Understanding Solidity: Investigating Knowledge of a Functional Object Property in Brown Capuchin Monkeys (*Sapajus apella*) and Common Squirrel Monkeys (*Saimiri sciureus*)

Eleanor J. Jordan[1], Luke A. J. Townrow[2], Chloë India Wright[3,4], and Amanda M. Seed[1,*]

[1]School of Psychology and Neuroscience, University of St. Andrews
[2]School of Psychology, Cardiff University
[3]School of Philosophy, Psychology, and Language Sciences, University of Edinburgh
[4]School of Life Sciences, University of Lincoln (current affiliation)

*Corresponding author (Email: ams18@st-andrews.ac.uk)

Abstract – Humans form abstract representations about the physical properties of objects, with very young infants having 'core knowledge' about solidity and continuity. Whether nonhuman primates also form abstract representation of physical properties is debated. Despite studies showing that some nonhuman primates can discriminate between functional and non-functional tools, whether they achieve this by recognizing an object's physical properties or via associative learning of perceptual cues remains contested. One method for exploring an individual's physical cognition is the trap-tube (where subjects push a reward out of a pipe whilst avoiding a trap). The trap-tube investigates whether participants understand that rewards cannot pass through solid objects (solidity) and that unsupported objects will fall (gravity). Initial research reported that chimpanzees and capuchin monkeys solved the test using associative rules; however, modifying the task by removing the tool revealed an understanding of solidity in chimpanzees. In this series of studies, we tested 12 squirrel monkeys and 14 capuchin monkeys on a modified trap-box, where the monkeys could move the reward using their fingers rather than a tool. No individual of either species passed the task at above chance levels. In Experiment 2, the same trap blocked one of two cups: both species learned to avoid it. In Experiment 3, only capuchins generalized this solution to novel materials, performing better when cues had a functional versus a non-functional relationship to the outcome. Squirrel monkeys appeared to learn via association, while results from capuchins indicated a sensitivity to the concept of solidity.

Keywords – Object knowledge, Solidity, Physical cognition, Trap task, Capuchin monkeys, Squirrel monkeys

The world is full of objects with differing physical properties. It is these properties that dictate how different objects will interact with one another. For example, the rigidity of an object will determine whether it is able to hold its shape when force is applied to it. In humans, it has been suggested that, from just a few months old, infants have some understanding of basic object properties such as solidity and continuity (Spelke & Kinzler, 2007). Humans likely acquire this object knowledge through multiple routes: certain object properties may be understood from birth, 'core knowledge,' whilst other

components of our object knowledge are built up via our life experiences. Irrespective of the process through which object knowledge is acquired, it is accepted that humans form abstract representations of object properties, which they use during problem solving to predict how objects will behave in relation to both themselves and other objects.

Despite the acquisition of object knowledge being accepted by most researchers working on the early cognitive development of children, its application to nonhuman primate cognition is controversial (Penn & Povinelli, 2007). This is especially true for new world monkeys where their performance in tests designed to investigate object knowledge has been notably inconsistent. Within the literature, there is evidence that both non-tool-using and tool-using new world primates are able to learn to identify and select the most appropriate or efficient tools for retrieving food, based on perceptual tool properties (capuchin monkeys: Evans & Westergaard, 2004; Fujita et al., 2003, 2011; Judge & Bruno, 2012; Manrique et al., 2011; Mayer et al., 2014; Painter et al., 2019; Sabbatini et al., 2014; Schrauf et al., 2008; Visalberghi & Neel, 2003; Yocom & Boysen, 2010; common marmosets: Gagne et al., 2012; Spaulding & Hauser, 2005; Yamakazi et al., 2011; cottontop tamarins: Hauser et al., 1999, 2002; Santos et al., 2005, 2006; Spaulding & Hauser, 2005; squirrel monkeys: Painter et al., 2019). However, although findings showing the preferential selection of functional tools provide support for the idea that new world monkeys are sensitive to the efficiency of tools, such studies do not confirm that the monkeys do so by forming abstract representations of the object properties involved in the tools' functionality. All four tested species are, after minimal experience, able to generalize their learning when non-functional changes to perceptual cues (e.g., color and texture) and novel tools are introduced (capuchin monkeys: Evans & Westergaard, 2004; Fujita et al., 2003, 2011; Manrique et al., 2011; Painter et al., 2019; Yocom & Boysen, 2010; common marmosets: Spaulding and Hauser, 2005; cottontop tamarins: Hauser et al., 1999, 2002; Santos et al., 2005, 2006; Spaulding & Hauser, 2005; squirrel monkeys: Painter et al., 2019). Although this is impressive, unless more complex controls are included, the monkeys could be solving the novel task via perceptual generalization: simply choosing the tool that looked the most similar to the initial training tool (Povinelli, 2003). For example, in the throwing paradigm (Evans & Westergaard, 2004), capuchins learned to throw a rock connected to a string to 'fish' for honey. In the test trials, they were presented with the same tool they had been trained with (or in their second experiment, an almost identical tool made of novel materials) versus a similar tool that was functionally different in some way and, therefore, was also perceptually different. The capuchins' ability to select the correct tool could thus be explained by choices based solely on visual similarity to the training tool, without the need for an appreciation of the functional properties of the tools (see Povinelli, 2003, for further discussion).

Similarly, when selecting a functional versus non-functional spatial arrangement, monkeys have been shown to learn using spatial rules: i.e., to "select the tool with food on top" in cloth pulling studies (capuchin monkeys: Yocom & Boysen, 2010; common marmosets: Yamazaki et al., 2011; cottontop tamarins: Hauser et al., 1999) and to "select the tool with food in front" in hook pulling studies (capuchin monkeys: Cummins-Sebree & Fragaszy, 2005; Fujita et al., 2003, 2011; Painter et al., 2019; cottontop tamarins: Hauser et al., 2002; Santos et al., 2006; squirrel monkeys: Painter et al., 2019). These studies successfully demonstrate that new world monkeys can learn to recognize appropriate tools and tool arrangements but are not convincingly able to show that they do so due to an appreciation of the object properties involved. In fact, when Painter et al. (2019) introduced a novel hook tool, which controlled for multiple spatial cues, both capuchin and squirrel monkeys failed to generalize their initial success to these novel tools. After initially passing a hook pulling task, the capuchin monkeys found it easy to generalize from hook tools to V tools, whereas the squirrel monkeys took just as long to learn to solve the task with the novel toolset. Next, when the monkeys were presented with H and F shaped tools, both species appeared to be solving the task using a hierarchy of spatial rules: "avoid inverted tools" followed by "choose tools without a barrier to the food." However, making the discrimination more subtle makes interpretation of the results more difficult, as even with the most subtle distinction between two objects, there remains (by necessity) some perceptual constant that correlates with the functional difference. Furthermore, tasks involving increasingly subtle discriminations place more demands on peripheral

abilities such as attention and increasing the demands on these abilities could become a limiting factor to performance (Seed et al., 2012; Seed & Mayer, 2017).

More compelling evidence for an understanding of object properties in capuchins was found by Manrique et al. (2011). In this study, capuchins received training to retrieve a food reward using a wooden stick and, in the following trials, needed to select the correct (rigid) tool (the one that would interact with the reward in the same way as their training tool) from a set of novel tools. The authors showed that the capuchins chose correctly only after experiencing how the tools behaved when manipulated; when forced to make their choice based only on visible cues, the capuchins did not perform above chance. This suggests that they were using the affordances of the tools to make their decisions. Furthermore, the capuchins solved a transfer task that required the use of a tool with the opposite property (i.e., a flexible tool), ruling out a learned preference for rigidity.

Another approach to disentangling object knowledge from associative learning of perceptual cues is to use the same or similar stimuli within two contexts: one in which the cue has functional relevance, and one in which it does not. Mayer et al. (2014) showed that capuchins, apes, and children all performed better when they needed to avoid a broken tool that was connected to a reward than they did when they simply needed to avoid pulling from the side where the same broken tool was placed as a cue, but not attached to the reward.

Acquiring knowledge about an object's properties is meaningful and of use only when these properties are placed into an appropriate context, and often this requires that they are incorporated into a causal framework. For example, knowing that a stick tool is rigid is only useful if you can reason that this enables it to rake in a reward. In nonhuman primates, one of the classic tests for this sort of knowledge is the trap-tube task of causality (Vislberghi & Limongelli, 1994). In this task, a reward is placed in the center of a tube and can be retrieved by using a stick to push it along the solid surface of the tube so that it falls out of one end. On one side of the tube, a trap is positioned such that the surface of the tube is no longer solid and hence, pushing the reward this way causes the reward to fall into the trap, rendering it unobtainable. Participants should learn to push the reward along the solid surface, away from the trap. Once an individual has successfully learned to avoid the trap, the tube is rotated so that the trap is in the roof of the tube and, hence, is non-operational. This transfer task aims to test what the participant has learned from the initial stage of the task. If they have not gained an understanding that they must push the reward along a continuous solid surface in order to retrieve it, and instead have learned to solve the task using a simple rule such as "push away from the trap," then they will continue to avoid the trap, despite it now being non-operational and possible to retrieve the reward from either direction. This task has previously been presented to only one species of new-world monkey: the capuchin monkey. Only one of the four capuchin monkeys tested solved the initial stage of the test and avoided the trap at above chance rates, but then continued to avoid the trap once it was inverted, suggesting that she had not learned the properties of the trap-tube (Visalberghi & Limongelli, 1994). However, when given the same task, human adults will also continue to avoid the inverted trap (Silva et al., 2005), and so these results must be taken with caution.

Since its initial use with capuchin monkeys, the trap-tube test has undergone numerous modifications. The modified trap-task involves the reward being placed into the center of a shelf within a transparent box, with holes in the front of the box allowing participants to insert their fingers and move the reward left or right along the shelf (Seed, Call, et al., 2009). Once pushed to the end of the shelf, the reward can fall out of the box either through vertical channels at either edge of the shelf, or through side exits in-line with the shelf (Figure 1). 'Trap' pieces can be inserted into the channels to block the reward from falling out of the box at any of the four possible exits and 'bridge' pieces can be inserted at the top of the channel, which allow the participant to move the reward towards the side exits. Different configurations of these trap and bridge pieces are presented to participants to investigate their understanding of solidity and support. This modified trap-task has three main advantages over the original trap-tube test. First, the new setup facilitates the presentation of more informative transfer tasks (Seed, Call, et al., 2009). As mentioned above, understanding an object's properties is useful only in reference to other things, and in the real world, objects with a certain property (e.g., solidity) can be either good or bad

depending on the context. In the classic trap-tube, the solid surface is always the successful option, and so the participants' understanding cannot be fully investigated as any preference for the solid surface could be due to a learned association between the solid surfaces and gaining rewards, rather than an understanding that solid surfaces prevent the reward from falling. In the modified trap-task, if a participant successfully learns to retrieve rewards from the box, the trap can be altered in such a way that the perceptual cue (and its related property) is still present in the box but no longer instantiates the correct solution. By presenting the participant with a variety of these transfer tasks, it is possible to investigate in detail what the participants have understood about the properties of the trap (see Seed, Call et al., 2009 for a detailed description).

Figure 1

The Two-Trap-Box the Trap-Box from the Monkeys' Perspective, With All Traps in Place, Installed in the Window of the Research Cubicle

Second, within the modified trap-task, the position of the tube is modified such that participants are no longer required to push the reward away from themselves, and instead must push the reward left or right. Pushing food away from yourself to retrieve it is not a natural behavior for most animals, and it is argued that this unnatural set-up is likely to mask the knowledge of some participants (Mulcahy & Call, 2006). This may well be true for capuchin monkeys, as their performance is better when presented with the aforementioned raking or pulling tasks (Evans & Westergaard, 2004; Fujita et al., 2003, 2011; Mayer et al., 2014; Yocom & Boysen, 2010).

Thirdly, the modified trap-task removes the tool component of the task. As well as enabling non-tool using species to be tested, this should also increase the performance of some tool-using species, as using a tool to manipulate a reward whilst also monitoring how the reward is interacting with the environment requires an individual to simultaneously coordinate two dynamic relationships (Fragaszy & Cummins-Sebree, 2005). In a hook-pulling study with capuchin monkeys, participants failed to select the correct tool when obstacles were added to the environment (Fujita et al., 2003). The authors suggested that this failure was due to difficulties in processing second order relations: i.e., relationships between three entities, A (tool), B (reward), and C (obstacle). This is discussed in detail in a review paper by

Fragaszy and Cummins-Sebree (2005) where, in concurrence with the hook pulling study, the authors detail that, when capuchins had to anticipate where along a path a ball would stop, they were unable to anticipate that it would fall into a gap in the pathway. In their review, the authors showed that capuchins often rely on single relations and, in more complex studies, will attempt to find strategies that avoid simultaneous reasoning about multiple relations. However, after mastering a single relation, capuchin monkeys are then able to integrate a second relation into their problem solving, as shown in the case of capuchins learning to use a joystick to move a cursor across a display to a rewarded goal box (Leighty & Fragaszy, 2003). Similar results occur in the raking tasks: once the capuchins had received experience of how the rewards interacted with obstacles, they were subsequently able to choose the correct hook-tool to avoid them (Fujita et al., 2011). When applied to the trap task, these results suggest that either providing abundant tool experience or removing the tool component of the task would increase the performance of capuchin monkeys, as they would no longer have to simultaneously coordinate multiple relationships.

Use of the modified trap-task has enabled chimpanzees, corvids and children to pass the task (avoiding the trap) in larger numbers, with some individuals also managing to pass some of the transfer tasks (chimpanzees: Limongelli et al., 1995; Mulcahy & Call, 2006; Seed, Call et al., 2009; corvids: Seed et al., 2006; but see Taylor et al., 2009; Tebbich et al., 2007; children: Seed & Call, 2014). However, no monkey species has yet been tested with a modified trap-task. Therefore, in Experiment 1, we presented the modified trap-task from Seed, Call et al. (2009) to two species of new world primates: capuchin monkeys and squirrel monkeys. Squirrel monkeys are not tool users in the wild, and there is only one reported instance of squirrel monkey tool use in captivity (Buckmaster et al., 2015). Therefore, the removal of the tool component of the task enables us to test these new world primates. However, as they have never been tested with this type of apparatus, we could not predict how they would perform. Previous work has shown that capuchins struggle with multi-level relationships (Fragaszy & Cummins-Sebree, 2005; Fujita et al., 2003). Therefore, we expected that the removal of the tool component could allow the monkeys to solve the task in greater numbers, just as it did in studies with chimpanzees and preschool-aged children (Seed, Call et al., 2009; Seed et al., 2014; Völter & Call, 2014). Subjects that solved the modified trap-task, would receive transfer configurations to further investigate their understanding of the functional relevance of object properties and control for the use of spatial cues.

Experiment 1—Two-Trap-Box

Method

Subjects & Housing

All the monkeys participating in the study were housed at the University of St Andrews' "Living Links to Human Evolution" research center located within the Royal Zoological Society of Scotland's Edinburgh Zoo. At the center, the monkeys live in two mixed-species communities made up of common squirrel monkeys (*Saimiri sciureus*) and brown tufted capuchin monkeys (*Sapajus* sp.) Both groups' enclosures consist of an indoor capuchin area (7 m x 4.5 m x 6 m), to which both species have access; an indoor squirrel monkey enclosure (5.5 m x 4.5 m x 6 m), to which only the squirrel monkeys have access; and a large shared outdoor area (approximately 900 m^2), consisting of natural vegetation and climbing structures. Situated between the indoor areas is a research room where, at specified research times, the monkeys have access to their testing cubicles. Research sessions took place up to five days a week, twice a day at 11.15 a.m.–12.45 p.m. and 2.15 p.m.–4 p.m., Monday to Friday. All monkeys were tested in these familiar testing cubicles, which they entered voluntarily and were able to leave at any time. Participants came from both groups at the research center: The East group and the West group. The groups live in adjacent enclosures, which are mirror images of each other, under identical housing conditions and in similarly sized social groups. The monkeys are fed a variety of fruits, vegetables, cereals, and insects several times per day. They are never food-deprived, and water is available *ad*

libitum. All participation in experiments was voluntary and all food rewards provided (peanuts, raisins, yoghurt and oat mix, and sunflower seeds) were supplemental to the monkeys' daily diet.

Between June and July 2016, we tested 12 squirrel monkeys (3 from the West group and 9 from the East group). We also tested 7 capuchin monkeys from each group, with half of the monkeys undergoing testing between October and December 2016 and the other half between November 2017 and March 2018. Tables 1 and 2 detail which individuals of each species were involved in each experiment.

Table 1

Capuchin Monkey Participants' Details and the Experiments in which They Participated

Monkey	Group	Sex	Experiment			
					3	
			1	2	Functional	Non-functional
Anita	East	F	.54	**.76**	x	x
Junon	East	F	.53	.59	.73	.6
Rosa	East	F	x	x	.5	.63
Carlos	East	M	.49	**.61**	**.8**	.57
Chico	East	M	.5	**.7**	**.93**	.6
Flojo	East	M	.52	**.8**	.63	.67
Kato	East	M	.51	**.63**	**.73**	.57
Reuben	East	M	.49	**.72**	.5	.37
Alba	West	F	x	x	**.8**	.53
Hazel	West	F	.51	**.71**	x	x
Luna	West	F	x	x	**.9**	**1**
Pedra	West	F	.56	.56	**.67**	.5
Pixie	West	F	.55	.52	**.83**	.53
Bear	West	M	.58	**.71**	**.77**	.47
Diego	West	M	x	x	.63	.57
Figo	West	M	x	x	**.9**	.53
Fudge	West	M	x	x	.57	.47
Inti	West	M	.46	**.7**	.57	.47
Mekoe	West	M	x	x	**.87**	.73
Rufo	West	M	x	x	**.87**	.7
Torres	West	M	.59	**.66**	.6	.63
Ximo	West	M	.53	**.68**	**.8**	.6

Note. Proportion of trials correct in each experiment is detailed, with the score for monkeys that reached the criteria for significance presented in bold. The x indicates that the monkey did not participate in the experiment. For Experiment 3, scores are separated into the functional and non-functional conditions.

Apparatus

The modified trap-task (the two-trap-box; 30 cm x 30 cm x 10 cm) was mounted in the window of one of the testing cubicles (Figure 1). The two-trap-box has a transparent Perspex front so that the monkeys can see into the box. Inside the box, there is an 18.5 cm long grey shelf 4 cm below the top, onto which the food is placed centrally from the back of the box. A 4.5 cm-wide vertical channel runs down both sides of the box, allowing the food to fall into the cubicle when it is pushed to either edge of the shelf, if the channel is not blocked with a trap. Just below the level of the shelf, there are 2 cm x 8 cm

horizontal openings on each side of the box through which the food reward can be retrieved if the grey shelf pieces are in place. Along the front of the box there are three slots in the Perspex (5 cm x 1 cm, positioned 0.5 cm apart) that run in line with the complete length of the shelf. The slots are large enough for the monkeys to insert their fingers into the box, but too small for the food reward to pass through. Subjects can insert their fingers into the slots and push the reward left or right along the shelf to make it either fall down one of the side channels or, if the shelf piece is in place, fall out of the side exits. The configuration of the box can be changed by the insertion or removal of shelf, blocker, and trap pieces through the back of the box (Figures 1 & 2). All inserted pieces can be seen by the monkeys from the front of the box.

Table 2

Squirrel Monkey Participants' Details and the Experiments in which They Participated

Monkey	Group	Sex	Experiment			
					3	
			1	2	Functional	Non-functional
Amarilla	East	F	.46	**.69**	.27	.23
Cali	East	F	.46	.62	.47	.27
Ciara	East	F	.52	**.61**	.43	.63
Dora	East	F	.46	**.65**	.57	.53
Elie	East	F	x	x	.27	.33
Flora	East	F	.48	**.74**	x	x
Gabriella	East	F	.55	.61	.43	.43
Lexi	East	F	x	.63	.47	.37
Maya	East	F	x	x	.53	.43
Pelusa	East	F	.5	.57	**.77**	0.5
Roca	East	F	.57	.54	.5	.43
Sipi	East	F	x	**.78**	x	x
Valencia	East	F	.42	.51	.4	.47
Gisele	West	F	x	x	**.6**	.47
Guacamole	West	F	x	x	.53	.67
Inca	West	F	x	x	.5	.53
Jasmin	West	F	x	x	.6	.43
Loki	West	F	.53	x	x	x
Orla	West	F	x	x	.53	**.63**
Sancha	West	F	.5	**.73**	x	x
Toomi	West	F	.5	.66	.57	.57

Note. Proportion of trials correct in each experiment is detailed, with the score for monkeys that reached the criteria for significance presented in bold. The x indicates that the monkey did not participate the experiment. For Experiment 3, scores are separated into the functional and non-functional conditions.

Figure 2

Removable Trap Pieces from the Two-Trap-Box
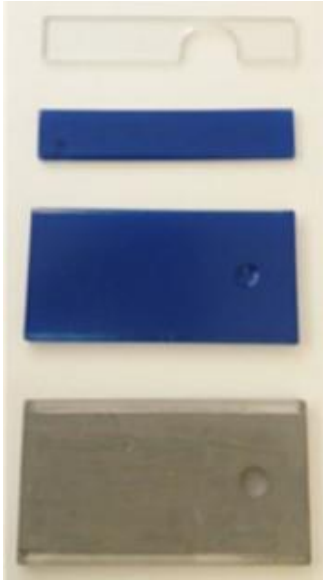
Note. From top to bottom: the transparent side piece used to stop the monkeys reaching in through the side of the box; the blue side blocker piece; the blue trap piece; and the grey shelf piece.

There are four possible exits for the food reward: one at the bottom of each of the vertical channels (4.5 cm x 10 cm) and one in each side of the box, in line with either side of the shelf (2 cm x 8 cm). Each of the exits can be blocked with a blue blocker piece (Figure 2). A blue trap piece (5.5 cm x 11 cm x 0.5 cm) can be inserted either halfway up or at the bottom of each vertical channel to trap the reward as it falls off the shelf, and a smaller blue blocker piece (0.3 cm x 11 cm x 2 cm) can be inserted through the side of the box to block the side exits. The grey shelf pieces (6 cm x 11 cm x 0.5 cm) can be inserted either side of the shelf. These stop the reward falling down the channel but allow it to pass over them and out of the side exits. Part-way through testing the monkeys, some subjects began to reach their arms in through the side exits of the box to grab the reward before it fell off the shelf. To prevent this, two additional side pieces were introduced. Transparent side pieces (0.3 cm x 11 cm x 2 cm) with a small hole (2 cm in diameter) were inserted into both side exits (Figure 2). These prevented the monkeys from reaching their arms into the box, but still allowed the reward to fall out of the box and into the cubicle if it was moved in the correct direction. Holes positioned strategically in the back of the box allowed the experimenter to remove trapped food from any of the possible trapped positions (Figure 1).

There were four trap-box designs (Figure 3), which correspond to tubes A–D in Seed et al. (2006) and which were identical to those used with chimpanzees in Seed, Call et al. (2009). Configuration A uses one shelf piece and one trap piece. To retrieve the reward, the monkey must push the food away from the trap, over the shelf, and out of the side exit. Configuration B uses only one trap piece: the monkey must push the food away from the trap so that it falls down the unblocked channel and out of the bottom exit. Configuration C uses one shelf piece and one side blocker piece: the monkey must push the food away from the shelf and side blocker pieces so that it falls down the vertical channel and out of the bottom exit. Configuration D uses one shelf piece and both bottom blockers: the monkeys must push the food along the shelf piece to the side exit. A remedial training configuration was used if a subject pushed the food reward in only one direction for 30 consecutive trials to give the monkey experience of retrieving the reward from the alternative side of the box. Configurations A and B were used in phase I, and configurations C and D were used in phase II. However, as none of the monkeys passed phase I, none continued to phase II; therefore, configurations C and D will not be discussed any further.

Figure 3

Box Configurations
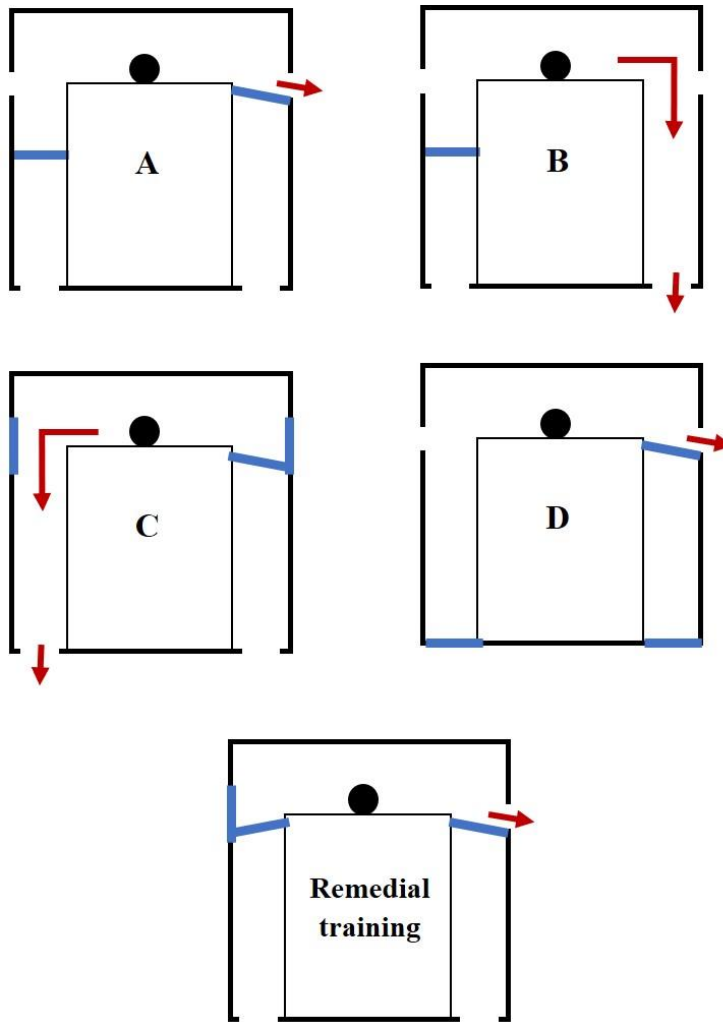
Note. Each box configuration is labelled by the letter in its center. Arrows depict the correct exit locations for each configuration. For each configuration, the removable parts can be set up on the opposite side of the box so that the reward must be pushed the alternative direction. The remedial training box was used to attempt to correct any side bias that the subjects developed.

Procedure

Subjects were presented with up to 10 sessions of 10 trials each to learn to solve the two-trap-box. To account for individual differences in the monkeys' attention spans, subjects received between five and twenty trials per day, over two testing sessions. This was given as either two sessions of five trials (a block of 10 per day) or as two sessions of 10 trials (two blocks of 10 per day).

In phase I, four squirrel monkeys (Pelusa and Amarilla from the East group; Sancha and Toomi from the West group) and seven capuchin monkeys (Pixie, Hazel, Bear, and Torres from the West group; Flojo, Reuben, and Chico from the East group) were presented with configuration A, whilst four squirrel monkeys (Roca, Gabriella, and Dora from the East group; Loki from the West group) and seven capuchin monkeys (Kato, Junon, Carlos, and Anita from the East group; Pedra, Inti, and Ximo from the West group) were presented with configuration B. Four additional squirrel monkeys were presented with a modified version of configurations A and B, where the blue blocker was placed at the bottom of the

channel so that they could gather tactile information about the trap (A: Flora and Valencia; B: Cali and Ciara).

In all cases, the trap appeared five times on the left and five times on the right in a random sequence (with the constraint that the trap was presented in the same orientation no more than twice in a row). In an attempt to prevent habit formation, a criterion for passing was used. The criterion was set at the same level as that used for the rooks in Seed et al. (2006) and the chimpanzees in Seed, Call et al. (2009): monkeys were deemed to have solved the problem if they made 16 or more correct responses over two consecutive blocks of 10 trials. This is significant according to a binomial test with alpha set at .05. If, after 30 trials, a subject had a persistent side bias (100% response to one side), they were presented with the remedial training box (Figure 3). If, after 30 trials on the remedial configuration, their side bias persisted, testing for this subject was ended. Similarly, if a subject had not reached the threshold after 100 trials on either configuration A or B in phase I, testing for this subject was ended.

At the beginning of a session, the monkey voluntarily entered the testing cubicle and was isolated in one of the testing cubicles by the experimenter. Then, at the beginning of each trial, the experimenter gained the attention of the subject and ensured they watched while the pieces were inserted. To prevent the monkeys from using the position of the experimenter's hand as a cue for which way to move the reward, piece insertion always involved activity on both sides of the box. For configurations with just one piece, at the same time as inserting the piece, the experimenter held their hand up on the opposite side of the box, parallel to where the piece was being inserted. For configurations with two pieces, the pieces were inserted at the same time. The food reward was then inserted through a channel in the back of the box, which placed it in the center of the shelf. On some occasions, the monkey required some encouragement to manipulate the reward; this was achieved by tapping the cubicle window directly above the center of the shelf to draw the subject's attention back to the reward. The experimenter then took a step back while the monkey worked on the reward (see the supplementary videos of each species working on the two-trap-box; capuchin monkey: Video S2; squirrel monkey: Video S3). To prevent any cueing from the experimenter and to minimize distractions, the monkey was unable to see the experimenter through the box and vice versa. Therefore, the experimenter watched the monkey on a video camera set up to record trials. This was located either to the side of the testing cubicle or on the ceiling at the back of the cubicle, looking forward over the monkey's shoulder. In every condition, moving the food one way resulted in successful retrieval, and movement in the opposite direction resulted in the food becoming trapped. If subjects were incorrect, the reward was removed through the back of the box and discarded into a separate food pot by the experimenter.

Coding

Trials were live coded by the experimenter and scored as correct (1) if the subject successfully retrieved the food reward and incorrect (0) if the food became trapped. A second coder scored 25% of all trials from the recorded video material to establish inter-observer reliability. Fleiss' kappa was calculated and, according to Landis and Koch (1977), inter-observer reliability showed "almost perfect agreement" (correct choice: $K = .98, p < .001$).

Analyses

All analyses were carried out in R studio (RStudio Team, 2019) using R (version 3.6.2; R Core Team, 2019). We looked at the proportion of trials in which individuals pushed the reward in the correct direction. To analyze this, we performed one-sample *t*-tests to look at the average score compared to chance for each species. Additionally, we ran an independent-samples *t*-test to compare the performance of squirrel monkeys who were able to gather both tactile and visual information about the trap with those who had received only visual information. Following this, to look for effects of box configuration and species on performance, we conducted a generalized linear mixed model (GLMM; Baayen, 2008) using the glmer function of the R package lme4 (Bates et al., 2015). The GLMM had a binomial error structure

and logit link function (McCullagh & Nelder, 1989) with the trial score (1 or 0) as the dependent variable (DV) and the test predictor variables box configuration, species, and trial number (*z*-transformed to a mean of zero and a standard deviation of one). We included monkey ID as a random effect and all possible random slope components (Barr et al., 2013). As an overall test of the effect of the predictor variables, we compared the full model with a null model lacking the test predictors but comprising the same random effect structure as the full model (Forstmeier & Schielzeth, 2011), using a likelihood ratio test (Dobson, 2002). *P*-values for the individual effects were based on likelihood ratio tests comparing the full model with the respective reduced models (Barr et al., 2013) using R function drop1 with argument 'test' set to "Chisq". Using the R package car (Fox & Weisberg, 2011), we calculated variance inflation factors (VIF) for the standard linear model, excluding the random effects, and confirmed that there was no collinearity between predictors (for all test predictors, VIF = 1; Field, 2005).

Results

As a group, the squirrel monkeys did not avoid the trap at a rate significantly different to chance (one-sample *t*-test: $t(11) = -0.28$, $p = .79$, 95% CI [0.47, 0.52]), and there was no significant difference in performance between individuals that were able to gather tactile and visual information and those that were able to gather only visual information (independent-samples *t*-test: $t(6) = 1.45$, $p = .20$, 95% CI [-0.03, 0.10]). As a group, the capuchin monkeys pushed the reward in the correct direction significantly more often than chance (one-sample *t*-tests: $t(13) = 2.65$, $p = .02$, 95% CI [0.50, 0.55]). However, when looking at individual performance, none of the monkeys met the criteria for successfully passing their first configuration of traps. Before reaching 10 sessions, two of the squirrel monkeys and five of the capuchin monkeys developed a side bias that was not corrected by presenting them with the remedial training box and so were dropped from testing (squirrel monkeys: Toomi and Sancha; capuchin monkeys: Carlos, Flojo, Hazel, Kato, and Reuben), while the remaining 10 squirrel monkeys and nine capuchin monkeys never reached the criteria for passing, even after 100 trials. See Table S1 for the raw data and videos of each species working on the two-trap-box (capuchin monkey: Video S2; squirrel monkey: Video S3).

A GLMM with score as the DV and the test predictor variables, species, box configuration, and trial number was not significant when compared to a null model lacking these predictors (LRT: $\chi^2 = 2.45$, $df = 3$, $p = .49$). There was no effect of either species or box configuration on the likelihood of avoiding the trap.

Discussion

Monkeys of both species seemed unable to learn to avoid the traps. For the squirrel monkeys, this was not influenced by being able to gather additional tactile information about the trap piece. Although, as a group, the capuchin monkeys performed above chance, all 14 individuals failed to avoid the traps. The result was surprising, as there is strong evidence that capuchin monkeys can learn functional discriminations (Evans & Westergaard, 2004; Fujita et al., 2003, 2011; Judge & Bruno, 2012; Manrique et al., 2011; Painter et al., 2019; Sabbatini et al., 2014; Schrauf et al., 2008; Yocom & Boysen, 2010), with some evidence of this in squirrel monkeys (Painter et al., 2019). One explanation for their failure is the planning involved in the task. The trap-box requires an individual to first visually locate the position of the trap and then to hold this information in mind whilst manipulating a reward through the box. This requires an additional cognitive ability, as the individual must plan their action ahead of time. However, in previous tasks, capuchins have been able to select the correct tool in advance of being given access to an apparatus (Manrique et al., 2011) and have even been shown to travel 8 m to a different room to collect the appropriate tool after viewing an apparatus (Evans & Westergaard, 2004; Judge & Bruno, 2012). The planning involved in the two-trap task is no more demanding than in these previous tasks, and thus, is unlikely to be the only factor limiting their ability.

Another possible explanation is that, although removing the tool component reduced the number of relations involved in the task, it still required a fairly demanding motor skill. From observations, it

appeared that manipulating the reward in a controlled manner was unexpectedly difficult for the monkeys (see video of each species working on the box; capuchin monkey: Video S2; squirrel monkey: Video S3). Therefore, one potential reason for their failure could be that, rather than attending to the position of the blue trap piece (which could be used as a perceptual cue), the monkeys were concentrating on trying to exhibit demanding motor coordination. The blue trap piece may not have been a salient enough cue for the monkeys to attend to and so they were unable to use the position of this trap piece to solve the box. This suggestion is investigated in Experiment 2.

Experiment 2—Blocked Cups

In Experiment 1, all 26 monkeys of both species failed to learn even the simplest configuration of the two-trap-box, which could be solved via simple discrimination, and learning the associative rule: "avoid the blue piece." The failure of the monkeys to learn this discrimination was surprising as, as discussed above, both species of monkey have passed other discrimination tasks by forming similar associative rules. In Experiment 2, we aimed to investigate whether the failure of the monkeys to learn this discrimination was due to the cue they were required to attend to within the box (namely the blue trap piece). We presented the monkeys with the same cue they learned to avoid in Experiment 1 in a more simplistic and familiar context where subjects were required to choose between two cups, both containing food, one of which was blocked by a trap (the same trap piece used in Experiment 1). If the monkeys were unable to use the trap piece as a cue for which cup to avoid, then this would likely be the reason for their failure in Experiment 1. However, due to previous work where both species learned to solve simple discrimination tasks, we expected that the familiar set-up of Experiment 2 would enable them to successfully learn to avoid the trap.

Another potential reason for the monkeys' failure in Experiment 1 could be the distance between the reward and the relevant discriminatory feature of the task, which has been shown to affect (social) learning in capuchins (Wood & Whiten, 2017). In the trap box, the monkeys were required to look ahead at the pathway the reward would take after being moved in either direction and then plan their motor actions accordingly. The set-up in Experiment 2 addressed this by bringing the trap closer to the location where the monkeys made their choices (lessening the attentional demands) and removed the intricate motor actions required to move the reward. Instead, the monkeys could see the relation between the reward and the trap before choosing and then were simply required to point to their preferred choice; a motor action with which all individuals were experienced.

In this experiment, we hoped to explore whether the blue trap piece was a salient enough feature for the monkeys to learn to avoid in a simplified context.

Method

Subjects

Tables 1 and 2 detail the individuals of each species participating in each experiment. For Experiment 2, in July 2016 and February–April 2018, we tested all nine of the squirrel monkeys from the East group that had participated in Experiment 1, as well as two additional monkeys from this group. At these times, we also tested two of the squirrel monkeys from the West that had participated in Experiment 1, as the third had been moved to another collection. Between November 2017 and March 2018, we tested all 14 capuchin monkeys that participated in Experiment 1.

Apparatus

The monkeys had considerable experience of reaching out through the cubicle window to one side or another to indicate a choice; thus, this experiment made use of this action. Two transparent cups with which the monkeys were familiar (from a previous discrimination task) were used, and the monkeys

had to learn to avoid a blue blocker piece from the trap-box, which blocked them from retrieving a reward from the cup.

Procedure

Subjects were given up to 10 sessions of 10 trials each to learn to solve the task. As in Experiment 1, the subjects received between five and twenty trials per day over two testing sessions: either two sessions of five trials (a block of 10 per day), or as two sessions of 10 trials (two blocks of ten per day). As before, subjects were deemed to have solved the task if they made 16/20 correct responses over two consecutive blocks, which is significant according to a binomial test with alpha set at .05.

At the beginning of a session, a monkey voluntarily entered the testing cubicle and was isolated in one of the testing cubicles by the experimenter. At the beginning of each trial, the experimenter turned on the camera (positioned centrally at the back of a table) and gained the subject's attention to ensure they were watching as the cups were set up. The blue blocker piece was held up for the monkey to see and then placed on one side of the table just out of their reach. The position of the blue blocker piece followed the same random sequence as in the trap-box trials, appearing five times on the left and five times on the right in a random sequence (with the constraint that the trap was not presented in the same orientation more than twice in a row). Two food rewards (raisins) were simultaneously placed at each side of the table, in line with the two holes in the cubicle window. One reward was placed on top of the blue blocker piece, and one onto the table itself. Next, two transparent cups were simultaneously placed over the rewards. This meant the table now had a cup on either side, both containing a reward; however, one cup was sat directly on the table, while the other was sat on top of a blue blocker piece (Figure 4). All participants had substantial experience with forced choice tasks that required them to reach out of one of the holes in the cubicle window to indicate a choice. Once the cups were in position, the table was pushed forward and the experimenter called "choosing" to signal to the monkey to make a choice. The monkey was then given 30 s to reach out and make their choice, after which the experimenter lifted the chosen cup. In all trials, all monkeys made a choice within this 30 s window. If the monkey chose the cup directly on the table, the experimenter lifted the cup and passed the reward to the monkey. If the monkey chose the cup positioned on top of the blue blocker, the experimenter lifted both the cup and the blocker piece together, such that the reward was trapped inside, and shook the cup so that the monkey saw that the raisin was trapped. The cup was then held up so only the blue blocker was visible to the monkey. Then, as the monkey watched, the correct cup was lifted, and the experimenter took the reward and discarded it into a pot located at the back of the table. The table was then moved back and reset for the next trial (See Video S4 of the procedure).

Coding

Trials were live coded by the experimenter and scored as correct (1) if the subject chose the open cup without the blue blocker piece beneath it and incorrect (0) if the subject chose the trapped cup with the blue blocker piece trapping the food reward. A second coder scored 25% of all trials from the recorded video material to establish inter-observer reliability. Fleiss' kappa was calculated and, according to Landis and Koch (1977), inter-observer reliability showed "almost perfect agreement" (correct choice: $K = .95$, $p < .001$).

Figure 4

Schematic of the Trap Cups Set-up
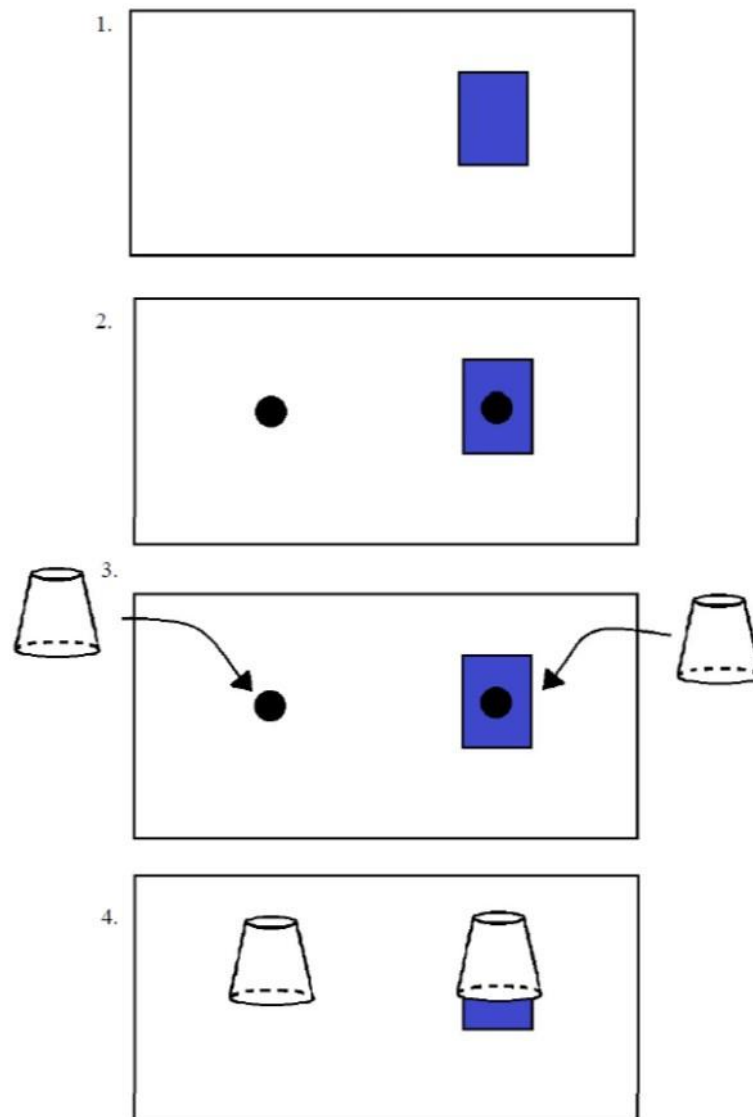
Note. The numbers signal the order of events: the blue trap piece is positioned randomly to the left or right, then the rewards are placed on the table before being covered by two transparent cups.

Results

As a group, both species of monkey managed to avoid the trap at a rate significantly above chance (one-sample *t*-test: squirrel monkeys: $t(12) = 6.40$, $p < .001$, 95% CI [0.59, 0.69]; capuchin monkeys: $t(13) = 8.04$, $p < .001$, 95% CI [0.62, 0.71]). When looking at individual performance, six of the 13 squirrel monkeys and 11 of the 14 capuchin monkeys reached the criteria for passing within 10 sessions, and only one monkey developed a side bias (Pedra—a capuchin monkey). The remaining seven squirrel monkeys and two capuchin monkeys never reached the criteria for passing, even after 100 trials (squirrel monkeys: Cali, Gabriella, Pelusa, Roca, Toomi, Valencia, and Lexi; capuchin monkeys: Junon and Pixie). The raw data can be found in Table S1).

Discussion

When the same discriminatory cue as used in Experiment 1 was presented to the monkeys in a more familiar and simplified context, both species of new world monkey learned to avoid the trap piece and successfully retrieve a reward. This result shows that the blue trap piece is a salient enough cue for the monkeys to learn to avoid and, therefore, not the reason for their failure in Experiment 1.

To solve this task, monkeys needed to learn to attend to the position of the blue trap piece and avoid this cup when making their choice. This simple discrimination task could be solved by learning the associative rule of "avoid the blue piece" without requiring any previous object knowledge. Had the monkeys learned this simple discrimination within the context of the trap-box (Experiment 1), it would have allowed for transfer tasks to be undertaken to disentangle the role of associative learning from the monkeys' pre-existing object knowledge (i.e., an understanding that rewards cannot pass through solid objects). Experiment 2 was specifically designed to rule out the possibility that the blue trap piece was not a salient enough cue for the monkeys, so a transfer task to disentangle associative learning and object knowledge is not possible within the setup. However, within this setup, the relationship between the cue (the trap piece) and the outcome (losing the reward) was functional as choosing the cup with the trap piece meant that the reward was trapped inside the cup when it was lifted and so was irretrievable. It is possible that the presence of this functional relationship played a role in the monkeys' ability to solve the task.

Research with apes has shown that problem-solving tasks based on functional relationships are easier to learn compared with those where the relationship is arbitrary (Albiach-Serrano & Call, 2014; Albiach-Serrano et al., 2012, 2015; Hanus & Call, 2011; Haun & Call, 2009; Mayer et al., 2014;). When presented with a pulling task, where participants were required to choose between an intact versus a broken paper strip, chimpanzees, orangutans, and 2-, 3-, and 4-year-old children were all able to learn to correctly select an intact paper strip (Albiach-Serrano et al., 2015). However, in the non-functional condition, when the paper strips were removed and lines were painted onto the table to look identical to the paper strips, only the 4-year-olds were able to pass. This suggests that the chimpanzees, orangutans, 2- and 3-year-old children were using something other than perceptual cues to solve the task. The presence of a functional relationship between the discriminatory cue and the outcome in the functional condition seems to have made the discrimination easier to learn (Albiach-Serrano et al., 2015). Similarly, when a modified trap-box, identical to the one used in Experiment 1, was presented to children, those as young as 2.5 years were able to pass the initial configuration (avoiding the trap and retrieving the reward). However, when the box was covered so that the traps could no longer be seen and instead lines were painted onto the box-cover to indicate the position of the traps, only children aged 4.5 years old or above were able to solve the task (Seed & Call, 2014).

In Experiment 3, we investigate the role of a functional relationship in discrimination learning by presenting the monkeys with a similar—but novel—task where the perceptual cue does not always have a functional relationship with the retrieval of rewards.

Experiment 3—Stuffed Cups

In Experiment 2, we showed that the monkeys could learn a simple discrimination in a task requiring them to use the blue trap piece as a cue for which cup to avoid. Although this could be solved by learning the associative rule of "avoid the blue piece," the relationship between the cue (the trap piece) and the outcome (losing the reward) was functional: if the cup with the trap piece was chosen, the reward was trapped inside the cup when it was lifted and so was irretrievable. It has been well documented that apes are better able to learn associations when the relationship between the cue and the outcome is functional compared to when it is arbitrary (Albiach-Serrano & Call, 2014; Albiach-Serrano et al., 2012, 2015; Hanus & Call, 2011; Haun & Call, 2009; Mayer et al., 2014). Only one of these studies included capuchin monkeys, but showed that the monkeys performed better when they needed to avoid a broken tool that was connected to a reward than they did when they simply needed to avoid pulling from the side

where the same broken tool was placed as a cue, but not attached to the reward (Mayer et al., 2014). These results suggest that, like apes, the capuchin monkeys are better able to learn a discrimination when the cue has a functional relationship to the retrieval of rewards compared to when the cue is purely symbolic.

In Experiment 3, we investigate whether the monkeys' ability to learn to avoid a cue could be aided by the functional relationship between the cue and the outcome. We present the monkeys with a new task involving the same relationship (namely that a reward cannot pass through a solid object: the 'blocker'). Additionally, we introduce a control task where the 'blocker' has an arbitrary (non-functional) relationship to the retrieval of food. As discussed above, if the functional relationship between the cue and the outcome is playing a role in learning then we expect this control task to be harder for the monkeys to learn as there is no functional reason to avoid the blocker other than it being 100% associated with the incorrect choice.

Method

Subjects

For Experiment 3, we tested a total of 17 squirrel monkeys and 20 capuchin monkeys from November–December 2019. Tables 1 and 2 detail which of the monkeys of each species from Experiments 1 and 2 were tested. We also tested eight squirrel monkeys and seven capuchin monkeys that had not had any experience with the trap cups of Experiment 2 (although three of these capuchin monkeys had experience of cups with lids in a previous study; Jordan, unpublished data).

Apparatus

Both conditions involved two transparent cups that both contained a raisin, with one of the cups also containing a balled up green paper towel (Figures 5 and 6). In the functional condition, the cups were presented upright with the raisins in the bottom and the paper towel in the top of one of the cups, blocking the raisin from being tipped out when turned upside down (Figure 5). In this condition, the cup without the paper towel was the correct choice, as it allowed the raisin to fall out when the cup was tipped, and the cup with the paper towel was the incorrect choice as it blocked the raisin from being tipped out. Presentation of the cups was as visually similar as possible in both conditions, so that the most salient difference between the two conditions was the functionality of the relationship between the green paper towel and the reward. To achieve this, the cups were presented upside down over the raisins, which were placed on the table, with the paper towel in the bottom of one of the cups (but appearing above the raisin, and thus not blocking it; Figure 6). In the non-functional condition—identical to the functional condition—the cup without the paper towel was the correct choice. However, unlike in the functional condition, the relationship between the paper towel and a correct choice was arbitrary. When presented to the monkeys, the two conditions looked almost identical, with the only difference being whether the opening of the cup was above or below the raisins; and therefore, whether the paper towel was or was not blocking the opening of the cup.

Procedure

Subjects were presented with three sessions of 10 trials each of a functional and non-functional condition; totaling six sessions. Half of the monkeys received three functional sessions followed by three non-functional sessions, with the other half receiving them in the opposite order. As in the previous experiments, monkeys received between five and twenty trials per day over two testing sessions. This was given as either two sessions of five trials (a block of 10 per day), or as two sets of 10 (two blocks of 10 per day). All monkeys participated in a total of 30 trials of each condition.
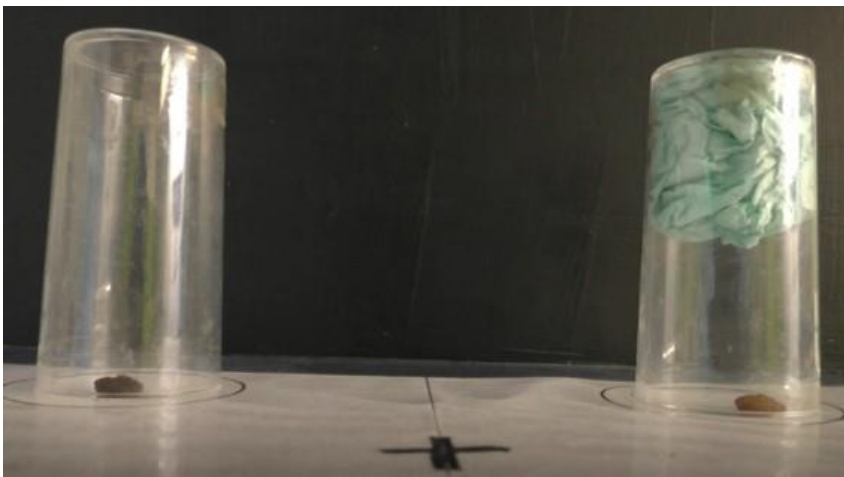
Figure 5

Photo of the Blocked Cups Set-Up for the Functional Condition

Note. The cups are positioned the right way up, with raisins in the bottom, and one of them with a green paper towel ball in the top which blocks the raisin from being tipped out.

Figure 6

Photo of the Blocked Cups Set-Up for the Non-Functional Condition

Note. The cups are positioned upside down over the raisins, with a green paper towel ball placed in the top of one of the cups to act as an arbitrary cue for which cup to avoid.

At the beginning of a session, the monkey voluntarily entered the testing cubicle and was isolated in one of the testing cubicles by the experimenter. At the start of each trial, a large occluder was placed in front of the cubicle to block the monkey's view whilst the cups were set up. The cups (containing the raisins and, in one cup, a paper towel ball) were then placed at equidistant points from the center of the table, in line with the holes in the cubicle windows (Figures 6 and 7 show the setup of the cups for each condition). The occluder was then raised, and the monkey's name was called as the experimenter moved the table forward so that the cups were within the monkey's reach (Figure 7 illustrates the basic procedure). Upon moving the table, the experimenter called "choosing" to signal to the monkey to make a

choice indicated by reaching out of one of the holes in the cubicle window (an action with which all participants had substantial experience). The monkey was then given 30 s to make their choice, after which the experimenter lifted the chosen cup. In all trials, all monkeys made a choice within this 30 second window. If the monkey chose the correct cup, in functional sessions, the reward was tipped out and given to the monkey; in non-functional sessions, removing the cup left the reward uncovered on the table so that the monkey could take it. If the monkey chose the incorrect cup, in functional sessions, the experimenter tipped the blocked cup and emphasized that the raisin was stuck inside; in non-functional sessions, as the experimenter lifted the cup, they also removed the raisin from the table. In non-functional sessions, the green paper towel ball had no functional relationship with retrieval of rewards and so the removal of the reward by the experimenter was an arbitrary, but consistent, outcome after lifting the cup (see Video S5 of the procedure following both a correct and incorrect choice in each condition). The cups were then moved to the back of the table and the occluder was replaced to start the next trial. The position of the green paper towel ball followed the same random sequence as in the previous experiments, appearing five times on the left and five times on the right in a random sequence (with the constraint that it was presented in the same orientation no more than twice in a row).

Coding

Trials were live coded by the experimenter and scored as correct (1) if the subject chose the cup without the green paper towel ball and incorrect (0) if the subject chose the cup containing the green paper towel ball. A second coder scored 25% of all trials from the recorded video material to establish inter-observer reliability. Fleiss' kappa was calculated and, according to Landis and Koch (1977), inter-observer reliability showed "almost perfect agreement" (correct choice: $K = .98$, $p < .001$).

Analyses

We performed one-sample *t*-tests to look at the average scores compared to chance for each species. To further analyze the data, we conducted multiple GLMMs (Baayen, 2008) using the R packages lme4 (Bates et al., 2015), and car (Fox & Weisberg, 2011). Prior to each model, trial number was *z*-transformed (to a mean of zero and a standard deviation of one) to make the estimates easier to interpret. For all models, as an overall test of the effect of the predictor variables, we compared the full model with a null model lacking the test predictors but comprising the same random effect structure as the full model (Forstmeier & Schielzeth, 2011), using a likelihood ratio test (Dobson, 2002). *P*-values for the individual effects were based on likelihood ratio tests comparing the full model with the respective reduced models (Barr et al., 2013) using the R function drop1 with argument 'test' set to "Chisq."

To look at any effect of an interaction between condition (functional vs. non-functional) and species on the monkeys' performance, we ran a GLMM with binomial error structure and logit link function (McCullagh & Nelder, 1989), with the DV set to trial score (1/0). The test predictor variables were condition order, trial number, experience (whether they had participated in Experiment 2), and an interaction between species and condition. The model also included the random effect of monkey ID and all possible random slope components (Barr et al., 2013). We calculated variance inflation factors (VIF) for the standard linear model excluding the random effects and confirmed that there was no collinearity between predictors (for all test predictors VIF = 1; Field, 2005).

To follow up on these results, we performed two further GLMMs to look at the effect of condition, experience, trial number, and condition order on each monkey species separately. First, we separated the data by species and then ran a separate GLMM for each dataset with binomial error structure and logit link function with score (1/0) as the DV and the test predictor variables trial number, experience, and an interaction between condition and condition order. The models also included the random effect of monkey ID, as well as all possible random slope components. The variance inflation factors (VIF) confirmed there was no collinearity between predictors (for both models, all test predictors VIF = 1; Field, 2005).

Figure 7

An Illustration of the Procedure for Presenting Monkeys with the Cups in Experiment 3
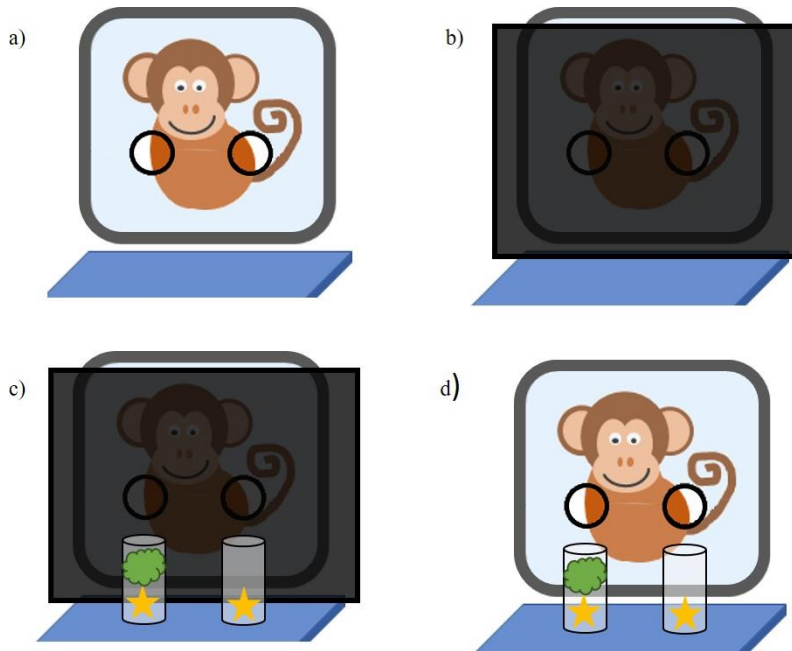
Note. The grey box represents the cubicles, with black circles showing the position of the holes in the cubicle window. The yellow star represents the rewards. a) a monkey is isolated in the research cubicle; b) a large occluder is placed across the window to block the monkeys' view of the table; c) the cups are positioned on the table; d) the occluder is removed and the monkey is given 30 s to make their choice.

Results

As a group, over all 60 trials, the capuchin monkeys avoided the green blocker at a rate significantly above chance (one-sample *t*-test: $t(19) = 6.08$, $p < .001$, 95% CI [0.60, 0.71]). In contrast, the squirrel monkeys did not choose the unblocked cup more often than expected by chance (one-sample *t*-test: $t(16) = -0.73$, $p = .47$, 95% CI [0.43, 0.54]). Looking at individual scores, in the functional condition, 12 out of the 20 capuchins monkeys and two out of 17 squirrel monkeys passed the task, scoring 16/20 over two consecutive sessions (significant according to a binomial test with alpha set at .05; capuchins: Alba, Bear, Carlos, Chico, Figo, Kato, Luna, Mekoe, Pedra, Pixie, Rufo, and Ximo; squirrel monkeys: Gisele and Pelusa). In contrast, in the non-functional condition, only one monkey of each species passed the task (capuchin monkey: Luna; squirrel monkey: Orla). The raw data table can be found in Table S1).

A GLMM with choosing the correct cup as the DV and the test predictor variables condition order, trial number, experience, and an interaction between species and condition was significant when compared to a null model lacking these predictors (LRT: $\chi^2 = 40.57$, $df = 6$, $p < .001$; see Table 3 for GLMM2 model output). There was a significant interaction between species and condition (LRT: $\chi^2 = 8.67$, $df = 1$, $p = .003$). Figure 8 illustrates these results.

Table 3

GLMM2: The Effect of an Interaction Between Species and Condition on Selecting the Correct Cup in Experiment 3

Term	Estimate	SE	X^2	df	p
(Intercept)	1.033	0.166			
Condition[a]	−0.705	0.136	−5.184	1	<.001
Species[b]	−1.088	0.216	−5.044	1	<.001
Experience[c]	0.379	0.183	2.076		.038
Condition order[d]	−0.208	0.153	−1.355	1	.176
Trial number	−0.058	0.045	−1.296	1	.195
Condition by Species Interaction	0.5686	0.193	3.042	1	.002

Note. Reference categories: [a]non-functional; [b]squirrel monkeys; [c]naïve; [d]non-functional – functional. The covariate trial number was *z*-transformed to a mean of zero and a standard deviation of one. The sample for GLMM2 consisted of 2,220 trials from 37 monkeys.

Figure 8

The Average Performance in the Functional Versus Non-Functional Conditions of Experiment 3 for Both Monkey Species
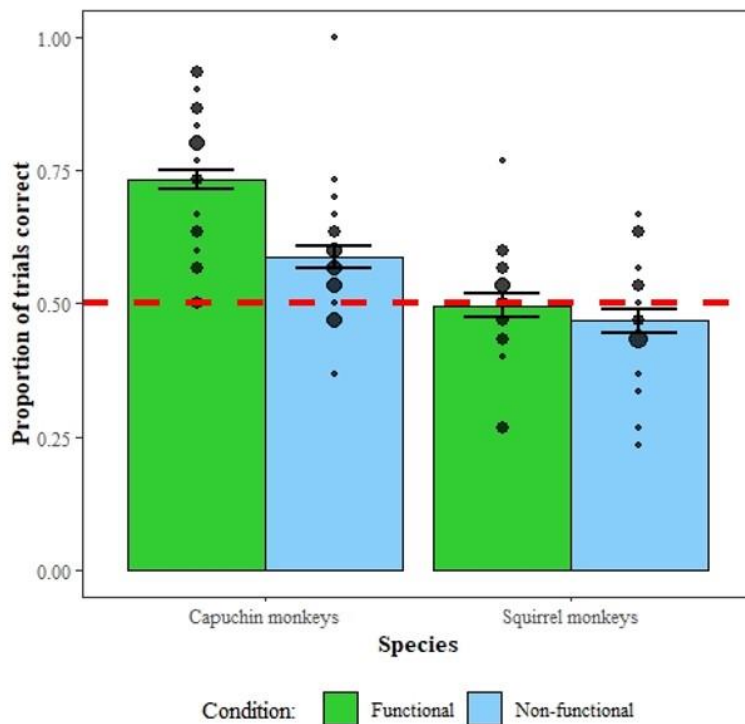
Note. The error bars show the mean ± standard error, with the red dashed line indicating chance. The points indicate the average score for individual participants, with the size of the point reflecting the number of individuals at any one point.

There was also a significant effect of experience (LRT: χ^2 = 4.00, *df* = 1, *p* = .046), with naïve monkeys performing better than experienced monkeys. There was no significant effect of trial number suggesting that monkeys were not learning during the test.

We then split the data by species and ran further GLMMs. For the squirrel monkeys, a GLMM with choosing the correct cup as the DV and the test predictor variables experience, trial number, and an interaction between condition and condition order, was significant when compared to a null model lacking these predictors (LRT: χ^2 = 12.29, *df* = 5, *p* = .031; see Table 4. for GLMM3 model output). There was a significant effect of condition order (z = -2.29, *df* = 1, *p* = .02) with monkeys that received the functional condition first performing better than those that received the non-functional condition first;

however, there was no significant effect of condition and the interaction between condition and condition order was not significant. This means that the monkeys that received the functional condition first scored highly in both conditions, and the monkeys that received the non-functional condition first scored significantly lower in both functional and non-functional conditions; suggesting a cohort effect.

Table 4

GLMM3: The Effect of Condition, Experience, and Condition Order on the Squirrel Monkeys Selecting the Correct Cup in Experiment 3

Term	Estimate	SE	X^2	df	p
(Intercept)	0.065	0.149			
Condition[a]	–0.106	0.174	–0.609	1	.542
Experience[b]	0.378	0.177	2.126	1	.034
Condition order[c]	-0.506	0.221	-2.288	1	.022
Trial number	-0.128	0.066	-1.931	1	.054
Interaction between condition and condition order	-0.034	0.256	-0.131	1	.900

Note. Reference categories: [a]non-functional; [b]naïve; [c]non-functional – functional. The covariate trial number was z-transformed to a mean of zero and a standard deviation of one. The sample for GLMM3 consisted of 1,020 trials from 17 monkeys.

For the squirrel monkeys, there was also a significant effect of experience (LRT: $\chi^2 = 4.24$, $df = 1$, $p = .039$) with naïve monkeys performing better then experienced monkeys. The predictor variables condition and trial number were not significant, suggesting that the monkeys were not learning the correct solution during the test.

For the capuchin monkeys, a GLMM with choosing the correct cup as the DV and the test predictor variables experience, trial number, and an interaction between condition and condition order was significant when compared to a null model lacking these predictors (LRT: $\chi^2 = 17.92$, $df = 5$, $p = .003$; see Table 5. for GLMM4 model output). The interaction between condition and condition order was not significant. However, although the capuchins performed above chance in both conditions (one sample t-tests: functional: $t(19) = 7.38$, $p < .001$, 95% CI [0.67, 0.80]; non-functional: $t(19) = 2.96$, $p = .008$, 95% CI [0.53, 0.65]), they performed significantly better in the functional condition ($z = -3.05$, $df = 1$, $p = .002$), suggesting that the presence of a functional relationship between the paper towel and an incorrect choice had a significant positive effect on their performance (Figure 8 illustrates these results). The predictor variables, trial number, condition order, and experience were not significant, suggesting that these did not have a significant effect on the monkeys' performance.

Table 5

Glmm4: The Effect of Condition, Experience, and Condition Order on the Capuchin Monkeys Selecting the Correct Cup in Experiment 3

Term	Estimate	SE	X^2	Df	p
(Intercept)	0.912	0.237			
Condition[a]	–0.731	0.240	–3.048	1	.002
Experience[b]	0.462	0.316	1.463	1	.143
Condition order[b]	0.023	0.316	0.074	1	.941
Trial number	0.010	0.063	0.166	1	.868
Interaction between condition and condition order	0.016	0.313	0.052	1	.959

Note. Reference categories: [a]non-functional; [b]naïve; [c]non-functional – functional. The covariate trial number was z-transformed to a mean of zero and a standard deviation of one. The sample for GLMM4 consisted of 1,200 trials from 20 monkeys.

Discussion

The squirrel monkeys did not avoid the blocked cup more often than expected by chance, and results from the GLMM suggest a cohort effect, with the squirrel monkeys that received the functional condition first scoring better overall. Naïve monkeys that had not participated in Experiment 2 also scored better; however, individual scores (Table 2) show that only one squirrel monkey scored above chance in the non-functional condition, and two monkeys scored above chance in the functional condition. These findings suggest that they may have learned to solve the task by learning a simple rule (e.g., "avoid the paper towel"), rather than by learning the functional properties of materials (i.e., that the reward could not pass through a solid object). Therefore, once the blue trap piece was changed to a green paper ball, they were unable to use anything learned during Experiment 2 to guide their decisions in the current task. In fact, it appears that having previous experience of a similar task where they had learned to "avoid the blue piece" actually made it more difficult to form a new associative rule when presented with a similar task involving different cues. However, due to the small sample size of naïve monkeys, these results should be interpreted with caution.

In contrast, the capuchin monkeys solved the task and chose the correct cup at a rate above chance in both the functional and non-functional conditions. Furthermore, their performance was significantly better in the functional condition than in the non-functional condition, suggesting that the presence of a functional relationship between the materials and the outcome helped the monkeys to solve the task. As pointed out by an anonymous reviewer, there is one confounding explanation for our results: the cue is made more obvious in the functional condition than it is in the non-functional condition, as the trapped reward is emphasized to the monkeys when the cup is tipped. This difference in emphasis may have caused the participants to pay more attention to it and hence led them to learn to avoid it faster. However, the same cue is present in both conditions and we found no effect of condition order on performance. If the monkeys had been solving the task because their attention was being disproportionately drawn to the cue in the functional sessions, we would have expected those that received the functional condition first to have transferred this learning into their non-functional sessions and continue to avoid the cue. Additionally, performance was not affected by trial number, suggesting that the monkeys were not learning the solution over the course of the test. Moreover, their performance was not significantly affected by experience, suggesting that experiencing the trap-cup procedure of Experiment 2 did not influence their ability to solve the task. Irrespective of their participation in Experiment 2, within just 30 trials of Experiment 3, the monkeys were avoiding the cup with the paper towel ball, and when this blocker was playing a functional role (in preventing them retrieving the reward) their ability to avoid it increased. Although these results support the suggestion that monkeys are faster at learning associations based on functional relationships than those based in arbitrary relationships, the reason for this remains unknown. Despite evidence that experience played no role in aiding the monkeys to solve this task, it remains likely that their previous experience made the functional condition easier to learn. One possibility is that, via their previous experiences, they have acquired knowledge about solidity that aided them in this task. Therefore, though we are unable to disentangle the role of core knowledge from the role of previous experience, we can conclude that the monkeys have some sensitivity to the property of solidity.

General Discussion

All 26 monkeys of both species failed to solve the initial configuration of the two-trap-box in Experiment 1, despite 11/14 capuchin monkeys and 8/13 squirrel monkeys successfully learning to avoid the same 'trap piece' within 10 sessions when discriminating between two cups in the simplified set-up of Experiment 2. In Experiment 3, we demonstrated that for the capuchin monkeys, learning to use the presence of a tissue ball as a cue for which cup to avoid was easier when the relationship between the object and the loss of a reward was functional rather than arbitrary (non-functional). The squirrel monkeys were unable to solve the task in Experiment 3 and performed at chance in both the functional

and non-functional conditions. These results suggest a possible species difference in understanding of object properties; the squirrel monkeys likely used perceptual or spatial rules to solve the trap cups task in Experiment 2, whilst— in agreement with results from previous work (Painter et al., 2019)—the capuchin monkeys did appear to have some understanding of the functional properties of the trap.

To solve the initial box configuration of Experiment 1, participants could have used the blue trap piece as a perceptual cue and simply learned the rule "avoid the blue piece." The performance of the monkeys in Experiment 1 mimics the results found in the original study where only one capuchin was able to pass the initial stage after more than 90 trials (Visalberghi & Limongelli, 1994). Even with the removal of the tool component of the task, the monkeys failed to avoid the trap, leaving our prediction that removing the tool would facilitate their performance unsupported. The failure of all the monkeys to solve this initial stage, despite it being solvable via this simple discriminatory rule, may have been due to the trap piece not being a salient enough cue. This concern was addressed in Experiment 2, where we confirmed that both species of monkey were able to learn to avoid the blue trap piece, therefore, indicating that their failure in Experiment 1 was more likely due to task constraints. Experimenter observations during testing revealed that this could be due to the motor demands of the task. Removing the tool component may have inadvertently introduced a different difficulty factor into the task: reward manipulation. Both species of monkey found it surprisingly difficult to master moving the reward along the shelf to make it fall out of the box (see videos; capuchin: Video S2; squirrel: Video S3). Mastering this motor action required a lot of concentration and appeared to be the most challenging part of the task. Previous work with capuchin monkeys has shown that, when presented with computer mazes where they must use a joystick to navigate a cursor through a maze and into a goal box, they first master the motor skill of moving the joystick before considering the relation between the cursor and the goal box (Leighty & Fragaszy, 2003). When this pattern of cognition is applied to the two-trap-box in Experiment 1, it highlights the possibility that the fine motor skills required to simply perform the task may have been too cognitively demanding for the monkeys. However, by the final trials, even those monkeys that had become relatively skilled at manipulating the reward (see videos; capuchin: Video S2; squirrel: Video S3) were still unable to solve the task, suggesting that the task may pose further difficulties.

Capuchin monkeys have been shown to preferentially rely on single relations and, even for tasks involving multiple relations, they adopt strategies that allow them to avoid considering multiple relations simultaneously (Fragaszy & Cummins-Sebree, 2005). In Experiment 1, removal of the tool component in the modified trap task lowers the number of relations that must be considered but does not turn the task into a single relation task; in order to move the reward in the right direction, the monkeys must still consider the relationship between themselves, the reward and the trap. Furthermore, the trap placed in the vertical channel involves multiple object-object relations: first the reward falls when unsupported by the shelf, and then it fails to pass through the barrier. When capuchin monkeys were required to anticipate the path of movement of a ball, they were unable to predict where the end point would be for a ball moving along a surface with a gap (Fragaszy & Cummings-Sebree, 2005). Similarly, when presented with raking tasks involving obstacles, capuchin monkeys consistently found a trap more difficult to avoid than a barrier (Fujita et al., 2011). This suggests that an obstacle that causes a change in the plane of movement may be more difficult to reason about than one that stops the movement of an object—and the modified trap-box involves both.

Additionally, it could be that capuchins find object-object relationships more difficult to consider than object-self relationships. This appears to be true of the one successful capuchin in the original trap-tube study, as she used a distance-based rule to pass the initial stage of the task; pushing the reward towards the end of the tube to which it was closest (and therefore closer to herself) rather than considering the relationship between the trap and the reward (Visalberghi & Limongelli, 1994). In the functional condition of Experiment 3, the monkeys may have been able to make a correct choice by considering the relationship between themselves and the paper towel ball (in that it was blocking them having access the reward) or, alternatively, the relationship between themselves and the reward (in that one option allowed them direct access to the reward if they could have reached into the cups). This would have enabled them to solve the task without having to consider the relationship between the reward and the object. In

contrast, in the non-functional condition, neither option provided them with direct access to the rewards and so, when considering the relationship between themselves and the reward, both options were equal (as both options blocked their direct access to the reward) and the monkeys struggled to consider the arbitrary and purely symbolic relationship between the paper towel and the reward. Capuchin monkeys' understanding of object-object and object-self relations is an area that could benefit from further investigation.

The capuchin monkeys but not the squirrel monkeys were able to solve the task in Experiment 3, a species difference not found in Experiment 2, where we provided more trials for learning. We tentatively suggest that the capuchins may have a more abstract understanding of solidity, whilst the squirrel monkeys likely solved Experiment 2 by forming an associative rule that could not be applied to the novel materials and so they were unable to pass the task. Finding that capuchins performed significantly better in the functional versus the non-functional condition demonstrates not only that they are readily able to pass a basic discrimination task between a rewarded versus an unrewarded cup, but that this learning is much faster if the relationship between the perceptual cue and the release of a reward is functional. This supports previous research suggesting that capuchin monkeys are capable of incorporating the functional role of an object into their learning and decision making (Manrique et al., 2011; Mayer et al., 2014). The squirrel monkeys were unable to pass the stuffed cups design of Experiment 3 in either the functional or non-functional conditions. Given more trials and further experience, we expect that the squirrel monkeys would have been able to pass, but this suggests that they would need to learn a new perceptual-based rule for novel materials, and we would not predict an effect of condition.

The significant difference in performance between the capuchin and squirrel monkeys is in line with the only other previous study that compares these two species in a similar task. When presented with a hook pulling task, capuchin monkeys—but not squirrel monkeys—were able to generalize to novel hook tools (Painter et al., 2019). This difference in performance could reflect a difference in causal understanding that is mirrored in their natural tool use behaviors. Capuchin monkeys are extractive foragers and some sub-species are natural tool users in the wild (e.g., Moura & Lee, 2004; Westergaard & Sumoi, 1995), whereas squirrel monkeys have never been documented using tools in the wild. It may be that there is a cognitive difference between the species that allows capuchins to understand causal relations and therefore facilitates their tool use.

In conclusion, this group of experiments adds to the current literature suggesting that capuchins—but not squirrel monkeys—are sensitive to the functional properties of objects, and specifically to solidity. Their failure on the modified trap task opens up new avenues for research, and further work should look at capuchin monkeys' ability to reason about changes in the axes of movement and investigate their understanding of multi-relational situations and object-self relationships in more detail, to further delineate the extents and limits of their problem-solving abilities and the possible differences with other primates such as great apes (who succeed in the two-trap-box task). Additionally, the literature would benefit from studies of object knowledge in a wider range of new world primates to investigate the phylogenetic history of the ability to form abstract concepts: for example, capuchin monkeys may represent a case of convergent evolution with great apes similar to that of corvids (Seed, Emery et al., 2009).

Acknowledgements

References

Albiach-Serrano, A., Bugnyar, T., & Call, J. (2012). Apes (*Gorilla gorilla, Pan paniscus, P. troglodytes, Pongo, abelii*) versus corvids (*Corvus corax, C. corone*) in a support task: The effect of pattern and functionality. *Journal of Comparative Psychology*, *126*(4), 355–367. http://doi.org/10.1037/a0028050

Albiach-Serrano, A., & Call, J. (2014). A reversed-reward continency task reveals causal knowledge in chimpanzees (*Pan troglodytes*). *Animal Cognition*, *17*(5), 1167–1176. http://doi.org/10.1007/s10071-014-0749-9

Albiach-Serrano, A., Sebastián-Enesco, C., Seed, A., Colmenares, F., & Call, J. (2015). Comparing humans and nonhuman great apes in the broken cloth problem: Is their knowledge causal or perceptual? *Journal of Experimental Child Psychology, 139*, 174–189. http://doi.org/10.1016/j.jecp.2015.06.004

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press. http://doi.org/10.1017/CBO9780511801686

Barr, J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. http://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. http://doi.org/10.18637/jss.v067.i01

Buckmaster, C. L., Hyde, S. A., Parker, K. J., & Lyons, D. M. (2015). Cup tool-use by squirrel monkeys. *American Journal of Primatology, 77*(12), 1323–1332. http://doi.org/10.1002/ajp.22486

Cummins-Sebree, S. E., & Fragaszy, D. M. (2005). Choosing and using tools: Capuchins (*Cebus apella*) use a different metric than tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*, *119*(2), 210–219. http://doi.org/10.1037/0735-7036.119.2.210

Dobson, A. (2002). *An introduction to generalized linear models*, 2nd ed. Chapman and Hall. http://doi.org/10.1201/9781420057683

Evans, T. A., & Westergaard, G. C. (2004). Discrimination of functionally appropriate and inappropriate throwing tools by captive tufted capuchins (*Cebus apella*). *Animal Cognition, 7*(4), 255–262. http://doi.org/10.1007/s10071-004-0220-4

Field, A. (2005). *Discovering statistics using SPSS*, 2nd ed. http://doi.org/10.1111/j.1365-2648.2007.04270_1.x

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology, 65*(1), 47–55. http://doi.org/10.1007/s00265-010-1038-5

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*, 2nd ed. Sage Publications.

Fragaszy, D. M., & Cummins-Sebree, S. E. (2005). Relational spatial reasoning by a nonhuman: The example of capuchin monkeys. *Behavioral and Cognitive Neuroscience Reviews*, *4*(4), 282–306. http://doi.org/10.1177/1534582306286573

Fujita, K., Kuroshima, H., & Asai, S. (2003). How do tufted capuchin monkeys (*Cebus apella*) understand causality involved in tool use? *Journal of Experimental Psychology: Animal Behavior Processes*, *29*(3), 233–242. http://doi.org/10.1037/0097-7403.29.3.233

Fujita, K., Sato, Y., & Kuroshima, H. (2011). Learning and generalization of tool use by tufted capuchin monkeys (*Cebus apella*) in tasks involving three factors: Reward, tool, and hindrance. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 10–19. http://doi.org/10.1037/a0020274

Gagne, M., Levesque, K., Nutile, L., & Locurto, C. (2012). Performance on patterned string problems by common marmosets (*Callithrix jacchus*). *Animal Cognition*, *15*(5), 1021–1030. http://doi.org/10.1007/s10071-012-0511-0

Hanus, D., & Call, J. (2011). Chimpanzee problem-solving: Contrasting the use of causal and arbitrary cues. *Animal Cognition*, *14*(6), 871–878. http://doi.org/10.1007/z10071-011-0421-6

Haun, D. B. M., & Call, J. (2009). Great apes' capacities to recognize relational similarity. *Cognition*, *110*(2), 147–159. http://doi.org/10.1016/j.cognition.2008.10.012

Hauser, M. D., Kralik, J., & Botto-Mahan, C. (1999). Problem solving and functional design features: Experiments on cotton-top tamarins, *Saguinus oedipus oedipus*. *Animal Behaviour*, *57*(3), 565–582. http://doi.org/10.1006/anbe.1998.1032

Hauser, M., Pearson, H., & Seelig, D. (2002). Ontogeny of tool use in cottontop tamarins, *Saguinus oedipus*: Innate recognition of functionally relevant features. *Animal Behaviour*, *64*(2), 299–311. http://doi.org/10.1006/anbe.2002.3068

Judge, P. G., & Bruno, S. (2012). Transport of functionally appropriate tools by capuchin monkeys (*Cebus apella*). *American Journal of Primatology*, *74*(3), 199–209. http://doi.org/10.1002/ajp.21987

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. http://doi.org/10.2307/2529310

Leighty, K. A., & Fragaszy, D. M. (2003). Joystick acquisition in tufted capuchins (*Cebus apella*). *Animal Cognition, 6*(3), 141–148. http://doi.org/10.1007/s10071-003-0176-9

Limongelli, L., Boysen, S. T., & Visalberghi, E. (1995). Comprehension of cause effect relations in a tool-using task by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 109*(1), 18–26. http://doi.org/10.1037/0735-7036.109.1.18

Manrique, H. M., Sabbatini, G., Call, J., & Visalberghi, E. (2011). Tool choice on the basis of rigidity in capuchin monkeys. *Animal Cognition, 14*(6), 775-786. http://doi.org/10.1007/s10071-011-0410-9

Mayer, C., Call, J., Albiach-Serrano, A., Visalberghi, E., Sabbatini, G., & Seed, A. (2014). Abstract knowledge in the broken-string problem: Evidence from nonhuman primates and pre-schoolers. *PloS One, 9*(10). e108597. http://doi.org/10.1371/journal.pone.0108597

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman and Hall. http://doi.org/10.1201/9780203753736

Moura, A. D. A., & Lee, P. C. (2004). Capuchin stone tool use in Caatinga dry Forest. *Science, 306*(5703), 1909–1909. http://doi.org/10.1126/science.1102558

Mulcahy, N. J., & Call, J. (2006). How great apes perform on a modified trap-tube task. *Animal Cognition, 9*(3), 193–199. http://doi.org/10.1007/s10071-006-0019-6

Painter, M. C., Russell, R. C., & Judge, P. G. (2019). Capuchins (*Sapajus apella*) and squirrel monkeys (*Saimiri sciureus*) fail to attend to the functional spatial relationship between a tool and a reward. *Journal of Comparative Psychology, 133*(4), 463–473. http://doi.org/10.1037/com0000179

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology, 58*, 97–118. http://doi.org/10.1146/annurev.psych.58.110405.085555

Povinelli, D. J. (2003). *Folk physics for apes: The chimpanzee's theory of how the world works.* Oxford University Press. http://doi.org/10.1093/acprof:oso/9780198572190.001.0001

R Core Team (2019). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

RStudio Team (2019). *RStudio: Integrated development for R* [Computer software]. RStudio, Inc. http://www.rstudio.com/

Sabbatini, G., Manrique, H. M., Trapanese, C., Vizioli, A. D. B., Call, J., & Visalberghi, E. (2014). Sequential use of rigid and pliable tools in tufted capuchin monkeys (*Sapajus* spp.). *Animal Behaviour, 87*, 213–220. http://doi.org/10.1016/j.anbehav.2013.10.033

Santos, L. R., Rosati, A., Sproul, C., Spaulding, B., & Hauser, M. D. (2005). Means-means-end tool choice in cotton-top tamarins (*Saguinus oedipus*): Finding the limits on primates' knowledge of tools. *Animal Cognition, 8*(4), 236–246. http://doi.org/10.1007/s10071-004-0246-7

Santos, L. R., Pearson, H. M., Spaepen, G. M., Tsao, F., & Hauser, M. D. (2006). Probing the limits of tool competence: Experiments with two non-tool-using species (*Cercopithecus aethiops* and *Saguinus oedipus*). *Animal Cognition, 9*(2), 94–109. http://doi.org/10.1007/s10071-005-0001-8

Schrauf, C., Huber, L., & Visalberghi, E. (2008). Do capuchin monkeys use weight to select hammer tools? *Animal Cognition, 11*(3), 413–422. http://doi.org/10.1007/s10071-007-0131-2

Seed, A. M., & Call, J. (2014). Space or physics? Children use physical reasoning to solve the trap problem from 2.5 years of age. *Developmental Psychology, 50*(7), 1951–1962. http://doi.org/10.1037/a0036695

Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology: Animal Behavior Processes, 35*(1), 23–34. http://doi.org/10.1037/a0012925

Seed, A., Emery, N., & Clayton, N. (2009). Intelligence in corvids and apes: A case of convergent evolution? *Ethology, 115*(5), 401–420. http://doi.org/10.1111/j.1439-0310.2009.01644.x

Seed, A., & Mayer, C. (2017). *Problem solving*. In J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, & T. Zentall (Eds.), *APA handbook of comparative psychology: Perception, learning, and cognition* (pp. 601–625). American Psychological Association. http://doi.org/10.1037/0000012-027

Seed, A., Seddon, E., Greene, B., & Call, J. (2012). Chimpanzee 'folk physics': Bringing failures into focus. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1603), 2743–2752. http://doi.org/10.1098/rstb.2012.0222

Seed, A. M., Tebbich, S., Emery, N. J., & Clayton, N. S. (2006). Investigating physical cognition in rooks, *Corvus frugilegus*. *Current Biology, 16*(7), 697–701. http://doi.org/10.1016/j.cub.2006.02.066

Silva, F. J., Page, D. M., & Silva, K. M. (2005). Methodological-conceptual problems in the study of chimpanzees' folk physics: how studies with adult humans can help. *Learning & Behavior. 33*(1), 47-58. http://doi.org/10.3758/bf03196049

Spaulding, B., & Hauser, M. (2005). What experience is required for acquiring tool competence? Experiments with two callitrichids. *Animal Behaviour*, *70*(3), 517–526. http://doi.org/10.1016/j.anbehav.2004.11.024

Taylor, A. H., Hunt, G. R., Medina, F. S., & Gray, R. D. (2009). Do New Caledonian crows solve physical problems through causal reasoning? *Proceedings of the Royal Society B: Biological Sciences, 276*(1655), 247–254. http://doi.org/10.1098/rspb.2008.1107

Tebbich, S., Seed, A. M., Emery, N. J., & Clayton, N. S. (2007). Non-tool-using rooks, *Corvus frugilegus*, solve the trap-tube problem. *Animal Cognition*, *10*(2), 225–231. http://doi.org/10.1007/s10071-006-0061-4

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause effect relations in tool-using capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, *108*(1), 15–22. http://doi.org/10.1037/0735-7036.108.1.15

Visalberghi, E., & Neel, C. (2003). Tufted capuchins (*Cebus apella*) use weight and sound to choose between full and empty nuts. *Ecological Psychology*, *15*(3), 215–228. http://doi.org/10.1207/S15326969ECO1503_2

Westergaard, G. C., & Suomi, S. J. (1995). The stone tools of capuchins (*Cebus apella*). *International Journal of Primatology*, *16*(6), 1017–1024. http://doi.org/10.1007/BF02696114

Wood, L. A., & Whiten, A. (2017). Visible spatial contiguity of social information and reward affects social learning in brown capuchins (*Sapajus apella*) and children (*Homo sapiens*). *Journal of Comparative Psychology, 131*(4), 304-316. http://doi.org/10.1037/com0000076

Yamazaki, Y., Iriki, A., & Watanabe, S. (2011). Modulation of physical understanding by common marmosets (*Callithrix jacchus*). *Animal cognition*, *14*(2), 175–186. http://doi.org/10.1007/s10071-010-0351-8

Yocom, A. M., & Boysen, S. T. (2010). Capuchins (*Cebus apella*) can solve a means-end problem. *Journal of Comparative Psychology*, *124*(3), 271–277. http://doi.org/10.1037/a0019369