Evolution and Human Behavior xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens

Rhesus macaques use probabilities to predict future events

Francesca De Petrillo[a,b,d,*], Alexandra G. Rosati[a,c]

[a] Department of Psychology, University of Michigan, 530 Church St, Ann Arbor, MI 48109, USA
[b] Institute for Advanced Study in Toulouse, Manufacture des Tabacs, 21, Allée de Brienne, 31015 Toulouse, France
[c] Department of Anthropology, University of Michigan, USA
[d] Istituto di Scienze e Tecnologie della Cognizione, CNR, Rome, Italy

ARTICLE INFO

Keywords:
Intuitive statistics
Probabilistic reasoning
Logical inferences
Non-human primates

ABSTRACT

Humans can use an intuitive sense of statistics to make predictions about uncertain future events, a cognitive skill that underpins logical and mathematical reasoning. Recent research shows that some of these abilities for statistical inferences can emerge in preverbal infants and non-human primates such as apes and capuchins. An important question is therefore whether animals share the full complement of intuitive reasoning abilities demonstrated by humans, as well as what evolutionary contexts promote the emergence of such skills. Here, we examined whether free-ranging rhesus macaques (*Macaca mulatta*) can use probability information to infer the most likely outcome of a random lottery, in the first test of whether primates can make such inferences in the absence of direct prior experience. We developed a novel expectancy-violation looking time task, adapted from prior studies of infants, in order to assess the monkeys' expectations. In Study 1, we confirmed that monkeys (n = 20) looked similarly at different sampled items if they had no prior knowledge about the population they were drawn from. In Study 2, monkeys (n = 80) saw a dynamic 'lottery' machine containing a mix of two types of fruit outcomes, and then saw either the more common fruit (*expected trial*) or the relatively rare fruit (*unexpected trial*) fall from the machine. We found that monkeys looked longer when they witnessed the unexpected outcome. In Study 3, we confirmed that this effect depended on the causal relationship between the sample and the population, not visual mismatch: monkeys (n = 80) looked equally at both outcomes if the experimenter pulled the sampled item from her pocket. These results reveal that rhesus monkeys spontaneously use information about probability to reason about likely outcomes, and show how comparative studies of nonhumans can disentangle the evolutionary history of logical reasoning capacities.

1. Introduction

One of the hallmarks of human cognition is the ability to form rational predictions about future events, a cognitive skill that underpins logical and mathematical reasoning (Cesana-Arlotti, Téglás, & Bonatti, 2012; Fontanari, Gonzalez, Vallortigara, & Girotto, 2014; Gopnik, 2012). Accordingly, humans appear to be particularly talented at drawing conclusions and inferring generalizable knowledge from limited data (e.g., Denison & Xu, 2012; Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Tenenbaum, Griffiths, & Kemp, 2006). For example, scientists can make generalizations about the laws underpinning a phenomenon on the basis of an observed sample of data, and use estimates of probability to make predictions about what will happen in the face of uncertain outcomes (Cesana-Arlotti et al., 2012). This ability for estimating probabilities to predict likely outcomes can support complex inferences about what is likely to occur in the future. What are the origins of these foundational human abilities? In the current study, we examined whether rhesus monkeys (*Macaca mulatta*) can make such inferences in order to trace the evolutionary history of statistical inferences.

A traditional view has been that the ability to make logical inferences that incorporate knowledge about probabilities is dependent on language and formal education (Piaget & Inhelder, 1975; Tversky & Kahneman, 1981). Indeed, research with adults suggests that engaging in these sorts of logical inferences can be very challenging, and adults routinely use heuristics when reasoning about probability information (Kahneman & Tversky, 2013; Tversky & Kahneman, 1974). For example, people often misconstrue the probability of events when they occur in sequence, judging a sequence of head-tail-head-tail-tail to be more likely than a sequence of head-head-head-tail-tail, even though each outcome has a 50% probability to occur in each independent toss (Tversky & Kahneman, 1974; Nickerson, 2002; Burns & Corpus, 2004).

* Corresponding author at: Institute for Advanced Study in Toulouse, Manufacture des Tabacs, 21, Allée de Brienne, 31015, Toulouse, France
E-mail address: fdepetri@umich.edu (F. De Petrillo).

https://doi.org/10.1016/j.evolhumbehav.2019.05.006
Received 12 October 2018; Received in revised form 22 February 2019; Accepted 20 May 2019
1090-5138/ © 2019 Elsevier Inc. All rights reserved.

Please cite this article as: Francesca De Petrillo and Alexandra G. Rosati, Evolution and Human Behavior, https://doi.org/10.1016/j.evolhumbehav.2019.05.006

This kind of evidences suggests that sophisticated statistical reasoning develops late in human ontogeny and may still depend on simple heuristics throughout adulthood. However, recent evidence indicates that people without relevant formal education can nonetheless demonstrate some sophisticated abilities. For example, comparisons of indigenous Mayan groups have shown that individuals can solve fairly complex statistical problems by considering prior and posterior information to predict the occurrence of random outcomes, and even combine multiple probabilities to accurately predict outcomes (Fontanari et al., 2014). This suggests that while language and formal education may facilitate abilities to reason about probabilities, they are not necessary for such abilities to emerge. Rather, language and education may build upon pre-existing cognitive abilities.

Another line of evidence that abilities to reason about probabilities have deep roots comes from studies of preverbal infants. Recent work shows that even young infants are 'intuitive statisticians,' and can make many inferences about probability prior to the emergence of language (Téglás, Girotto, Gonzalez, & Bonatti, 2007; Xu & Garcia, 2008; Téglás et al., 2011; Denison, Reed, & Xu, 2013; Xu & Denison, 2009; Denison & Xu, 2014). Much of this infant work has employed violation-of-expectation looking time methods. These methods capitalize on the fact that infants tend to look longer at events that violate their expectations because they make rational predictions about how the world typically works (Sim & Xu, 2019). In one such study (Xu & Garcia, 2008), eight-month-old infants were first presented with a box containing a population of mostly red ping-pong balls and just a few white ping-pong balls. Infants watched an adult remove a sample of either four red and one white ping-pong balls, or one red and four white ping-pong balls while the particular contents of the box were occluded, and then the original population was revealed. In fact, infants looked longer when it turned out that a statistically unlikely sample was drawn from the population, indicating they expected the demonstrator to pull out more of the common item in the population. These expectations about probability further appear to guide active choices, as 12-month-old infants can both use proportions to predict the outcome of a random draw and also use this prediction to approach and obtain a preferred reward when actively searching (Denison & Xu, 2014).

A final piece of evidence that cognitive abilities for reasoning about probabilities can emerge in the absence of language or formal schooling comes from comparisons of nonhuman animals. Recent work has examined whether animals exhibit an intuitive sense of statistics similar to those seen in babies, and revealed that at least some primates can make similar judgments. For example, chimpanzees (*Pan troglodytes*), gorillas (*Gorilla gorilla*), orangutans (*Pongo pygmaeus*), bonobos (*Pan paniscus*) and capuchin monkeys (*Sapajus spp.)* can use information about probability to select options that are most likely to provide a high-value reward (Eckert, Call, Hermes, Herrmann, & Rakoczy, 2018a, 2018b; Eckert, Rakoczy, & Call, 2017; Rakoczy et al., 2014; Tecwyn, Denison, Messer, & Buchsbaum, 2017). In the basic setup for these studies, primates are presented with two containers of food which differ in their ratio of preferred and non-preferred foods. One item is then drawn from each container, without the primate seeing its exact identity. Across several trials, primates, can use information about the proportion of their preferred item in the container to select the container most likely to provide a high-value reward. Other work suggests that primates are also capable of using some forms of probabilistic reasoning in other contexts (Drucker, Rossa, & Brannon, 2016; Hanus & Call, 2014). For example, when rhesus macaques are presented with arrays with different ratios of positive and negative stimuli, they successfully choose the array with the higher ratio of positive stimuli in order to gain a reward (Drucker et al., 2016). Overall, this evidence indicates that at least some primate species use expectations about probability to guide their active choices and optimize rewards.

However, all comparative work to date has involved the experience of rewarded outcomes over multiple trials. This kind of experiential learning could potentially depend on learning mechanisms that are distinct from the more spontaneous forms of statistical demonstrated by young children. Infants' skills for drawing inferences from and about probability information also allows them to make accurate predictions about future events across multiple contexts, even ones that they have never previously experienced. For example, in one study (Téglás et al., 2007), infants saw a novel animated lottery machine device containing three yellow objects and one blue object moving around inside. They then observed either one of the yellow objects or the blue object exit the machine. Infants correctly predicted that the more common object was more likely to come out of the machine than the less common object. Thus, preverbal infants could infer that when a single object is randomly drawn from a larger population, it is likely to be part of the most numerous kind—without ever seeing such an event previously. Infants' predictions in these contexts are also fairly flexible, as they can integrate new physical and numerical information into their expectations (Téglás et al., 2011). For example, when infants were presented with a similar lottery machine task, but the spatial position of the objects as well as duration that the machine was occluded from their sight before the object exited, they could appropriately use this information to update their predictions. When the occlusion time was long, infants preferentially used information about the ratio of the objects in the machine, whereas if this time was brief they instead considered the original *proximity* of the objects to the exit point when forming their expectations. While infants can make rapid statistical inferences in the absence of extensive experience, it is currently unclear if nonhumans can do the same, versus whether they require more extensive experience with outcomes to engage in probabilistic reasoning.

More broadly, the increasing evidence for some forms of intuitive statistical reasoning across several primate species sets the stage to ask new questions about the evolutionary processes building these skills. Are statistical inferential abilities widely shared across primates, or they independently evolved in a few distinct taxa? In fact, it is plausible that abilities for statistical reasoning in apes and capuchins represent instances of independent cognitive convergence. New World monkeys (including capuchins) and Old World primates (including apes and macaques) diverged approximately 40–45 million years ago (Perelman et al., 2011). Yet despite this evolutionary distance, capuchins show several striking behavioral similarities with apes that are not seen in other New World monkey species (De Petrillo & Rosati, 2019), such as tool use and risky hunting behavior (Fragaszy, Izar, Visalberghi, Ottoni, & de Oliveira, 2004; Rose, 1997; Visalberghi, Sirianni, Fragaszy, & Boesch, 2015). This behavioral convergence with chimpanzees mirrors the similarities between capuchin and chimpanzee cognition. In matched comparison with chimpanzees and bonobos, capuchin monkeys and chimpanzees both exhibit a similar preference for risky outcomes compared to bonobos (De Petrillo, Ventricelli, Ponsi, & Addessi, 2015; Heilbronner, Rosati, Stevens, Hare, & Hauser, 2008). As such, it might be that these inferential abilities are seen in only a few highly specialized primate species living in especially variable and complex foraging environments.

On the other hand, the ability to make some sort of statistical inferences is thought to be a crucial component of foraging, and therefore might be widely shared in some form. In particular, many wild animals search for food in variable, heterogeneous environments, so it is plausible that cognitive mechanisms for tracking statistical regularities concerning where or when food is likely to be found will allow animals to be more effective foragers. Indeed, extensive research has revealed that many diverse animal species are sensitive to probabilistic variation in rewards distributions in other contexts (e.g., McCoy & Platt, 2005; Hayden & Platt, 2007; Heilbronner et al., 2008; Hayden & Platt, 2009; Stevens, 2010; Haun, Nawroth, & Call, 2011; Rosati & Hare, 2012, 2013; MacLean, Mandalaywala, & Brannon, 2012; De Petrillo et al., 2015; see also Kacelnik & Bateson, 1996,Rosati, 2017, and Santos and Rosati, 2015 for reviews). For example, when presented with choices between a safe option, yielding a constant food reward, and a risky option, yielding a reward that varies probabilistically around the mean,

many different primate species including great apes (Haun et al., 2011; Heilbronner et al., 2008; Rosati & Hare, 2012, 2013), capuchin monkeys (De Petrillo et al., 2015), rhesus macaques (Hayden & Platt, 2009; McCoy & Platt, 2005), mangabey monkeys (*Cercocebus torquatus torquatus*; Rivière, Stomp, Augustin, Lemasson, & Blois-Heulin, 2018) and various lemurs (MacLean et al., 2012) are sensitive to risk and incorporate information about probabilistic variation in rewards into their decision preferences. Thus, statistical inferences may play a crucial role in foraging behaviors more generally.

Examining whether a broader set of primate species exhibit an intuitive sense of probability, and linking variation in this cognitive ability to species' phylogeny and natural history, is crucial to understand the evolutionary contexts that favors the emergence of some kinds of cognitive capacities over the others. In the current study, we therefore examined whether a human-like ability to make flexible and spontaneous statistical inferences is shared with rhesus macaques (*Macaca mulatta*). Rhesus macaques are an important animal model for human decision-making, and they also exhibit strategies that characterize human decision-making when faced with decisions about risk (Hayden & Platt, 2009; McCoy & Platt, 2005; Yamada, Tymula, Louie, & Glimcher, 2013). However, no prior work has examined whether rhesus macaques exhibit an 'intuitive' sense of probability. While there is robust evidence that capuchins and apes can make statistical inferences to acquire food rewards over repeated trials (Eckert et al., 2017; Eckert, Call, et al., 2018a, 2018b; Rakoczy et al., 2014; Tecwyn et al., 2017), there is more mixed evidence from macaque species. In particular, when tested with a similar procedure used with apes and capuchins, a few individual long-tailed macaques (*Macaca fascicularis*) appeared to use statistical reasoning to select options that are most likely to provide the high-value reward, but they did not exhibit this preference at the group level (Placì, Eckert, Rakoczy, & Fischer, 2018). Thus, there is not currently strong evidence for intuitive statistical abilities in macaques.

Here, we specifically examined if rhesus monkeys can use intuition about probability to form rational expectations about the outcome of a single future event based upon its likelihood in the absence of direct experience, an ability that has not been demonstrated in nonhumans. To do so, we tested a large sample of semi-free-ranging monkeys, using a new expectancy looking time task modeled on prior work with infants (Téglás et al., 2007). This allowed us to examine statistical inference without providing rewarded experiences over trials. In the basic setup, monkeys observed an artificial 'lottery' machine containing a 3:1 distribution of two types of artificial fruit. The machine consisted of a clear plastic cube that could be rotated, similar to a lottery or raffle drum. Then they saw a single item emerge from the machine which matched either the more common fruit in the machine's population of items (*expected condition*) or the relatively rare fruit (*unexpected condition*). If monkeys are able to make expectations about a future single event based on its likelihood, then they should look longer at the unexpected sample.

In Study 1, we first confirmed that monkeys looked similarly at two types of fruit in the absence of information about the population from which they were sampled, as a basic validation of our setup. In Study 2, we then examined whether monkeys formed expectations about the most likely outcome of a sample randomly drawn from the lottery machine, the crucial test of whether they can infer most likely future outcomes without direct experience. We further examined the responses of monkeys who witnessed statistically improbably events, versus those who witnessed statistically impossible events. When infants observe the outcome of a lottery that is impossible (because the sampled object was physically blocked by a barrier), they also exhibit longer times, similar to responses in conditions where infants see statistically improbable events (Téglás et al., 2007). Finally, in Study 3, we examined whether expectations depend on a causal relationship between the sampled items and the lottery population. In particular, we tested whether macaques were only sensitive to the perceptual mismatch between the sampled item and the population (Xu & Garcia, 2008), versus whether their expectations specifically depended on the sample being causally related to the population. This last study is crucial to provide evidence that monkeys understand that the sample is drawn from a specific population and that they actually use this information, rather than simple perceptual heuristics, to form predictions about future outcomes.

2. Study 1: fruit preference validation

As our main lottery task involved monkeys' reasoning about different types of fruit outcomes, we first wanted to choose outcome stimuli that were approximately of equal interest to monkeys. Prior work has shown that monkeys are very interested in looking at similar artificial fruits (e.g., Hughes & Santos, 2012; Marticorena, Ruiz, Mukerji, Goddu, & Santos, 2011), and generally motivated to approach and search for such fruits (e.g., Rosati & Santos, 2016). However, it was unclear if they would exhibit a greater intrinsic interest in one type of fruit over another, which was a crucial issue for the lottery procedure in Studies 2 and 3. Here, we confirmed that monkeys look similarly at different fruits, in the absence of knowledge about the original population. In particular, monkeys observed a single fruit emerging from a rotating lottery machine (see Fig. 1). While they could hear that the machine contained several items, in this study they could not actually see the distribution of items in the machine, as it was visually obscured by a curtain (see Video S1). Thus, we could assess monkeys' relative looking time to our two possible outcomes in the same setup used in subsequent studies, but in the absence of any information about the distribution of fruit in the population. Monkeys were tested in a within-subjects design such that each individual witnessed both fruit types across two successive trials.

2.1. Ethics statement

All non-invasive behavioral tests were approved by the Institutional Animal Care and Use Committee (IACUC) for the University of Puerto Rico Medical Sciences Campus (protocol #A140116) and adhered to site guidelines for animal research.

2.2. Subjects

We tested 20 free-ranging rhesus macaques (10 males and 10 females; ranging from 1.9 to 12.4 years old, M = 5.8 years) living at the Cayo Santiago Field Station in Puerto Rico. At the time of testing, this field site had a population of approximately 1700 rhesus macaques, who free-range through the island's 38 acres. This population is well-habituated to human observers, and monkeys are also accustomed to participating in a variety of cognitive experiments (e.g., Drayton & Santos, 2017; Hughes & Santos, 2012; Marticorena et al., 2011; Martin & Santos, 2014; Rosati, Arre, Platt, & Santos, 2016; Rosati & Santos, 2016). Individual monkeys are identifiable through a combination of ear-notches and tattoos. The monkeys tested in this study were naive to the lottery machine apparatus. To be included in the study, monkeys had to successfully complete two test trials. An additional 7 monkeys were approached by the experimenters but were excluded because they approached the apparatus to try to steal the fruit (n = 4) or because of apparatus failure during the test (n = 3).

2.3. Apparatus

The apparatus consisted of a black box made of poster board (60 cm high x 50 cm deep x 42 cm wide), containing a transparent cube (16.5 cm^3). A population of 20 artificial (plastic) fruit items was located inside the cube, which could be rotated with a stick to make the fruit visibly and audibly move around (see Fig. 1d). In this study, the lottery cube was always occluded with a blue curtain drawn across the front, so monkeys could hear the fruits moving inside while it rotated, but they

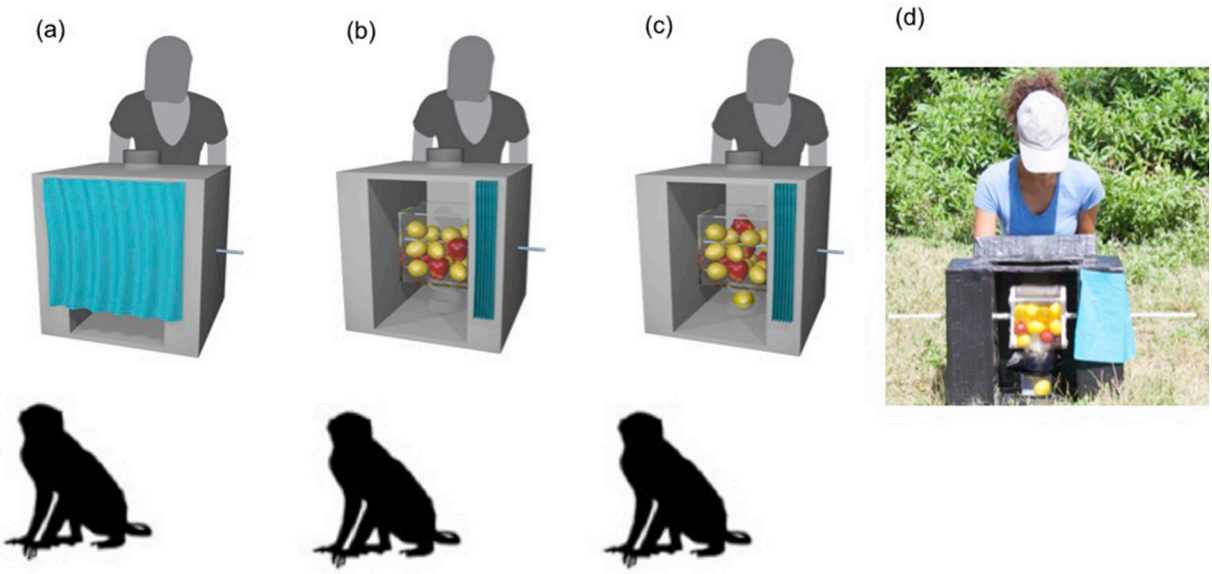

Fig. 1. Apparatus and procedure for Study 1 (Fruit Preference) and Study 2 (Probability Inference). (a) A curtain initially hides the population from the monkey's view. In Study 1 this curtain stays closed for the duration of the task. (b) In the procedure for Study 2, the experimenter opens the curtain. In both studies, the experiment rotates the cube three times. (c) A sample that either matches (*expected trial*) or does not match (*unexpected trial*) the population drops from the machine; in Study 1 the curtain remains closed such that animals did not have information about the contents of the population, whereas in Study 2 the curtain was open as pictured. (d) Picture of the actual setup.

could not directly see the distribution inside the machine; in subsequent studies, this curtain was opened to reveal the population inside the lottery machine. At the bottom of the box, directly below the cube, was a small transparent cup (9 cm in diameter × 11.5 cm high). This allowed the sample to drop out of the machine into a controlled location and then remain visible to the monkey. In this study, the machine itself was located behind a blue curtain such that monkeys could only hear it rotate, and then see a single item drop below the curtain into the transparent cup.

To control the outcome of the random draw, the experimenter could manipulate a peg mechanism on the back of the apparatus out of the monkey's sight, such that a specific fruit would appear to drop from the machine based on condition (e.g., either an apple or a lemon). This was accomplished because, unbeknownst to the monkey, the clear tube emerging from the lottery cube was actually divided in two compartments, one with each kind of fruit, that were blocked by a peg. The experimenter could manipulate these outcomes from the back of the apparatus, without their hands being visible to the monkey. Thus, it was possible for the experimenter to select the outcome for each trial by opening the corresponding compartment by removing that peg (see Video S1 for an example). While in the infants' studies using this setup (Téglás et al., 2007, 2011) the total population was made of only 4 items, in our study we increased the number of items in the population to 20 to ensure that the hidden mechanism controlling which fruit dropped from the lottery cube was not visible to the monkeys while the lottery cube was rotated (in subsequent studies where the curtain was opened). Importantly, we kept the ratio between the items of the population equal to that used with infants (e.g., 3:1).

2.4. Procedure

In tests, two experimenters located a calmly-sitting monkey. Experimenter 1 (E1), the demonstrator, presented the stimuli to the monkey sitting 1-2 m away, and Experimenter 2 (E2), the cameraperson, filmed the monkey's looking response. Subjects experienced two trials, each consisting of a *rotation event* whereby E1 manipulated a lottery machine, and a *sampling event* whereby a single item dropped out from the lottery machine into a clear container below it (see Fig. 1a and Video S1). E1 used a stick behind the apparatus (e.g., out of the monkey's view) to rotate the lottery cube three times, a process that took approximately 6 s. As the curtain was drawn across the front of the apparatus in this study, the monkey could hear the rotation event but could not see the contents of the population.

Then E1 stopped rotating the cube, called the monkey to ensure she was looking, and then released one item (a lemon or an apple). To ensure that the demonstrator was as blind as possible to conditions while running the live tests, E2 assigned the condition and told E1 which condition to run using a verbal code immediately before E1 released the fruit outcome (e.g. the demonstrator was blind to condition up to that final point). Only E1 knew the correspondence between the code and the condition; for example, E2 would announce that the condition was “left”, without knowing that “left” corresponded to the apple outcome (importantly, E2 could not see the fruit released from their filming position). That is, the experimenter assigning the condition did not manipulate the apparatus and was blind to the condition they were assigning. As E1 released the peg, she said “now” and looked downwards to avoid direct eye contact with the monkey during the rest of the trial. E2 video recorded the monkey's looking response for 10 s starting when E1 said “now”, calling “stop” when the 10s trial had completed over. E1 repeated the same basic procedure during the second trial (with order of trials was counterbalanced across subjects).

2.5. Data coding and statistical analysis

All sessions were videotaped and scored from video by two independent coders. To do so, we clipped individual trials out from longer test sessions and assigned each trial a random clip ID so they could be coded blind to condition and subject. Each trial was clipped to start a few seconds before E1 said “now” and end just after E2 said “stop”. The two coders then examined the videos frame-by-frame using the program MPEG Streamclip (30 frames per second). All monkeys' looking were coded for the 10-s immediately following the start of the demonstrator's “now” signal, following standard procedures for looking time task in this population (Drayton & Santos, 2016; Drayton & Santos, 2018; Marticorena et al., 2011; Martin & Santos, 2014). Thus, we could determine the total duration of time the monkey spent looking at the apparatus during the 10s long trial. Inter-rater reliability between the two coders was high (Pearson's $r = 0.98$).

ARTICLE IN PRESS

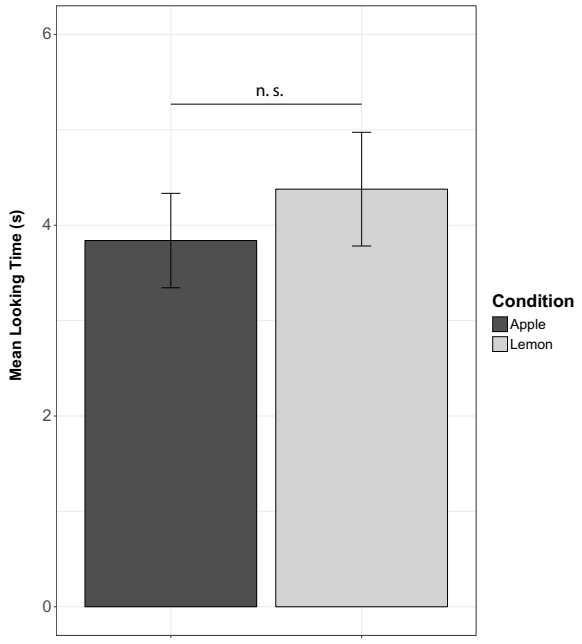

Fig. 2. Mean looking at outcomes in the absence of population information in Study 1 (Fruit Preference validation). Monkeys watched either a lemon or an apple fall from the lottery machine, without knowledge of the distribution of the lottery population. Graph shows the overall mean duration of looking time towards the different outcome types; error bars indicate standard error.

To analyze the data, we implemented linear mixed models (Baayen, 2008) in R version 3.4.1 (R Development Core Team, 2017). We used the *lmer* function from the LME4 software package (Bates, 2010) to examine looking times within subjects, and compared the fit of different models using likelihood ratio tests (LRT; Bolker et al., 2009). Models were refit using maximum likelihood for model comparisons, and reported parameter significance was calculated using the lmerTest (Kuznetsova, Brockhoff, & Christensen, 2015).

2.6. Data availability

Data from all studies is available from Dryad Digital Repository: https://doi.org/10.5061/dryad.22mv245.

2.7. Results and discussion

Overall, monkeys looked an average of 3.8 ± 0.49 s at the apple, and 4.4 ± 0.60s at the lemon, which was not significantly different overall (paired samples *t*-test: $t_{19} = -1.016$, $p = 0.32$; Cohen's $d = 0.2$; see Fig. 2). To examine what predicted the monkeys' looking responses, we first created a base linear mixed model that included *subject* (as a random factor accounting for repeated measures), *age* (linear in years), *sex* (male or female), and *trial number* (1 or 2) as factors. We included age and sex to account for any overall individual variation in looking responses (but note that this was a within-subjects design, so each individual experienced both fruit types in counterbalanced order), and trial number because monkeys tend to decrease their looking over successive events in similar contexts. This model indicated that the overall looking time decreased with age, and monkeys generally tended to look longer at the first trial than at the second. Then, to test whether monkeys showed different intrinsic interest in these two types of fruit, in the second model we added *trial condition* (apple outcome or lemon outcome) as an additional predictor. We found that including this predictor did not improve model fit compared to the base model [LRT: $\chi^2 = 1.469$, $p = 0.23$], indicating that monkeys looked at both fruit types for similar durations (see Table 1 for parameters from the full model).

Table 1
Parameters from the full model predicting monkey looking responses in Study 1 (Fruit Preference validation).

Predictor	Estimate	S.E.	t value	*p* value
Age (linear in years)	−0.355	0.138	−2.569	= **0.02**
Sex (reference: female)	0.641	0.876	0.732	= 0.47
Trial number (reference: trial 1)	−1.238	0.460	−2.690	= **0.01**
Condition (reference: apple)	0.539	0.460	1.171	= 0.26

Significant value are highlighted in bold.

The results from this study indicated that monkeys looked for a similar amount of time at lemons and apples in the absence of other information about the population they were drawn from. This suggests that the monkeys did not show a strong intrinsic preference for one type of fruit over the other. We also found that the overall duration of looking time decreased with age. This is consistent with previous studies in this population, which have shown that adult and older monkeys spend less time looking at various stimuli (Hughes & Santos, 2012; Rosati et al., unpublished data). Overall, this suggests that monkeys did not find one fruit type to be exceedingly more salient than the other in the absence of information about the population they came from. We therefore validated that the basic setup was appropriate to test the monkeys' inferences about probability in the subsequent studies.

3. Study 2: probability inference

In Study 2, we used the same basic procedure to examine whether monkeys could use probability information to predict the most likely outcome of a random draw. Here, monkeys directly observed the lottery machine (the population) and then saw a single item (the sample) drop out. The sample either matched the more frequent item in the machine (*expected trial*), or was rarer or absent in the population (*unexpected trial*). Monkeys were tested in a within-subjects design such that each individual witnessed both the expected and unexpected outcomes across two trials (order counterbalanced across subjects). In addition, we manipulated whether the unexpected event was improbable (e.g., the rarer item emerged) versus actually impossible (as the item that emerged was not present in the original population; likelihood condition was counterbalanced across subjects). As described earlier, infants show similar looking responses to such 'impossible' outcomes to merely 'improbable' outcomes based on the probability of a given item existing the container based on proportions (Téglás et al., 2007; Téglás et al., 2011). Here we tested whether monkeys also showed similar looking time responses to both kinds of unexpected situations.

3.1. Subjects

We tested 80 rhesus macaques (40 males and 40 females; ranging from 1.7 to 21.4, M = 7.0 years). Monkeys were randomly assigned to experience either the *improbable likelihood condition* or the *impossible likelihood condition*, as well as to either view a lottery machine containing a population composed primarily of lemons (*lemon machine*) or a population of primarily apples (*apple machine* condition).

All of the monkeys were naive to the apparatus, and none had participated in Study 1. Additional monkeys were approached for testing but not included in the study; most of these exclusions were consequences of testing animals in this free-ranging context. In particular, we excluded monkeys because they experienced interference from other monkeys who displaced them or approached the apparatus during testing ($n = 6$); the subject themselves closely approached the apparatus during the test ($n = 9$; note that many fruits were visible in this study, likely leading to more such approaches); the subject walked away from the testing area before completing both trials ($n = 12$); or they stopped attending the scene and did not complete both trials

ARTICLE IN PRESS

($n = 3$). In 11 other cases the test was stopped because of apparatus failure (such the fruit falling out of the bottom cup and rolling onto the ground). Finally, an additional 7 sessions were later excluded because the monkey was judged to be not looking at the scene during the start of a trial ($n = 5$) or they jumped or moved out of the camera frame during the trial ($n = 2$), and therefore their responses could not be coded. These monkeys were not included in the final sample of 80 monkeys. Note that the overall exclusion rate of this experiment was lower than or similar to those of previous looking time studies conducted with this population (e.g., Drayton & Santos, 2017; Marticorena et al., 2011; Martin & Santos, 2014).

3.2. Test conditions

In this study, we examined how three main factors shaped monkeys' expectations about future outcomes. First, we manipulated whether the outcome was *expected* (the fruit that dropped out matched the common fruit in the apparatus) or *unexpected* (a rare or absent fruit dropped out); each monkey experienced both trial types in counterbalanced order in a within-subject comparison. Second, we manipulated whether monkeys observed a statistically unlikely outcome (*improbable condition*) such that a rarer fruit dropped from the machine in the unexpected trial, versus whether they saw a physically impossible outcome (*impossible condition*) such that the machine only contained only one kind of fruit initially, but then a different type of fruit appeared to drop out in the unexpected trial. In particular, half of the monkeys were randomly assigned to the improbable condition where they experienced an apparatus containing a 3:1 distribution of two artificial fruits (i.e., 15 lemons and 5 apples, or vice versa, see video S1), whereas the other half of monkeys in the impossible condition saw a population of only a single type of fruit (e.g., 20 lemons and 0 apples or vice versa). Finally, even though Study 1 did not reveal any differences in looking times to lemons versus apples, we ensured that any intrinsic fruit preferences did not affect our results in the current study by manipulating the contents of the lottery machine across monkeys: half of the monkeys in each likelihood condition were randomly assigned to see a population consisting of predominantly or only apples (*apple machine condition*), and the other half saw a population of predominantly or only lemons (*lemon machine condition*), such that lemons represented the expected outcome.

3.3. Procedure and apparatus

We used the same basic procedure and apparatus as in Study 1, with two main exceptions. As in Study 1, the trial began with a curtain occluding the lottery machine (Fig. 1a), but E1 then opened the curtain revealing the population of fruits inside the lottery machine to the monkey before rotating the drum (Fig. 1b). As in Study 1, she rotated the cube three times and released the corresponding fruit sample (a lemon or an apple according to the trial; see Fig. 1c and Video S1). Again, before each trial, E2 announced the trial type (expected or unexpected) to E1 using a verbal code where only E1 knew the correspondence, such that monkeys were randomly assigned to trial type and E1 was blind to condition during most of the demonstration. Both the likelihood condition (*improbable* and *impossible*) and the type of fruit that was predominant in the population (*apple machine* and *lemon machine*) were counterbalanced across subjects. As the fruit contents of the apparatus could not quickly swapped out (e.g., in the way the trial order could be easily reversed in the course of testing different monkeys) due to the physical structure of the apparatus, the machine's fruit contents were rather switched across different days of testing over the entire study period. The apparatus was identical to that in Study 1 with a minor exception: here a second small transparent container, of the same dimension as the container on the bottom (9 cm in diameter × 11.5 cm high), was also on the top of the box directly above the cube, along with a small black stage behind it (such that a fruit placed in this container would have similar contrast to the backdrop as a fruit in the bottom container, which also had a dark backdrop; see Fig. 1d). This new container was used in Study 3, but was not utilized in the current study.

3.4. Data coding statistical analysis

Videos were processed and coded in the same way as in Study 1: individual trials were clipped from longer sessions and assigned a new random ID, so they could be coded blind to condition and subject. Each trial was coded frame-by-frame for duration of the monkey's looking by two independent coders, who had high inter-rater reliability (Pearson's $r = 0.96$). We used the same basic statistical analysis approach as in Study 1.

3.5. Results and discussion

Our first question was whether monkeys looked longer at unexpected outcomes versus expected outcomes, indicating sensitivity to likely future events using statistical reasoning. Looking at the entire sample of 80 individuals, monkeys looked an average of 4.19 ± 0.30s in the expected trial (where the fruit that emerged from the machine matched the original population), and 4.89 ± 0.30s in the unexpected trial (where the fruit that emerged from the machine was rare or absent in the original population). This was a significant difference (paired samples *t*-test: $t_{79} = -2.165$, $p = 0.03$, two-tailed; Cohen's $d = 0.24$; see Fig. 3a), indicating that monkeys looked longer at outcomes that were unexpected when they were statistically improbably or impossible, compared to those outcomes that were more likely.

We then examined what factors predicted the monkeys' looking responses by implementing linear mixed models. In a basic linear mixed model, we accounted for *subject* (as a random factor accounting for repeated trials within individuals), *age* (linear in years), sex (male or female), *machine* (lemon or apple machine) and *trial number* (1 or 2) as factors. This model indicated that the overall looking time decreased with age, and monkeys also generally tended to look longer at the first trial, as in Study 1. We then added *trial type* (expected or unexpected outcome) as a predictor in the second model. Aligning with our overall comparison, we found that this improved model fit compared with the base model [LRT: $\chi^2 = 5.178$, $p = 0.02$], as monkeys looked longer at the unexpected trial. In the third model, we then examined whether *likelihood* of the sample (improbable or impossible) influenced monkeys' looking time. However, this did not improve model fit [$\chi^2 = 0.173$, $p = 0.68$; see Table 2 for parameters from the full model]. Model fit was also not improved by including in the full model a *trial type X likelihood condition* interaction [LRT compared to the second model: $\chi^2 = 0.286$, $p = 0.86$]; monkeys exhibited similar responses to both impossible and improbable outcomes (see Fig. 3b). Indeed, post-hoc comparisons examining monkeys separately in each likelihood condition showed that both sub-samples of monkeys exhibited strong trends towards longer looking durations to the unexpected outcome, despite the reduced sample sizes in these comparisons. In the *improbable condition*, monkeys looked an average of 4.25 ± 0.43 s at the expected outcome and 5.05 ± 0.43 s to the unexpected outcome (paired samples t-test: $t_{39} = -1.64$, $p = 0.054$, one-sided). In the *impossible condition*, monkeys looked 4.13 ± 0.41 s on average at the expected outcome, and 4.72 ± 0.42 s on average at the unexpected outcome (paired sample t-test: $t_{39} = -1.39$, $p = 0.09$, one-sided).

These results are similar to what has been found in 12-month-old infants (Téglás et al., 2007), suggesting that rhesus monkeys, like infants, can make inferences about the most likely future outcome without direct experience of past sampling. In our study, we also directly manipulated the likelihood of the unexpected outcome within the same study, and found that monkeys showed similar increased looked responses to improbable events (when a rare fruit emerged) and impossible events (when an absent fruit emerged). However, an

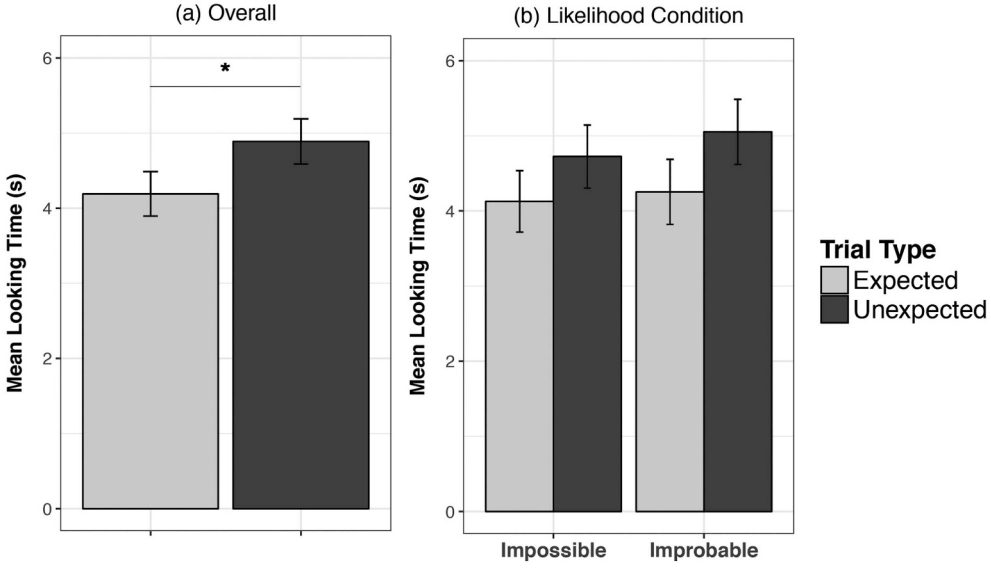

Fig. 3. Mean looking at outcomes in Study 2 (Probability Inference). Monkeys watched either the expected or the unexpected outcome fall from the lottery machine; the likelihood the unexpected event was improbable (3:1 distribution of two types of fruit) or impossible (only one type of fruit was present in the population). (a) Overall mean duration towards the unexpected and expected outcomes; error bars indicate standard error; * $p < 0.05$. (b) Mean duration of looking towards the different outcome types according to the likelihood.

Table 2
Parameters from the full model predicting monkey looking responses in the Probability Inference experiment (Study 2).

Predictor	Estimate	S.E.	t value	p - value
Age (linear in years)	−0.210	0.044	−4.780	< **0.0001**
Sex (reference: female)	−0.046	0.445	−0.103	0.92
Machine (reference: apples)	−0.374	0.449	−0.832	0.41
Trial number (reference: trial 1)	−0.962	0.306	−3.147	= **0.002**
Trial type (reference: expected)	0.698	0.306	2.284	= **0.02**
Likelihood (reference: impossible)	0.179	0.445	0.403	0.69

alternative explanation for these results is that the monkeys were using simple perceptual heuristics, rather than making inferences about likely events. In particular, in this setup monkeys could be detecting the perceptual difference between the (unexpected) sample and the majority contents of the original population, a cognitive shortcut sometimes termed *perceptual mismatch* (Xu & Garcia, 2008). For example, when presented with the lemon machine (mostly yellow lemons), it is possible that monkeys looked longer when a red apple dropped not because it was an unlikely outcome, but because it visually differed in the appearance from the original population. Study 3 addressed this possibility.

4. Study 3: non-causal control

In Study 3, we addressed the possibility that monkeys were responding to a perceptual mismatch in Study 2. Here, we presented a new set of monkeys with a population and sampled item that were not causally linked. Monkeys saw the experimenter rotate the lottery machine as in Study 2, but rather than releasing a fruit from the machine, she then visibly pulled the fruit out of her pocket (see Video S1). This fruit either matched the more common item in the machine (*expected trial*) or the rarer/absent item (*unexpected trial*). As in Study 2, we also counterbalanced the likelihood of these outcomes (*impossible* versus *improbable*) and the primary contents of the lottery machine within the same study design. This therefore provided a strong test of whether monkeys were only responding to the visual incongruency of the unexpected outcomes in Study 2, versus whether they were reasoning about the relationship between the sampled fruit and the original population.

4.1. Subjects

We tested 80 rhesus macaques (40 males and 40 females; ranging from 1.7 to 18.4 years old, M = 6.5 years), none of whom participated in either Study 1 or 2. Study 2 and 3 were conducted concurrently and monkeys were randomly assigned to one of the studies over the course of testing to ensure we did not accidentally sample different subsets of monkeys across the two studies. Within Study 3, monkeys were randomly assigned to witness the improbable or impossible likelihood condition, as well as the apple or lemon lottery machine. We approached additional monkeys for testing that were not included in the final sample because they approached the apparatus ($n = 4$), left the testing arena without completing both trials ($n = 6$), or were interfered with by another monkey ($n = 1$). Finally, two sessions were excluded during the video processing because subjects were not looking at the start of a trial.

4.2. Procedure and test conditions

In this study, we used the same basic setup as in Study 2 with one major difference: the sampled fruit did not fall out of the lottery machine. Rather, after rotating the cube three times (as in Study 2), E1 then visibly pulled the corresponding fruit sample (a lemon or an apple according to the trial; see Video S1) from her pocket and then placed it into the container above the cube (Fig. 4). E1 said “now” as she dropped the fruit, so monkeys' looking times were recorded for 10s starting from when the item was placed in the container as in the prior studies.

We placed the fruit above the cube in this condition to ensure that monkeys could clearly observe that there was no causal link between the lottery apparatus and the sampled item; we therefore ensured that the relative placement of the container above the cube was identical to the container below the cube. As in Study 2, each monkey experienced both the *expected* and *unexpected* trial in counterbalanced order, and the demonstrator was blind to the condition until the sampling event occurred (with trial type assigned by E2). We also again counterbalanced whether monkeys experienced the *probable likelihood condition* versus the *possible likelihood condition*, as well as and the contents of the lottery machine (*apple* vs. *lemon machine*) across subjects, using the same procedures described for Study 2.

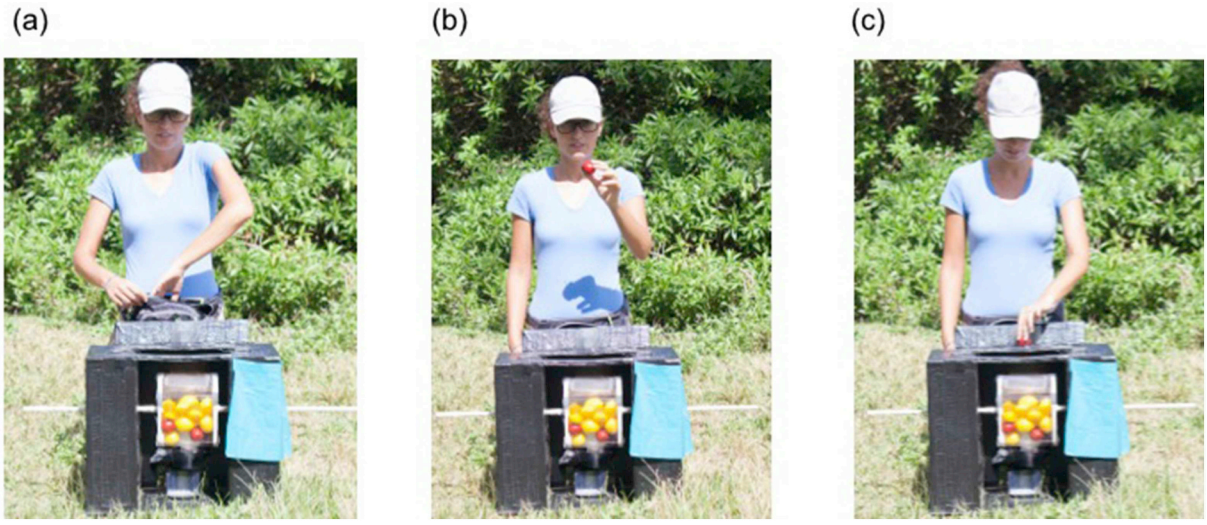

Fig. 4. Apparatus and procedure for Study 3 (Non-Causal Control). (a) After rotating the cube, the experimenter visibly removed the sampled item from her pocket; the sample either matched (expected) or did not match (unexpected) the population in the lottery machine. (b) The experimenter showed the sample to the monkey, and then (c) placed it into the container on the top of the box. Then, monkeys' looking response was recorded.

4.3. Data coding and analysis

Videos were processed and coded in the same way as Studies 1–2. Two independent coders coded the duration of the monkey's looking across trials and had a high inter-rater reliability (Pearson's $r = 0.96$). We used the same approach to analyze this data as in Study 2.

4.4. Results and discussion

Our first question was whether monkeys looked longer at unexpected (non-matching) outcomes versus expected (matching) outcomes, the key test of whether increased looking in Study 2 was driven by statistical reasoning or perceptual mismatch. Overall, monkeys looked an average of 4.82 ± 0.29 s to the expected trial, and 4.81 ± 0.27 s to the unexpected (non-matching) trial, which did not significantly differ (paired samples *t*-test: $t_{79} = 0.035$, $p = 0.97$, n.s.; see Fig. 5a). As in Study 2, we then implemented a linear mixed model accounting for *subject* (as a random factor), *age* (linear in years), sex (male or female), *machine* (lemon or apple machine) and *trial number* (1 or 2). To assess whether monkeys use perceptual heuristics to form expectations for the outcome of events, in the second model we then added *trial type* (expected or unexpected) as an additional predictor. In contrast with Study 2, we found that monkeys did not look longer at the unexpected trial than at the expected trial, as this predictor did not improve model fit compared with the base model [LRT: $\chi^2 = 0.0013$, $p = 0.97$]. In the third model, we further included the *likelihood* of the outcome (probable or possible), but this also did not improve model fit [$\chi^2 = 0.0088$, $p = 0.93$]. This result contrast with those of Study 2, where monkeys looked longer at the unexpected outcome (see Table 3 for parameters from the full model). Moreover, examining monkeys separately in each likelihood condition showed that neither subset showed any tendency to look longer at the unexpected versus the expected outcome, unlike in Study 2 (*improbable condition*: 5.01 ± 0.39 s in the expected outcome, 4.65 ± 0.40s to the unexpected outcome; paired samples t-test: $t_{39} = 0.79$, $p = 0.78$, one-sided; *impossible condition*: 4.63 ± 0.44 s to the expected outcome; 4.95 ± 0.37 s to the unexpected outcome; paired samples t-test: $t_{39} = -0.69$, $p = 0.25$, one-sided).

These results show that monkeys do not just look longer at mere perceptual mismatches in the absence of a causal link between the

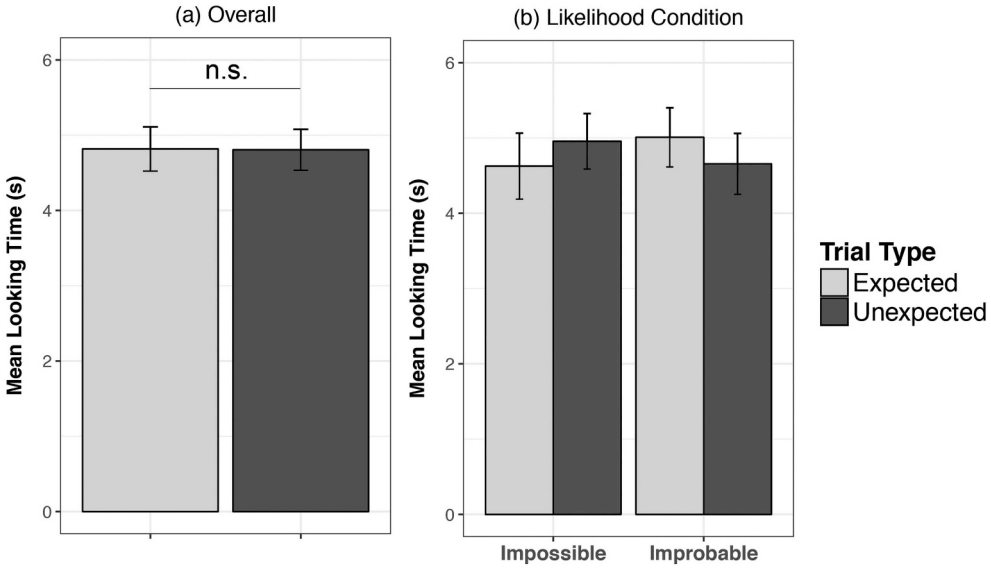

Fig. 5. Mean looking at outcomes in Study 3 (Non-causal control). Monkeys watched the experimenter pull either the expected or unexpected outcome from her pocket; the likelihood the unexpected outcome was either improbable (3:1 distribution of two types of fruits) or impossible (only one type of fruit was present in the population). (a) Overall mean duration towards the unexpected and expected outcomes; error bars indicate standard error. (b) Mean duration of looking towards the different outcome types according to the likelihood.

ARTICLE IN PRESS

Table 3
Parameters from the full model predicting monkey looking responses in Study 3 (Non-Causal control).

Predictor	Estimate	S.E.	t value	p - value
Age (covariate)	−0.225	0.049	−4.517	< **0.0001**
Sex (reference: female)	0.932	0.407	2.287	= **0.02**
Machine (reference: apple)	0.549	0.407	1.348	0.18
Trial number (reference: trial 1)	−0.599	0.321	−1.866	0.07
Trial type (reference: expected)	−0.011	0.321	−0.036	0.97
Likelihood (reference: impossible)	−0.037	0.408	−0.091	0.93

sample and the original population. When there was no relationship between the lottery machine and the sample because the experimenter pulled the sample from her pocket (rather than the sample falling from the lottery drum), monkeys looked similarly at both expected and unexpected trials. This result indicates that monkeys in Study 2 were unlikely to have been responding purely to the perceptual mismatch between unexpected and expected trials. Rather, they seem to have appreciated the causal link between the lottery machine and the sample in that situation, and used this link to guide their expectations about what would happen. Importantly, the setup in this study was identical to that in Study 2 except that the experimenter pulled the fruit outcome from her pocket and positioned it above the machine, rather than the item dropping out of the machine through a causal mechanism. That is, the degree of perceptual mismatch (or match) in both studies was identical: the only difference was that statistical knowledge could inform the monkeys' inferences in Study 2.

5. General discussion

Our findings show that rhesus monkeys can use their intuitions of probability to predict the likely outcome of future events in the absence of direct prior experience, like humans. We created a novel dynamic task to assess monkeys' statistical reasoning, where they observed a population of fruit in a rotating lottery machine and then saw a single item drop out. In Study 1, we first confirmed that monkeys look similarly at different kinds of fruits in the absence of any direct information about the population that they came from. In Study 2, the main test of whether monkeys can use probability information to predict future outcomes, we found that monkeys looked longer when they observed a rare or absent fruit dropping from a lottery machine, compared to when the common fruit type dropped out. Interestingly, macaques responded similarly to impossible events and improbable events, mirroring some prior work with human infants (Téglás et al., 2007, 2011). Finally, the results from Study 3 ruled out the possibility that monkeys were using simple perceptual heuristics to accomplish this inference. In particular, we found that when the experimenter pulled the sample from her pocket such that the causal relationship between the sample and the population was removed, monkeys looked at matching and non-matching outcomes at similar rates. Overall, this set of results indicates that monkeys can infer what is statistically likely or unlikely based on the distribution of a population. Importantly, our study utilized a violation-of-expectation looking time method similar to prior infant studies (Téglás et al., 2007, 2011), allowing us to examine their spontaneous expectations in a one-shot context in the absence of prior direct experience. This differs from past work with other primate species using experiential tasks.

We developed a novel procedure that could be used with free-ranging monkeys, so there were also some differences from prior work with infants. For example, Téglás et al. (2007) examined infants' expectations about improbable and impossible events in two separate experiments with subtly different setups, and used a barrier to establish that certain objects exiting the apparatus was physically impossible. In contrast, here monkeys were randomly assigned to the improbable versus impossible condition within the same study involving the same setup, and we manipulated likelihood by altering the contents of original population (e,g. entirely one kind of fruit in the impossible condition, and mostly one kind of fruit in the improbable condition). Both infants and monkeys showed similar increases in looking to impossible and improbable events, but these could be due nonetheless to distinct psychological processes. In the case of the monkeys in particular, our setup may have made it more difficult for monkeys to tell if there were truly no exemplars of the rare fruit in the impossible condition, given the number of items rotating inside the drum, so this might have resulted in them responding to both contexts as examples of improbable outcomes.

Another important issue concerns how magnitudes may shape probability expectations. We used a 3:1 ratio of items in the probable condition, as in the infant study we drew upon for our setup (Téglás et al., 2007), but had a total of 20 items rather than four inside the machine. The impact of the absolute size of the population, versus the ratio between the items, is currently unclear. For example, infants did not show increased looking to unexpected outcomes when tested with a lottery machine consisting of 16 objects with the same 3:1 ratio (Téglás, Ibanez-Lillo, Costa, & Bonatti, 2015), suggesting that probabilistic reasoning may be constrained by magnitude. Conversely, there is clear evidence that infants can reason about single-case probabilities even with larger set sizes up to 50 items (Denison & Xu, 2010a, 2010b, 2012). Several studies have further shown that nonhuman primates can successfully reason about probabilities when confronted with populations made of > 20 items (Rakoczy et al., 2014; Tecwyn et al., 2017; Eckert et al., 2017; Eckert, Call, et al., 2018a, 2018b, Eckert, Rakoczy, et al., 2018), and one recent comparison specifically showed that ratios rather than absolute magnitude drives chimpanzees' ability to make predictions about the most favorable outcome from a random draw (Eckert, Call, et al., 2018a, 2018b, Eckert, Rakoczy, et al., 2018). In our current study, macaques were in fact able to from expectations of probabilities with a population of 20 items, more similar to the prior findings with chimpanzees showing success in such a situation.

Our work provides also provides evidence that an intuitive sense of statistics is may be present across several different primate groups. While there was prior evidence for such abilities in several great ape species and capuchins, a species of New World monkey (Eckert et al., 2017; Eckert, Call, et al., 2018a, 2018b; Rakoczy et al., 2014; Tecwyn et al., 2017), prior evidence from Old World monkeys has been more mixed (Placì et al., 2018). One important question for future comparative work concerns whether these species exhibit statistical inferential abilities due to shared common descent, whether they independently evolved across these different taxa. As mentioned earlier, capuchins show evidence for striking behavioral and cognitive convergence with chimpanzees (Fragaszy et al., 2004; Rose, 1997; Visalberghi et al., 2015), including in terms of a preference for risky outcomes (De Petrillo et al., 2015; De Petrillo & Rosati, 2019; Heilbronner et al., 2008). A similar risk-prone tendency has also been found in rhesus macaques (Hayden & Platt, 2007; McCoy & Platt, 2005; Xu & Kralik, 2014), who diverged from the ape lineage approximately 30 mya (Perelman et al., 2011). This pattern of risk preferences in these three species contrasts with extensive work with many other species tested to date, who tend to exhibit some degree of risk aversion (Kacelnik & Bateson, 1996; Rosati, 2017). Thus, rhesus macaques, capuchins, and chimpanzees may have adopted common decision-making strategies that are not necessarily shared with other primates, and the emergence of intuitive statistics abilities in these taxa might be tied to these kinds of behavioral strategies. Thus, future work testing a wider range of species is crucial to understand the evolutionary history of this ability.

The current study also expands on prior comparative work by demonstrating that at least some non-human primates can not only reason using statistical information about probabilities, but also appreciate the relationship between a sample and a population in the absence of direct prior experience. In particular, monkeys in the current studies only

ARTICLE IN PRESS

experienced one trial each of the expected and unexpected outcomes, yet even with this minimal direct experience they could make rational inferences about future outcomes. As their intuitions seem to occur without experiencing the frequency of past outcomes, this suggests the possibility that their expectations may be derived from inductive logic, much like in infants (Cesana-Arlotti et al., 2018). This issue of whether other species can make inductive inferences has important implications for what mechanisms support learning in humans. In particular, recent proposals have argued that inductive inference mechanisms are based on Bayesian principles, such that infants make predictions about the likelihood of a set of hypotheses and then update their hypotheses when additional pieces of information are accumulated (Denison & Xu, 2012; Téglás et al., 2011). In this view, Bayesian updating allows infants and children to draw general conclusions from a small amount of information—and therefore rapidly acquire new knowledge across many different contexts, ranging from learning about the physical properties of objects to learning new words (e.g., Griffiths et al., 2010; Tenenbaum et al., 2006; Xu & Griffiths, 2011). The current study, together with prior work on non-human primates, suggests that non-human primates may also be able to generalize from a small amount of data (Eckert et al., 2017; Rakoczy et al., 2014; Tecwyn et al., 2017). If these sorts of inductive learning mechanisms are in fact shared across species, one question concerns what is special about the learning mechanisms underlying knowledge acquisition in humans, given the salient differences in human behavior and cognition compared to that of other species.

One possible difference between the statistical inference mechanisms seen in primates and those seen in humans is whether or how this intuitive sense of statistics is flexibly applied to other domains, such as decision-making and social cognition. Human infants are able to integrate other kinds of prior knowledge, such as their understanding of others' desires or how physical objects work, into their probabilistic inferences (Denison & Xu, 2012; Gweon, Tenenbaum, & Schulz, 2010; Téglás et al., 2011; Denison & Xu, 2014). Can other primates do the same? While rhesus monkeys can reason about some mental states in others (Drayton & Santos, 2017; Drayton & Santos, 2018; Flombaum & Santos, 2005; Martin & Santos, 2014; Santos, Nissen, & Ferrugia, 2006), it is unclear if they can harness their statistical reasoning capacities to make social predictions. Along the same lines, rhesus macaques and many other primate species are sensitive to probabilistic variation in rewards when making decisions about value (McCoy & Platt, 2005; Heilbronner et al., 2008; Rosati & Hare, 2012, 2013; De Petrillo et al., 2015), but it is unclear if these sorts of decisions under uncertainty rely upon their intuitive sense of statistics. In fact, some theories argue that the ability to combine rules and concepts across different domains is a novel feature of human cognition (Gopnik, 2012; Carey, 2004; Spelke & Kinzler, 2007; Shettleworth, 2012; Buckner & Krienen, 2013). Conversely, a recent study demonstrated that chimpanzees can account for an agent's preferences when making statistical inferences (Eckert, Call, et al., 2018a). This evidence suggests that the ability to integrate intuitive statistics with other cognitive domains might be shared at least with our closest relative, although it is currently unclear if this ability extends to other primates more generally.

In conclusion, our work demonstrates that rhesus macaques, a nonhuman primate species without language or pedagogy, can use an intuitive sense of probability to make predictions about a single future event in the absence of prior experience. Our findings from monkeys mirror prior results with human infants, and we modeled our dynamic lottery task on this infancy work to make the setup more comparable to human studies. Therefore, these results converge with accumulating evidence from cross-cultural research, developmental psychology, and comparative cognition supporting the possibility that this ability does not require language, and formal logical thinking and may in fact be more evolutionarily ancient. More broadly, we show that species other than humans can predict single future events based on information about probability, suggesting that one of the hallmarks of human logic is shared with other non-human primates.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.evolhumbehav.2019.05.006.

Acknowledgements

We thank Megan Cole, Hayoung Chang, Emily Walco, and Laura Lewis for assistance with data collection and coding, and Felix Warneken for comments on the manuscript. The authors are grateful to the Cayo Santiago Santiago Field Station and staff including Angelina Ruiz Lambides, Nahiri Rivera Barreto, Giselle Caraballo Cruz, and Bianca Giura for their research support. This work was supported by a National Center for Research Resources CM-5-P40RR003640-13 award to the Caribbean Primate Research Center and the University of Puerto Rico, and an Office of Research Infrastructure Programs (ORIP) of the National Institutes of Health (NIH) through Grant Number 5P40OD012217 to the Caribbean Primate Research Center and the University of Puerto Rico. FDP was supported by ANR Labex IAST and AR was supported by the Sloan Foundation.

Declaration of interest

None.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Bates, D. (2010). The LME4 package: Linear mixed-effects models using S4 classes. See http://www.r-project.org/.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution, 24*, 127–135. https://doi.org/10.1016/j.tree.2008.10.008.

Buckner, R. L., & Krienen, F. M. (2013). The evolution of distributed association networks in the human brain. *Trends in Cognitive Sciences, 17*, 648–665. https://doi.org/10.1016/j.tics.2013.09.017.

Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: "Gambler's fallacy" versus "hot hand". *Psychonomic Bulletin & Review, 11*(1), 179–184. https://doi.org/10.3758/BF03206480.

Carey, S. (2004). Bootstrapping & the origin of concepts. *Daedalus, 133*(1), 59–68. https://doi.org/10.1162/001152604772746701.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science, 359*(6381), 1263–1266. https://doi.org/10.1126/science.aao3539.

Cesana-Arlotti, N., Téglás, E., & Bonatti, L. L. (2012). The probable and the possible at 12 months: Intuitive reasoning about the uncertain future. *Advances in child development and behavior. Vol. 43. Advances in child development and behavior* (pp. 1–25). . https://doi.org/10.1016/B978-0-12-397919-3.00001-0.

De Petrillo, F., & Rosati, A. G. (2019). Ecological rationality: Convergent decision-making in apes and capuchins. In: *Behavioural Processes, 164*, 201–213. https://doi.org/10.1016/j.beproc.2019.05.010.

De Petrillo, F., Ventricelli, M., Ponsi, G., & Addessi, E. (2015). Do tufted capuchin monkeys play the odds? Flexible risk preferences in *Sapajus* spp. *Animal Cognition, 18*(1), 119–130. https://doi.org/10.1007/s10071-014-0783-7.

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5-and 6-month-olds. *Developmental Psychology, 49*(2), 243. https://doi.org/10.1037/a0028278.

Denison, S., & Xu, F. (2010a). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science, 34*, 885–908. https://doi.org/10.1111/j.1551-6709.2010.01111.x.

Denison, S., & Xu, F. (2010b). Twelve-to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*(5), 798–803. https://doi.org/10.1111/j.1467-7687.2009.00943.x.

Denison, S., & Xu, F. (2012). Probabilistic inference in human infants. *Advances in Child Development and Behavior, 43*, 27–58. https://doi.org/10.1016/B978-0-12-397919-3.00002-2.

Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition, 130*(3), 335–347. https://doi.org/10.1016/j.cognition.2013.12.001.

Drayton, L. A., & Santos, L. R. (2016). A decade of theory of mind research on Cayo Santiago: Insights into rhesus macaque social cognition. *American Journal of Primatology, 78*, 106–116. https://doi.org/10.1002/ajp.22362.

Drayton, L. A., & Santos, L. R. (2017). Do rhesus macaques, *Macaca mulatta*, understand what others know when gaze following? *Animal Behaviour, 134*, 193–199. https://doi.org/10.1016/j.anbehav.2017.10.016.

Drayton, L. A., & Santos, L. R. (2018). What do monkeys know about others' knowledge? *Cognition, 170*, 201–208. https://doi.org/10.1016/j.cognition.2017.10.004.

Drucker, C. B., Rossa, M. A., & Brannon, E. M. (2016). Comparison of discrete ratios by rhesus macaques (*Macaca mulatta*). *Animal Cognition, 19*, 75–89.

ARTICLE IN PRESS

Eckert, J., Call, J., Hermes, J., Herrmann, E., & Rakoczy, H. (2018a). b. Intuitive statistical inferences in chimpanzees and humans follow Weber's law. *Cognition, 180*, 99–107. https://doi.org/10.1016/j.cognition.2018.07.004.

Eckert, J., Call, J., Hermes, J., Herrmann, E., & Rakoczy, H. (2018b). c. Intuitive statistical inferences in chimpanzees and humans follow Weber's law. *Cognition, 180*, 99–107. https://doi.org/10.1016/j.cognition.2018.07.004.

Eckert, J., Rakoczy, H., & Call, J. (2017). Are great apes able to reason from multi-item samples to populations of food items? *American Journal of Primatology, 79*. https://doi.org/10.1002/ajp.22693.

Eckert, J., Rakoczy, H., Call, J., Herrmann, E., & Hanus, D. (2018a). Chimpanzees consider humans' psychological states when drawing statistical inferences. *Current Biology, 28*, 1959–1963. https://doi.org/10.1016/j.cub.2018.04.077.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology, 15*(5), 447–452. https://doi.org/10.1016/j.cub.2004.12.076.

Fontanari, L., Gonzalez, M., Vallortigara, G., & Girotto, V. (2014). Probabilistic cognition in two indigenous Mayan groups. *Proceedings of the National Academy of Sciences, 111*(48), 17075–17080. https://doi.org/10.1073/pnas.1410583111.

Fragaszy, D., Izar, P., Visalberghi, E., Ottoni, E. B., & de Oliveira, M. G. (2004). Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. *American Journal of Primatology, 64*(4), 359–366. https://doi.org/10.1002/ajp.20085.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science, 337*(6102), 1623–1627. https://doi.org/10.1126/science.1223416.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences, 14*, 357–364. https://doi.org/10.1016/j.tics.2010.05.004.

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107*(20), 9066–9071. https://doi.org/10.1073/pnas.1003095107.

Hanus, D., & Call, J. (2014). When maths trumps logic: Probabilistic judgements in chimpanzees. *Biology Letters, 10*(12), 20140892. https://doi.org/10.1098/rsbl.2014.0892.

Haun, D. B., Nawroth, C., & Call, J. (2011). Great apes' risk-taking strategies in a decision making task. *PLoS ONE, 6*(12), e28801. https://doi.org/10.1371/journal.pone.0028801.

Hayden, B. Y., & Platt, M. L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Current Biology, 17*(1), 49–53. https://doi.org/10.1016/j.cub.2006.10.055.

Hayden, B. Y., & Platt, M. L. (2009). Gambling for Gatorade: Risk-sensitive decision making for fluid rewards in humans. *Animal Cognition, 12*(1), 201–207.

Heilbronner, S. R., Rosati, A. G., Stevens, J. R., Hare, B., & Hauser, M. D. (2008). A fruit in the hand or two in the bush? Divergent risk preferences in chimpanzees and bonobos. *Biology Letters, 4*(3), 246–249. https://doi.org/10.1098/rsbl.2008.0081.

Hughes, K. D., & Santos, L. R. (2012). Rotational displacement skills in rhesus macaques (*Macaca mulatta*). *Journal of Comparative Psychology, 126*(4), 421. https://doi.org/10.1037/a0028757.

Kacelnik, A., & Bateson, M. (1996). Risky theories—The effects of variance on foraging decisions. *American Zoologist, 36*(4), 402–434. https://doi.org/10.1093/icb/36.4.402.

Kahneman, D., & Tversky, A. (2013). Choices, values, and frames. *Handbook of the fundamentals of financial decision making: Part I* (pp. 269–278). . https://doi.org/10.1142/9789814417358_0016.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmertest'. *R package version, 2*(0).

MacLean, E. L., Mandalaywala, T. M., & Brannon, E. M. (2012). Variance-sensitive choice in lemurs: Constancy trumps quantity. *Animal Cognition, 15*(1), 15–25. https://doi.org/10.1007/s10071-011-0425-2.

Marticorena, D., Ruiz, A. M., Mukerji, C., Goddu, A., & Santos, L. R. (2011). Monkeys represent others' knowledge but not their beliefs. *Developmental Science, 14*, 1406–1416. https://doi.org/10.1111/j.1467-7687.2011.01085.x.

Martin, A., & Santos, L. R. (2014). The origins of belief representation: Monkeys fail to automatically represent others' beliefs. *Cognition, 130*, 300–308. https://doi.org/10.1016/j.cognition.2013.11.016.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience, 8*(9), 1220–1227.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review, 109*(2), 330.

Perelman, P., Johnson, W. E., Roos, C., Seuánez, H. N., Horvath, J. E., Moreira, M. A., ... Schneider, M. P. C. (2011). A molecular phylogeny of living primates. *PLoS Genetics, 7*(3), e1001342. https://doi.org/10.1371/journal.pgen.1001342.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children (LJ Leake & PD Burrell & HD Fischbein, trans.)*.

Placì, S., Eckert, J., Rakoczy, H., & Fischer, J. (2018). Long-tailed macaques (*Macaca fascicularis*) can use simple heuristics but fail at drawing statistical inferences from populations to samples. *Open Science, 5*(9), 181025. https://doi.org/10.1098/rsos.181025.

R Development Core Team (2017). A language and environment for statistical computing. *Vienna, Austria*.

Rakoczy, H., Clüver, A., Saucke, L., Stoffregen, N., Gräbener, A., Migura, J., & Call, J. (2014). Apes are intuitive statisticians. *Cognition, 131*(1), 60–68. https://doi.org/10.1016/j.cognition.2013.12.011.

Rivière, J., Stomp, M., Augustin, E., Lemasson, A., & Blois-Heulin, C. (2018). Decision-making under risk of gain in young children and mangabey monkeys. *Developmental Psychobiology, 60*(2), 176–186. https://doi.org/10.1002/dev.21592.

Rosati, A. G. (2017). Decisions under uncertainty: preferences, biases, and choice. In J. Call (Vol. Ed.), *APA handbook of comparative psychology. Vol. 2. APA handbook of comparative psychology* (pp. 329–357). The American Psychological Association.

Rosati, A. G., Arre, A. M., Platt, M. L., & Santos, L. R. (2016). Rhesus monkeys show human-like changes in gaze following across the lifespan. *Proceedings of the Royal Society B, 283*, 20160376. https://doi.org/10.1098/rspb.2016.0376.

Rosati, A. G., & Hare, B. (2012). Decision making across social contexts: Competition increases preferences for risk in chimpanzees and bonobos. *Animal Behaviour, 84*(4), 869–879. https://doi.org/10.1016/j.anbehav.2012.07.010.

Rosati, A. G., & Hare, B. (2013). Chimpanzees and bonobos exhibit emotional responses to decision outcomes. *PLoS ONE, 8*(5), e63058. https://doi.org/10.1371/journal.pone.0063058.

Rosati, A. G., & Santos, L. R. (2016). Spontaneous metacognition in rhesus monkeys. *Psychological Science, 27*, 1181–1191. https://doi.org/10.1177/0956797616653737.

Rose, L. M. (1997). Vertebrate predation and food-sharing in *Cebus* and *Pan*. *International Journal of Primatology, 18*(5), 727–765.

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, *Macaca mulatta*, know what others can and cannot hear. *Animal Behaviour, 71*(5), 1175–1181. https://doi.org/10.1016/j.anbehav.2005.10.007.

Santos, L. R., & Rosati, A. G. (2015). The evolutionary roots of human decision making. *Annual Review of Psychology, 66*, 321–347. https://doi.org/10.1146/annurev-psych-010814-015310.

Shettleworth, S. J. (2012). Modularity, comparative cognition and human uniqueness. *Philosophical Transactions of the Royal Society B, 367*, 2794–2802. https://doi.org/10.1098/rstb.2012.0211.

Sim, Z. L., & Xu, F. (2019). Another look at looking time: Surprise as rational statistical inference. *Topics in Cognitive Science,* (1), 154–163. https://doi.org/10.1111/tops.12393.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x.

Stevens, J. R. (2010). Rational decision making in primates: the bounded and the ecological. *Primate Neuroethology,* 96–116.

Tecwyn, E. C., Denison, S., Messer, E. J., & Buchsbaum, D. (2017). Intuitive probabilistic inference in capuchin monkeys. *Animal Cognition, 20*(2), 243–256. https://doi.org/10.1007/s10071-016-1043-9.

Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences, 104*(48), 19156–19159. https://doi.org/10.1073/pnas.0700271104.

Téglás, E., Ibanez-Lillo, A., Costa, A., & Bonatti, L. L. (2015). Numerical representations and intuitions of probabilities at 12 months. *Developmental Science, 18*(2), 183–193. https://doi.org/10.1111/desc.12196.

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*(6033), 1054–1059. https://doi.org/10.1126/science.1196404.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*(7), 309–318. https://doi.org/10.1016/j.tics.2006.05.009.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458. https://doi.org/10.1126/science.7455683.

Visalberghi, E., Sirianni, G., Fragaszy, D., & Boesch, C. (2015). Percussive tool use by Taï Western chimpanzees and Fazenda Boa Vista bearded capuchin monkeys: A comparison. *Philosophical Transactios of the Royal Society B, 370*(1682), 20140351. https://doi.org/10.1098/rstb.2014.0351.

Xu, E. R., & Kralik, J. D. (2014). Risky business: Rhesus monkeys exhibit persistent preferences for risky options. *Frontiers in Psychology, 5*(258), https://doi.org/10.3389/fpsyg.2014.00258.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month old infants. *Cognition, 112*, 97–104. https://doi.org/10.1016/j.cognition.2009.04.006.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences, 105*(13), 5012–5015. https://doi.org/10.1073/pnas.0704450105.

Xu, F., & Griffiths, T. L. (2011). In F. Xu (Ed.). *Probabilistic models of cognitive development: Towards a rational constructivist approach to the study of learning and development* (pp. 2011). Elsevier.

Yamada, H., Tymula, A., Louie, K., & Glimcher, P. W. (2013). Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. *Proceedings of the National Academy of Sciences*201308718. https://doi.org/10.1073/pnas.1308718110.