CrossMark

ORIGINAL PAPER

Comparing species decisions in a dichotomous choice task: adjusting task parameters improves performance in monkeys

Laurent Prétôt[1] · Redouan Bshary[2] · Sarah F. Brosnan[1,3]

Abstract In comparative psychology, both similarities and differences among species are studied to better understand the evolution of their behavior. To do so, we first test species in tasks using similar procedures, but if differences are found, it is important to determine their underlying cause(s) (e.g., are they due to ecology, cognitive ability, an artifact of the study, and/or some other factor?). In our previous work, primates performed unexpectedly poorly on an apparently simple two-choice discrimination task based on the natural behavior of cleaner fish, while the fish did quite well. In this task, if the subjects first chose one of the options (ephemeral) they received both food items, but if they chose the other (permanent) option first, the ephemeral option disappeared. Here, we test several proposed explanations for primates' relatively poorer performance. In Study 1, we used a computerized paradigm that differed from the previous test by removing interaction with human experimenters, which may be distracting, and providing a more standardized testing environment. In Study 2, we adapted the computerized paradigm from Study 1 to be more relevant to primate ecology. Monkeys' overall performance in these adapted tasks matched the performance of the fish in the original study, showing that with the appropriate modifications they can solve the task. We discuss these results in light of comparative research, which requires balancing procedural similarity with considerations of how the details of the task or the context may influence how different species perceive and solve tasks differently.

Introduction

In comparative psychology, researchers compare species and use their similarities and differences to improve understanding the evolution and development of their cognition and behavior (Kamil 1988, 1998). One challenge with this is that when differences between species are found (e.g., one species outperforms another in a same task), it is often difficult to determine what were the underlying factors that caused it (Bitterman 1960, 1965). Was it a difference in cognitive architecture? Social environment? Ecology? Was it an artifact of the species' anatomy or physiology? Some aspect of the procedure? Or a combination of some or all of these? Therefore, a single comparison, whether the behavior is shown to be similar or different, is almost never sufficient for a full understanding of that behavior.

Recently, Salwiczek et al. (2012) conducted a cross-taxon study designed to compare the performance of cleaner fish with several non-human primate species (capuchin monkeys, chimpanzees, and orangutans) on a seemingly simple two-choice discrimination task derived from the cleaner's natural habitat (Bshary and Grutter 2002). The goal of this study was to compare two distantly related taxa that have both evolved complex social and

✉ Laurent Prétôt
  lpretot1@student.gsu.edu

[1] Department of Psychology, Language Research Center, Georgia State University, Atlanta, GA 30303-5010, USA

[2] Department of Behavioral Ecology, University of Neuchâtel, 2000 Neuchâtel, Switzerland

[3] Department of Philosophy and Neuroscience Institute, Georgia State University, Atlanta, GA, USA

Springer

foraging behaviors (especially those related to cooperation between individuals) to see whether the primates could solve it as well as the fish despite the lack of ecological relevance. If this had been the case, it would have indicated that the primates were able to use their relatively more developed cognitive ability to overcome their lack of experience. However, it was not; the adult fish outperformed all of the primate species (and juvenile cleaner fish). This was surprising and indicated that the task might not be as easy as it seemed. For example, the primates might have failed at the task because of the general tendency animals show to be strongly influenced by the first of two or more rewards that they encounter (e.g., Timberlake et al. 1987; Davison and Baum 2000; also see Salwiczek et al. 2012). In light of this literature, it is also surprising that the cleaner fish performed as well as they did on the task. Unfortunately, the study design did not allow us to further explore which factors made the task difficult for the primates. Thus, the goal of the current study was to test two potential explanations, first, that the primates would do better in a computerized task that minimized extraneous cues and the presence of the experimenter, and second, that they would do better in a task designed to more closely mimic their ecology.

The paradigm we used was based on a mutualistic interaction between the cleaner fish *Labroides dimidiatus* and its "client" species. Cleaner fish remove parasites and other material from client reef fish, which visit them at "cleaning stations." Clients have been categorized as either residents, which have small territories or home ranges that allow them to access only one cleaning station, or as choosy clients, which have larger home ranges that cover several cleaning stations. Cleaners typically compete against one another over access to choosy clients, providing faster and better service, while they each have exclusive access to their resident clients. Choosy clients are expected to use their choice options by visiting stations where the service is better. Field observations have found that choosy clients have priority of cleaning access over the residents (Bshary 2001). They also typically switch to another cleaner if ignored, but are more likely to return to the same cleaning station if they are inspected (Bshary and Schäffer 2002), indicating that the clients' choosiness is the cause of priority of access.

Bshary and Grutter (2002) simulated this interaction in the laboratory, replacing client fish with plates. In the experiment, one plate simulated the choosy client, while the other represented the resident. Fish could feed on the choosy plate only if they started to feed on it before they foraged on the resident plate, otherwise the choosy plate was withdrawn while the fish was eating from the resident plate, just as choosy clients leave if they are not inspected rapidly. The resident plate, however, always stayed in the testing area until the fish had stopped feeding on it, just as resident clients often queue for service while the cleaner fish is inspecting another client. Crucially, both plates offered the same amount and type of food and hence were equally attractive as food patches. Within just a few trials, cleaner fish as a group showed a significant tendency to visit the choosy plate first, supporting previous field observations that client choice drives cleaners' decisions to give visitors priority of access.

In their study, Salwiczek et al. (2012) tested two species of non-human primates, capuchin monkeys and chimpanzees, that converge with cleaner fish on both their tendency to cooperate with one another (e.g., capuchins: Brosnan 2010; chimpanzees: Boesch and Boesch 1989) and their propensity to eat both immobile and mobile food sources, which may roughly correlate with the stable resident and mobile choosy clients (e.g., plant materials vs. hunting for insects and smaller vertebrates; Goodall 1986; Fragaszy et al. 2004). Additionally, they tested orangutans, which eat fruit (Galdikas 1988) and, less frequently, insects or other mobile animal protein sources (Rijksen 1978), but who do not cooperate to the same degree in natural situations (but do in captivity: see Chalmeau et al. 1997; Dufour et al. 2009). All of these primates have in common, though, that they do not cooperate with their food patches and that food patches of the same type do not behave differently according to what they look like (e.g., their color). This is in stark contrast to cleaners, which cooperate with clients that differ in color, such that color can be used as a recognition cue and associated with the species' strategic options.

In Salwiczek et al. (2012), fish outperformed all three primate species, but notably, the majority of monkeys (but not apes) who learned the task subsequently performed well on a reversal test (i.e., when the role of each plate was suddenly reversed). This demonstrated that, once acquired, the primates were able to flexibly use the rules needed in the task. Although these findings imply that at least part of the difference in decision-making could be due to the difference in the ecological relevance of the task to the different species, we could not rule out other factors that might play a role. One way to address this is to test successive adaptations of the task that have been modified to change a variable of interest. In the current paper, we used this approach to test two potential explanations in the species of primates that performed best on the previous task, capuchin monkeys, on two adapted versions of the task to see whether they would learn the task more rapidly with these alternate procedures. In addition, we compared their performance to that of rhesus macaques, another non-human primate species that cooperates and is very successful at cognitive tasks but had not had previous experience with the task, to see the degree to which our results generalized across primates.

First, we hypothesized that the primates' lack of success could be due to extraneous cues that were part of the procedure itself that may have more heavily influenced primates than fish (e.g., the presence of the experimenter). Specifically, in Study 1, we used a computerized paradigm that replaced plates with virtual icons on a computer screen. Computerized testing is familiar to our primate subjects of both species and has the benefit that it minimizes extraneous cues that may influence the primates' performance, such as side biases (Masataka 1990; Anderson et al. 1996; Westergaard and Suomi 1997; de Waal et al. 2008) and the presence of visible foods (see Boysen and Berntson 1995; Boysen et al. 1996, 1999, 2001; Murray et al. 2005). While we could not determine which, if any, of these cues influenced their behavior, if any of these cues were the key challenge to solving the task, we predicted that (1) capuchin monkeys would reach learning criterion more quickly in this version than they did in the plate task and (2) rhesus macaques would perform similarly to the capuchins (because they did not participate in the previous study, there was no previous performance to compare them to).

Second, we hypothesized that adapting the procedure to be more ecologically relevant to the primates would improve their performance. As discussed above, cleaners' ephemeral food patches only leave if the cleaner does not choose them, whereas monkeys' ephemeral food patches do not wish to be consumed and may try to escape prior to being chosen. Therefore, in our second study, the ephemeral stimulus moved from the beginning of the trial, as if "escaping." We predicted that this change would allow primates to reach criterion more rapidly. Additionally, to rule out that any change was simply due to the increased attention to a moving icon, for a second version of the task, both stimuli were moving (vibrating) from the beginning of the task. Note that the primates could have been influenced by one or both of the factors in these studies, or those that we did not test, but this is a first step at beginning to understand the factors driving the monkeys' performance.

General methods

Subjects and housing

All experiments were approved by the Georgia State University Institutional Animal Care and Use Committee (IACUC; capuchins: A13022; macaques: A13021) and met the standards of the USA. Georgia State University is fully approved by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC).

Capuchins

We tested 11 captive born brown capuchin monkeys (*Cebus apella*; 6 males, average age 12 years, range 5–22 years, and 5 females, average age 11 years, range 4–16 years) from two stable social groups at the Language Research Center of Georgia State University, USA. We initially tested 10 of these subjects and added one who became old enough to participate in testing prior to Study 2. Subjects were separated from their social groups only for behavioral and cognitive testing. Subjects were fed a diet according to their species-specific needs, consisting of primate chow and fresh fruits and vegetables. They also received enrichment foods several times per day. Animals were never food or water deprived. Running water was available ad libitum, including during testing.

The monkeys lived in two large indoor/outdoor enclosures. Each enclosure contained ample three-dimensional climbing space as well as trapezes, perches, and enrichment items. The enclosure for each social group was divided into an indoor area (approximately one-half of their total space) and an outdoor area. The subjects had previously been trained to enter test boxes attached to their indoor area, which allowed us to separate individuals from their group for testing. Subjects could choose not to participate at any time by walking away from the experimenter. The monkeys were tested 3–5 days per week. No subject was ever involved in more than one session of this study on any given day (see detail in General Procedure, below).

Macaques

We tested eight captive born adult male rhesus macaques (*Macaca mulatta*; average age 18 years, range 9–29 years) at the Language Research Center of Georgia State University, USA. They had continuous access to water and worked for fruit-flavored primate pellets. They also received a daily diet of fruits and vegetables to supplement the food they received from testing and were never food or water deprived.

Apparatus

Subjects were tested using the Language Research Center's Computerized Test System comprising a personal computer, digital joystick, color monitor, and pellet dispenser (Richardson et al. 1990; Washburn and Rumbaugh 1992; Evans et al. 2008). Monkeys manipulated the joystick to produce isomorphic movements of a computer-graphic cursor on the screen. Contacting appropriate computer-generated stimuli with the cursor provided them a 45-mg (capuchins) or 94-mg (macaques) banana-flavored chow

pellet (Bio-Serv, Frenchtown, NJ) using a pellet dispenser that interfaced to the computer through a digital I/O board (PDISO8A; Keithley Instruments, Cleveland, OH). All monkeys had previously participated in multiple cognitive and behavioral experiments involving this computerized test system (e.g., Beran 2007, 2008; Beran and Parrish 2012; Beran et al. 2008a, 2012; Brosnan et al. 2012; Evans and Beran 2012). Monkeys were tested on the computerized apparatus while in their testing enclosures.

General procedure

The experimental design was based on the studies of Bshary and Grutter (2002) and Salwiczek et al. (2012). Two icons, one representing an ephemeral choice and one a permanent choice (see Table 1), were presented on a computer screen, and subjects had to choose between them. If the permanent stimulus was chosen first, the subject could not make another choice, but if the subject chose the ephemeral stimulus first, they could then choose the permanent one as well, resulting in one reward for each choice. The optimal outcome was always to select the ephemeral stimulus first, which ultimately gave the subject two rewards. The location of each stimulus was randomly determined, but in conditions in which only two locations were used (e.g., left/right), the stimulus position was counterbalanced within a session so that they were presented on each side an equal number of times, with the constraint that there were no more than three trials in a row on the same side (as in Salwiczek et al. 2012; see details in the procedures, below). Each trial ended once the monkeys made a choice between the options. Each subject received 10 sessions of 10 trials each, unless otherwise noted. The intertrial time interval varied depending on the task (see detail in the procedures, below).

Subsequent to the initial learning test, subjects completed a reversal test (Rumbaugh 1971, 1997; Salwiczek et al. 2012). The methodology was identical to that in the initial test, but the role of each stimulus was reversed (i.e., the stimulus that was ephemeral in the initial learning test became permanent and vice versa). These tests are frequently used to demonstrate flexibility in cognitive processing. If subjects initially preferred the ephemeral stimulus and then reversed to prefer the newly ephemeral (formerly permanent) stimulus, this indicated that they were consistent in their preference. The outcomes are more difficult to interpret if subjects initially preferred the permanent stimulus. Reversing to maintain a preference for the newly permanent (formerly ephemeral) stimulus indicated a consistent preference for the permanent stimulus; however, maintaining a preference for the newly ephemeral (formerly permanent) stimulus was uninformative because we cannot disentangle two possible explanations: Did they prefer that physical stimulus, or did they recognize the utility of choosing the ephemeral stimulus first when they chose the now-ephemeral stimulus to do so after the stimuli switched? We report data on reversal tasks in all cases, but do not attempt to interpret them in this latter situation. For each study, stimuli were counterbalanced such that each was the ephemeral stimulus for approximately half of the subjects.

Learning criterion and statistics

In Study 1, we initially used the same criterion as was used in Salwiczek et al. (2012). This criterion was based on a two-tailed signs test, such that significance was reached when a subject developed a preference for the ephemeral source in 10/10 or 9/10 trials on one session, in ≥8/10 on two successive sessions, or in ≥7/10 trials on three consecutive sessions; however, unlike primates in the previous study, some monkeys developed a preference for the permanent over the ephemeral stimulus, so for Study 2, developing a preference for *either* stimulus met criterion. Two additional changes to Study 2 were that (1) we still used the 10/10 criterion on one unique session, but only if the subject selected each of the two stimuli in at least one trial of a previous session to ensure that they had experience with both outcomes (note that this did not apply to the reversal phase, because a subject had experienced both options in the initial phase) and (2) we required two sessions of 9/10 because it came to our attention that, while a

Table 1 Computer tasks

Version	ITI (s)	Location	Study 1A	Study 1B	Study 2A	Study 2B
Version 1	300	Left/right	⬠ ✚	No version 1	♥ ⊙	⚑ ◗
Version 2	60	Left/right	⬠ ✚	⌂ �묘	No version 2	No version 2
Version 3	60	Random	⬠ ✚	⌂ �묘	No version 3	No version 3
Version 4	30	Random	⬠ ✚	⌂ ▮	No version 4	No version 4

The table describes the version, intertrial time interval (ITI), spatial location of the stimuli, and the stimuli used for each study and version

single 9/10 session is used frequently as a criterion in cognitive and behavioral testing, Monte Carlo simulations show that this criterion is not sufficiently strict (see Grant 1946).

We used a two-tailed exact Wilcoxon signed-ranks test to analyze differences in subjects' performance between the two learning phases and across the studies. We also used a two-tailed exact Mann–Whitney U test to compare performance between species. Subjects that failed were assigned 110 trials for the purposes of analysis, which was the minimum number of trials at which testing ceased. Sometimes subjects did receive more than this because if on their final session the subjects were close to meeting criterion (i.e., at least 7 choices for one option), they were given another set (or two, in the case of sessions with 7/10 correct) of 10 trials to give them the opportunity to do so. For all of our cross-study and cross-species comparisons, we used the data from subjects who either preferred the ephemeral icon or did not develop any preference, but did not use the data from the subjects who significantly preferred the permanent stimulus, because our goal was to determine what changes allowed the monkeys to meet the same performance criteria as the fish.

Due to our small sample sizes, we also provided a measure of the effect size (Pearson's correlation coefficient, r; see Rosenthal 1991). Cohen (1988)'s classification of effect size magnitude was used, whereby $r \leq 0.09$ mean that there is "no effect" of the treatment, $r = 0.10–0.29$ is considered a "small effect," $r = 0.30–0.49$ is a "moderate effect," and $r \geq 0.50$ is a "large effect."

Study 1

Procedure

Study 1 was carried out between July and December 2012 with the same eight capuchin subjects tested in the original plate study (Salwiczek et al. 2012) and two additional subjects who had no experience with the original paradigm, and between August and November 2012 with macaques. At the beginning of each trial, two stimuli that differed in shape and color appeared simultaneously on the computer screen (the orientation of the stimuli differed across the versions of the task; see Table 1). The monkeys used a cursor (a red dot on the screen) to select one of the stimuli.

Study 1A

Subjects were tested on four versions of the task that varied on intertrial time interval (ITI) and stimulus location until they met criterion (see Table 1 for a summary of procedures in the order in which the subjects experienced them). Once

they met criterion, they were not tested on subsequent versions. Additionally, once they met criterion, they were given a reversal test using the same version of the task on which they had just passed criterion. We used a green pentagon and a purple cross-stimulus in all the versions (see Table 1). Note that the ITI was the interval between stimulus presentations, and thus, the subject could have taken more time to actually make a choice. Rhesus monkeys sometimes failed to re-engage in the task after an ITI (this is why trial counts were not always multiples of 10), and thus, we added a tone concurrent with the presentation of a new stimulus to attract their attention in an attempt to ameliorate this problem. Capuchin monkeys typically completed their sessions and did not have a tone in any condition.

Version 1 had an ITI of 5 min and a left–right presentation of the stimuli (e.g., one stimulus was on each side of the screen); if they failed to meet criterion, subjects were tested in subsequent versions that were altered in ways that we hypothesized would improve performance. Version 2 had an ITI of only 1 min, because in the previous study a decreased ITI increased performance on the task (see Salwiczek et al. 2012). For this version, we used a maximum of 50 trials because the three capuchin subjects tested on that version developed a side bias, which did not occur in other conditions. To avoid the subjects developing side biases, Version 3 was the same as Version 2, except that the stimuli were randomized across four positions that were lined up on either a left–right or up–down axis. Version 4 replicated Version 3, except with a shorter ITI of 30 s.

The reversal test was identical, except that the role of each stimulus was reversed. As with the regular test, subjects who did not meet criterion on the reversal test were subsequently moved to the next condition set available until they either met criterion or completed all possible conditions. An issue with how counterbalancing was calculated in the computer program meant that for the first two sessions of the capuchins' Version 1, the same stimulus was presented six times on one side (rather than five). To be thorough, the two subjects who passed criterion in the first two sessions were given an additional session with correct counterbalancing. Also, due to a technical problem with the food dispenser dispensing incorrectly, one capuchin (Lily) was excluded from the analysis because we could not know what she was basing her decisions on. One capuchin (Logan) and one macaque (Chewie) inadvertently received two reversal sessions of Version 1 on 1 day.

Study 1B

Each capuchin subject who passed both the initial and reversal learning phases of Study 1A was subsequently tested in a generalization task in which both ephemeral and permanent stimuli were replaced with new ones to see how

quickly the subjects were able to apply the same rules to a new situation (Harlow 1949; Schrier 1984; Barros et al. 2002). For this task, subjects were tested in the same version of Study 1A on which they had met criterion, with the exception that we did not use Version 1, with the ITI of 300 s, to maintain subjects' motivation to participate. Similarly to Study 1A, subjects who did not succeed in the first version on which they were tested progressed to the next version in the order indicated in Table 1. To be consistent with Study 1A, we used a maximum of 50 trials in Version 2 of Study 1B. We used a blue circle-top rectangular and a brown L-shape stimulus in all versions (see Table 1). Note that one subject (Nkima) always chose the permanent icon after 60 trials on the reversal phase of Version 3, at which point we decided to terminate the study. Rhesus monkeys were not run on Study 1B.

Results

Individual data for all subjects are in Tables 2 (capuchins) and 3 (macaques). Subjects' performance across the conditions is shown in Figs. 1 (initial phase) and 2 (reversal phase).

Capuchin Study 1A

Initial learning phase Six out of the nine subjects tested in the initial learning phase of Version 1 reached criterion within 100 trials (or 10 sessions; range 20–100 trials, $X = 45$ trials, SD ± 32.71). The three subjects who did not meet criterion in Version 1 also did not in Version 2, but all did so in Version 3 (range 30–96 trials, $X = 62$ trials, SD ± 33.05).

Reversal learning phase Two of the six subjects who had passed the initial learning phase of Version 1 reached criterion on the reversal phase within 110 trials (range 60–110 trials, $X = 85$ trials, SD ± 35.36). The other four subjects met criterion in Version 2 (range 10–40 trials, $X = 25$ trials, SD ± 12.91). One of the three subjects who reached criterion on the initial learning phase of Version 3 reached criterion on the Version 3 of the reversal phase (110 trials). Of the other two subjects, one met criterion in Version 4 (30 trials) and one never reached criterion.

Initial versus reversal learning phases Subjects who met criterion in Version 1, and therefore were the quickest in the initial learning phase, performed better on the initial learning phase than on the reversal phase (Wilcoxon signed-ranks test: $W = 0$, $Z = -2.20$, $N = 6$, $P = 0.03$, $r = -0.64$). This was different than in the original plate task, where capuchins met criterion faster on the reversal phase than the initial phase (see Salwiczek et al. 2012).

Considering all subjects who passed in all versions, however, subjects showed no significant difference in speed of learning between the two phases (Wilcoxon signed-ranks test: $W = 11$, $Z = -0.98$, $N = 8$, $P = 0.33$, $r = -0.25$).

Computer versus plate tasks (within-subjects) We compared the performance of the eight capuchin subjects who participated in the initial learning phase of the original plate task (Salwiczek et al. 2012) with their own data on the computerized task (note that there is an order confound as all plate testing was completed prior to computerized testing, although 3 years separated the two experiments). Subjects performed only marginally better on the computerized learning than the previous plate task, although the effect size is at the high end of medium, indicating that this may be a real effect (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.84$, $N = 7$, $P = 0.13$, $r = -0.49$; Fig. 1). Interestingly, the monkeys' performance on the reversal phase of the computerized task was marginally worse than the reversal phase of the previous plate task (Wilcoxon signed-ranks test: $W = 2.50$, $Z = -1.70$, $N = 7$, $P = 0.13$, $r = -0.45$; Fig. 2).

Macaque Study 1A

Initial learning phase Four out of the eight subjects tested in the initial learning phase of Version 1 reached criterion, preferring the ephemeral icon, within 50 trials (range 10–50 trials, $X = 36.5$, SD ± 18.05), while the other half reached criterion with a preference for the permanent icon, within 40 trials (range 20–40 trials, $X = 27.25$, SD ± 9.50). Taken as a group, macaques did not show any significant preference for either icon (binomial/sign test: $P = 1.00$).

Reversal learning phase The four subjects who chose the ephemeral icon on the learning phase of Version 1 reached criterion on the reversal phase within 88 trials (range 40–88 trials, $X = 65.75$, SD ± 20.21). The four subjects who initially chose the permanent icon reached a significant preference for their previously preferred icon, which now signaled the ephemeral reward, within 90 trials (range 10–90 trials, $X = 37.5$, SD ± 35.94). In those cases, however, we cannot say whether the subjects learned that the ephemeral option ultimately provided additional rewards or that they simply preferred that particular icon.

Initial versus reversal learning phases Too few subjects initially preferred the ephemeral task for statistical significance; however, we note that all four performed better on the initial phase than on the reversal phase (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.83$, $P = 0.13$, $r = -0.65$).

Table 2 Capuchin Study 1

Subject	Capuchin Study 1A								Capuchin Study 1B					
	Version 1		Version 2		Version 3		Version 4		Version 2		Version 3		Version 4	
	Initial	Reversal	Initial	Reversal	Initial	Reversal	Initial	Reversal	Initial	Reversal	Initial	Reversal	Initial	Reversal
Nala	20	60							30	56[d]				
Logan	30	110[b]							30	30				
Liam	20	>110		10					10	30				
Nkima[a]	30	>120		20					>50[c]		30	>60[e]		
Gambit[a]	100	>100		30					30	50				
Wren	70	>100		40					10	100				
Griffin	>120		>50[c]		30	110					30	100		
Drella	>130		>50[c]		60	>120		30						
Gabe	>120		>51[c,d]		96[d]	>43[d]		>132[d]					80	80

Number of trials for each subject in each version of Study 1A (initial learning) and Study 1B (generalization). For each version, the left column indicates the number of trials for the initial learning phase, while the right column indicates the number of trials for the reversal phase. Note that we did not use Version 1 in Study 1B

[a] New subject (not tested in Salwiczek et al. 2012)

[b] Subject received two sessions in 1 day

[c] Subject developed a side bias

[d] Subject did not complete the 10 trials in one or more sessions

[e] Subject always chose the permanent icon

Ephemeral versus permanent choice There was no significant difference in learning speed between subjects who preferred the ephemeral icon and those who preferred the permanent one (Mann–Whitney U test: $U = 4$, $Z = -1.16$, $N_{ephemeral} = 4$, $N_{permanent} = 4$, $P = 0.29$, $r = -0.41$; see Table 3).

Table 3 Macaque Study 1A

Subject	Initial	Reversal 1	Reversal 2
Macaque Study 1A			
Chewie	10	40[c]	
Han	41[a]	62[a]	
Murphy	45[a]	88[a]	
Obi	50	73[a]	
Luke	20[b]	30	60
Gale	20[b]	20	133[a,d]
Lou	29[a,b]	90	80
Hank	40[b]	10	86[a]

The left column indicates the number of trials for the initial learning phase, the middle column indicates the number of trials for the reversal phase, and the right column indicates the number of trials for the second reversal phase

[a] Subject did not complete the 10 trials in one or more sessions

[b] Subject developed an initial preference for the permanent stimulus

[c] Subject received two sessions in 1 day

[d] Subject was tested using a reduced ITI (60–30 s)

Capuchins versus fish

Monkeys' performance in Study 1A did not significantly differ from fishes' performance in the plate task (Mann–Whitney U test: $U = 24.50$, $Z = -0.30$, $N_{capuchins} = 9$, $N_{fish} = 6$, $P = 0.81$, $r = -0.08$).

Capuchins versus macaques

Unlike the capuchins, the macaques were equally likely to prefer the ephemeral or the permanent icon. When only comparing the subjects who initially chose the ephemeral icon, however, the two species did not significantly differ in performance in the learning phase (Mann–Whitney U test: $U = 12$, $Z = -0.93$, $N_{capuchins} = 9$, $N_{macaques} = 4$, $P = 0.39$; $r = -0.26$), but the macaques reached criterion in the reversal phase faster than did capuchins (Mann–Whitney U test: $U = 3$, $Z = -2.22$, $N_{capuchins} = 8$, $N_{macaques} = 4$, $P = 0.02$, $r = -0.64$).

Capuchin Study 1B

Initial learning phase We tested the subjects who passed both the initial and the reversal learning phases in any version of Study 1A on a generalization task designed to determine whether they would extrapolate to a novel situation. Five out of six subjects who met criterion on Version 1 of Study 1A met criterion within 30 trials in Version 2 of Study 1B (we did not use Version 1, so subjects started on
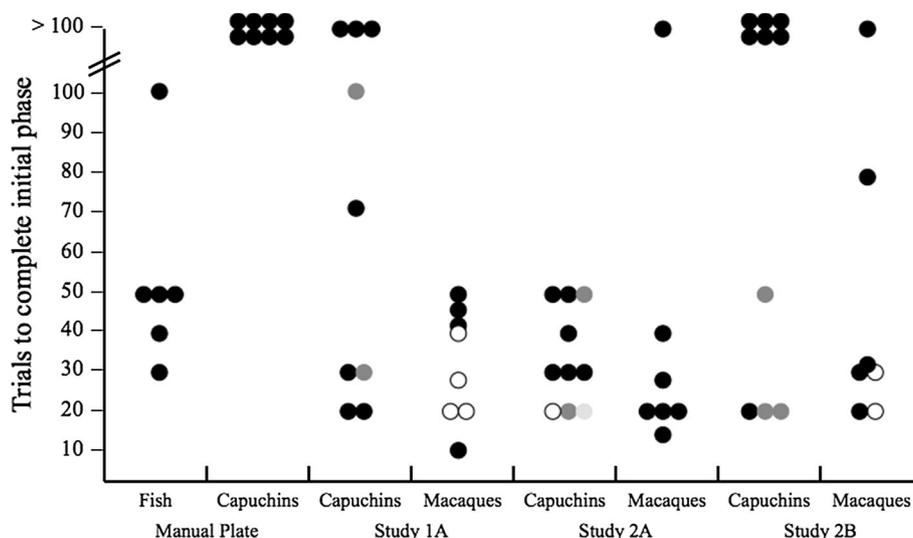
Fig. 1 Initial learning phase of the plate and computer tasks. The number of trials required for individuals to select either the ephemeral or the permanent stimulus on the original manual plate task (data from Salwiczek et al. 2012), computer Study 1A and Study 2, as compared to the fish. The *black dots* designate subjects that developed a preference for the ephemeral stimulus, whereas the *white dots* designate those who preferred the permanent stimulus. The *gray dots* designate new capuchin subjects who were not tested in the original plate task (Salwiczek et al. 2012): the *dark gray dots* are for those who developed an initial preference for the ephemeral icon, whereas the *light gray dot* is for the subject who developed an initial preference for the permanent icon. Note that one capuchin subject was dropped from Study 1A, and one capuchin and one macaque subjects were not tested subsequent to Study 1A

Fig. 2 Reversal learning phase of the plate and computer tasks. The number of trials required for individuals to reverse their preference on the original manual plate task (data from Salwiczek et al. 2012), computer Study 1A and Study 2, as compared to the fish. We only included data from those subjects who preferred the ephemeral stimulus on the initial learning phase (see Fig. 1). The *black dots* designate subjects who were tested in all possible conditions (capuchins: manual + computer tasks; macaques: computer tasks), whereas the *gray dots* designate new capuchin monkeys who were not tested in the original plate task (Salwiczek et al. 2012)

Version 2; range 10–30 trials, $X = 22$ trials, SD $\pm$ 10.95). The only subject who did not reach criterion in Version 2 did so in Version 3 (30 trials). The two subjects who did not reach criterion on the learning phase until Versions 3 and 4 met criterion on this generalization task in 30 trials (subject who met criterion on Version 3) and 80 trials (subject who met criterion on Version 4), respectively.

Reversal learning phase The five subjects tested in Version 2 met criterion within 100 trials on the reversal phase (range 30–100 trials, $X = 53.20$ trials, SD $\pm$ 28.66). Among the two subjects who started in Version 3, one reached criterion in 100 trials, while the other did not meet criterion within 60 trials (this subject always chose the permanent icon, at which point we terminated the study). Finally, the subject who started Version 4 reached criterion in 80 trials.

Study 1A versus Study 1B (within-subjects) As a group, subjects performed only marginally better on the initial learning phase of Study 1B than Study 1A (Wilcoxon signed-ranks test: $W = 4.50$, $Z = -1.61$, $N = 8$, $P = 0.13$, $r = -0.40$).

Study 1 discussion

Overall, capuchins were somewhat more likely to meet criterion on the computerized task than on the original plate task. Moreover, their learning speed on the computerized task did not differ from the fishes' performance on

the plate task. Finally, seven of eight monkeys met criterion on their first version of the generalization task, indicating that once they understood the procedure, they could generalize to novel stimuli. Thus, based on this evidence, it seems that the capuchin monkeys were better able to learn the same task when it was presented on a computer as compared to the manual procedure. This was presumably due to the fact that there were fewer extraneous cues that may have influenced the subjects' performance, either by distracting them from the important cues or by providing additional (but unnecessary) information that needed to be considered. While we cannot identify which cue(s) was relevant, our adapted task changed primates' perception of the task so that they were now able to learn it at equal speed to the fish.

These results are very important as they again emphasize the importance of how a task is presented in the subjects' ability to learn it. On the one hand, these results negate those from Salwiczek et al. (2012), as the monkeys were able to learn the task, but on the other hand, they emphasize that the monkeys could not learn using the same task parameters as the fish, but required alternate ones. This complicates comparative research as it requires studies *both* to test the different species on identical parameters [for a direct comparison, as we did in Salwiczek et al. (2012)] and *also* to test species using parameters that are adapted to that species' preferences and abilities (as we did here) to rule out the possibility that the difference is not in the task itself, but in how the task is presented. Of course, variations that allow a species to learn to the best of their

ability could also derive from differences in cognitive architecture, ecology, or how they interact with the task, and in most cases will likely differ on more than one dimension. Thus, it is important to test several of these possibilities in an attempt to pin down which feature(s) led to the species difference, which is what we do in the current paper.

One potential confound with our result is that this was many of the monkeys' second time to do this basic procedure. However, we consider it unlikely that their performance was strongly influenced by their previous experience for several reasons. First, approximately 3 years had elapsed between the subjects participating in the plate task and the first computer task (Study 1A), during which time they had participated in many other cognitive and behavioral tasks unrelated to this one. Moreover, two of the subjects in the computer task did not participate in the previous plate task, but performed at the same level as the more experienced monkeys, showing that experience was not necessary for this level of performance. The reversal learning results also provide evidence that subjects' success was not entirely based on previous experience; if experience was substantially influencing their performance, we would expect subjects to show increased speed of reversal learning across multiple presentations of the same problem (rhesus: Harlow 1949; Harlow and Warren 1952; capuchins: Gossette and Inman 1966), which they did not.

In an attempt to address this confound, we also tested rhesus monkeys, who had not previously been tested. All macaques reached preference criterion; however, while the capuchins were more likely to prefer the payoff maximizing choice, the macaques were equally likely to prefer to maximize payoffs or to avoid what likely appeared to be an unreliable stimulus (e.g., the ephemeral was often not available for them to choose). This was a surprising finding, given these rhesus macaques' higher performance in cognitive tasks as compared to capuchin monkeys (e.g., Beran et al. 2009, 2014; Beran and Smith 2011). It is difficult to know exactly what subjects based their decision on, but we discuss this further in the General Discussion, including data from both Studies 1 and 2.

Study 2

Study 2, which was completed subsequent to Study 1, was designed to be more ecologically relevant for primates. We predicted that this would increase the speed at which the monkeys met criterion above and beyond their performance on the computerized task. This study consisted of two experiments, A and B. In both tests, the ephemeral stimulus was moving from the beginning of the trial, which more closely mimics primates' natural ephemeral food sources (e.g., insects; see Introduction for more discussion of this). Study 2A had a single change from Study 1: From the beginning of the trial, the ephemeral icon moved toward the edge of the screen (where it escaped), while the permanent icon remained stationary. In Study 2B, we added another change; in this case, both icons vibrated on the spot before the ephemeral icon (still vibrating) moved away, while the permanent icon kept vibrating in the same location. This was done to control for the increased visual salience of a moving icon (e.g., attention effects).

We predicted subjects would do better in Study 2 as compared to Study 1. If performance increased only in Study 2A, this would indicate that it was due to an attention/salience effect, whereas if it increased in both Study 2A and 2B, this would more likely indicate that the change was due to the structure of the task, implicating ecological salience.

Procedure

Study 2A was carried out between September and October 2014 with capuchins and between February and March 2015 with macaques, while Study 2B was carried out between November and December 2015 with both monkey species. We tested eight of the nine capuchins used in Study 1, plus an additional subject with no previous experience with any version of the paradigm, and seven of the eight macaques used in Study 1. In order to compare the performance of monkeys in both Study 1 and Study 2, subjects were again tested in Version 1 of Study 1, using novel stimuli, except that when the subject touched the start button to begin the trial, both stimuli appeared and the ephemeral stimulus began moving from its initial position toward a lower corner of the screen. In Study 2B, both icons vibrated on the spot for 200 ms before the ephemeral icon (still vibrating) began moving. The subjects had to track and touch the ephemeral stimulus first, before it moved off the screen and could not be "caught" (the cursor moved only slightly faster than the ephemeral stimulus). We used a yellow heart and a purple-pink ring stimuli for Study 2A, and a blue waive and an orange moon for Study 2B (see Table 1). Note that two macaques (Obi and Luke) received two reversal sessions on 1 day (respectively in Study 2A and Study 2B). Also, one macaque (Luke) received an additional session because the data of the second session were lost.

Results

Individual data for all subjects tested on Study 2A and Study 2B are in Table 4. Subjects' performance across the conditions is shown in Figs. 1 and 2.

Table 4 Study 2

Subject	Study 2A		Study 2B	
	Initial	Reversal	Initial	Reversal
Capuchin Study 2				
Nala	50	60	>100	/
Logan	30	50	20	40
Liam	20[b]	50	>100	/
Nkima	20	30	20	80
Gambit	20[b]	20	50	>100
Wren	30	>100	>100	/
Griffin	30	>100	>100	/
Gabe	50	>100	>110	/
Lily	40	30	>100	/
Widget[a]	50	50	20	60
Macaque Study 2				
Chewie	20	>109[c]	30	>100
Han	40	50	30[b]	60
Murphy	14[c]	33[c]	32[c]	>115[c]
Obi	28[c]	55[d]	80	>100
Luke	20	76[c]	20[e]	>100[d]
Lou	20	10	20[b]	20
Hank	107[c]	20	>111[c]	/

For each version of the study (A and B), the left column indicates the number of trials for the initial learning phase, while the right column indicates the number of trials for the reversal phase

[a] New subject (not tested in Study 1)

[b] Subject developed an initial preference for the permanent stimulus

[c] Subject did not complete the 10 trials in one or more sessions

[d] Subject received two sessions in 1 day

[e] Subject received an additional session

Capuchin Study 2A

Initial learning phase All subjects met criterion in fewer than 50 trials; however, while eight out of the ten subjects met criterion with a preference for selecting the ephemeral icon first (range 20–50 trials, $X = 37.50$ trials, SD $\pm$ 11.65), two preferred the permanent icon (20 trials). Taken as a group, capuchins did not show any significant preference for either icon (binomial/sign test: $P = 0.11$).

Reversal learning phase Five out of the eight subjects who preferred the ephemeral icon in the initial phase reversed their preference in fewer than 60 trials (range 30–60 trials, $X = 44$ trials, SD $\pm$ 13.42), while the other three subjects did not meet criterion within 100 trials. The two subjects who initially preferred the permanent icon reversed their preference in 20 and 50 trials, although it is again difficult to interpret these data.

Initial versus reversal learning phase As in Study 1, there was a trend for capuchins to learn the initial learning phase in fewer trials than the reversal phase (Wilcoxon signed-ranks test: $W = 2$, $Z = -2.05$, $N = 8$, $P = 0.06$, $r = -0.51$).

Within-subjects comparison with previous studies There was no significant difference between Study 1A and Study 2A in monkeys' learning performance in either the initial (Wilcoxon signed-ranks test: $W = 2$, $Z = -1.48$, $N = 6$, $P = 0.19$, $r = -0.43$) or the reversal phase (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.34$, $N = 6$, $P = 0.50$; $r = -0.39$; Figs. 1, 2). Unlike in Study 1A, however, subjects were faster to learn the initial phase of Study 2A than in the previous plate task (Wilcoxon signed-ranks test: $W = 0$, $Z = -2.23$, $N = 6$, $P = 0.03$, $r = -0.64$).

Capuchin Study 2B

Initial learning phase Four out of the ten subjects met criterion in fewer than 50 trials, with a preference for selecting the ephemeral icon first (range 20–50 trials, $X = 27.50$ trials, SD $\pm$ 15). Five subjects never met criterion within 100 trials, and the last one never did so within 110 trials. Despite the fact that all monkeys who developed a preference did so for the ephemeral icon, taken as a group, capuchins did not show a significant preference for either icon (binomial/sign test: $P = 0.75$).

Reversal learning phase Three out of the four subjects who preferred the ephemeral icon in the initial phase reversed their preference within 80 trials (range 40–80 trials, $X = 60$ trials, SD $\pm$ 20), while the last subject did not meet criterion within 100 trials.

Initial versus reversal learning phase Too few subjects initially preferred the ephemeral task for statistical significance; however, we note that all four subjects performed better on the initial phase than on the reversal phase (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.84$, $P = 0.13$, $r = -0.65$).

Within-subjects comparison with previous studies Although nonsignificant, subjects showed a trend toward being faster on the initial phase of Study 2A than Study 2B (Wilcoxon signed-ranks test: $W = 3$, $Z = -1.87$, $N = 8$, $P = 0.08$, $r = -0.47$; Figs. 1, 2). Like in Study 2A, subjects did not differ in performance between Study 2B and Study 1A (Wilcoxon signed-ranks test: $W = 7$, $Z = -0.74$, $N = 8$, $P = 0.59$, $r = -0.19$); however, unlike in Study 2A, they did not differ in performance between Study 2B and the previous plate task (Wilcoxon

signed-ranks test: $W = 1$, $Z = -0.45$, $N = 7$, $P = 1.00$, $r = -0.12$).

Macaque Study 2A

Initial learning phase All seven subjects met criterion within 107 trials, with a preference for selecting the ephemeral icon first (range 14–107 trials, $X = 35.57$ trials, SD ± 32.59). Unlike in Study 1A, macaques showed a preference for the ephemeral icon over the permanent one (binomial/sign test: $P = 0.02$).

Reversal learning phase All but one subject reversed their preference in fewer than 76 trials (range 10–76 trials, $X = 40.67$ trials, SD ± 24.36). The last subject did not reach criterion within 109 trials.

Initial versus reversal learning phase Subjects' learning speed did not differ between the initial and reversal phases (Wilcoxon signed-ranks test: $W = 7.50$, $Z = -1.10$, $N = 7$, $P = 0.31$; $r = -0.29$).

Study 1A versus Study 2A (within-subjects) Although all seven subjects performed better in Study 2A than in Study 1A, there were too few subjects who initially preferred the ephemeral icon in both studies for statistical significance; we note, however, that all three subjects who had an initial preference for the permanent icon in Study 1A preferred the ephemeral icon in Study 2A, and three of the four subjects who developed a preference for the ephemeral icon in both tasks learned the initial task more rapidly in Study 2A than in Study 1A (Wilcoxon signed-ranks test: $W = 2$, $Z = -1.10$, $N = 4$, $P = 0.38$; $r = -0.39$; Figs. 1, 2). Three of these four subjects also learned the reversal phase of Study 2A more rapidly than in Study 1A (Wilcoxon signed-ranks test: $W = 4$, $Z = -0.37$, $N = 4$, $P = 0.88$; $r = -0.13$).

Macaque Study 2B

Initial learning phase Six of the seven subjects met criterion within 80 trials; however, while four out of the six subjects met criterion with a preference for selecting the ephemeral icon first (range 20–80 trials, $X = 40.50$ trials, SD ± 26.85), two preferred the permanent icon (20 and 30 trials). The last subject never reached criterion within 111 trials. Unlike in Study 2A, subjects did not show any significant preference for either icon (binomial/sign test: $P = 1.00$).

Reversal learning phase None of the four subjects who preferred the ephemeral icon in the initial phase reversed their preference within 100 trials. The two subjects who

initially preferred the permanent icon reached a significant preference for their previously preferred icon, which now signaled the ephemeral reward, in 20 and 60 trials (although again, we cannot interpret this result).

Initial versus reversal learning phase Too few subjects initially preferred the ephemeral task for statistical significance; however, we note that all four subjects learned the initial task more rapidly than the reversal (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.83$, $P = 0.13$; $r = -0.65$).

Within-subjects comparison with previous studies Too few subjects initially preferred the ephemeral icon in the studies for statistical significance; however, we note that three of the four subjects who learned the initial task in both studies learned it more rapidly in Study 2A than in Study 2B, while the last one performed equally in both studies (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.60$, $P = 0.25$; $r = -0.57$; Figs. 1, 2). Three of the four subjects who performed the reversal task in both studies learned it more rapidly in Study 2A than in Study 2B, while the last one failed to learn it in Study 2B (Wilcoxon signed-ranks test: $W = 1$, $Z = -1.46$, $P = 0.25$; $r = -0.52$). Two of the three subjects who learned the initial task in both studies learned it more rapidly in Study 1A than in Study 2B, while the last one showed the opposite (Wilcoxon signed-ranks test: $W = 1$, $Z = -1.07$, $P = 0.50$; $r = -0.44$). The three subjects learned the reversal task more rapidly in Study 1A than in Study 2B (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.60$, $P = 0.25$; $r = -0.65$).

Capuchins versus fish

As in Study 1A, there was no significant difference in learning performance between fish in the plate task and capuchins in either Study 2A (Mann–Whitney U test: $U = 13.50$, $Z = -1.43$, $N_{capuchins} = 8$, $N_{fish} = 6$, $P = 0.18$, $r = -0.38$) or Study 2B (Mann–Whitney U test: $U = 20.50$, $Z = -1.07$, $N_{capuchins} = 10$, $N_{fish} = 6$, $P = 0.28$, $r = -0.27$).

Capuchins versus macaques

As in Study 1A, there was no significant difference in performance between monkey species in either Study 2A (Mann–Whitney U test: $U = 15$, $Z = -1.53$, $N_{capuchins} = 8$, $N_{macaques} = 7$, $P = 0.16$; $r = -0.40$) or Study 2B (Mann–Whitney U test: $U = 21.50$, $Z = -0.44$, $N_{capuchins} = 10$, $N_{macaques} = 5$, $P = 0.69$; $r = -0.11$). Also, although the two species did not differ on the reversal phase in Study 2A (Mann–Whitney U test: $U = 20.50$,

$Z = -0.88$, $N_{capuchins} = 8$, $N_{macaques} = 7$, $P = 0.40$, $r = -0.23$), capuchins showed a trend toward being faster than macaques in Study 2B (Mann–Whitney U test: $U = 1.50$, $Z = -2.00$, $N_{capuchins} = 4$, $N_{macaques} = 4$, $P = 0.11$; $r = -0.71$).

Study 2 discussion

Overall more capuchin monkeys found the payoff maximizing solution in Study 2A than in Study 1A and, unlike in Study 1A, their performance in Study 2A was better than in the original plate task. Similarly to Study 1A, the capuchin subject with no previous experience did as well as her conspecifics who were more familiar with the paradigm, indicating that their performance in Study 2A was not necessarily due to more experience with the task. Like capuchins, more macaques solved the task in Study 2A than in Study 1A. This finding was even more striking as all subjects, including those who preferred the permanent icon in Study 1A, developed a preference for the ephemeral option. One reason for the success in Study 2A may have been the fact that the moving icon was more prone to catching the monkeys' attention than in any of the other studies. The monkeys' performance was no different in Study 2B, when both icons were vibrating, than in Study 1A (and not much different from the original plate task).

Finally, both macaque and capuchin monkeys took longer overall to meet criterion on the reversal phase than the initial phase in both studies. Our finding was consistent with previous research showing that capuchins may have difficulty with reversal tests (e.g., Brosnan and de Waal 2004; Beran et al. 2008b), and contrasts with the capuchins' results in the original plate task (Salwiczek et al. 2012) in which subjects more rapidly learned the reversal phase than the initial phase.

General discussion

In the current paper, our goal was to replicate a previous study in which monkeys and apes failed to learn a two-choice discrimination task as rapidly as cleaner fish, in order to tease apart some potential explanations for the fishes' superior performance. In particular, we tested two possible explanations for their failure: first, that the unintentional extraneous cues present in a manual (as opposed to computerized) testing environment may have made the task more challenging for the primates, and second, that adding a movement component that more closely replicated natural stimuli from the monkeys' ecology would improve their performance. As we discussed in detail in each Study's Discussion section, both of these modifications allowed the monkeys to pass the task that they had previously failed. We do not repeat the discussions of why these modifications succeeded here, but instead consider the overall implications.

One interesting finding in the current study, not seen previously, was that several monkeys developed a preference for choosing the permanent icon. There are several possible explanations for this. For instance, they may have preferred the option that never disappeared, or found the disappearing one frustrating or confusing. Alternately, it may be that this was related to some sort of superstitious behavior (Blanchard et al. 2014). It is also worth noting that while both species showed this suboptimal preference, it was particularly notable in the macaques. In particular, while the computerized methodology of Study 1 increased capuchin monkeys' preference for the ephemeral icon over the original plate task, half of the rhesus macaques developed a preference for the permanent icon, indicating that the task might be perceived differently between the two species. Understanding this difference may shed light on why the permanent icon ever came to be preferred.

Before we can be certain that the changes in procedure were responsible for the improved performance, there are other possible explanations that must be ruled out. For instance, one possibility is that this is an experience effect, resulting from the fact that these animals were tested on these paradigms sequentially. However, we think this is unlikely for two main reasons. First, the better performance of the capuchin monkeys on the computerized tasks was also found for novice monkeys who had no previous experience with the task. While this does not rule out an experience effect among the experienced individuals, it does show that novices could do much better on this task than (other) novice monkeys did on the previous test. Also in support of this, half of our macaque subjects performed at a level comparable to the capuchins in both studies, again despite having no previous experience with the paradigm. Second, the monkeys did not improve their performance in the reversal tests, indicating that even among the experienced individuals, there was not a great influence of previous experience on their results.

A second potential challenge to our results is that we did not also test fish. In particular, it may be important to see whether their performance decreased, which would be expected given that the task was designed to be ecologically relevant for monkeys, not fish. Then, again, depending on the degree to which fish can generalize, they might have maintained a high level of performance. Of course, the current version, which uses a computerized procedure, would be very challenging to test using identical procedures with the fish (although see Saverino and Gerlai 2008; Siebeck et al. 2009, 2010), but an adapted paradigm could be used. Additional monkey–fish comparisons are a line of

research that we are pursuing using a different procedure that is more amenable to underwater testing.

Our results highlight a tension present in any comparative work, particularly when testing species that differ substantially in body plan, ecology, or cognitive ability; when designing the task, the researchers must not only develop procedures that meet the needs of each species, but also make a choice between keeping the task *parameters* identical and keeping the *goals* of the task identical. In fact, we originally chose this task for our comparison explicitly because both primates and fish are capable of two-choice tasks with minimal modifications required (i.e., primates choose with their hands, whereas fish choose with their mouths). Our current results show that even in identical tasks (e.g., the earlier plate task), details in how subjects experience it may make the task substantially different between species.

Therefore, we recommend that direct comparisons take a two-step approach. First, researchers should directly compare the subjects on a task in which parameters are identical, as we did in Salwiczek et al. (2012), to establish a baseline for comparing further results. Second, researchers should then, when possible, compare the subjects based on whether they meet certain criteria that respect the spirit of the decision being tested, but also address species-specific strengths by using different procedures that may or may not be identical, as we tried to do here. Neither approach is ideal in isolation, but combining the two across a long-running research program allows for a very strong comparative test.

We note that this is becoming more common among comparative psychologists. For instance, Pepperberg and Hartsfield (2014) replicated Salwiczek et al. (2012)'s study with parrots and found that they performed at levels comparable to the cleaner fish (and therefore better than the primates) in the plate task. They argued that this task may be more natural for species that are physically limited to grabbing only one food item at a time (e.g., using a mouth or beak) because these animals are accustomed to sequential acquisition of food, as in this task, rather than simultaneous acquisition, as is possible with two hands. This artificial limitation may have been frustrating to the bimanual monkeys. In this context, it is interesting that in the current task, our monkeys were using a joystick to control a single cursor and were able to learn more rapidly. We believe that this supports Pepperberg & Hartsfield's hypothesis, because the sequential task was easier for the (bimanual) monkeys to learn when they were limited to sequential acquisition by the testing modality (to which they were already accustomed) rather than artificially being forced to choose one option at a time when they could easily have grabbed both if not limited by the experimenter. This may indicate that it is not only the behavior and presentation of the stimuli that is important, but also the format of the test, and provides a possible explanation for other results, such as the finding that primates more easily pass the trap-tube task when they have a single tool that they must decide where to use than when they are presented with two tubes, each baited with its own tool (Mulcahy and Call 2006; Girndt et al. 2008). Continued work on questions like these across species that vary on many dimensions will help to clarify how ecology, cues, and cognitive abilities influence decision-making across variety of contexts.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All applicable international, national, and/or institutional guidelines for the care and use of animals in this study were followed. All procedures performed in this study were in accordance with the Georgia State University Institutional and Animal Care and Use Committee (capuchins: A13022, rhesus: A13021) and met the standards of the USA.

References

Anderson JR, Degiorgio C, Lamarque C, Fagot J (1996) A multi-task assessment of hand lateralization in capuchin monkeys (*Cebus apella*). Primates 37:97–103

Barros RDS, Galvão ODF, Mcilvane WJ (2002) Generalized identity matching-to-sample in *Cebus apella*. Psychol Rec 52:441–460

Beran MJ (2007) Rhesus monkeys (*Macaca mulatta*) enumerate sequentially presented sets of items using analog numerical representations. J Exp Psychol Anim B 33:42–54

Beran MJ (2008) Monkeys (*Macaca mulatta* and *Cebus apella*) track, enumerate, and compare multiple sets of moving items. J Exp Psychol Anim B 34:63–74

Beran MJ, Parrish AE (2012) Sequential responding and planning in capuchin monkeys (*Cebus apella*). Anim Cogn 15:1085–1094

Beran MJ, Smith JD (2011) Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). Cognition 120:90–105

Beran MJ, Harris EH, Evans TA, Klein ED, Chan B, Flemming TM, Washburn DA (2008a) Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*) and rhesus monkeys (*Macaca mulatta*): the effects of differential and nondifferential reward. J Comp Psychol 122:52–61

Beran MJ, Klein ED, Evans TA, Chan B, Flemming TM, Harris EH, Washburn DA, Rumbaugh DM (2008b) Discrimination reversal learning in capuchin monkeys (*Cebus apella*). Psychol Rec 58:3–14

Beran MJ, Smith JD, Coutinho MVC, Couchman JJ, Boomer J (2009) The psychological organization of "uncertainty" responses and "middle" responses: a dissociation in capuchin monkeys (Cebus apella). J Exp Psychol Anim B 35:371–381

Beran MJ, Evans TA, Klein ED, Einstein GO (2012) Rhesus monkeys (Macaca mulatta) and capuchin monkeys (Cebus apella) remember future responses in a computerized task. J Exp Psychol Anim B 38:233–243

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

Bitterman ME (1960) Toward a comparative psychology of learning. Am Psychol 15:704–712

Bitterman ME (1965) Phyletic differences in learning. Am Psychol 20:396–410

Blanchard TC, Wilke A, Hayden BY (2014) Hot-hand bias in rhesus monkeys. J Exp Psychol Anim Learn Cogn 40:280–286

Boesch C, Boesch H (1989) Hunting behavior of wild chimpanzees in the Taï National Park. Am J Phys Anthropol 78:547–573

Boysen ST, Berntson GG (1995) Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees (Pan troglodytes). J Exp Psychol Anim B 21:82–86

Boysen ST, Berntson GG, Hannan MB, Cacioppa JT (1996) Quantity-based interference and symbolic representations in chimpanzees (Pan troglodytes). J Exp Psychol Anim B 22:76–86

Boysen ST, Mukobi KL, Berntson GG (1999) Overcoming response bias using symbolic representations of number by chimpanzees (Pan troglodytes). Anim Learn Behav 27:229–235

Boysen ST, Berntson GG, Mukobi KL (2001) Size matters: impact of size and quantity on array choice by chimpanzees (Pan troglodytes). J Comp Psychol 115:106–110

Brosnan SF (2010) What do capuchin monkeys tell us about cooperation? In: Forsyt DR, Hoyt CL (eds) For the greater good of all: perspectives on individualism, society, and leadership. Palgrave Macmillan, Basingstoke, pp 11–28

Brosnan SF, de Waal FBM (2004) Socially learned preferences for differentially rewarded tokens in the brown capuchin monkey (Cebus apella). J Comp Psychol 118:133–139

Brosnan SF, Wilson B, Beran MJ (2012) Old World monkeys are more similar to humans than New World monkeys when playing a coordination game. Proc R Soc Lond B Biol Sci 279:1522–1530

Bshary R (2001) The cleaner fish market. In: Noë R, van Hooff JARAM, Hammerstein P (eds) Economics in nature. Cambridge University Press, Cambridge, pp 146–172

Bshary R, Grutter AS (2002) Asymmetric cheating opportunities and partner control in the cleaner fish mutualism. Anim Behav 63:547–555

Bshary R, Schäffer D (2002) Choosy reef fish select cleaner fish that provide high-quality service. Anim Behav 63:557–564

Chalmeau R, Lardeux K, Brandibas P, Gallo A (1997) Cooperative problem solving by orangutans (Pongo pygmaeus). Int J Primatol 18:23–32

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, Hillsdale

Davison M, Baum WM (2000) Choice in a variable environment: every reinforce counts. J Exper Anal Behav 74:1–24

de Waal FBM, Leimgruber K, Greenberg AR (2008) Giving is self-rewarding for monkeys. Proc Natl Acad Sci USA 105:13685–13689

Dufour V, Pelé M, Neumann M, Thierry B, Call J (2009) Calculated reciprocity after all: computation behind token transfers in orangutans. Biol Lett 5:172–175

Evans TA, Beran MJ (2012) Monkeys exhibit prospective memory in a computerized task. Cognition 125:131–140

Evans TA, Beran MJ, Chan B, Klein ED, Menzel CR (2008) An efficient computerized testing method for the capuchin monkey (Cebus apella): adaptation of the LRC-CTS to a socially housed nonhuman primate species. Behav Res Methods 40:590–596

Fragaszy DM, Visalberghi E, Fedigan LM (2004) The complete capuchin: the biology of the Genus Cebus. Cambridge University Press, Cambridge

Galdikas BMF (1988) Orang-utan diet, range, and activity at Tanjung Puting, Central Borneo. Int J Primatol 9:1–35

Girndt A, Meier T, Call J (2008) Task constraints mask great apes' ability to solve the trap table task. J Exp Psychol Anim B 34:54–62

Goodall J (1986) The chimpanzees of Gombe: patterns of behavior. Harvard University Press, Cambridge

Gossette RL, Inman N (1966) Comparison of spatial successive discrimination reversal performances of two groups of New World monkeys. Percept Motor Skill 23:169–170

Grant DA (1946) New statistical criteria for learning and problem solution in experiments involving repeated trials. Psychol Bull 43:272–282

Harlow HF (1949) The formation of learning sets. Psychol Rev 56:51–65

Harlow HF, Warren JM (1952) Formation and transfer of discrimination learning sets. J Comp Physiol Psychol 45:482–489

Kamil AC (1988) Synthetic approach to the study of animal intelligence. In: Leger DW (ed) Comparative perspectives in modern psychology: Nebraska symposium on motivation. University of Nebraska Press, Lincoln, pp 230–257

Kamil AC (1998) On the proper definition of cognitive ethology. In: Balda RP, Bednekoff PA, Kamil AC (eds) Animal cognition in nature. Academic Press, New York

Masataka N (1990) Handedness of capuchin monkeys. Folia Primatol 55:189–192

Mulcahy NJ, Call J (2006) How great apes perform on a modified trap-tube task. Anim Cogn 9:193–199

Murray EA, Kralik JD, Wise SP (2005) Learning to inhibit prepotent responses: successful performance by rhesus macaques, Macaca mulatta, on the reversed-contingency task. Anim Behav 69:991–998

Pepperberg IM, Hartsfield LA (2014) Can Grey parrots (Psittacus erithacus) succeed on a "complex" foraging task failed by nonhuman primates (Pan troglodytes, Pongo abelii, Sapajus apella) but solved by wrasse fish (Labroides dimidiatus)? J Comp Psychol 128:298–306

Richardson WK, Washburn DA, Hopkins WD, Savage-Rumbaugh ES, Rumbaugh DM (1990) The NASA/LRC computerized test system. Behav Res Methods Instrum Comput 22:127–131

Rijksen HD (1978) A field study on Sumatran orangutans (Pongo pygmaeus abelii Lesson 1827): ecology, behaviour and conservation. H. Veenman Zonen BV, Wageningen

Rosenthal R (1991) Meta-analytic procedures for social research, 2nd edn. Sage, Newbury Park, p 19

Rumbaugh DM (1971) Evidence of qualitative differences in learning processes among primates. J Comp Physiol Psychol 76:250–255

Rumbaugh DM (1997) Competence, cortex, and primate models: a comparative primate perspective. In: Krasnegor NA, Lyon GR, Goldman-Rakic PS (eds) Development of the prefrontal cortex: evolution, neurobiology, and behavior. Paul H. Brookes, Baltimore, pp 117–139

Salwiczek LH, Prétôt L, Demarta L, Proctor D, Essler J, Pinto AI, Wismer S, Stoinski T, Brosnan SF, Bshary R (2012) Adult cleaner wrasse outperform capuchin monkeys, chimpanzees and orang-utans in a complex foraging task derived from cleaner-client reef fish cooperation. PLoS ONE 7:e49068

Saverino C, Gerlai R (2008) The social zebrafish: behavioral responses to conspecific, heterospecific, and computer animated fish. Behav Brain Res 191:77–87

Schrier AM (1984) Learning how to learn: the significance and current status of learning-set formation. Primates 25:95–102

Siebeck UE, Litherland L, Wallis GM (2009) Shape learning and discrimination in reef fish. J Exp Biol 212:2112–2118

Siebeck UE, Parker AN, Sprenger D, Mäthger LM, Wallis G (2010) A species of reef fish that uses ultraviolet patterns for covert face recognition. Curr Biol 20:407–410

Timberlake W, Gawley DJ, Lucas GA (1987) Time horizons in rats foraging for food in temporally separated patches. J Exper Anim Behav Proc 13:302–309

Washburn DA, Rumbaugh DM (1992) Testing primates with joystick-based automated apparatus: lessons from the Language Research Center's Computerized Test System. Behav Res Methods Instrum Comput 24:157–164

Westergaard GC, Suomi SJ (1997) Capuchin monkey (Cebus apella) grips for the use of stone tools. Am J Phys Anthropol 103:131–135