Do lemurs know when they could be wrong? An investigation of information seeking in three species of lemur (*Lemur catta, Eulemur rubriventer,* and *Varecia variegata*)

Heather K. Taylor
Clare L. Cunningham
Scott Hardie

https://doi.org/10.1037/com0000211

Do lemurs know when they could be wrong? An investigation of information seeking in three species of lemur (*Lemur catta*, *Eulemur rubriventer*, *Varecia variegata*).

Heather K. Taylor, Clare L. Cunningham and Scott Hardie

Evolutionary and Biological Approaches to Behaviour Research Group

Division of Psychology, Abertay University

Footnotes:

Heather K. Taylor, Clare L. Cunningham and Scott Hardie are affiliated to the Evolutionary and Biological Approaches to Behaviour Research Group, Abertay University, Bell Street, Dundee, DD1 1HG, Scotland.

Acknowledgements:

We wish to thank the staff at Camperdown Wildlife Centre for their support and assistance during data collection.

Correspondence concerning this article should be addressed to Dr Clare Cunningham, Division of Psychology, Abertay University, Bell Street, Dundee, DD1 1HG, Scotland, Email: c.cunningham@abertay.ac.uk

ABSTRACT (223 words)

Sixteen lemurs, including representatives from three species (*Lemur catta*, *Eulemur rubriventer*, *Varecia variegata*), were presented with a food seeking task where information about the rewards location, in one of two plastic tubes, was either known or not known. We evaluated whether lemurs would first look into the tube prior to making a choice. This information-seeking task aimed to assess whether subjects would display memory awareness, seeking additional information when they became aware they lacked knowledge of the rewards location. We predicted lemurs would be more likely to look into the tube when they had insufficient knowledge about the rewards position. Lemurs successfully gained the reward on most trials. However, they looked on the majority of trials regardless of whether they had all the necessary information to make a correct choice. The minimal cost to looking may have resulted in checking behaviour both to confirm what they already knew and to gain knowledge they did not have. When the cost of looking increased (elevating end of tube requiring additional energy expenditure to look inside - Experiment 2), lemurs still looked into tubes on both seen and unseen trials; however, the frequency of looking increased when opaque tubes were used (where they could not see the rewards location after baiting). This could suggest they checked more when they were less sure of their knowledge state.

Keywords: Prosimians; lemurs; memory awareness; information seeking; metacognition

Introduction

An awareness of the contents of memory allows humans to react appropriately when confronted with uncertain situations; they can differentiate between knowing and not knowing, and subsequently seek more information when necessary (Hampton, Zivin & Murray, 2004). Memory awareness represents one component of metacognition, which can be functionally defined as the ability to reflect on the contents of one's own cognition (see Nelson, 1996). Research investigating this ability in humans frequently uses methods reliant on language, with participants making subjective, verbal statements of 'knowing' and 'uncertainty' (Coutinho et al., 2015). As these paradigms cannot be used with non-verbal organisms, Smith et al. (1995) developed a paradigm that both induced a state of uncertainty and presented an objective, behaviourally measurable way of dealing with this uncertainty for use with non-human animals. Utilising psychophysical tests, Smith et al. (1995, 1997) presented subjects with two primary discrimination responses and a third 'uncertain' response that allowed subjects to 'escape' any tests they chose, hypothetically when they were uncertain of the correct response. This escape option allowed subjects to decline trials, instead progressing to an easy, guaranteed win trial. Including this option introduces a meta-level response, allowing subjects to report on their thoughts about whether they know or not.

Assessment of *uncertainty responding* has produced support for metacognitive capabilities in apes and old world monkeys. Apes reliably chose to select an escape response during a searching task when ignorant of a rewards location (Suda-King, 2008; Suda-King, Bania, Stromberg & Subiaul, 2013). Rhesus macaques (*Macaca mulatta*) frequently chose to escape difficult trials (Smith et al., 1997), displayed the ability to generalise this response over a range of tasks (Brown, Templer & Hampton, 2017; Couchman, Coutinho, Beran & Smith, 2010; Morgan, Kornell, Kornblum & Terrace, 2014; Smith, Redford, Beran & Washburn, 2010;

Templer & Hampton, 2012) and continued to do so when this response received no reward (Beran, Smith, Redford & Washburn, 2006). Findings from New World monkeys are less conclusive, with capuchin monkeys (*Cebus apella*) often failing to use the escape response as effectively as Old World monkeys, resulting in increased errors during difficult trials (Beran, Perdue, Church & Smith, 2016; Beran, Perdue & Smith, 2014; Fujita, 2009, see Smith, Smith & Beran, 2018, for a review).

Many uncertainty-monitoring tasks require extensive training. Smith (2009) proposed more naturalistic paradigms would provide less trial-intensive methods, reducing the opportunity for associations to form between particular trial types and uncertain responses, more closely aligning with situations animals may encounter in their natural habitats. To address these concerns, Call and Carpenter (2001) developed an 'information seeking' paradigm incorporating a search task where subjects (*Pan troglodytes*, *Pongo pygmaeus*, *Homo sapiens*) either witnessed the baiting of one of several opaque tubes with a reward, or did not. Subjects displaying meta-memory, and consequently the capacity to reflect on memories of what they had seen, ought to seek more information by looking into the tubes when ignorant of the rewards location i.e. after 'unseen baiting' trials. In contrast, on 'seen baiting' trials, subjects were presented with all the information necessary to locate the reward, and should be able to make a selection without seeking further information (i.e. without looking). Subjects displayed differential looking behaviour across trial type, seeking more information during unseen baiting trials. Call and Carpenter (2001) propose these results support the hypothesis that subjects knew when they did not know where the reward was, and acted to rectify this uncertainty.

All species of great ape that have been tested have since been found to *seek more information* when presented with an uncertain situation (Beran, Smith & Perdue, 2013; Call, 2010; Call & Carpenter, 2001; Marsh & MacDonald, 2012). Investigation of memory awareness in Old World

monkeys has produced support for information seeking in baboons (*Papio papio*), lion-tailed macaques (*Macaca silenus*) and rhesus macaques (*Macaca mulatta*), with investigation in New World monkeys being restricted to a single species (*Cebus apella*), and providing more equivocal results (Basile, Hampton, Suomi & Murray, 2009; Basile, Schroeder, Brown, Templer & Hampton, 2015; Beran & Smith, 2011; Hampton et al., 2004; Malassis, Gheusi & Fagot, 2015; Marsh, 2014; Paukner, Anderson & Fujita, 2006; Rosati & Santos, 2016; Vining & Marsh, 2015).

The decreased performance seen in New World monkeys across both uncertainty monitoring and information seeking tasks is of importance when considering the emergence of metacognition across the Primate order. The more limited success of capuchins (*Cebus apella*) may suggest metacognition evolved selectively, or more strongly, following the divergence of New World and Old World lineages. Alternatively, given that assessments to date are rather species-limited, this may reflect a deficit in methodological approach, rather than metacognitive abilities (see Smith et al., 2018 for further discussion). Assessment of metacognition in non-primate species has likewise presented inconclusive findings; for example, in birds (Fujita, Nakamura, Iwasaki & Watanabe, 2012; Goto & Watanabe, 2012), dolphins (Smith et al., 1995) and rats (Foote & Crystal, 2007). Further comparative research is needed to map out the phylogenetic distribution of metacognition across animal specials, providing a clearer view of the evolutionary emergence of these capacities. Lemurs are ideal candidates for this evaluation, representing an early divergence within primates with a unique evolutionary history (see Martin, 2000; Yoder & Yang, 2004; Herrera & Dávalos, 2016), and so providing an important comparative baseline for understanding cognitive evolution in the Primate lineage.

In order to succeed during an information seeking task, individuals are required to mentally represent objects that are removed from view, keeping in mind the location of rewards after

witnessing baiting procedures. Previous assessments have found support for this ability in lemurs, with Deppe, Wright and Szelistowski (2009) finding lemurs (*Eulemur fulvus*, *Eulemur mongoz*, *Lemur catta*, *Hapalemur griseus*) displayed proficiency with visible object displacement tests of object permanence. Black (*Eulemur macaco*, N = 3) and brown (*E. fulvus*, N = 3) lemurs have also displayed some proficiency during inferential reasoning tasks. Maille and Roeder (2012) found three lemurs successfully selected a baited cup when offered either visual or auditory information regarding the contents of two opaque cups, one that contained a reward. When presented with auditory information regarding the contents of only the empty cup, one subject selected the correct cup above chance levels. Thus, it may be expected lemurs would be able to keep in mind a rewards location during an information seeking task when baiting of the apparatus was seen or unseen, displaying less looking behaviour when they witnessed the placing of the reward than when baiting was unseen and they needed to obtain further knowledge for success.

The present study aimed to assess whether three species of lemur (*Lemur catta*, *Eulemur rubriventer*, *Varecia variegata*) could discriminate between knowing and not knowing, seeking more information when required and displaying memory awareness. The use of an information seeking paradigm similar to that developed by Call and Carpenter (2001) presented a more naturalistic and less trial intensive situation than previous uncertainty monitoring tasks (Smith et al., 1995; Smith, 2009). Subjects' looking behaviour was recorded when presented with two tubes, one that was baited with a reward, and either allowed direct visual access to the contents or where contents could not be viewed without looking down the tubes' length. By varying visual access to the baiting procedure, subjects were either given the information needed to select the correct tube or were not. In all trials, subjects were free to seek more information about the rewards location by either looking down the length of the tube to seek visual affirmation, or using olfactory cues to ascertain which tube contained food. Although we cannot

distinguish which cue type was used, in all cases, we would expect lemurs to lower their head towards the tube before making a choice in unseen trials. It is head lowering to align with the tube opening that we refer to as a 'look', with the caveat that information gained from looking could be either visual, olfactory, or both.

Methods

Subjects

Subjects were six red-bellied lemurs (*E. rubriventer*), seven ring-tailed lemurs (*L. catta*) and three black and white ruffed lemurs (*V. variegata*). All subjects were born in captivity and housed in three enclosures at Camperdown Wildlife Centre (Dundee, Scotland), where they had access to both indoor and outdoor areas. Two enclosures contained mixed species groups during training, with red-bellied and ring-tailed lemurs housed in enclosure 1, and red-bellied and black and white lemurs in enclosure 3 (see Table 1). A second group of ring-tailed lemurs were housed in a single species group in enclosure 2. Due to relocation of subjects following training, this became one mixed species enclosure and two single species enclosures during Experiments 1 and 2. During these experiments, a mixed group of red-bellied and black and white lemurs were housed in enclosure 3, with a second group of red-bellied lemurs in enclosure 1, and a group of ring-tailed lemurs in enclosure 2. Each indoor enclosure contained a partition splitting the space into two separate areas, both with an outdoor access hatch that could be closed. As each species in the mixed groups naturally favoured one half of the enclosure, subjects were tested throughout in their species-specific groups. Subjects were fed a daily assortment of fresh fruit, vegetables and barley rings, all of which were freely available during testing, along with fresh water *ad libitum*.

Apparatus

Figure 1: (A) Clear tubes used during training and Experiment 2, showing wooden stoppers used throughout testing. (B) Opaque tubes used in Experiment 1 and 2. (C) Wooden ramp used during Experiment 2 to raise the end where subjects made a choice, shown with opaque tubes mounted on top.

The apparatus consisted of a pair of either clear or opaque acrylic tubes (50cm x 5cm) mounted parallel to one another on top of a wooden board (50cm x 25.5cm) (Figure 1). Two stoppers consisting of a wooden dowel with a block attached to the end (50cm) were used to slide the reward within reach of the subject and prevent it from being knocked back along the tube during the test (Figure 1). A wooden ramp was used during Experiment 2 to raise the height of the tubes at the end where subjects made their selection. The wooden boards holding the tubes could be mounted on top of the ramp, allowing easy switching between either opaque or clear tubes dependent on the trial type. A stopwatch was used to record trial length and inter-trial interval during training, with a handheld video camera used to record all trials during Experiments 1 and 2 for later behavioural coding.

Training with clear tubes

Subjects

All 16 lemurs took part in initial training (*E. rubriventer*, N = 6, *L. catta*, N = 7, *V. variegata*, N = 3). Two ring-tailed lemurs in one enclosure (see Table 1) were restricted to their indoor enclosure during testing in order to deny access to the red-bellied lemurs sharing the enclosure who often attempted to enter and disrupt performance. All other groups were tested in their

respective indoor areas, with free range between indoor and outdoor enclosures available throughout testing.

Procedure

Clear tubes were used during training, allowing subjects to easily see the whole length of the tube giving direct visual access to the reward (Figure 1a). This familiarised subjects with the task demands, with the direct visual access aiming to present an easier task than being required to attend to either the baiting procedure or to seek more information when presented with opaque tubes during later experiments. Tubes were presented at floor level meaning the lemurs could easily see the reward from a seated position. If subjects chose to seek either visual or olfactory information of the tube's contents, they needed to lower their head near to the floor until it was level with the entrance of the tubes. This would be unnecessary effort not required to make a correct selection during training with clear tubes.

Individuals were not separated from their group for testing. We continued to present the apparatus to the group until all those motivated to participate had completed the necessary trials to reach criterion. This may have resulted in over-training of the 'correct' response in individuals who continued to interact with the apparatus after reaching criterion performance (for further explanation of potential impact, see results). This was however preferable to unnecessary stress induced by separation from the group. Testing in the group also represents a more ecologically valid environment for skill acquisition and learning. The total number of trials individuals completed are reported in the results. If more than one individual was present within the testing area, the test subject was considered as the individual situated in front of the apparatus, at the open end of the tubes, with the experimenter ensuring this individual visually attended to the apparatus at the onset of each trial.

The experimenter was located within the lemur enclosures throughout testing but as far away from the testing area as space would allow. Subjects were presented with two clear tubes, one of which was baited with a reward and one that was not. Red grapes were used as the reward for all lemurs except one, who after displaying a lack of interest in multiple trials using grapes showed a preference for barley rings. Seen trials consisted of the subject being shown the reward before the experimenter used a closed hand containing the food to place it in one of the clear tubes in full view of the subject. Using only one hand to bait the tube in the seen trials reduced task demands as subjects did not need to understand 'transparency', with subjects able to use hand movements to select the correct location independently of visual exploration of the tubes. In contrast, during unseen trials, the subject was not shown the reward before baiting. The experimenter placed both closed hands into the clear tubes simultaneously. One hand contained the reward that was surreptitiously deposited into either the left or right tube. From the lemurs' point of view, when they had not seen which tube was baited, both tubes could potentially contain a reward. In unseen trials, visual or olfactory exploration was therefore necessary to distinguish which tube contained the reward. Although inclusion of seen and unseen baiting was not essential during the training phase as lemurs could 'see' the reward through the clear tube, both trial types were given to maintain consistency with later experimental conditions.

Subjects were required to select a tube following the baiting procedure, which was either seen or unseen by the subject, with the location of the baited tube varying from left to right. Trial type (left/right, seen/unseen) was randomised, with the only contingency being that the same trial type was not presented more than twice in a row to prevent facilitation of side bias. Subjects were given a maximum trial length of 60 seconds, and a minimum inter-trial interval of 20 seconds, following which the apparatus were re-baited and a new trial began. Subjects were considered to have made a selection once they placed their hand inside a tube, reaching for the

reward. They were classified as correct if they reached into the baited tube, and incorrect if they reached into the tube that did not contain a reward. If the correct tube was selected, the food reward was slid to within reach allowing subjects to retrieve it, whereas if the incorrect tube was selected the subject was unrewarded. If a subject placed their hand in to one tube and then the other, the tube they selected first was taken as their choice. There were no instances of subjects placing both hands into both tubes simultaneously. If a selection was made within the trial period, regardless of correct or incorrect, the trial ended and the inter-trial interval began. If subjects made no selection within the maximum trial length of 60 seconds, the inter-trial interval occurred and then the next trial began with the tubes being re-baited according to the trial schedule. For each trial, subject ID, trial type and whether the subject selected the correct or incorrect tube were noted.

Analyses

In order to pass the training phase, subjects were required to achieve 9/10 correct responses over two consecutive blocks of 10 trials. All subjects received 10 trials per block, regardless of response type (correct or incorrect). For example, if a lemur gave two incorrect responses on the first two trials, they could not reach criterion performance on that block. However, they would still be given the remaining eight trials of that block, and then proceed to the next set of 10 trials where 9/10 correct responses were needed to be on track for criterion performance. Wilcoxon signed rank tests evaluated whether trial type (seen or unseen) affected number of correct responses and if subjects were more likely to make a correct selection when the reward was presented on the left or right side.

Results and discussion

All subjects that completed this phase (N = 12) successfully passed 10/10 trials in the first two blocks (Table 1). Four subjects failed to complete the task, and therefore did not progress to the testing phase. BW3 and RB4 showed little interest in participating, rarely approaching the apparatus or attempting trials. The youngest subject tested, RB3, showed interest but failed to understand the task requirements, never placing her hand inside a tube. In order to give a fair assessment of these subjects' abilities, they were each presented with 50 trials; however, none obtained a reward on any presentation. They were therefore excluded from further testing. The remaining unsuccessful subject, RT3, was unable to access the materials due to monopolisation by other subjects and so did not participate in testing.

Table 1: Subjects' age, sex, enclosure number, experiment participation and trials to criterion performance (9/10 correct responses in two blocks of 10 trials) in the training phase.

The total number of trials presented to each subject in the training phase ranged from 20-46 (Mdn = 26.5) (although criterion was reached by all those that passed (N = 12) in the first 20 trials). This variation was because subjects were not separated for testing and were free to participate in all presented trials. Although separation would have been preferable, the enclosure design did not allow for easy manoeuvring of individuals and would have caused unnecessary stress that would be detrimental to task performance and welfare of the lemurs. For illustrative purposes, this meant the individual that completed the most trials (46) completed 20 trials to reach criterion and then made a response on a further 26 presentations of the apparatus. We hoped to only bait the apparatus when the target individual was in the vicinity to remove the possibility of other subjects receiving more trials. However, due to the initial novelty of the apparatus, lemurs tended to crowd making it impossible to restrict access to one individual on all trials. This was rectified in later experiments where it was possible to limit access by non-tested individuals. In total, 352 trials were given across all twelve subjects, of which 351 were

correct. Lemurs who participated in testing made a response on all trials. Only one incorrect response was given by subject (RB2), who selected the non-baited tube after completing the first two blocks where she successfully reached criterion performance.

Considering all trials subjects completed, whether the trial was 'seen' or 'unseen' did not significantly affect number of correct choices made by subjects (Wilcoxon signed rank: $W = 10.0$, $p = 0.916$, $N = 12$). There were no significant differences in number of correct choices dependent on side of presentation (Wilcoxon signed rank: $W = 14.5$, $p = 0.400$, $N = 12$). That there were no effects on accuracy is not surprising given that subjects were almost at ceiling, with only one incorrect response overall.

Despite the use of clear tubes permitting direct visual access to the reward, subjects bent down to look along the tube's length, or potentially to seek olfactory information about the tubes contents, in almost all trials (although we did not quantify how often this occurred). It is possible that they struggled with the transparent nature of the tubes, failing to understand that the object seen through the tube was the same object seen when looking into the tube. The black and white lemur group particularly seemed to struggle in early trials with the relationship between reaching inside the tube and obtaining the reward, initially attempting to gain access to the reward through the tubes' outside wall before making the correct response. This finding is in line with MacLean et al. (2013) who found performance improved over testing when assessing lemurs (*Eulemur macaco*, *Eulemur mongoz*, *Lemur catta*, *Propithecus coquereli*, *Varecia variegata*) ability to reach around to the opening of a transparent tube to retrieve a food reward, rather than attempting to access the reward through the side of the tube.

Experiment 1: Looking behaviour to supplement knowledge

Subjects

Before Experiment 1 began, two ring-tailed lemurs were relocated to an alternative enclosure at the zoo, with two individuals dying shortly after this move (see Table 1). Consequently, 10 subjects who successfully completed training took part in this experiment. This included four ring-tailed lemurs, four red-bellied lemurs and two black and white ruffed lemurs. The issue of dual species occupation in one enclosure was eliminated by the relocation of the two ring-tailed subjects. Therefore, all subjects were tested in species groups in their indoor enclosures, with access to the outdoor enclosure freely available.

Procedure

Brown opaque tubes were used in order to obstruct subjects' view of the reward (Figure 1b). In conjunction with the presentation of materials at floor level, this aimed to ensure subjects could not see the reward's location and encountered a physical cost (effort needed to bend down until their head was level with the tubes entrance) in order to seek the reward. The experimenter was again situated within the enclosure, presenting subjects with two tubes in the same way as for the training phase, one of which contained a reward. Tubes were only baited when the target individual entered the vicinity of the apparatus. The positioning of the reward, reduced novelty of the task and smaller numbers in the enclosures allowed us to control access by individuals not being tested so all subjects received the same amount of trials in Experiment 1. Red grapes were used as rewards for all lemurs.

The baiting procedure was either seen or unseen by subjects, with the reward's location being either the left or right tube. The method of baiting was the same as for the training phase; thus, lemurs had the knowledge needed to make a correct response on the seen trials, but not on

unseen trials. Only by seeking more information, either looking down the tube's length or smelling the contents of the tube, could they determine the location of the reward on the unseen trials. Trial type was once more randomly assigned, with the same trial type presented no more than twice in succession. Subjects completed 20 trials in total, consisting of equal numbers of seen/unseen and left/right trials. Trials were recorded using a video camera for later coding, with subjects given a maximum of 60 seconds to attempt a trial, and a minimum inter-trial interval of 20 seconds. Trials ended following a correct or incorrect selection (insertion of hand into a tube in an attempt to retrieve the reward), or after 60 seconds if no selection was made. If subjects performed looking behaviour but failed to select a tube within the 60-second trial period, they were marked as making no selection. The apparatus were removed from reach as soon as selection was made to prevent subjects being able to reach into both tubes.

Experiment 1 aimed to assess the subjects' information seeking behaviour; therefore the number of looks into each tube, and correct or incorrect selection were all coded from video recordings. A 'look' was defined as the subject lowering their head until the tube opening was at eye level. On some trials, lemurs would look into the tube multiple times and so we recorded total number of looks per trial. To be scored as a repeated look, the subject needed to make an initial look and then lift its head away from the tube opening, before bending down to look for a further time. If subjects followed the baiting process and were capable of discriminating between knowing the reward's location or not, they were predicted to display more frequent looking behaviour on unseen than seen trials. In addition, trial duration was measured, assessing whether subjects' latency to selection increased during unseen trials where they were required to seek more information. Intra-rater reliability (IRR) was calculated using intra-class correlation coefficient (ICC) on number of looks per trial, with 20% of trials re-coded by HKT. ICC estimates and their 95% confidence intervals were calculated using SPSS v25, based on a single rater, absolute

agreement, two-way mixed-effects model (Koo & Li, 2016). The ICC was found to be excellent at 0.98 (CI's = 0.96, 0.99).

Analyses

Individual performances were evaluated using binomial tests to determine whether lemurs were selecting the correct tube significantly above chance across all 20 presentations (seen and unseen). An extension of GZLM, Generalized Estimating Equations (GEE), was used to build models to assess the effect of trial type (seen or unseen) on the number of correct responses and whether subjects looked inside the tube before making a selection. GEE is suited to analyses of small (10 or more subjects), non-normally distributed, dependant data and can support many different types of variable such as binary or counts (Garson, 2013). All analyses were conducted in SPSS V25.

A series of models were run to evaluate the effect of trial type and looking behaviour on performance. The first model set were specified as binary logistic (for yes/no responses) with a logit link function and response (correct or incorrect) as the dependent variable and subject identity included as a within subject factor, and trial number as the between subject factor. Trial type (seen or unseen) and presence or absence of looking behaviour used as factors. Full details of goodness of fit and parameter estimates of all models can be found in the Electronic Supplementary Materials (ESM); only results from the best fitting models are reported here. For the GEE analyses, trials where the subject did not make a choice were removed as no response does not necessarily represent a lack of knowledge. A model-based estimator was used due to small sample size (Garson, 2013) and the working correlation matrix was specified as 'independent' meaning trial response outcomes were assumed to be independent of each other in all models. A second series of models were specified using a Poisson distribution (for counts),

with a log link function, with number of looks before selection as the dependent variable. Trial type (seen or unseen) was included as a factor, along with subject identity as a within subject factor and trial number as the between subject factor, with a model-based estimator and independent correlation matrix. Details of all models can be found in the ESM with results from the best fitting models (when more than one model was run) reported here based on QIC values.

A final set of GEE models assessed response time differences by trial type (seen or unseen) with the assumption that not knowing the rewards location would increase trial duration as they sought further information. For this analysis, trials where no selection occurred were again omitted to avoid skewing the data when trial time reached its maximum. A linear model with identity link function was specified with a model-based estimator procedure due to small sample size. Trial type was added as a factor, and subject identity as the within subject factor and trial number as the between subject factor (see ESM for full details of all models). All tests were two-tailed and alpha was set at 0.05 throughout.

Results and discussion

Collectively, subjects made a response on all but 11 trials where no choice was made within the maximum trial duration. Looking at individual responses, 9/10 subjects performed significantly above chance (when considering only the trials where a response was made), consistently selecting the correct tube across all trials (seen and unseen) (Table 2). The poorer performance of RB1 may have been due to the young age of this subject who turned two years old during testing.

Table 2: Number of correct and incorrect responses, no selections and binomial results by subject in Experiment 1 (* = significant at 0.05 level, ** = significant at 0.01 level). If subjects

436 did not make a response (no selection), we did not include this in the binomial test or GEE
437 analyses.

438

439 There was no significant effect of trial type on number of correct responses given (GEE: β = -
440 0.655, χ^2 = 1.760, df = 1, p = 0.185), suggesting being unable to see the reward placed did not
441 have a detrimental effect on performance. Collectively, lemurs performed looking behaviour on
442 all but six trials where a selection was made, therefore, unsurprisingly presence or absence of
443 looking behaviour was not a significant predictor of performance (GEE: β = 0.819, χ^2 = 0.508,
444 df = 1, p = 0.476). However, lemurs obtained 50% of rewards on trials where they did not look
445 (representing chance performance) compared to 86% of rewards on trials where they did look,
446 suggesting they were benefitting from confirming their knowledge state.

447

448 Although lemurs obtained more rewards in unseen trials when they did look before making a
449 choice, there was no evidence that they were looking into the tube more in the unseen than in the
450 seen trials (GEE: β = -0.087, χ^2 = 0.477, df = 1, p = 0.490). These results suggest that lemurs
451 were not using looking behaviour to supplement their knowledge when they did not know for
452 certain where the reward was. However, that there was no increase in looking behaviour when
453 trials were unseen was due to the high number of looks performed in the seen trials; subjects
454 looked on most trials regardless of whether they had seen the reward placed or not (in 189 trials
455 where a response was made, lemurs performed looking behaviour in 183).

456

457 There was no main effect of trial type on trial duration (GEE: β = -0.276, χ^2 = 0.032, df = 1, p =
458 0.995), indicating subjects did not take longer to reach a decision about which tube to select
459 when they had not observed placement of the reward. We are cautious in our interpretation of
460 these data however, as latency to make a choice is conflated with number of looks in our

method. It takes longer to make more looks and so we could be measuring motor behaviour rather than cognitive processing. Even if longer latencies are indicative of longer cognitive processing times, the mechanism (memory search versus vacillation/hesitation) is unclear. That we did not find an effect could be due to the use of total trial duration as response latency and perhaps 'time to first look' would have been a more informative measure in this regard.

Trials where subjects did not make a selection may be informative. If they do not have the necessary information to make a successful response, for example, in unseen trials, they may be more likely to avoid failure through a null response. If this is the case, we would expect an increased number of trials where no selection was made in the unseen condition. However, subjects selected a tube at near identical levels across seen and unseen baiting, suggesting the lack of information given about the rewards location during unseen baiting did not hinder subjects' motivation to participate.

Experiment 2: Increased cost of looking

Experiment 2 incorporated an increased cost of looking in order to further assess subjects' information seeking behaviour. A ramp was used to tilt the tubes up from the ground at the end where subjects made a selection. This meant subjects had to stand up in order to either look down the length of the tube or gain clear olfactory information about the tubes contents, not simply lower their head down as in previous trials. By increasing the energy expenditure required to look, subjects were expected to seek more information only when they lacked knowledge about the rewards location (unseen/opaque trials). Subjects displayed a drop in the level of interest shown in the task between the training phase and Experiment 1, likely due to the removal of the direct visual aspect afforded by clear tubes during training. In an effort to regain

levels of interest, Experiment 2 therefore also incorporated a mix of clear and opaque tubes, presented in pairs, with one 'clear tube' trial presented for every two 'opaque tube' trials.

Subjects

The same ten subjects that took part in Experiment 1 participated in Experiment 2. Subjects were again tested in their respective indoor enclosures, in species-specific groups, with access to outdoor enclosures available throughout testing.

Procedure

The procedure for Experiment 2 was consistent with Experiment 1, with the following exceptions. Subjects were presented with either two clear or two opaque tubes, placed on top of the ramp (Figure 1c). Subjects completed 10 clear trials and 20 opaque trials, at a ratio of one clear to every two opaque to maintain motivation to participate. Equal numbers of seen/unseen and left/right trials were given for both clear and opaque tubes. A maximum trial duration of 60 seconds was used to maintain consistency with Experiment 1 with an inter-trial interval a minimum of 20 seconds. All trials were video recorded for later coding with the same measures used as in the previous experiment. Intra-rater reliability (IRR) was again calculated using ICC on number of looks per trials, with 20% of trials re-coded by HKT. ICC estimates and their 95% confidence intervals were calculated as in Experiment 1, with ICC found to be excellent at 0.94 (CI's = 0.90-0.97). Subjects were again classified as making a correct selection, incorrect selection or no selection in line with Experiment 1. Trials ended after a selection was made with the apparatus being removed from reach after the first tube was chosen, or after 60 seconds if no selection was made.

Analyses

Analyses for Experiment 2 were consistent with Experiment 1 with the following exceptions. Individual performance was evaluated using binomial tests for trials using the opaque tubes only, as all lemurs made a correct response on all presentations with clear tubes. The first series of GEE models evaluated number of correct responses with trial type (seen or unseen) as a factor (subject identity was a within subject factor and trial number a between subject factor throughout). As before, only trials where a response was made were included in the models (and binomial tests). For these analyses, trials with clear tubes were excluded to prevent singularity in the Hessian matrix due to constant data for one predictor (all lemurs responded correctly an all trials when clear tubes were presented) (Garson, 2013). Unlike the GEE on performance in Experiment 1, presence or absence of looking behaviour was not included as a factor as all subjects looked at least once on all trials. For GEE models with frequency of looking and response time as dependent variables, trial type (see or unseen) and tube type (clear or opaque) were included as factors. As in Experiment 1, all tests were two-tailed and alpha was set at 0.05 throughout.

Results and discussion

Of 300 trials (30 to each subject) the reward was obtained on 276 trials, with 16 incorrect choices and 8 trials where no response was made (Table 3). On presentations of the clear tube, all lemurs responded correctly on all presentations, regardless of whether the baiting procedure was seen or unseen (Table 3). For the opaque tubes, 9/10 subjects performed above chance across seen and unseen trials, with the number of correct responses when the baiting procedure was unseen being lower compared to seen (99 correct in seen trials compared to 92 in unseen

trials). This difference did not reach statistical significance. (GEE: $\beta = -1.171$, $\chi^2 = 3.770$, df = 1, p = 0.052).

Table 3: Number of correct and incorrect responses, no selections and binomial results by subject in Experiment 2. Binomial test results relate to performance on opaque tubes only as all subjects responded correctly on all clear tube presentations, with seen and unseen trials pooled for these analyses (* = significant at 0.05 level, ** = significant at 0.01 level). If subjects did not make a response (no selection), we did not include the trial in the binomial tests tor GEE analyses.

Lemurs performed looking behaviour in the majority of trials, regardless of tube type (clear or opaque) and whether they observed the baiting procedure or not. The frequency of looking behaviour did significantly increase in unseen trials (GEE: $\beta = -0.168$, $\chi^2 = 4.543$, df = 1, p = 0.033). To evaluate whether the type of tube influenced the number of looks, GEEs were run on all trials (clear and opaque), revealing type of tube (clear or opaque) did not predict number of looks. However, there was a significant interaction with seen trials with clear tubes producing less looking behaviour than unseen trials using opaque tubes (Table 4). In general, lemurs looked more when tubes were opaque (Figure 2).

Table 4: GEE model parameters for Experiment 2 with **number of looks** as the dependent variable and trial type (seen or unseen) and tube type (clear or opaque) as factors. Only trials where a response was made are considered. Significant results are highlighted in bold text (*p < 0.05, ** p < 0.01). Trial type did significantly influence number of looks with more looks occurring in the unseen trials. Tube type did not impact looking behaviour, however there were

more looks displayed when the tube was opaque and baiting unseen compared to see, clear trials.

Figure 2: Mean number of looks observed in seen and unseen trials using clear (solid bars, N = 10) and opaque (patterned bars, N = 20) tubes. Error bars represent 95% confidence intervals. With both clear and opaque tubes, more looks occurred in the unseen condition.

Response times were not affected by trial type (GEE: $\beta = -1.179$, $\chi^2 = 0.040$, df = 1, $p = 0.841$), suggesting lemurs did not take more time making a decision when they had not seen the baiting procedure. Response times were not significantly influenced by tube type (GEE: $\beta = -0.979$, $\chi^2 = 1.083$, df = 1, $p = 0.298$).

General discussion

Despite clear tubes granting direct visual access to the reward location, lemurs frequently performed unnecessary looking behaviour during clear tube trials. Subjects ought to have made correct selections without seeking more information in all trials with clear tubes, regardless of whether they saw the baiting or not. Their unnecessary looking behaviour may therefore represent a deficit in the metacognitive abilities of Prosimians. Alternatively, given the current subjects' lack of experience with transparent testing apparatus, it is possible their looking represents difficulties in comprehending transparency. Lemurs may have failed to understand that the reward seen *through* the tube would be present when looking *into* the tube. This was suggested to explain failure on a similar task in *C. apella*, with the visual presence of the bait perhaps acting as a marker designating a particular tube as correct, rather than being the physical

reward (Paukner et al., 2006). Subjects may therefore have looked into the tubes in order to seek visual confirmation of the rewards presence within a 'marked' tube.

Although lemurs did tend to look into the clear tubes before making a selection, they did so less frequently than when opaque tubes obscured the rewards location. There were 17/100 clear tube trials where subjects made selections without first looking in to the tubes during Experiment 2, with 12 of these selections occurring in seen baiting trials. This suggests that following the baiting process is within the capacity of lemurs. It could be the cost of looking was not sufficiently high to avoid subjects double checking the location of the reward to avoid error, similar to the 'passport effect' seen in apes where individuals check something they already know when the cost of making a mistake is high (Call, 2010).

When assessed with opaque tubes, lemurs continued to perform looking behaviour. Subjects gathered information when required on unseen trials, but continued to do so during seen trials, when the costs of looking were low (Experiment 1). This could suggest subjects lacked an awareness of knowledge gained during baiting in this experiment, where there was no visual marker of the correct tube once baited. Although the experimenter aimed to ensure subjects were attending to the task at the onset of each trial, it was unfortunately not possible to reliably discern whether subjects followed the baiting procedure from the current video footage of testing. Lemurs' unnecessary looking may therefore have been the result of a general failure to attend to the baiting, or a more active avoidance of this difficult task (Basile et al., 2009). When costs of looking are low, it may be easier to pay this minimal cost and look, rather than expend cognitive resources keeping track of the baiting procedure. If subjects fail to attend to the baiting, they would lack any awareness of the reward's location on seen trials with looking necessary to select correctly, much like unseen trials. If this was the case, subjects' initial looks would have been directed into the correct tube first at chance levels. However, subjects looked

611 first into the correct tube on the majority of seen trials. It is therefore unlikely that they gained
612 no information from the baiting procedures.
613
614 Hampton et al. (2004) proposed excessive looking may represent an implicit memory of the
615 reward's location, which subjects are not aware of possessing, guiding behaviour. Alternatively,
616 subjects may have looked during seen trials due to difficulty suppressing impulsive foraging
617 behaviour, or simply to confirm the reward's location observed during baiting. This may have
618 served the purpose of visual confirmation, or potentially allowed subjects to supplement their
619 visual knowledge with olfactory information. As Call and Carpenter (2001) proposed, if the cost
620 of looking is too small, then the benefit to be gained from doing so, even if it is just to check the
621 reward's location, outweighs the cost incurred. The introduction of an increased cost of looking
622 in Experiment 2 did result in a decreased number of unnecessary looks on seen trials, when
623 subjects had the necessary knowledge to respond correctly. Lemurs displayed the highest
624 number of looks during unseen trials with opaque tubes, suggesting some awareness of the
625 security of their knowledge. However, the increased cost of looking did not eliminate looking
626 behaviour on seen trials, providing little support for subjects possessing an awareness of their
627 knowledge.
628
629 The findings of the current study highlight potential differences in levels of memory awareness
630 across primates. Lemurs performed more poorly than great apes, Old World or New World
631 monkeys tested on similar paradigms, failing to display trial-appropriate looking behaviour
632 (Hampton, 2009; Vining & Marsh, 2015). It is possible the unnecessary looking behaviour
633 displayed here was due to the low cost of looking in the Training Phase and Experiment 1. The
634 looking behaviour demonstrated during Experiment 2 suggests lemurs may seek information at a
635 greater rate when ignorant to a reward's location, as they decreased unnecessary looking
636 behaviour when there was an increased energy cost to repeatedly checking. Further research

with a more significant cost is needed to fully evaluate memory awareness in Strepsirrhine primates; however, the current findings provide only limited support for metacognitive abilities. Given the lack of consistent support for metacognition in New World monkeys, taken together with the current findings, this may suggest some cognitive features prerequisite to metacognition are present within these species, with metacognition emerging more fully within Catarrhine primates.

Compliance with Ethical Standards: All applicable national and institutional guidelines for the care and use of animals in research were followed. The research study was approved by the Research Ethics Committee of the School of Social and Health Sciences at Abertay University and Camperdown Park.

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

650 References

651

652 Basile, B. M., Hampton, R. R., Suomi, S. J., & Murray, E. A. (2009). An assessment of memory
653 awareness in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*, 169-180.

654

655 Basile, B. M., Schroeder, G. R., Brown, E. K., Templer, V. L., & Hampton, R. R. (2015).
656 Evaluation of seven hypotheses for metamemory performance in rhesus monkeys. *Journal of*
657 *Experimental Psychology: General, 144*, 85-102.

658

659 Beran, M. J., Perdue, B. M., Church, B. A., & Smith, J. D. (2016). Capuchin monkeys (*Cebus*
660 *apella*) modulate their use of an uncertainty response depending on risk. *Journal of*
661 *Experimental Psychology: Animal Learning and Cognition, 42*, 32-43.

662

663 Beran, M. J., Perdue, B. M., & Smith, J. D. (2014). What are my chances? Closing the gap in
664 uncertainty monitoring between rhesus monkeys (*Macaca mulatta*) and capuchin monkeys
665 (*Cebus apella*). *Journal of Experimental Psychology: Animal Learning and Cognition,*
666 *40*, 303-316.

667

668 Beran, M. J., & Smith, J. D. (2011). Information seeking by rhesus monkeys (*Macaca mulatta*)
669 and capuchin monkeys (*Cebus apella*). *Cognition, 120*, 90-105.

670

671 Beran, M. J., Smith, J.D., & Perdue, B. M. (2013). Language-trained chimpanzees (*Pan*
672 *troglodytes*) name what they have seen but look first at what they have not seen.
673 *Psychological Science, 24*, 660- 666.

674

675 Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques
676 (*Macaca mulatta*) monitor uncertainty during numerosity judgements. *Journal of*
677 *Experimental Psychology: Animal Behavior Processes, 32*, 111-119.
678
679 Brown, E. K., Basile, B. M., Templer, V. L., & Hampton, R. R. (2019). Dissociation of memory
680 signals for metamemory in rhesus monkeys. *Animal Cognition, 22*, 331-341.
681
682 Brown, E. K., Templer, V. L., & Hampton, R. R. (2017). An assessment of domain-general
683 metacognitive responding in rhesus monkeys. *Behavioural Processes, 135*, 132-144.
684
685 Call, J. (2010). Do apes know that they could be wrong? *Animal Cognition, 13*, 689-700.
686
687 Call, J., & Carpenter, M. (2001). Do chimpanzees and children know what they have seen?
688 *Animal Cognition, 4*, 207-220.
689
690 Couchman, J. J., Coutinho, M. V. C., Beran, M. J., & Smith, J. D. (2010). Beyond stimulus cues
691 and reinforcement signals: a new approach to animal metacognition. *Journal of Comparative*
692 *Psychology*, *124*, 356-368.
693
694 Coutinho, M. V. C., Redford, J. S., Church, B. A., Zakrzewski, A. C., Couchman, J. J., & Smith,
695 J. D. (2015).0 The interplay between uncertainty monitoring and working memory: Can
696 metacognition become automatic? *Memory & Cognition, 43*, 990–1006.
697
698 Deppe, A. M., Wright, P. C., & Szelistowski, W. A. (2009). Object permanence in lemurs.
699 *Animal Cognition, 12*, 381-388.
700

Foote, A. L., & Crystel, J. D. (2007). Metacognition in the rat. *Current Biology, 17*, 551-555.

Fujita, K. (2009). Metamemory in tufted capuchin monkeys (*Cebus apella*). *Animal Cognition, 12*, 575- 585.

Fujita, K., Nakamura, N., Iwasaki, S., & Watanabe, S. (2012). Are birds metacognitive? In M. J. Beran, J. L. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of Metacognition* (pp. 50-61). Oxford University Press: Oxford.

Garson, G. D. (2013). *Generalized linear models/Generalized estimating equations*. Asheboro, NC: Statistical Publishing Associates.

Goto, K., & Watanabe, S. (2012). Large-billed crows (*Corvus macrorhynchos*) have retrospective but not prospective metamemory. *Animal Cognition, 15*, 27-35.

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

Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition, 7*, 239-246.

Herrera, J. P., & Dávalos, L. M. (2016). Phylogeny and divergence times of lemurs inferred with recent and ancient fossils in the tree. *Systematic Biology, 65*, 772-791.

727 Koo, T. R., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation
728 coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155-163.
729
730 MacLean, E. L., Sandel, A. A., Bray, J., Oldenkamp, R. E., Reddy, R. B., & Hare, B. A. (2013).
731 Group size predicts social but not non-social cognition in lemurs. *PLoS One, 8*, 1-9.
732
733 Maille, A., Roeder, J. J. (2012). Inferences about the location of food in lemurs (*Eulemur*
734 *macaco* and *Eulemur fulvus*): a comparison with apes and monkeys. *Animal Cognition, 15*,
735 1075-1083.
736
737 Malassis, R., Gheusi, G., & Fagot, J. (2015). Assessment of metacognitive monitoring and
738 control in baboons (*Papio papio*). *Animal Cognition, 18*, 1347-1362.
739
740 Marsh, H. L. (2014). Metacognitive-like information seeking in lion-tailed macaques: a
741 generalized search response after all? *Animal Cognition, 17*, 1313-1328.
742
743 Marsh, H. L., & MacDonald, S. E. (2012). Information seeking by orangutans: a generalized
744 search strategy? *Animal Cognition, 15*, 293-304.
745
746 Martin, R. D. (2000). Origins, diversity and relationships of lemurs. *International Journal of*
747 *Primatology, 21*, 1021-1049.
748
749 Morgan, G., Kornell, N., Kornblum, T., & Terrace, H. S. (2014). Retrospective and prospective
750 metacognitive judgements in rhesus macaques (*Macaca mulatta*). *Animal Cognition, 17,*
751 249- 257.
752

753 Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102-116.
754
755 Paukner, A., Anderson, J. R., & Fujita, K. (2006). Redundant food searches by capuchin
756 monkeys (*Cebus apella*): a failure of metacognition? *Animal Cognition, 9*, 110-117.
757
758 Roberts, W. A., McMillan, N., Musolino, E., & Cole, M. (2012). Information seeking in
759 animals: metacognition? *Comparative Cognition & Behavior Reviews, 7*, 85-109.
760
761 Rosati, A. G., & Santos, L. R. (2016). Spontaneous metacognition in rhesus monkeys.
762 *Psychological Science, 27*, 1181-1191.
763
764 Smith, J. D. (2009). The study of animal metacognition. *Trends in Cognitive Science, 13*, 389-
765 396.
766
767 Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain
768 response in the bottlenosed dolphin (*Tursiops truncatus*). *Journal of Experimental*
769 *Psychology: General, 124*, 391-408.
770
771 Smith, J. D., Shields, W. E., Schull, J., & Washburn, D. A. (1997). The uncertain response in
772 humans and animals. *Cognition, 62*, 75-97.
773
774 Smith, J. D., Redford, J. S., Beran, M. J., & Washburn, D. A. (2010). Rhesus monkeys (*Macaca*
775 *mulatta*) adaptively monitor uncertainty while multi-tasking. *Animal Cognition, 13*, 93-
776 101.
777

778 Smith, T. R., Smith, J. D., & Beran, M. J. (2018). Not knowing what one knows: A meaningful
779 failure of metacognition in capuchin monkeys. *Animal Behavior and Cognition, 5*, 55-67.
780
781 Smith, J. D., & Washburn, D. A. (2005). Uncertainty monitoring and metacognition by animals.
782 *Current Directions in Psychological Science, 14*, 19-24.
783
784 Suda-King, C. (2008). Do orangutans (*Pongo pygmaeus*) know when they do not remember?
785 *Animal Cognition, 11*, 21-42.
786
787 Suda-King, C., Bania, A. E., Stromberg, E. E., & Subiaul, F. (2013). Gorillas' use of the escape
788 response in object choice memory tests. *Animal Cognition, 16*, 65-84.
789
790 Templer, V. L., & Hampton, R. R. (2012). Rhesus monkeys (*Macaca mulatta*) show robust
791 evidence for memory awareness across multiple generalization tests. *Animal Cognition, 15*,
792 409-419.
793
794 Vining, A. Q., & Marsh, H. L. (2015). Information seeking in capuchins (*Cebus apella*): A
795 rudimentary form of metacognition? *Animal Cognition, 18*, 667-681.
796
797 Yoder, A. D., & Yang, Z. (2004). Divergence dates for Malagasy lemurs estimated from
798 multiple gene loci: geological and evolutionary context. *Molecular Ecology, 13*, 757-773.
799

800 Table 1

801 **Table 1**: Subjects' age, sex, enclosure number, experiment participation and trials to criterion
802 performance (9/10 correct responses in two blocks of 10 trials) in the training phase.
803

Subject ID	Species	Age (years)	Sex	Enclosure	Experiments Completed (T = Training)	Number of trials to criterion	Training trials (number received***)
RB1	Red Bellied	2	F	1	T,1,2	20	29
RB2	Red Bellied	9	F	1	T,1,2	20	41
RB3	Red Bellied	1	F	1	None	-	-
RB4	Red Bellied	14	M	1	None	-	-
RB5	Red Bellied	18	F	3	T,1,2	20	20
RB6	Red Bellied	3	F	3	T,1,2	20	20
RT1	Ring Tailed	11	M	1/2*	T,1,2	20	20
RT2	Ring Tailed	6	F	1/2*	T**	20	20
RT3	Ring Tailed	3	F	2	None	-	-
RT4	Ring Tailed	3	F	2	T,1,2	20	24
RT5	Ring Tailed	3	F	2	T,1,2	20	46
RT6	Ring Tailed	3	F	2	T,1,2	20	36
RT7	Ring Tailed	3	F	2	T**	20	44
BW1	Black & White	14	F	3	T,1,2	20	30
BW2	Black & White	14	F	3	T,1,2	20	22
BW3	Black & White	15	F	3	None	-	-

804 *Subjects RT1 and RT2 were moved from enclosure 1 to enclosure 2 after completion of training.
805 ** Subjects RT2 and RT7 died before Experiment 1 commenced.
806 *** As subjects were not separated during testing, access could not be restricted to the apparatus.
807 Thus, all individuals who reached criterion level did so in the first two blocks (Trials = 20), but continued
808 to make responses on other trials. We include how many responses each individual made in total for
809 transparency (see text for further explanation).

810

811

812 Table 2

813

814 **Table 2**: Number of correct and incorrect responses, no selections and binomial results by
815 subject in Experiment 1 (* = significant at 0.05 level, ** = significant at 0.01 level). If subjects did
816 not make a response (no selection), we did not include this in the binomial test or GEE analyses.

817

Subject ID	Species	Sex	Correct responses	Incorrect responses	No selection	Binomial results
RB1	Red Bellied	F	11	5	4	0.067
RB2	Red Bellied	F	20	0	0	0.000**
RB5	Red Bellied	F	15	1	4	0.000**
RB6	Red Bellied	F	19	1	0	0.000**
RT1	Ring Tailed	M	19	1	0	0.000**
RT4	Ring Tailed	F	15	5	0	0.015*
RT5	Ring Tailed	F	20	0	0	0.000**
RT6	Ring Tailed	F	18	1	1	0.000**
BW1	Black & White	F	16	3	1	0.002**
BW2	Black & White	F	16	3	1	0.002**

818

819

820 Table 3

821

822 Table 3: Number of correct and incorrect responses, no selections and binomial results by
823 subject in Experiment 2. Binomial test results relate to performance on opaque tubes only as
824 all subjects responded correctly on all clear tube presentations, with seen and unseen trials
825 pooled for these analyses (* = significant at 0.05 level, ** = significant at 0.01 level). If subjects
826 did not make a response (no selection), we did not include the trial in the binomial tests tor
827 GEE analyses.

828

829

Subject ID	**Correct responses (clear N = 10)**	**Correct responses (opaque N = 20)**	**Incorrect responses (opaque N = 20)**	**No selection responses (opaque N = 20)**	**Binomial results on opaque tubes**
RB1	10	16	1	3	0.002**
RB2	10	16	1	3	0.002**
RB5	10	19	1	0	0.000**
RB6	10	18	1	1	0.000**
RT1	10	16	3	1	0.001**
RT4	10	19	0	1	0.000**
RT5	10	20	0	0	0.000**
RT6	10	18	1	1	0.000**
BW1	10	14	6	0	0.369
BW2	10	18	2	0	0.000**

830

831

832

833 Table 4

834 **Table 4:** GEE model parameters for Experiment 2 with **number of looks** as the dependent variable and
835 trial type (seen or unseen) and tube type (clear or opaque) as factors. Only trials where a response was
836 made are considered. Significant results are highlighted in bold text (*p < 0.05, ** p < 0.01). Trial type
837 did significantly influence number of looks with more looks occurring in the unseen trials. Tube type did
838 not impact looking behaviour, however there were more looks displayed when the tube was opaque
839 and baiting unseen compared to seen, clear trials.

Parameter	**Label**	**ß**	**df**	**X^2**	**P-value**
Trial type	Seen	-0.168	1	4.543	**0.033***
	Unseen	.		.	.
Tube type	Clear	-0.127	1	2.275	0.253
	Opaque	.		.	.
Trial type*tube type	Seen*Clear	-0.319		6.732	**0.009****
	Seen*Opaque	-0.158		2.799	0.094
	Unseen*Clear	-0.132		1.305	0.253
	Unseen*Opaque	.		.	.

840

841

842

843 Figure 1

844

845

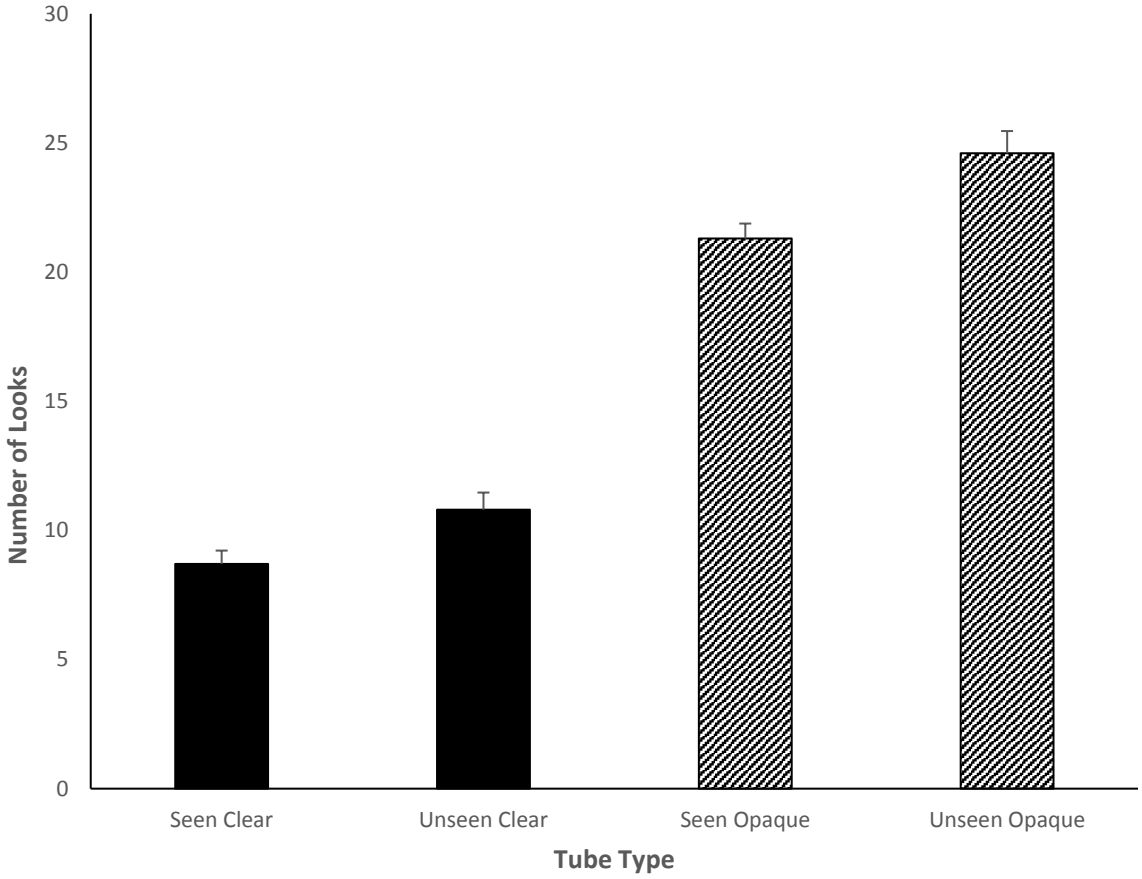
846 **Figure 1:** (a) Clear tubes used during training and Experiment 2, showing wooden stoppers used
847 throughout testing. (b) Opaque tubes used in Experiment 1 and 2. (c) Wooden ramp used during
848 Experiment 2 to raise the end where subjects made a choice, shown with opaque tubes mounted on
849 top.

850

851

852 Figure 2

853



854
855

856

857 **Figure 2**: Mean number of looks observed in seen and unseen trials using clear (solid bars, N = 10) and
858 opaque (patterned bars, N = 20) tubes. Error bars represent SE. With both clear and opaque tubes,
859 more looks occurred in the unseen condition.

860

861

862

863

864

865

Do lemurs know when they could be wrong?: An investigation of information seeking in three species of lemur (*Lemur catta*, *Eulemur rubriventer*, *Varecia variegata*).

Heather K. Taylor, Clare L. Cunningham and Scott Hardie

Evolutionary and Biological Approaches to Behaviour Research Group

Division of Psychology, Abertay University

Electronic Supplementary Material:

In all models, subject identity is the within subject factor and trial number, the between subject factor. Alpha is set at 0.05 for all significance values. Only trials where a response was made are included in the analyses as a null response does not necessarily imply a lack of knowledge.

Table S1: GEE models for Experiment 1 with **number of correct responses** as the dependent variable and trial type and presence or absence of looking behaviour as factors, showing goodness of fit (QIC), parameter estimates and significance (*p < 0.05, ** p < 0.01). Bold text identifies best model (based on QIC) reported in manuscript and significant results.

Model	QIC	Parameter	Label	ß	X^2	P-value
1	130.032	Trial type	Seen	-0.655	1.760	0.185
			Unseen	.	.	.
2	131.504	Trial type	Seen	-0.704	0.1.961	0.161
			Unseen	.	.	.
		Looking behaviour	Present	0.819	0.508	0.476
			Absent	.	.	.

Table S2: GEE models for Experiment 1 with **number of looks** as the dependent variable trial type as a factor, showing goodness of fit (QIC), parameter estimates and significance (*p < 0.05, ** p < 0.01). Bold text identifies model reported in manuscript and significant results.

Model	QIC	Parameter	Label	ß	X^2	P-value
1	**71.356**	Trial type	Seen	-0.087	0.477	0.490
			Unseen	.	.	.

Table S3: GEE models for Experiment 1 with **trial duration** as the dependent variable with trial type as a factor, showing goodness of fit (QIC), parameter estimates and significance (*p < 0.05, ** p < 0.01). Bold text identifies model reported in manuscript and significant results.

Model	QIC	Parameter	Label	ß	X^2	P-value
1	29883.641	Trial type	Seen	-0.276	0.022	0.881
			Unseen	.	.	.

887

888 **Table S4:** GEE models for Experiment 2 with **number of correct responses** as the dependent variable
889 with trial type as a factor, showing goodness of fit (QIC), parameter estimates and significance (*p <
890 0.05, ** p < 0.01). Bold text identifies model reported in manuscript and significant results. Only
891 responses with opaque tubes are included in the analyses as all lemurs responded correctly on all
892 presentations with the clear tubes where reward placement was always visible at the time of choice.

Model	QIC	Parameter	Label	ß	X^2	P-value
1	**107.330**	Trial type	Seen	-1.171	3.770	0.052
			Unseen	.	.	.

893

894 **Table S5:** GEE models for Experiment 2 with **number of looks** as the dependent variable with trial type
895 as factor, showing goodness of fit (QIC), parameter estimates and significance (*p < 0.05, ** p < 0.01).
896 Bold text identifies model reported in manuscript and significant results. Only responses with opaque
897 tubes are included in the analyses as all lemurs responded correctly on all presentations with the clear
898 tubes where reward placement was always visible at the time of choice.

Model	QIC	Parameter	Label	ß	X^2	P-value
1	33.991	Trial type	Seen	-0.158	2.799	0.094
			Unseen	.	.	.

899

900 **Table S6:** GEE models for Experiment 2 with **number of looks** as the dependent variable with trial type
901 and tube type as factors, showing goodness of fit (QIC), parameter estimates and significance (*p <
902 0.05, ** p < 0.01). Bold text identifies models reported in manuscript and significant results. Both
903 opaque and clear tubes are included in these analyses.

Model	QIC	Parameter	Label	ß	X^2	P-value
1	**56.802**	Trial type	Seen	-0.168	4.543	**0.033***
			Unseen	.	.	.
		Tube type	Clear	-0.127	2.275	0.253
			Opaque	.	.	.
			Seen*Clear	-0.319	6.732	**0.009****

	Trial type*tube type	Seen*Opaque	-0.158	2.799	0.094
		Unseen*Clear	-0.132	1.305	0.253
		Unseen*Opaque	.	.	.

Table S7: GEE models for Experiment s with **trial duration** as the dependent variable with trial type and tube type as factors, showing goodness of fit (QIC), parameter estimates and significance (*p < 0.05, ** p < 0.01). Bold text identifies best model reported in manuscript and significant results.

Model	**QIC**	**Parameter**	**Label**	**ß**	**X^2**	**P-value**
1	16782.764	Trial type	Seen	-0.179	0.040	0.841
			Unseen	.	.	.
2	16721.271	Trial type	Seen	-0.195	0.048	0.827
			Unseen	.	.	.
		Tube type	Clear	-0.979	1.083	0.298
			Opaque	.	.	.