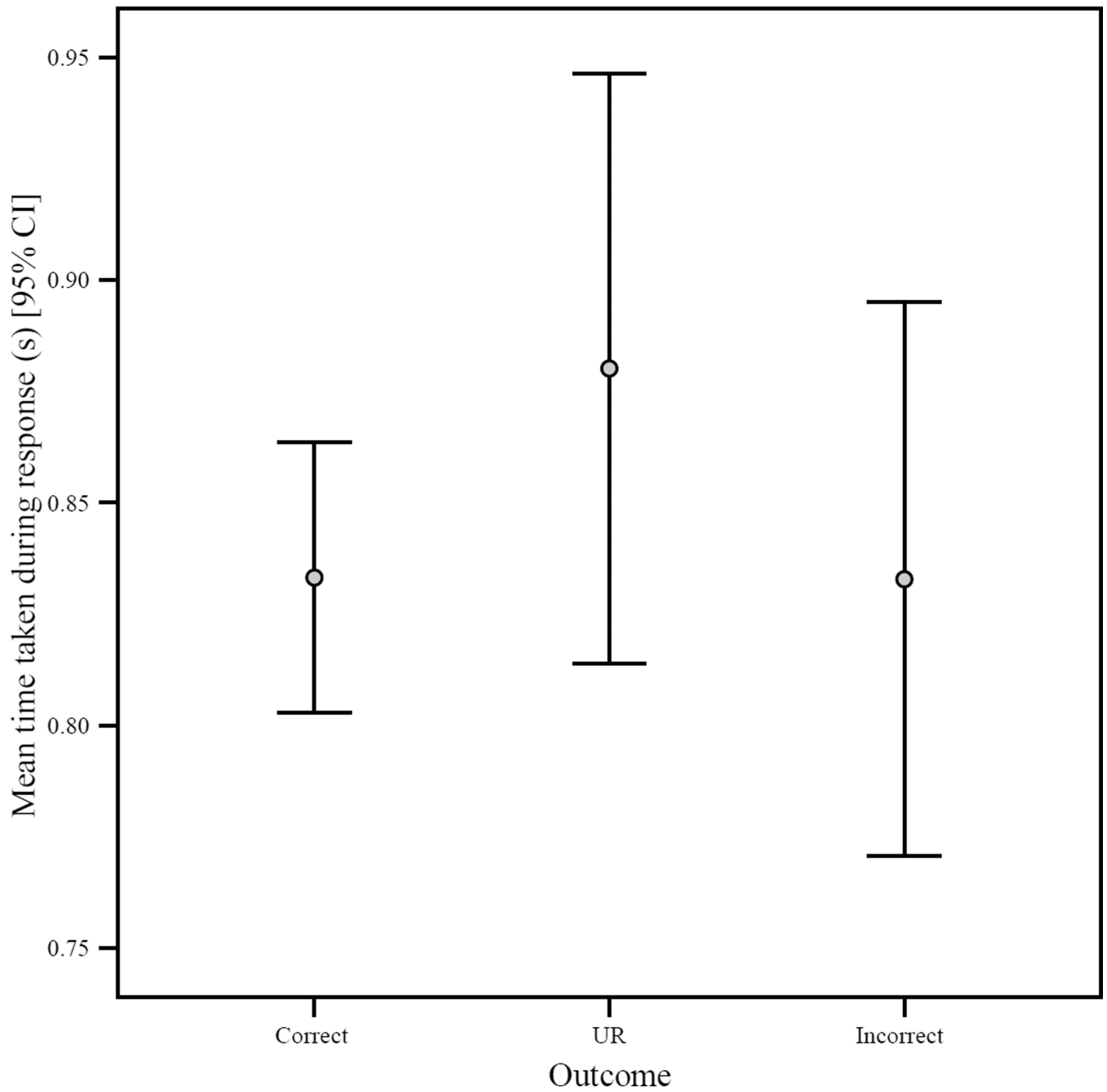
**Figure 1.** Percentage of sparse, dense, or uncertainty responses (UR) with respect to density level. Levels 30 and 31 are the most objectively difficult; objective difficulty decreases as level emanates from these points (see text).



**Figure 2.**  
The 95% confidence interval for mean objective difficulty per response/outcome type.

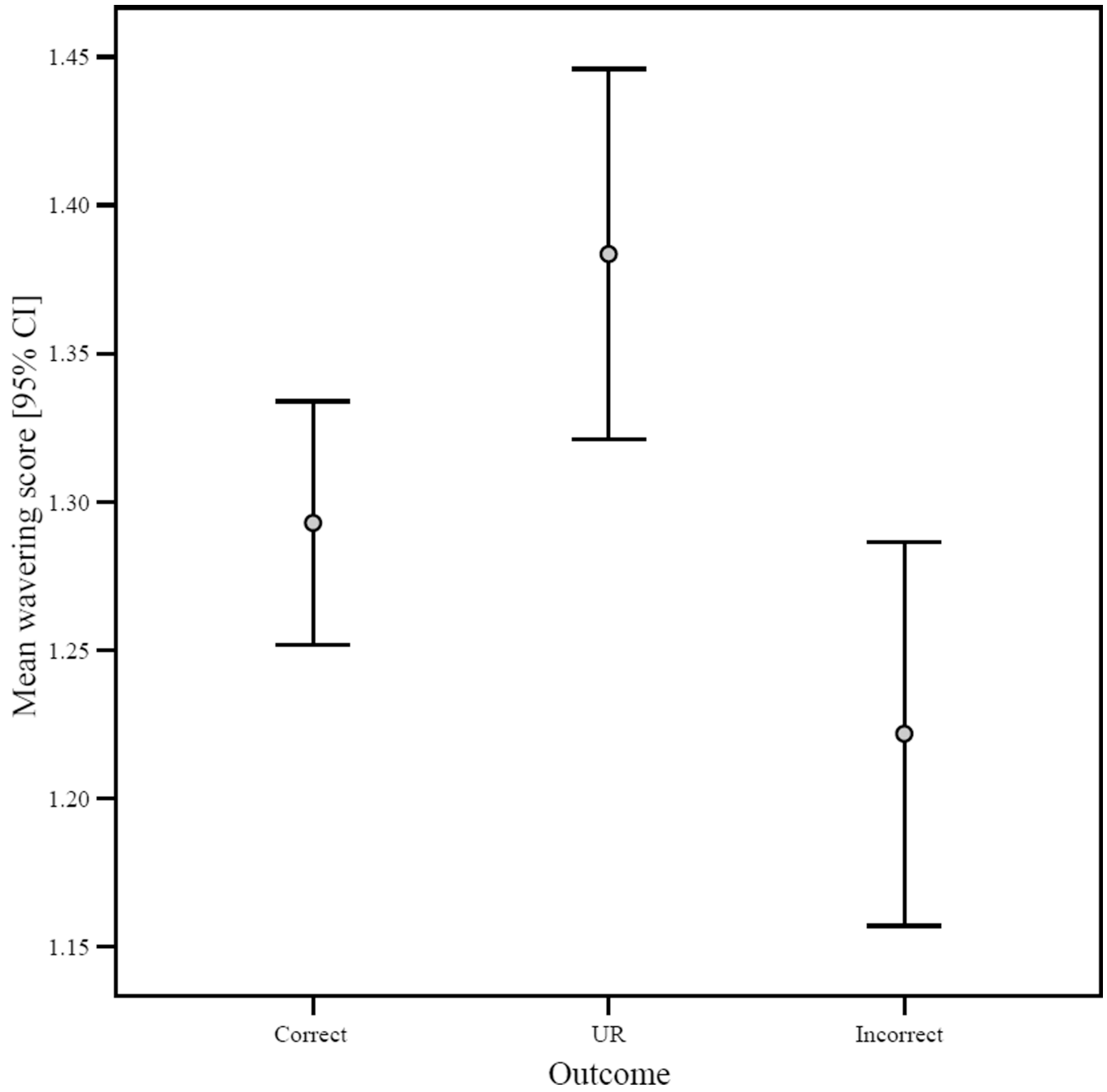


**Figure 3.** The 95% confidence interval for mean time before deflecting the joystick (seconds) per response/outcome type. “Time before deflection” is measured as the time between the monkey orienting towards a visible trial to when the cursor on the computer screen is first moved.

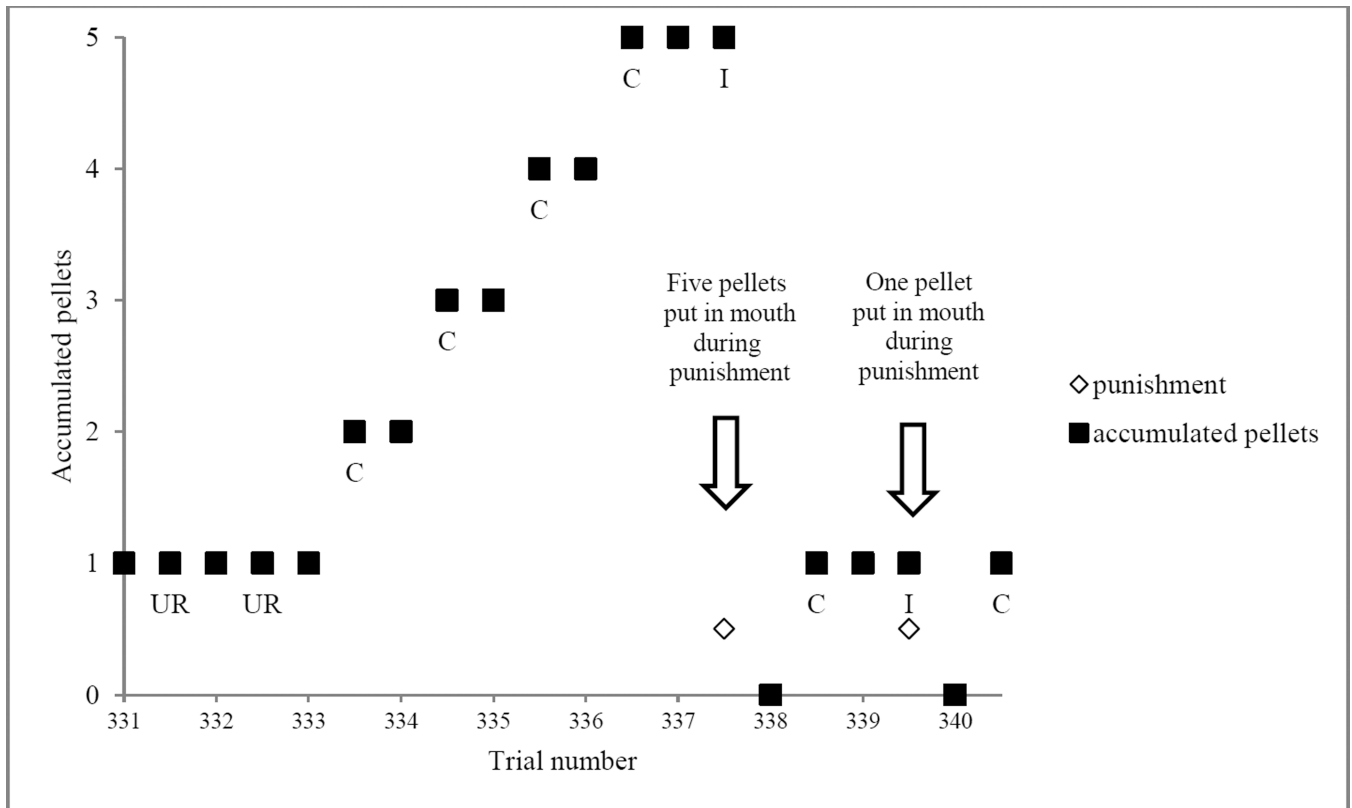


**Figure 4.** The 95% confidence interval for mean time taken for actual response (seconds) per response/outcome type. This is the time between initial cursor movement to contact with the selected icon.





**Figure 5.**  
The 95% confidence interval for mean wavering score (see text) per response/outcome type.



**Figure 6.**

Temporal accumulation and placement of pellets into the mouth by the rhesus macaque "Murph" over a 10-trial block. Individual trials are divided into: trial visible and response (whole numbers) and outcome (the 0.5's, not labeled on x axis). Pellets after correct responses are delivered at the 0.5's. This particular block was chosen for illustrative purposes, but is not exceptional for the data set in question (see Table 1). C = correct trial, I = incorrect trial, UR = uncertainty response.

**Table 1**

Pellets placed in mouth by trial outcome and phase

<i>Phase of trial</i>	<b>Pellet-in-mouth placement rates (pellets/trial)</b>		
	Correct responses ( <i>n</i> = 957 trials)	Incorrect responses ( <i>n</i> = 285)	Uncertainty responses ( <i>n</i> = 486)
Problem visible	0.27	0.32	0.11
Response	0.37	0.31	0.09
Blank screen	< 0.01	0.15 ["punishment"]	0.05
	Proportion of trials with cage-licking		
	Correct responses	Incorrect responses	Uncertainty responses
	0.40	0.67	0.85

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript