Factors Influencing Species Performance in a Cross-Taxon Comparative Research Program

Laurent Prétôt
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/psych_diss

Recommended Citation
doi: https://doi.org/10.57709/10463896

This Dissertation is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
FACTORS INFLUENCING SPECIES PERFORMANCE IN A CROSS-TAXON
COMPARATIVE RESEARCH PROGRAM

by

LAURENT PRETOT

Under the Direction of Sarah F. Brosnan, PhD

ABSTRACT

Species vary in the ease with which they can solve apparently similar problems. For instance, problems will be interpreted differently by different species due to differences such as how they process the world or their ecology. The latter is the focus of the ecological approach to cognition, which posits that ecology influences decision-making such that each species performs better on tasks that are naturally relevant to them. In a previous work, my colleagues and I compared the performance of cleaner fish and nonhuman primates, which differ substantially in brain size and ecology, on a dichotomous choice task derived from the cleaner fish ecology. In the task, subjects chose between two different plates, each containing a same food; if they chose the food from plate A, they could then choose the food from plate B as well, whereas if they
chose B first, A was no longer available. Fish were better than primates at solving this task, emphasizing the role of ecology in shaping one species’ decision-making. For my dissertation project, I explored possible explanations for the primates’ poor performance in the task. In a first series of studies, I investigated the possibility that species differed in the task because of differences in the capacity to recognize the relevant cues (i.e. the plate design), to solve it. I thus repeated the task with fish and nonhuman primates, using variations designed to be more salient to primates. In a first experiment, the foods were different colors, whereas in a second experiment, they were hidden to avoid the prepotent response. In a second series of studies, I tested monkeys in a computerized paradigm that differed from the plate task by removing interaction with the human experimenter, which may be distracting, and providing a more standardized testing environment. Finally, in a last study, I investigated one possible cognitive limitation to the primates in the plate task, the failure to use backwards induction to solve it. These studies allowed me to investigate the role of ecology in species’ decision-making, a perspective that is often neglected in studies of cognition.

INDEX WORDS: Ecology, Cognition, Decision-making, Dichotomous choice task, Fish, Primates
FACTORS INFLUENCING SPECIES PERFORMANCE IN A CROSS-TAXON COMPARATIVE RESEARCH PROGRAM

by

LAURENT PRETOT

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy
in the College of Arts and Sciences
Georgia State University
2017
FACTORS INFLUENCING SPECIES PERFORMANCE IN A CROSS-TAXON COMPARATIVE RESEARCH PROGRAM

by

LAURENT PRETOT

Committee Chair: Sarah F. Brosnan

Committee: David A. Washburn

Michael J. Beran

Heather Kleider-Offutt

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2017
DEDICATION

I dedicate my dissertation to my queen, Macyory, and my parents, Pascale and Bernard, who have always believed in me, and supported me continuously throughout my PhD program (and beyond!).

Merci! Gracias!
ACKNOWLEDGEMENTS

This work would not have been possible without the advice and support of many people. First, I would like to express my deep gratitude to my supervisors, Dr. Sarah F. Brosnan and Dr. Redouan Bshary, for their guidance, enthusiastic encouragement, and useful critiques of this research work and throughout my programs at Georgia State University and at the University of Neuchâtel. Also, thank you to my committee members for their constructive feedback and suggestions on this project. Finally, I want to thank the Cebus Lab team and Language Research Center staff for the monkey studies, the primate research department at Zoo Atlanta for the ape and monkey studies, and the University of California Berkeley Gump Field Station for the field support on fish.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. v

LIST OF TABLES ........................................................................................................................... x

LIST OF FIGURES ........................................................................................................................ xi

1 INTRODUCTION ............................................................................................................................ 1

1.1 Literature Review ....................................................................................................................... 3

1.1.1 Cognitive and Ecological Approaches to Cognition ............................................................ 3

1.1.2 Cross-Species Comparison Techniques ................................................................................. 5

1.1.3 Comparative Studies: The Fish Model .................................................................................... 8

1.1.4 Testing the Ecological Hypothesis: The Fish-Primate Model ................................................ 11

1.1.5 Potential Factors Influencing Species Performance in the Plate Task ............................. 13

1.2 Overview of this Dissertation ................................................................................................... 15

2 STUDY 1: FACTORS INFLUENCING THE DIFFERENT PERFORMANCE
OF FISH AND PRIMATES ON A DICHOTOMOUS CHOICE TASK ........................................ 20

2.1 Introduction .............................................................................................................................. 20

2.2 Methods ................................................................................................................................... 25

2.2.1 Subjects and Housing ........................................................................................................ 25

2.2.2 General Procedure ............................................................................................................. 27

2.3 Study 1: Plate Task in Fish ...................................................................................................... 32

2.4 Study 2: Color Task ................................................................................................................. 37
2.5 Study 3: Cup Task ................................................................. 40
2.6 General Discussion ............................................................... 44
2.7 Tables ................................................................................ 49
2.8 Figures .............................................................................. 52

3 STUDY 2: VARIATION IN PERFORMANCE OF ORANGUTANS,
GORILLAS AND DRILLS IN TWO VERSIONS OF A CHOICE TASK DERIVED
FROM THE CLEANER FISH ECOLOGY ............................................. 55
3.1 Introduction ........................................................................ 55
3.2 Methods ............................................................................. 59
  3.2.1 Subjects and Housing ....................................................... 59
  3.2.2 General Procedure ......................................................... 61
3.3 Study 1: Plate Task in Gorillas and Drills ............................... 64
3.4 Study 2: Color Task ............................................................. 66
3.5 Study 3: Cup Task ............................................................... 69
3.6 General Discussion ............................................................... 72
3.7 Tables ................................................................................. 75
3.8 Figures .............................................................................. 79

4 STUDY 3: COMPARING SPECIES DECISIONS IN A DICHOTOMOUS
CHOICE TASK: ADJUSTING TASK PARAMETERS IMPROVES PERFORMANCE IN
MONKEYS ..................................................................................... 81
4.1 Introduction ........................................................................ 81
4.2 Methods.............................................................................................................. 86

4.2.1 Subjects and Housing....................................................................................... 86

4.2.2 Apparatus.......................................................................................................... 87

4.2.3 General Procedure............................................................................................ 88

4.3 Study 1.................................................................................................................. 91

4.3.1 Procedure........................................................................................................... 91

4.3.2 Results................................................................................................................. 93

4.3.3 Study 1 Discussion............................................................................................ 97

4.4 Study 2.................................................................................................................. 99

4.4.1 Procedure........................................................................................................... 100

4.4.2 Results................................................................................................................. 101

4.4.3 Study 2 Discussion............................................................................................ 106

4.5 General Discussion............................................................................................... 106

4.6 Tables.................................................................................................................. 111

4.7 Figures.................................................................................................................. 115

5 STUDY 4: MONKEYS SHOW BACKWARDS INDUCTION IN A MANUAL

MAZE TASK............................................................................................................. 117

5.1 Introduction............................................................................................................ 117

5.2 Methods............................................................................................................... 124

5.2.1 Subjects and Housing....................................................................................... 124
5.2.2 General Procedure .................................................................................................................. 125

5.3 Results ........................................................................................................................................ 130

5.4 Discussion ..................................................................................................................................... 132

5.5 Tables .......................................................................................................................................... 135

5.6 Figures .......................................................................................................................................... 139

6 CONCLUSIONS ............................................................................................................................ 141

6.1 Summary of Results ....................................................................................................................... 141

6.2 Major Take-Home Messages ......................................................................................................... 143

   6.2.1 Importance of Cues .................................................................................................................. 143

   6.2.2 Importance of Experience ....................................................................................................... 144

   6.2.3 Importance of Designing Careful Procedures ......................................................................... 145

   6.2.4 Unexpectedly Good Performance of Fish .............................................................................. 148

6.3 Implications ................................................................................................................................... 149

   6.3.1 Scientific Contribution ............................................................................................................ 149

   6.3.2 Interface Between Ecology and Cognition ............................................................................. 152

6.4 Future Directions .......................................................................................................................... 159

   6.4.1 Refining Methods to Test the Ecological Hypothesis ............................................................ 159

   6.4.2 Children’s Cooperative Decision-Making in Socially Valid Context .................................. 161

REFERENCES ...................................................................................................................................... 163
LIST OF TABLES

Table 1 Two-Choice Moorean Fish Plate Task................................................................. 49
Table 2 Two-Choice Monkey-Fish Cup Task.................................................................. 50
Table 3 Two-Choice Monkey-Fish Cup Task.................................................................. 51
Table 4 Orangutan-Gorilla-Drill Subjects General Information...................................... 75
Table 5 Two-Choice Gorilla-Drill Plate Task.................................................................. 76
Table 6 Two-Choice Orangutan-Gorilla-Drill Color Task .............................................. 77
Table 7 Two-Choice Orangutan-Gorilla-Drill Cup Task.................................................. 78
Table 8 Computer Tasks .................................................................................................. 111
Table 9 Capuchin Computer Study 1 ............................................................................. 112
Table 10 Macaque Computer Study 1A........................................................................... 113
Table 11 Capuchin-Macaque Computer Study 2............................................................ 114
Table 12 Paddle-Box Task: Probability to Retrieve Food ............................................ 135
Table 13 Paddle-Box Task: Summary of the Trials......................................................... 136
Table 14 Paddle-Box Task: Trials to Reach Criterion.................................................... 137
Table 15 Paddle-Box Task: Correct Unfamiliar Trials................................................... 138
LIST OF FIGURES

**Figure 1** Monkey-Fish Experimental Set-Up................................................................. 52

**Figure 2** Monkey-Fish Sets of Stimuli................................................................. 53

**Figure 3** Monkey-Fish Initial Learning Phase .................................................. 54

**Figure 4** Orangutan-Gorilla-Drill Sets of Stimuli .......................................... 79

**Figure 5** Orangutan-Gorilla-Drill Initial Learning Phase ......................................... 80

**Figure 6** Capuchin-Macaque Initial Learning Phase .................................... 115

**Figure 7** Capuchin-Macaque Reversal Learning Phase ..................................... 116

**Figure 8** Paddle-Box Apparatus ................................................................................. 139

**Figure 9** Paddle-Box Task: Phase Results ............................................................. 140
1 INTRODUCTION

Species differ in numerous factors, but perhaps the most surprising for the traditional view of comparative cognition is that different species reactions to the same experimental task can differ substantially. Why this is the case is a major question that must be addressed in cognitive psychology. Generally, the focus has been on two obvious factors that influence decisions, differences in species’ ecologies and cognitive abilities. To investigate which factors influence an individual’s decisions and how much, in what circumstances and how they interact, psychologists have used two nonmutually exclusive approaches that focus on different aspects of a species’ behavior. On one hand, the cognitive approach attempts to explain decision-making based on brain size and relative intelligence, and predicts that species that evolved larger brains (including primates) should demonstrate advanced cognitive skills as compared to species that did not. On the other hand, the ecological approach explains decision-making based on one species’ ecology. According to this view, each species is predicted to have developed its own particular abilities and outperform others in tasks for which they possess specific adaptations (e.g., Balda & Kamil, 1989; Kamil, 1988, 1998; Kamil & Mauldin, 1987; Shettleworth, 2009).

In a prior study, Salwiczek, Prêtôt, Demarta, et al. (2012) compared the performance of three primate species (Cebus [Sapajus] apella, Pongo spp. and Pan troglodytes) with the bluestreak cleaner wrasse (Labroides dimidiatus) on a dichotomous choice task derived from the natural cooperative interaction between the cleaner fish and its client fish species, who visit the cleaner’s territory to have their ectoparasites and dead or infected tissue removed. In the task, subjects chose between two different plates, each containing the same food and amount; if they chose the food from plate A, they could then choose the food from plate B as well, whereas if they chose the food from plate B first, plate A was no longer available. Therefore, the optimal
foraging strategy was to take the food from plate A first before the food from plate B. The fish were better than the primates at solving this task, emphasizing the role of ecology in shaping a species’ decision-making. However, to draw the general conclusion that it is the ecological relevance of a task that primarily influences species’ abilities to solve it, it was necessary to show that primates would improve, and match or outperform fish, in a similar task more relevant to primates, and also rule out possible alternative explanations for the primates’ poor performance.

To address the first issue, for this project, I investigated the possibility that species differed in the plate task because of differences in the capacity to recognize the relevant cues (i.e., the plate design), to solve it. To do this, I compared the performance of several nonhuman primate species with cleaner fish in versions of the plate task with cues more relevant to primates than fish. In a first experiment, the cues were changed from the design/color of the plate, which I hypothesized was more ecologically relevant to fish, to the color of the foods (the plates remained identical), which is presumably more ecologically relevant to primates. In a second experiment, the foods were hidden (by cups in primates and behind plates in fish), because primates are known to have difficulty making the correct choices when food is visible. Second, I tested monkeys in two computerized versions of the plate task. The first paradigm differed from the plate task by removing interaction with the human experimenter, which may be distracting, and providing a more standardized testing environment, whereas the second was an adaptation of the first paradigm designed to be more relevant to primate ecology. Finally, I investigated one possible cognitive limitation to the primates in the plate task, the failure to use backwards induction to solve it. In this study, I tested monkeys using the manual “paddle-box” task,
originally designed for testing planning skills in apes, which consisted of a box containing three levels of paddles that subjects rotated to drop a food into a tube that delivered it.

In this dissertation, I first provide a literature review, which goes over some major studies that have inspired and motivated my choices of dissertation projects. Then, I present each of these projects (two of which have been already published, one is currently in review, and one is still in preparation). Finally, I conclude with the overall implications and future directions of my research to date.

1.1 Literature Review

1.1.1 Cognitive and Ecological Approaches to Cognition

It is often assumed that measures of brain complexity, typically estimated based on brain size, reflect a species’ cognitive abilities, in particular the presence of more general cognitive abilities (e.g., Bitterman, 1965; Deaner, van Schaik, & Johnson, 2006; Deaner, Isler, Burkart, & van Schaik, 2007; Reader, Hager, & Laland, 2011). The cognitive approach proposes that species’ differences in cognitive capacities are due to relative distances in phylogeny. Consequently, taxa that evolved a large brain (including primates) should be more likely to demonstrate advanced cognitive skills as compared to species that did not. This approach supports the general-purpose intelligence hypothesis, which predicts that brain-sized differences in intelligence are unrelated to obvious features of ecology (Hare & Wrangham, 2002). Instead, abilities can be applied across contexts and have not been selected to solve any specific evolutionary problem. From a comparative perspective, a similar approach is the general process view, which holds that cognitive capacities involve processes that are widely distributed in
animal taxa and used for a wide range of problems (Bitterman, 2000; Macphail & Bolhuis, 2001).

Although brain complexity is clearly important, so too is the environment in which each species evolved. The ecological approach to cognition (also called the adaptive specialization approach, adapted cognition hypothesis or ecological intelligence hypothesis) recognizes this latter aspect and predicts that because cognitive abilities have been selected to solve evolutionary salient problems, a species’ most flexible cognitive abilities can be demonstrated in settings for which these abilities evolved (see Hare & Wrangham, 2002). Consequently, each species is considered to have developed its own particular abilities and outperform others in tasks for which they have specific adaptations (e.g., Balda & Kamil, 1989; Bshary, Salwiczek, & Wickler, 2007; Emery, 2006; Kamil, 1988; Kamil & Mauldin, 1987; Shettleworth, 2009). If “intelligence is seen as solving problems of ecological relevance in the environment in which the species evolved, then all species still extant are equally intelligent in their own ways and the question becomes what different species’ intelligence consist of.” (Shettleworth, 1998, p. 570).

A classic instance of the ecological approach is the comparative study of food-caching in birds. In early work, Kamil and Balda (1989) found that Clark’s nutcrackers (Nucifraga Columbiana) and pinyon jays (Gymnorhinus cyanocephalus), two species most dependent upon stored food for surviving winters, outperformed less cache-dependent scrub jays (Aphelocoma coerulescens) in a food-recovery experiment. Although these results could not be taken as proof of species differences in other aspects of cognition (e.g., spatial memory), they emphasized at least the importance of ecology in shaping specific aspects of species’ decision-making.

In a follow-up study, however, nutcrackers performed better than scrub jays in a spatial memory task, but not in a nonspatial, color memory task (Olson, Kamil, Balda, & Nims, 1995).
This finding strengthened Kamil and Balda (1989)’s suggestion that differences in performance on spatial tasks might reflect species’ dependence on food caching. Similar results were obtained for two populations of black-capped chickadees (Pravosudov & Clayton, 2002): Alaskan chickadees, which live in harsh environments and are highly dependent on food caching, performed better in spatial memory tests than Colorado chickadees, although the populations did not differ in a nonspatial version of the task.

Importantly, cognitive and ecological approaches are not contradictory. On the contrary, although they focus on different underlying causes of a species’ behavior, they presumably both influence behavior. As a result of the conflict between both views, Kamil (1988, 1998) proposed a synthetic multilevel approach to the study of animal cognition, one that integrates both cognitive and ecological questions and eliminates false dichotomies.

1.1.2 Cross-Species Comparison Techniques

Cross-species comparisons are helpful to explore questions about the adaptive function of cognitive abilities, the focus of the ecological approach (Kamil, 1988, 1998). Somewhat surprisingly, however, empirical evidence supporting this hypothesis is rare (Shettleworth, 2009); in particular, most of these studies have strictly limited the comparison to closely related species. One reason for this is that such studies often suffer from the so-called “learning-performance” problem (Bitterman, 1960, 1965): species’ differences in performance in a task cannot be explained exclusively by differences in cognitive ability per se, because of the influence of external factors associated with the procedure (see Balda & Kamil, 1989). For example, in their bird study on food-caching, Balda and Kamil (1989) pointed out that species’ differences in performance in the food-recovery task may not have been the exclusive result of
differences in cognitive abilities between species (i.e., spatial memory), because other variables, like site preferences, systematic movement patterns, differential use of space and other species-specific behavioral differences, might also have played a role in their behavior.

Different alternative strategies have been proposed to control for this problem. For example, Bitterman (1965) suggested a technique he called “control by systematic variation.” The idea was to develop a series of experiments that test each potential variable (called “contextual variables”; Macphail, 1985) in order to determine which one(s) may affect one species’ performance. This strategy presents, however, a major downside in that it is functionally impossible to test for all variables. An alternative technique to solve this problem is called “converging operations” (Kamil, 1988, 1998), in which species that are compared are tested with a battery of different experiments based on the same cognitive ability but across different settings. If the pattern of species differences holds across those tasks, this provides fairly convincing evidence that the difference is cognitive. For example, after finding differences in performance in the food-recovery task (Balda and Kamil, 1989), researchers tested the same species in a variety of spatial memory tasks (Kamil, Balda, & Olson, 1994; Olson et al., 1995). Species differences in performance remained similar across all tasks (but not in a nonspatial memory task; see Olson et al., 1995), indicating that these differences were likely due to difference in spatial memory.

Despite its potential, the technique of converging operations does not test for the hypothesis that the species differences observed are due to some general factor; that is, a species may outperform another not because both differ in specific cognitive abilities, but because one is simply more adapted to laboratory settings than the other. Consequently, an additional step, called “differentiation,” is required (see Kamil, 1998). This strategy involves designing an
experiment to be as similar as possible to those demonstrating species differences through converging operations, but testing a different cognitive ability. If there are differences in species’ performance in the new task, and these differences depend on the context of the task, then we have evidence that the differences in performance are due to specific cognitive abilities and not just some general factor. Additionally, this means that the differences observed in other tasks are unlikely to be due to contextual variables, because if these external factors were responsible for the differences in a specific task, those same factors should have resulted in species differences in other tasks, too (see Lefebvre & Giraldeau, 1996; Olson et al., 1995). Of course, it is important to keep in mind that although this technique (as with any other for that matter) strongly weakens the argument from contextual variables, it does not eliminate it. In addition, the major drawback of this approach is that it requires researchers to conduct a whole research program, which may be long and costly.

Despite these issues, taken together, following these techniques will provide strong support for the hypothesis that species’ differences in performance in a task is due to species-specific abilities rather than some aspect(s) of the procedure. For my dissertation, I took an approach similar to the one suggested by Kamil and colleagues (see Kamil, 1998), using converging operations, to investigate further species’ difference in performance on a dichotomous choice task derived from the cleaner-client fish cleaning interaction in the wild (Salwiczek et al., 2012). Unfortunately, due to time and resource constraints, I was not able to test species using the technique of differentiation. However, I address this issue in my conclusions and provide some future directions on how this could be implemented. Thus, despite this missing step, I believe that my research represents a nice example of the template promoted by Kamil (1998).
1.1.3 Comparative Studies: The Fish Model

Fish General Cognition

The ecological approach provides a general functional theoretical framework which allows for the integration of studies on any animal species, and has led to a great diversification of animals studied, and in particular to the appreciation that animal clades that lack particularly large and complexly structured brains may provide examples of impressive cognitive abilities (Salwiczek et al., 2012). This is in particular true for fishes, for which much evidence has emerged supporting greater cognitive ability than was previously recognized (e.g., Agrillo, Miletto Petrazzini, Tagliapietra, & Bisazza, 2012; Brown, Laland, & Krause, 2011; Ferrari, Trowell, Brown, & Chivers, 2005; Grosenick, Clement, & Fernald, 2007; Kendall, Rendell, Pike, & Laland, 2009; Piffer, Miletto Petrazzini, & Agrillo, 2013; Vail, Manica, & Bshary, 2013, 2014; for reviews, see Brown, 2015; Bshary, Gingins, & Vail, 2014; Bshary, Wickler, & Fricke, 2002). For example, the male cichlids (Astatotilapia burtoni) use transitive inference to predict fighting abilities of competitors (Grosenick et al., 2007), and sticklebacks (Pungitius pungitius) employ so-called hill climbing social learning strategies (Kendal et al., 2009), in which they compare their own foraging success with the success of observed individuals to update foraging decisions.

Importantly, many studies have shown that fish, at least as a taxon, demonstrate diverse skills previously attributed uniquely to primates (for a review, see Bshary et al., 2002; for an analog approach comparing birds and primates, see Emery, 2006; Marler, 1996). For example, the grouper (Plectropomus pessuliferus) and the coral trout (Plectropomus leopardus) hunt collaboratively with other species, including giant moray eels (Gymnothorax javanicus),
Napoleon wrasses (*Chelinus undulatus*) and even octopuses (*Octopus cyanea*; see Bshary, Hohner, Ait-el-Djoudi, & Fricke, 2006; Vail et al., 2013). Interestingly, the partner-choice abilities involved in these interactions are comparable to those of chimpanzees (see Vail et al., 2014; for the primate data, see Melis, Hare, & Tomasello, 2006a). Another behavior of interest for comparative psychologists and primatologists is the foraging decisions of cleaner wrasse (*Labroides dimidiatus*) with their so-called “client” fish species. This interaction involves many social complex behaviors, which have been the focus of attention in the Machiavellian intelligence hypothesis (Byrne & Whiten, 1988; also see Bshary, 2006; Bshary et al., 2002).

*The Cleaner-Client Fish Mutualism*

In the wild, cleaner fish occupy small territories (so-called “cleaning stations”) in which they interact with a variety of predatory and nonpredatory reef fish species (so-called “clients”) from which they remove ectoparasites, but also other material such as mucus and scales (Randall, 1958). Conflict occurs because cleaners prefer to eat mucus over ectoparasites (Grutter & Bshary, 2003), where eating the former constitutes cheating (for a review, see Bshary, 2010). Cleaners adjust levels of cooperation to the strategic options available to clients to react to cheating by cleaners. Predatory clients typically receive the highest service quality, whereas nonpredatory resident clients, who lack choice options, punish cleaners for cheating. Furthermore, cleaners pay attention to the presence of potential clients and are more cooperative to current clients if that allows them to access bystanders (Pinto, Oates, Grutter, & Bshary, 2011). Thus, cleaner wrasse show high adaptation to the specifics of an interaction in their foraging decisions, which are at the same time linked to interspecific social behavior.
Client fish have been categorized as either “resident” clients, which live in small territories and interact most frequently with the same cleaner fish (and are therefore a “permanent” food source for the cleaner), or as “choosy” clients, which have larger home ranges that cover several cleaning stations and can choose the cleaners with whom to interact (and are therefore an “ephemeral” food source for the cleaner). Cleaners typically compete among each other over access to choosy clients, while they each have exclusive access to their resident clients; consequently, choosy clients are expected to use their choice options by visiting stations where the service is better. Field observations found that choosy clients have priority for cleaning access over the residents (Bshary, 2001). This priority of access is likely because visitors will leave if they are not cleaned rapidly, and they are less likely to return to the same cleaner if they have been ignored in the past or have received a bad service (Bshary & Schäffer, 2002).

To further investigate this behavior, Bshary and Grutter (2002) conducted a lab study with Australian cleaner fish, in which they replaced client fish with plates; one plate mimicked the choosy client, while the other mimicked the resident client. Importantly, the two plates contained a same reward and only differed in size. In the task, if the subject picked the food from the choosy plate first, it could then take the food from the resident plate too (the same way as a client fish would prioritize cleaning access to a choosy client over a resident client). In contrast, if the resident plate was selected first, the choosy plate was withdrawn (the same way as a choosy client would leave and switch of cleaning partner if it is not cleaned first). In this case, the subject obtained only the food from the resident plate. Therefore, the optimal foraging strategy was to take the food from the choosy plate first before the food from the guaranteed, resident plate. Within just a few trials, cleaner fish learned to choose the choosy plate first,
supporting previous field observations of this behavior. The plate task represents a simple but
great example of how ecological questions can be further tested in a laboratory setting (for a
similar approach, see Bshary & Grutter, 2005, 2006; Bshary et al., 2008; Vail et al., 2014).

1.1.4 Testing the Ecological Hypothesis: The Fish-Primate Model

The ecological hypothesis is rather nonspecific with respect to the cognitive processes
that underlie the cleaners’ performance in the plate task. Therefore, an important remaining
question is whether success in the task involves widespread learning rules or rather specific
abilities (in which case it would support the ecological hypothesis). Could another species easily
solve the task at levels comparable to the fish? To answer these questions, my colleagues and I
subjected adult and juvenile Philippine cleaner fish (Labroides dimidiatus), capuchin monkeys
(Cebus [Sapajus] apella), chimpanzees (Pan troglodytes) and orangutans (Pongo spp.) to the
plate task (Salwiczek et al., 2012).

We chose primates as a group of comparison for three reasons. Most importantly, both
fish and primates (our three species in particular) possess complex foraging and social behaviors,
which have been used to explain the evolution towards large brain in primates (e.g., see Deaner
et al., 2006; Reader et al., 2011). In addition, at least chimpanzees and capuchin monkeys are
known to eat immobile food sources (like fruits and leaves) and hunt for meat and catch
ephemeral prey in the wild (e.g., Boesch, 1994; Boesch & Boesch, 1989; Fragaszy, Visalberghi,
& Fedigan, 2004; Rose, 1997), which may roughly correlate with the stable resident and mobile
choosy clients. Second, all three species cooperate, at least in laboratory settings, and
chimpanzees and capuchins also do so in the wild (e.g., chimpanzees: Boesch, 1994; Melis et al.,
2006a & 2006b; Suchak et al., 2014; orangutans: Chalmeau, Lardeux, Branditas, & Gallo, 1997;
Dufour, Pelé, Neumann, et al., 2009; capuchins: Brosnan, 2010; Perry, Manson, Dower, & Wikbert, 2003). Finally, all three primate species have a large brain-to-body ratio compared to other species (including most other primates, as capuchins have an atypically large brain-to-body ratio for New World monkeys; see Deaner et al., 2007), which has been used as a proxy for cognitive ability.

We hypothesized that if brain size is the key feature in determining whether species can solve this task, one would expect that primates outperform fish. However, if ecology is more important, then one would predict that fish outperform primates in this task derived from cleaner fish ecology. The comparison between adult and juvenile cleaners allowed them to address the potential role of individual experience. In addition to their longer experience of cooperating with client fish, adult cleaners usually interact about three times more frequently with visitors than do juveniles (Barbu, Guinand, Bergmüller, et al., 2011; Bshary, 2001). Thus, juveniles rarely experience the situation in which a visitor and a resident client seek cleaning simultaneously. If adult cleaners perform better than juveniles, it would indicate that individual experience with the task in the field helps to solve the experiment.

Subsequent to the initial learning test, subjects were tested on a reversal test (Rumbaugh, 1971, 1997; also see Gossette & Inman, 1966; Harlow, 1949; Harlow & Warren, 1952; Schusterman, 1962). The procedure was identical to the initial test, but the role of each stimulus was reversed (i.e., the former permanent plate now became the ephemeral plate, and vice versa). This test is frequently used to demonstrate flexibility in cognitive processing. In our task, we used this procedure as an additional way to test the role of learning in the task. In fact, reversal of roles never occurs in a fish natural context (e.g., a visitor client never turns into a resident client); therefore, success in the task could not be associated with fish specific adaptations. For the
primates, we hypothesized that if subjects were faster at solving the reversal than the learning test, then it would indicate that learning the task is what is challenging for them. Supporting the ecological hypothesis, adult cleaner fish performed better than all primates (and a closely related cleaner wrasse species; see Salwiczek et al., 2012) in learning the task. Interestingly, the majority of monkeys (but not apes) who eventually learned the task subsequently performed relatively well on the reversal procedure, like the fish, indicating that learning the rules was part of the difficulty of the task. In addition, we found that adult fish outperformed juveniles, which emphasized the important role of learning and experience to make optimal decisions in the wild.

1.1.5 Potential Factors Influencing Species Performance in the Plate Task

In this paper, we proposed four main (not mutually exclusive) reasons that might explain why the plate task is difficult to learn (Salwiczek et al., 2012). First, we hypothesized that it could relate to known reinforcement mechanisms: whichever plate was selected first, subjects always received one same immediate reward, therefore, unlike classic associative learning tasks, where subjects are rewarded once after selecting one of two stimuli (or a sequence of multiple stimuli), the plate task required subjects to select the second stimulus for an additional reward after it already received a reward from the first stimulus. In this case, the intermediate reward from the permanent plate may have interfered with learning mechanisms, in that it lowered the incentive value of the reward from the subsequent, ephemeral plate. That is, because the reward from the suboptimal, permanent option was received prior to the one from the optimal ephemeral option, blocking or overshadowing mechanisms may have lowered the incentive value of the ephemeral option.
Second, we hypothesized that it is possible that although both fish and primates presumably reacted to the removal of the ephemeral plate as a punishment, only fish may have perceived the task as a social interaction. In that case, they would have perceived the removal of the ephemeral plate as the loss of a cooperation partner, and hence as a punishment that reduced the likelihood that subjects would choose the permanent plate again in the future. The aversion to losing any client would make the ephemeral plate more attractive to cleaners and thus help in learning the task faster. Supporting this hypothesis, in previous work, the cleaner fish often responded in the task with tactile stimulation when the plate returned, a behavior cleaners typically use to reconcile and to make clients stay longer under natural conditions (Bshary & Würth, 2001).

Third, we hypothesized that the primates may have failed to use a higher cognitive mechanism, like backwards induction, to solve this task in lieu to the evolved predispositions seen in the fish. Backwards induction (first introduced by von Neumann & Morgenstern, 1944) is a problem-solving strategy that consists of analyzing a problem from back to front in order to determine what sequence of optimal actions is required to solve it. Evidence of this capacity in animals was first demonstrated in a chimpanzee, Julia, in a tool-use procedure (Döhl, 1968). In that study, Julia learned to open 10 boxes in the appropriate sequence, using specific tools that she could find inside each box, to retrieve a food at the end. In our plate task, although the two foods were presented simultaneously, they were always offered (or removed) sequentially. Therefore, after subjects experienced and learned the outcome associated with each option, they could have used backwards induction to determine what option they would have to choose next in order to obtain the additional reward from the subsequent (permanent) option. One big
difference between our study and this work, however, is that Julia received one reward at the end of the sequence, whereas in our study subjects received their rewards at each step.

Finally, we hypothesized that the primates’ poor performance in the task might have been due to frustration with the procedure. To illustrate this, some ape subjects would sometimes hit or grab the plates rather than choosing a reward, or even refused to participate. This frustration with the task might have caused the low performance in the task, in particular on the reversal learning procedure.

To begin to address these four possibilities, we still had to demonstrate that primates would improve their performance in a similar task derived from the primate ecology, and rule out possible alternative explanations for the primates’ poor performance. Also, wherever possible, we needed to test both fish and primates in the task, because we needed to make sure not only that primates improved, but also that if it was a task derived from their own ecology, they would outperform the fish. For my dissertation, I designed a series of adaptations to the task to test these hypotheses.

1.2 Overview of this Dissertation

In Study 1, my colleagues and I investigated the possibility that species differed in the plate task because of differences in the capacity to recognize the relevant cues to solve it (Lotem & Halpern, 2012; for a review, see Rowe & Healy, 2014). Thus, I tested Moorean fish and capuchin monkeys on a series of tasks designed to be more relevant to primates (Prêtôt, Bshary, & Brosnan, 2016b). In the original design that was based on cleaners, the plates differed but the food outcomes were identical, because cleaners always consume the identical foods from different food patches, in this case acquired from visitor and resident clients during social
interactions. This may have given the cleaners an advantage in the plate task simply because they were already predisposed to focus on the way the food was presented (i.e., the plate, which was indeed the relevant stimulus) rather than the food itself, which was uninformative as the foods were identical. For the primates, however, what is important is the food itself, not the food patch. Although foods may be associated with specific surroundings (e.g., a species of tree may provide hidden fruits), the general details of the source do not change (e.g., the fruits will not suddenly be found in a different species of tree) and the patch may not be informative about the quality of food (e.g., the position of the leaves will not tell whether the fruits are ripe; the fruits themselves must be inspected). Therefore, as food color is potentially a more ecologically relevant cue to primates than fish, I changed the color of the food item rather than the design/color of the plates.

In a second study, the foods were hidden, under cups for the monkeys and behind plates for the fish, as primates’ learning performance in choice tasks is generally decreased by the presence of visible rewards (see Boysen & Berntson, 1995; Boysen, Berntson, & Mukobi, 2001; Boysen, Mukobi, & Berntson, 1999; Boysen, Berntson, Hannan, & Cacioppa, 1996; Diamond, 1981; Murray, Kralik, & Wise, 2005). Although monkeys, as expected, improved their performance in both versions of the task as compared to the original plate task, the Moorean cleaner fish performed surprisingly well in all tasks, presumably because they generalized from their experience in the wild. However, they performed substantially worse at the reversal tasks than did the original Philippine fish, indicating possible differences in cognitive flexibility between these two populations.

In Study 2, I tested three additional primate species (orangutans, gorillas and drill monkeys) in the color and cup tasks (Prétôt et al., in prep), to see whether the prior results were capuchin-specific, or if they generalized across the primate taxon. Interestingly, orangutans
performed at levels comparable to capuchins in the color task, which supported my hypothesis. However, they did not solve the cup task, whereas neither gorillas nor drills solved either of the tasks. Although I can only speculate on why these primates in general did so poorly relative to the capuchins’ performance (excepting the orangutans in the color task), I see at least three possible explanations. First, it is possible that the cues were (inadvertently) more salient to capuchins than other species; therefore, capuchins might have been better than the other primates at finding the cues associated with success in the tasks. Second, and, I believe, most likely, the primates might have done poorly in the tasks for reasons that were external to the task. In particular, none of these species has a lot of experience with cognitive testing, especially relative to the capuchins. Perhaps more tellingly, the gorillas and drills, who did least well, have even less experience than the orangutans. Finally, capuchins may have done overall better in these tasks because of aspects of their ecology.

In Study 3, my colleagues and I investigated the possibility that the primates’ poor performance in the original plate task was due to unintentional extraneous cues associated with a manual testing environment, in particular the presence of an experimenter and, again, visible foods during the subjects’ choice (Prétôt, Bshary, & Brosnan, 2016a). To do this, I tested capuchin monkeys and a new primate species, the rhesus macaque (Macaca mulatta), in two computerized adaptations of the plate task. The first differed from the plate task by removing interaction with the human experimenter, which may be distracting, and providing a more standardized testing environment. The second was an adaptation of the first paradigm designed to be more relevant to primate ecology. To do this, the ephemeral stimulus was moving from the beginning of the trial, which more closely mimicked primates’ natural ephemeral food sources (e.g., insects or small vertebrates, who would presumably be moving towards escape from the
minute they realize that they have been seen). Monkeys’ overall performance in these tasks matched the performance of the fish in the original study. Again, these results showed that with the appropriate modifications, the monkeys could succeed in the task.

Relevant to this study, Pepperberg and Hartsfield (2014) proposed another limitation of the procedure that may have explained species’ difference in performance in the original plate task. They proposed that both fish and parrots might have excelled in the task because they are naturally constrained to making a single choice at a time (with the mouth or bill), whereas primates typically use both hands to reach multiple foods or objects simultaneously, but were forced by the procedure to make only one choice. As a result, they proposed, the primates may have found the limitations of the experimental task more difficult than the other species. Although this third task was not run to test this hypothesis, it did so because the monkeys’ computerized testing is run using a joystick paradigm, which by default limits subjects to a single choice at a time. The capuchins did indeed perform better on the computerized version in which they used a joystick to make their choice, which could support this hypothesis (Prétôt et al., 2016a). On the other hand, in more studies using an analogue of the procedure, two other species that are naturally constrained to single choices by their anatomy, pigeons (which choose with their bill) and rats (which choose with their mouth), failed to choose the food maximizing solution (Zentall, Case, & Berry, 2017; Zentall, Case, & Luong, 2016). Thus, this hypothesis, while intuitive and at least partially supported, needs additional testing using properly controlled procedures to determine the degree to which it influences subjects’ responses.

In Study 4, finally, I tested one possible cognitive limitation to the primates in the original plate task, the failure to use backwards induction (Prétôt & Brosnan, *in review*). Backwards induction is a problem-solving strategy that consists of analyzing a problem from
back to front in order to determine what sequence of optimal actions is required to solve it (von Neumann & Morgenstern, 1944). Although there is evidence that apes (and other animals) show some degree of backwards induction in various contexts and modalities (at least as per my definition; e.g., joystick-based computerized maze tasks: Beran, Parrish, Futch, et al., 2015; Fragaszy, Johnson-Pynn, Hirsh, & Brakke, 2003; Fragaszy, Kennedy, Murnane, et al., 2009; touchscreen-based computerized maze task: Iversen & Matsuzawa, 2001; manual “paddle-box” task: Tecwyn, Thorpe, & Chappell, 2013; manual finger-maze task: Völter & Call, 2014a; for a tool-use version, see Völter & Call, 2014b), few studies have investigated this ability in monkeys. In addition, most of these exceptions have used computerized procedures (e.g., Beran & Parrish, 2012; Beran et al., 2015; Fragaszy et al., 2003, 2009; Pan, Kennedy, Pickering, et al., 2011), whereas it is possible that this ability is best shown in manual tasks, which may provide a more intuitive interface. Therefore, in this study, I tested monkeys using the manual “paddle-box” task, originally designed for testing planning skills in apes, which consisted of a box containing different levels of paddles that subjects rotated to drop a food into a tube that delivered it (Tecwyn et al., 2013). To evaluate subjects’ degree of skill in the task, the locations of the food and tube were chosen according to the probability of retrieving the reward, starting with those that gave the highest probability, and moving sequentially to those with the lowest probability. Most subjects solved all levels of difficulty in the task, and monkeys succeeded as a group in a generalization test. Although it is still unclear why, if they possess backwards induction, the monkeys failed to use it in the original plate task, these findings demonstrate that species may not always use all of their cognitive abilities in every situation.
2 STUDY 1: FACTORS INFLUENCING THE DIFFERENT PERFORMANCE OF FISH AND PRIMATES ON A DICHOTOMOUS CHOICE TASK


2.1 Introduction

Animals’ decisions are constrained by their ecology, their cognitive ability, and the ways in which they can interact with the world, among other factors. The ecological approach to cognition posits that ecology influences decision making such that each species performs better on tasks that are naturally relevant to them (Balda & Kamil, 1989; Kamil, 1988; Kamil & Mauldin, 1987; Shettleworth, 2009). There are many possible mechanisms by which this could occur, but one likely possibility is that species have been selected to focus on cues that are relevant to them (Lotem & Halpern, 2012; for a review, see Rowe & Healy, 2014). Thus, species may be good at identifying problems that are relevant to their ecology and predisposed to look for some cues over others. For instance, research on food-caching birds has shown that nutcrackers, *Nucifraga columbiana*, which are highly dependent on stored food for surviving winters, outperform less cache-dependent species specifically in a spatial memory task, but not in a nonspatial, color memory task (Olson et al., 1995). Similar results were obtained for two populations of black-capped chickadees, *Poecile atricapillus* (Pravosudov & Clayton, 2002): Alaskan chickadees, which live in harsh environments and are highly dependent on food caching, performed better in spatial memory tests than Colorado chickadees, although the
populations did not differ in a nonspatial version of the task. Somewhat surprisingly, such comparisons of performance between ecologically relevant and nonrelevant tasks have remained rare (Shettleworth, 2009). Here, we extend work comparing two phylogenetically distant species, cleaner wrasses, *Labroides dimidiatus*, and brown capuchin monkeys, *Cebus [Sapajus] apella*, that converge on their tendency to cooperate with conspecifics but perform differently in a dichotomous choice task derived from a cleaner-specific cooperative situation (Salwiczek et al., 2012).

In the wild, cleaner fish remove parasites and other material from client reef fish, which visit them at their so-called cleaning stations. Clients have been categorized as either residents with small territories (or small home ranges) that allow them to access only one cleaning station, or as choosy clients, which have larger home ranges that cover several cleaning stations. Cleaners typically compete with each other over access to choosy clients, while they each have exclusive access to their resident clients; consequently, choosy clients are expected to use their choice options by visiting stations where the service is better. Field observations found that choosy clients have priority for cleaning access over the residents (Bshary, 2001); they also typically switch to another cleaner if ignored, but are more likely to return to the same cleaning station if they are inspected (Bshary & Schäffer, 2002), thus making the clients’ choosiness the likely cause of this priority of access.

To test this in the laboratory, Bshary and Grutter (2002) replaced client fish with plates; one plate simulated the choosy client, while the other represented the resident. Fish could feed on the choosy plate only if they started to feed on it before they went foraging on the resident plate, otherwise the choosy plate was withdrawn while the fish was eating from the resident plate, just as choosy clients leave if they are not inspected rapidly; the resident plate, however, always
stayed in the testing area until the fish had stopped feeding on it, just as resident clients often queue for service if the cleaner fish inspects another client. Crucially, both plates offered the same foods, in equal amount, and hence were equally attractive as food patches. Within just a few trials, cleaner fish inspected the choosy plate first, supporting previous field observations of this behavior.

In a subsequent study, Salwiczek et al. (2012) tested cleaner fish, capuchin monkeys, chimpanzees, *Pan troglodytes*, and orangutans, *Pongo spp.*, on this plate task. The goal of this study was to compare the performance of fish and primate species that converged on their tendency to cooperate with one another (e.g., capuchins: Brosnan, 2010; chimpanzees: Boesch & Boesch, 1989) and their propensity to eat both mobile and immobile food sources, which may roughly correlate with the stable resident and mobile choosy clients (e.g., plant materials vs. hunting for insects and smaller vertebrates; Fraga{}szy et al., 2004; Goodall, 1986), and to contrast this with orangutans, which primarily eat fruits (Galdikas, 1988) and less frequently insects or other mobile animal protein sources (Rijksen, 1978), but which do not cooperate to the same degree in natural situations (but do in captivity: Chalmeau et al., 1997; Dufour et al., 2009). In the task, fish outperformed all of the primate species. Although most of the monkeys (but not the apes) eventually learned how to solve the task, they did not do so as quickly as the fish.

Salwiczek et al. (2012)’s results may initially seem counterintuitive given the primates’ large brains and known problem-solving skills, but from the cue perspective they make sense. The fish were presented with a task that was derived from their own ecology, including the cues that were needed to solve it, whereas the primates needed to first learn which cues were relevant, and only then could they learn to solve the task. Of course, ecologically relevant cues are not the only possible causes for the differences; differences in cognition may generally be due to how
individuals perceive, process and/or act upon the available information, or to their motivation for the task itself (Shettleworth, 2009). Therefore, to understand this more fully, we must test the primates on alternative versions of the task that are designed to account for some of these other potentially mediating factors. Additionally, to truly test the hypothesis, it is essential to test the fish on the modified tasks as well, to see whether and how their performance changes across the tasks. In the current paper, we independently tested two nonexclusive reasons that could explain the poor performance of primates in the original plate task, namely whether the primates understood which cue held the relevant information for the decision, and the tendency of primates to be distracted by seeing food during the choice presentation.

Considering the first potential explanation for the superior performance of the fish, the task simulated a natural situation for the fish but not for the primates, so we hypothesized that only the fish would readily identify the relevant cue to solve the task (Lotem & Halpern, 2012). In the wild, cleaner fish consume small invertebrates on the surface of client reef fish (Côté, 2000; Randall, 1958), which only become visible at short range (i.e. that need to be searched for and found). Parasite abundance varies between species, partly as a correlate of client body size (Grutter, 1995); therefore, cleaners should prefer certain clients over others because of their quality as a food patch (Grutter, Glover, & Bshary, 2005). In other words, cleaners should focus on the way the food is presented, rather than on the food itself. This was reflected in the original plate task adaptation, where the plate color and pattern were the relevant stimuli, rather than the foods, which were identical and uninformative (Salwiczek et al., 2012).

For primates, what is important is the food itself, not the patch. Although foods may be associated with specific surroundings (e.g., a species of tree may provide hidden fruits), the general details of the source (e.g., leaf shape) do not change (e.g., the fruits will not suddenly be
found in a different species of tree) and the patch may not be informative about the quality of food (e.g., the position of the leaves will not tell whether the fruits are ripe; the fruits themselves must be inspected). Therefore, for this study, we tested to see whether a cue that was potentially more ecologically relevant to primates (and presumably less ecologically relevant to fish) would increase the primates’ performance. For this, we kept the plates identical, but used different colored food items. We predicted that if the difference found in Salwiczek et al. (2012) was simply due to a difference between species in where attention was focused, the monkeys would outperform the fish in this task.

Considering a second potential explanation for the superior performance of the fish, primates are known to have difficulty making the correct choices when food is present. While primates certainly can make rational choices when food is visible, and can learn to overcome the prepotent response with modifications (e.g., using symbols to represent foods; Boysen et al., 1999; Murray et al., 2005), the original task by Salwiczek et al. (2012) may have been challenging for the primates because of the presence of this extraneous cue (see Pepperberg & Hartsfield, 2014). Therefore, for our second study, we adapted the task to minimize any influence of having visible foods during the subjects’ choices. Note that because studies on the influence of food visibility on decision making in any context on cleaner wrasse are lacking, we had no prediction for whether visible vs. nonvisible food would affect their performance; on the one hand, if the plate design was the important cue, then in principle, this task should not have been more difficult, but on the other, swimming to the opposite side to claim food was presumably atypical for them, and therefore, this may have made the task more challenging.

An important aspect of our comparative approach concerned the choice of the experimental design. Because of the scarcity of nonhuman primate subjects, the capuchin
monkeys were, by necessity, tested in a within-subjects design; therefore, it was essential to collect within-subjects data for fish as well, and to give the fish equivalent experience with the paradigm. To do this, we tested cleaner fish on the original plate task prior to the two other studies. This also allowed us to compare our results for these cleaner fish (from Moorea) with those from our earlier study (from the Philippines; Salwiczek et al., 2012). To summarize, we predicted that (1) offering relevant information of the food (color) rather than some aspect of the plate (e.g., color, pattern) would be more ecologically relevant for primates, so they should outperform the fish in the first experiment, and (2) primates should solve the hidden-food task more quickly than the original plate task (we had no prediction for the effect of this alteration on the fish). If primates did notably better on one task than the other, it would provide evidence as to which of these factors were most important in driving primates’ outcomes in the earlier task, whereas a failure to improve performance in either task for primates would be difficult to interpret.

2.2 Methods

2.2.1 Subjects and Housing

Capuchin monkeys

We tested nine captive-born brown capuchin monkeys (5 males, average age: 12 years, range 7–17 years; 4 females, average age: 15 years, range 12–18 years) from two stable social groups at the Language Research Center of Georgia State University, Atlanta, Georgia, U.S.A. All subjects participated in both studies. Subjects were always housed with their social groups except when they separated voluntarily for behavioral and cognitive testing. Subjects were fed a diet according to their species-specific needs that included primate chow and fresh fruits and
vegetables. They also received enrichment foods several times per day. Animals were never deprived of food or water for testing purposes. Running water was available ad libitum, including during testing. All of the capuchin monkey experiments were approved by the Georgia State University IACUC (A12015) and met the standards of the United States. Georgia State University is fully accredited by AAALAC.

The monkeys lived in two large indoor/outdoor enclosures. Each enclosure contained ample three-dimensional climbing space as well as trapezes, perches and enrichment items. The enclosure for each social group was divided into an indoor area (approximately one-half of their total space) and an outdoor area. The subjects had previously been trained to voluntarily enter test boxes attached to their indoor area, which allowed us to separate individuals from their group for testing. Monkeys were tested in these testing enclosures. Subjects could choose not to participate at any time by walking away from the experimenter, and there were no consequences for the monkeys if they decided not to participate. No subject was ever involved in more than one testing session for any of the studies on any given day.

*Cleaner fish*

Fourteen adult wild cleaner wrasses of unknown sex were tested at the University of California Berkeley Gump Field Station in Moorea, French Polynesia. Subjects were caught with hand-nets from reefs surrounding the field station, and then housed individually in glass aquaria (approximately 50 × 40 × 40 cm) with a continuous flow of fresh sea water. All cleaners were supplied with an opaque Plexiglas shelter tube for hiding during the day and sleeping at night. Cleaners were first trained to feed off Plexiglas plates prior experiments. Individuals were fed mashed prawn flesh and kept for 1–5 weeks prior to commencing experiments. Once
experiments started, subjects were tested every day, and were not given any food except during testing. All studies were conducted during a 4-week visit to the field station by the first author. At the end of the study, all subjects were returned to the location they were caught. This research was approved by the Regional Delegation for Research and Technology (DRRT, Délegation Régionale à la Recherche et à la Technologie, Papeete, Tahiti).

### 2.2.2 General Procedure

The experimental set-ups for both species are illustrated in Figure 1.

**Capuchin monkeys**

The experimental design was based on the studies of Bshary and Grutter (2002) and Salwiczek et al. (2012). Subjects had to choose between two stimuli, each assigned to one of two specific roles, permanent or ephemeral. Choosing the permanent stimulus (which simulated a “resident” client) always resulted in an immediate reward, and this stimulus was available throughout the trial. The ephemeral stimulus (which simulated a “choosy” client) offered the same immediate reward, but only if it was the first one the subjects chose. If subjects chose the permanent stimulus first, the ephemeral stimulus was withdrawn out of reach and sight of the subject. Thus, the optimal outcome was to pick the ephemeral plate first for a first reward, which allowed the subject to also obtain the permanent stimulus for an additional reward. The side on which each stimulus was presented for each trial was randomly determined but counterbalanced within a session so that they were on each side an equal number of times, but with no more than three trials in a row on the same side (see Salwiczek et al., 2012). Each subject received 10 sessions of 10 trials each (unless otherwise noted noted).
Subsequent to the initial learning test, subjects were tested on a reversal test (Rumbaugh, 1971, 1997; Salwiczek et al., 2012). The procedure was identical to the initial test, but the role of each stimulus was reversed (i.e. the previous ephemeral stimulus now behaved like the permanent one, while the previous permanent stimulus now became the ephemeral one). If subjects initially preferred the ephemeral stimulus and then reversed to prefer the newly ephemeral (formerly permanent) stimulus, this would indicate that they are consistent in their preference. The outcomes are more difficult to interpret if subjects initially preferred the permanent stimulus. Reversing to maintain a preference for the newly permanent (formerly ephemeral) stimulus would indicate a consistent preference for the permanent stimulus; however, maintaining a preference for the newly ephemeral (formerly permanent) stimulus would be uninformative because we would not know whether they preferred that physical stimulus (e.g., the plate or food color), or whether they recognized the utility of choosing the ephemeral stimulus first when “forced” to do so after the stimuli were switched. Although we report data on reversal tasks in all cases, it is difficult to know what these latter data mean.

Choices were made using a choice apparatus designed for Salwiczek et al. (2012; see Figure 1). This apparatus was attached to the front of the monkey’s test box and was designed to limit the monkey to a single choice. It consisted of two Velcro doors attached to each other by a string that worked in a drawbridge-like fashion; that is, pulling one door closed when the other was pushed open (for further detail, see Salwiczek et al., 2012). Foods were presented to subjects on two plates placed on a single larger Plexiglas tray (to standardize the location of the plates) carried by the experimenter, who wore an opaque face shield at all times in order to minimize experimenter cueing (one subject was afraid of the face shield and so it was not worn for this subject). The larger tray had a central opaque barrier that kept the two choice plates clearly
separated. The trial started once the subjects faced the two food options. If a subject chose the permanent option first, the ephemeral option was removed from the larger Plexiglas tray and put on the top of the testing box, out of sight of the animal. Subjects were tested in four to five sessions per week, but never in more than one session per day. The intertrial interval (ITI) was 5 min, chosen because, in previous testing, it was the ITI at which the monkeys did the best (Salwiczek et al. (2012) started with a 15 min ITI, as with the fish, and tried a variety of different options until the monkeys succeeded). Rewards were 750 mg banana-flavor precision pellets (Bio-Serv, Frenchtown, NJ, U.S.A.), which could be dyed different colors, as needed, in Study 2 (see below). All monkey studies were carried out prior the fish studies.

**Cleaner fish**

Subjects were tested in their aquarium, in which a separation with an opaque central sliding door was introduced at approximately four-fifths of the aquarium length to create a large “experimental” compartment and a small “resting” compartment (see Figure 1; for previous training studies using a similar apparatus, see Bisazza, Agrillo, & Lucon-Xiccato, 2014; Gierszewski, Bleckmann, & Schluessel, 2013; Miletto Petrazzini, Agrillo, Izard, & Bisazza, 2015). The foods were placed on two plates that were attached to handles so that they could be moved towards subjects but also be retracted rapidly. A given trial started by confining the subject to the resting compartment of the aquarium (subjects quickly learned to swim behind the door before a trial started). The stimuli were then placed at the opposite end of the experimental compartment. After a few seconds, the door opened and the cleaner could enter the experimental compartment at will. Fish received two sessions per day, consisting of 10 trials each. The ITI was set at 15 min (as in Salwiczek et al., 2012). The two plates were placed far enough apart that
following a choice of the permanent option, the experimenter could remove the ephemeral option before the subject could take the food.

Prior to testing, a different set of fish from the same population were pretested for color preferences on a variety of colors to determine which ones to use during the testing. Pretesting was done on a different set of fish to avoid the possibility of inadvertently biasing the test subjects themselves. Colors were based upon those used in Salwiczek et al. (2012). Preferred colors for all fish were red and yellow, and nonpreferred colors were green and black.

*Learning Criterion and Statistics*

We based success on the behavior of cleaner fish in previous experiments, which was a preference for the outcome that maximized their food intake (Bshary & Grutter, 2002; Salwiczek et al., 2012). Therefore, subjects were considered to have solved the test when they met the preference criterion for choosing the payoff-maximizing ephemeral option. They failed if they either developed a preference for the permanent option (using the same criterion as for the ephemeral preference) or did not develop any preference within 100 trials. Subjects who developed a preference for either option were then tested on a reversal test so that all subjects had similar experience prior to subsequent studies. Subjects met the preference criterion when they showed a statistically significant preference for one of the options, which could be achieved by choosing the stimulus (1) 10/10 trials on one session, (2) 9/10 or 8/10 trials on two consecutive sessions or (3) 7/10 trials on three consecutive sessions. These criteria were more conservative than those used by Salwiczek et al. (2012) in two ways. First, we still used the 10/10 criterion on one unique session, but only if the subject selected each of the two stimuli in at least one trial of a previous session (to ensure that they had experience with both outcomes).
This did not apply to the reversal test, because a subject was considered to have already experienced both options in the initial phase. Second, the 9/10 criterion on one unique session was dropped because it came to our attention that, while used frequently as a criterion in cognitive and behavioral testing, Monte Carlo simulation showed that this criterion allowed for performance that exceeded chance levels (so we required two consecutive sessions of 9/10). Although we planned to limit subjects to 10 sessions, if a subject chose either option 9/10, 8/10 or 7/10 in the last session (i.e. session 10), it was given another set (or two, in the case of sessions with 7/10, followed by a second trial with 7 or more out of 10) of 10 trials to maximize its chances of reaching criterion (Prétôt et al., 2016a). Subjects then received reversal trials.

To compare subjects’ performance across species and conditions, we used two statistical tests. First, a two-tailed Fisher’s exact test allowed us to compare the number of individuals who met criterion for preferring the ephemeral option to those who did not; in this way, we could include data from all subjects, even those who developed a preference for the permanent option. We used a within-subjects design for the comparisons between tasks, and a between-subjects design to compare performance between species.

Our primary goal was to compare the capuchin monkeys’ responses to those of the Moorean fish across all three tasks (the original plate task and the two new studies reported here). However, to do this, we also needed to test the Moorean fish on the original task (Salwiczek et al., 2012) in order to give them equivalent experience to the monkeys for our within-subjects analysis. Repeating this test also allowed us to compare the outcomes of this Moorean population of cleaner fish to the previous Philippine population. Second, a two-tailed Mann–Whitney U exact test allowed us to analyze species’ differences in speed of learning. For
this, we compared the results only from the learning phase for the Moorean fish and the capuchin
monkeys because they were the only populations for which we had results from all three tests.

Study Counterbalancing

We tested two groups of fish and two groups of monkeys. The two groups of fish were
both first tested on the original plate task (Study 1) before being tested in either the color task
(Study 2) or the cup task (Study 3). We were able to test some subjects that completed one task
on the other task (i.e., some subjects participated in all three studies; the length of the field site
visit constrained how many subjects could be tested in both Study 2 and Study 3). All monkeys
had previous experience with the task from Salwiczek et al. (2012) and/or in an analogous
computerized format (Prétôt et al., 2016a); two monkeys had experience only with the
computerized format. Therefore, in monkeys, we did not replicate the first study. The first group
was tested in Study 2 and then Study 3, while the second group was tested in Study 3 before
Study 2.

2.3 Study 1: Plate Task in Fish

Procedure

We replicated the study of Salwiczek et al. (2012; original work: Bshary & Grutter,
2002), to allow for (1) the fish to have the same previous experience as most of the monkeys (all
but two subjects had previously experienced this study) and (2) a within-subjects design in
comparing how subjects did on the adapted tasks (studies 2 and 3), as we did for the monkeys.
This secondarily allowed us to compare results from the two different populations of cleaner
fish.
The 14 fish tested had a choice between two pieces of mashed prawn (0.001–0.005 g) placed on two different Plexiglas plates. Each plate was marked with a central black dot to help both experimenters and fish to locate the food on the plate. Both plates had the same surface, but differed in shape, color and pattern. Plate 1 was rectangular, with two yellow/green-striped edges and a black triangle, while plate 2 was square, with two edges in red and one in black (Figure 2). The colors were the same as the colors used in previous testing on the Philippine fish (see Salwiczek et al., 2012; we switched black for white because unlike in Salwiczek et al. (2012)’s task, the plates in the current study were white), and each contained one of the favored and one of the less favored colors based on our pretesting of the Moorean fishes’ color preferences (see General Procedure). Each plate was fixed to a wooden stick, which ended with a parafilm hook that allowed the experimenter to attach it to the inner glass surface of the aquarium. Both plates and food were presented at equal distance from the central sliding door. Half of the individuals were tested with plate 1 as the ephemeral choice, while the other half was tested with plate 2 as the ephemeral choice. As described above, if a subject picked the ephemeral plate first, it was allowed to take the food from the permanent plate as well. In contrast, if it picked the food from the permanent plate first, the ephemeral plate was withdrawn from the aquarium and placed out of sight of the subject. All subjects were then tested on the reversal learning task.

Results

Individual data for all subjects are in Table 1.

Initial learning phase
Nine out of the 14 subjects tested on the initial learning phase solved the task within 110 trials (range 20–110 trials, mean ±SD = 54.44 ± 34.68 trials). More subjects met the criterion when plate 2 was the ephemeral plate (seven of seven subjects) than when plate 1 was the ephemeral plate (two of seven subjects), possibly indicating a preference for that plate (Fisher’s exact test: \( P = 0.02 \)).

**Reversal learning phase**

Two of the nine subjects that solved the initial learning task solved the reversal task within 80 trials (range 40–80 trials, mean ±SD = 60 ± 28.28 trials). Six of the seven subjects that did not reverse their preference showed a side bias (binomial/sign test: all \( P \)s < 0.01), and the seventh one did so from trial 50 to trial 100 (binomial/sign test: \( P < 0.01 \)).

**Moorean vs. Philippine fish (between-subjects)**

We compared the fishes’ performance on the initial and reversal phases of the plate task to the performance of the Philippine fish population. There was no significant difference in the initial learning phase between Moorean fish (nine of 14 subjects succeeded in the task) and Philippine fish (all six subjects succeeded; Fisher’s exact test: \( P = 0.26 \)). However, more Philippine fish succeeded in the reversal phase (six of six subjects) than Moorean fish (two of nine subjects; Fisher’s exact test: \( P < 0.01 \)).

**Fish vs. monkeys (between-subjects)**

We compared the Moorean fishes’ performance on the initial and reversal phases of the plate task to the performance of the capuchin monkeys in the previous study (Salwiczek et al.,
More fish succeeded on the initial phase (nine of 14 subjects) than did monkeys (all eight subjects failed in the task; Fisher’s exact test: $P < 0.01$), while more monkeys succeeded on the reversal phase (seven of eight subjects) than fish (two of nine subjects; Fisher’s exact test: $P = 0.02$).

**Discussion**

As with previous work in cleaner fish, Moorean fish learned the plate task more rapidly than capuchin monkeys in Salwiczek et al. (2012). This allowed us to continue with our central comparison, between the Moorean cleaner fish tested in the current study, which were able to solve the initial learning task, and capuchin monkeys, which were not. Unless otherwise specified below, all comparisons between monkeys and fish were with the Moorean fish in the current study.

Interestingly, however, although the two populations did not differ in performance on the initial learning task, the Moorean fish were less likely to solve the reversal task than were the Philippine fish. Although our task cannot determine why this difference exists, there are several possible (nonmutually exclusive) reasons. First, recent evidence from Australia suggests that client species’ density and diversity as well as cleaners’ density may have important effects on cleaner performance in cognitive tasks (Wismer, Pinto, Vail, et al., 2014; also see Salwiczek et al., 2012). While we do not have measures for these two parameters, they do appear to differ on basic measures of client diversity and density; reports indicate that there are at least twice as many species of reef fish in the Philippines as in French Polynesia (FishBase: http://www.fishbase.org), making the Philippines the richest concentration of marine life on the
planet (Carpenter & Springer, 2005). It is possible that these (or other smaller-scale) parameters influence the two populations differently.

Second, the aquaria used in the current experiments were shorter than those used in the previous work with Philippine fish, which reduced the distance to make a choice. In primates, differences in the size of the enclosure (Bräuer, Call, & Tomasello, 2007), the orientation of the experimenter to the stimuli (Mulcahy & Call, 2009), and seemingly minor changes in procedure (e.g., providing one tool rather than two; Girndt, Meier, & Call, 2008) influence responses in cognitive tasks, and it is possible that the same was true here. Supporting this, all of the Moorean fish that failed to reverse their preferences showed robust side biases, something not seen in previous work (R. Bshary, personal observation).

Finally, despite our use of the same colors as in our previous work (Salwiczek et al., 2012) and our efforts to make both plates equally attractive, the Moorean fish showed a preference for one of the plates (the one with red stripes) in the initial task. This was somehow surprising, because none of the adult fish showed such biases in our previous study, although one juvenile did so (Salwiczek et al., 2012). Unfortunately, the apparent color preference makes it somewhat more difficult to interpret our fish data, and it may have influenced the fishes’ ability to solve the reversal task. We find it very interesting that one population should show a much stronger color preference than the other, and we hope to explore this topic further in the future.

Because of the inconsistency between the two populations on the reversal task, we do not compare these two populations of fish further on the reversal phase, although we note that these population differences are a very fruitful avenue for future research.
2.4 Study 2: Color Task

Procedure

The procedure was identical to the plate task described in Study 1, except that in this case, the plates had the same color, pattern and shape, and the foods on the plates differed in color.

Capuchin monkeys

The two plates were green and rectangular, and rewards were 750 mg banana-flavored precision pellets (Bio-Serv) that were colored pink or black with spray food coloring (Figure 2). Of the nine subjects, six (in one social group) were tested first in this study and subsequently on Study 3, while the other three (in a separate social group) were tested first in Study 3 and subsequently on this study.

Cleaner fish

The two plates were white and rectangular (presented vertically, as opposed to horizontally as in Study 1) and the foods were colored yellow-orange and purple-pink with liquid food coloring (Figure 2; these colors were chosen because both were favored by cleaners in pilot testing). Only nine of the 14 fish were tested in this task (two of these nine subjects were previously tested in Study 3; all subjects were previously tested in Study 1). For the initial learning phase, four subjects were tested using yellow-orange as the ephemeral choice, while five subjects started with purple-pink as the ephemeral choice. One of these five subjects (D12) chose purple-pink in 10/10 trials in each of the first four sessions and so was dropped from the study, leaving a sample size of eight fish. Because of time constraints on how long the
experimenter could stay at the field site, one subject (D1) received three sessions on one day, and another subject (D5) received three sessions on two different days.

Results

Individual data for all subjects are in Table 2.

Initial learning phase

All nine monkeys reached the preference criterion within 100 trials; seven individuals solved the task and preferred the ephemeral stimulus (range 20–100 trials, mean ±SD = 50 ± 27.08 trials), while two developed a preference for the permanent stimulus in 20 trials.

Seven of the eight fish tested reached the preference criterion within 100 trials; six individuals solved the task and preferred the ephemeral stimulus (range 50–100 trials, mean ±SD = 63.33 ± 19.66 trials), while one developed a preference for the permanent stimulus in 20 trials.

Reversal learning phase

Six of the seven monkeys that solved the initial learning task also solved the reversal task within 100 trials (range 40–100 trials, mean ±SD = 70 ± 21.91 trials). The two subjects who preferred the permanent stimulus in the initial task reached criterion for the ephemeral stimulus in 60 trials.

None of the fish that solved the initial task (N = 6) solved the reversal task, although this population of fish was also unlikely to reverse in the original plate task (see Study 1). The one subject that preferred the permanent stimulus in the initial task reached criterion for the ephemeral stimulus in 20 trials.
Plate vs. color tasks (within-subjects)

For the analysis, we only included subjects that completed both the plate and color tasks (seven monkeys and eight fish). Significantly more monkeys succeeded in the initial learning phase of the color task (five of seven subjects) than in the original plate task (none of seven subjects; Fisher’s exact test: \( P = 0.02 \); Figure 3).

The fish did not show any difference in performance between the color task (six of eight subjects succeeded) and the plate task (five of these eight subjects succeeded in the plate task; Fisher’s exact test: \( P = 1.00 \); Figure 3).

Monkeys vs. fish (between-subjects)

There was no significant difference in the initial learning phase between monkeys (seven of nine subjects succeeded) and Moorean fish (six of eight subjects succeeded; Fisher’s exact test: \( P = 1.00 \)). Considering only the subjects that solved the initial task, there was no difference in speed of learning between species (Mann–Whitney \( U \) exact test: \( U = 11.50, N_{\text{monkeys}} = 7, N_{\text{fish}} = 6, P = 0.20 \)).

Discussion

More monkeys succeeded in the color task than in the plate task, which supports our hypothesis that, for primates, food color is a more salient cue than plate design. In contrast to our predictions, however, cleaner fish did as well on the color task as on the original plate task. This could be an indication that the fish were able to generalize to a novel cue (or that the cue was
more salient than predicted), but we cannot rule out the possibility that their performance was due to a color preference.

2.5 Study 3: Cup Task

Procedure

The procedure was identical to Study 1, except that in this case, the foods were hidden (by cups in the case of capuchins, and placed behind the plates in the case of cleaners).

Capuchin monkeys

The food was placed under two opaque cups so that subjects could not see the food prior to the choice. We used two identical black plates to hold the two different containers, one of which was yellow with one red dot on each side and the other of which was purple with a blue triangle on each side (Figure 2). At the beginning of each trial, the subjects saw the experimenter hiding each food item under the container. Subjects were then presented the choice between the two containers. Subjects indicated their choice by touching a container, at which point, either they lifted the cup themselves, or the experimenter lifted it for them to access the food underneath. One social group completed this task prior to Study 2 and the other group completed this task subsequent to Study 2.

Cleaner fish

Instead of containers, which would not work under water, the food was placed on the reverse side of the plates (i.e. away from the direction from which the individual approached). Subjects made a choice by swimming behind the plate to obtain the food. All the fish learned
quickly (typically on the first trial) to swim to the reverse side of the plates. The two plates were white and triangular and only differed in color and pattern (both sides of the plates were colored/patterned). Plate 1 had two vertical red lines, while plate 2 had two diagonal yellow lines (Figure 2); for the initial learning phase, four subjects were tested using plate 1 as the ephemeral choice, while the other four started with plate 2 as the ephemeral choice. Eight of the 14 fish subjects were tested in this task (three of which had previous experience with Study 2). Note that, because of time constraints, one subject (D12) received three sessions on two different days, while another subject (D8) was not tested on the reversal learning phase.

Results

Individual data for all subjects are in Table 3.

Initial learning phase

All nine monkeys solved the task within 110 trials (range 20–110 trials, mean ±SD = 54.44 ± 32.06 trials), all preferring the ephemeral plate.

Seven out of the eight fish tested reached preference criterion within 110 trials; three individuals preferred the ephemeral stimulus (range 40–60 trials, mean ±SD = 53.33 ± 11.55 trials), while four individuals preferred the permanent stimulus (range 30–110 trials, mean ±SD = 62.50 ± 35.94 trials).

Reversal learning phase

Seven of the nine monkeys that succeeded in the initial learning phase reversed their preference within 100 trials (range 30–100 trials, mean ±SD = 74.29 ± 26.37 trials).
None of the fish that solved the initial task (\(N = 3\)) solved the reversal task. Among the three subjects that preferred the permanent stimulus in the initial task and were tested in the reversal task (one was not, due to time constraints, see above), two reached criterion for the ephemeral stimulus in 10 and 60 trials, while the remaining did not reach criterion within 100 trials.

*Plate vs. cup tasks (within-subjects)*

For the analysis, we only included subjects that completed both the plate and cup tasks (seven monkeys and eight fish). More monkeys succeeded in the initial learning phase of the cup task (seven of seven subjects) than in the original plate task (none of seven subjects; Fisher’s exact test: \(P < 0.01\); Figure 3).

The fish did not differ in performance between the cup task (three of eight subjects succeeded) and the plate task (five of these eight subjects succeeded in the plate task; Fisher’s exact test: \(P = 0.62\); Figure 3).

*Color vs. cup tasks (within-subjects)*

For the analysis, we only included subjects that completed both the cup and color tasks (nine monkeys and four fish). Monkeys did not show any difference in the initial learning phase between the color task (seven of nine subjects succeeded) and the cup task (nine of nine subjects succeeded; Fisher’s exact test: \(P = 0.47\)).

The fish did not show any difference in the initial learning phase between the color task (two of four subjects succeeded) and the cup task (one of four subjects succeeded; Fisher’s exact test: \(P = 1.00\)).
Monkeys vs. fish

Significantly more monkeys succeeded in the initial learning phase of the cup task (nine of nine subjects) than did fish (three of eight subjects; Fisher’s exact test: $P < 0.01$; Figure 3). Considering only the subjects that solved the initial task, there was no difference in speed of learning between species (Mann–Whitney $U$ exact test: $U = 11$, $N_{\text{monkeys}} = 9$, $N_{\text{fish}} = 3$, $P = 0.71$).

Discussion

As in the color task, monkeys improved their performance in the hidden-food task when compared to the original plate task, suggesting that the presence of food in the original task may have caused primates to act impulsively. The use of a method that obscured the food may have helped the monkeys to inhibit any possible prepotent response due to the presence of food rewards (see Boysen & Berntson, 1995; Boysen et al. 1999, 2001; Murray et al. 2005; Pepperberg & Hartsfield, 2014; Prétôt et al., 2016a).

Although fish did not show any difference in performance between the cup task and the two other tasks, their overall performance in the cup task was poorer than that of the monkeys. If anything, this is somewhat surprising given that ectoparasites consumed by cleaners are so small that they are only visible from relatively close range, indicating that cleaners are accustomed to not seeing food on their initial approach. One possibility is that the cleaners found the procedure more difficult; unlike the other tasks, this task required them to swim behind the plate to find food (although note that all subjects swam behind the plates to obtain food on their first session of exposure to it). Alternately, cleaners’ low performance might be due to a preference for one of the plates or some carryover effects resulting from the same colors (the two that they most
preferred) being used on multiple tests. Based on our analyses, we propose that carryover effects, if they existed, might have been small, because most subjects took a minimum of 40 trials to develop a preference for one of the plates, which indicates that their choice was not influenced by their previous exposure in the plate task. Nevertheless, subjects apparently had a preference for the red-striped plate (seven of eight individuals preferred that plate; binomial/sign test: $P = 0.07$), which may have hindered their overall performance in the task, independently of any carryover effects.

2.6 General Discussion

In the present paper, we tested two hypotheses about factors that may have caused primates to perform less well than cleaner wrasse on a previous dichotomous choice task based on the cleaners’ ecology (Salwiczek et al., 2012). In particular, we altered two factors that we predicted might have made the original task more difficult for primates than for fish. First, we changed the cue from being the plate surrounding the food to being the food itself, which we predicted would be more relevant to the monkeys. The primates’ performance improved, whereas the cleaners’ performance was unchanged (see details in Discussion of Study 2). Second, we hid the food (the cue was again the color/pattern of the plate), because visible foods are known to inhibit decision making in primates. Again, the primates’ performance improved, but the cleaners’ performance was unchanged (see details in Discussion of Study 3). Taken together, these results indicate that, not surprisingly, many factors play into determining what species may learn with more or less ease, and that, as predicted, the cues themselves are an important part of decision making.
Of course, while our results demonstrate the importance of these two factors in the primates’ decision making, they do not rule out the (very likely) possibility that other factors influence different species’ performance in such tasks. For example, previous work showed that parrots did as well as fish on the “fish” version of the task (Pepperberg & Hartsfield, 2014). The authors of that study proposed that this might be because fish and parrots are constrained to making a single choice at a time (e.g., with a mouth or bill), forcing them to prioritize, whereas monkeys are typically able to obtain two things at once (e.g., with both hands). As a result, the monkeys may have found the limitations of the experimental task difficult. In fact, in another study, we found that two different species of monkeys did better on a computerized version of the task, in which they chose an option with a joystick-controlled cursor, possibly because they were limited to a single choice (by necessity rather than by experimenter constraint; Prêtôt et al., 2016a). Continued work will help to identify all of the constraints on decision making, as well as interactions among them.

Turning now to the fish, we got two unexpected results. First, contrary to our hypothesis, neither of the cue-related changes led to decreased performance for the cleaner fish on the initial learning trials. Second, the Moorean fish were substantially worse at the reversal task than were the original Philippine fish. These are somewhat difficult to reconcile as they suggest two seemingly contradictory possibilities (on the one hand, that the fish were able to generalize across stimuli in a way that the primates were not, and on the other that the Moorean population was substantially less cognitively flexible than the Philippine fish). We discuss each of these in turn.

Considering the first outcome, cleaner fish performed similarly on the color and cup tasks as on the original plate task in Study 1, and did as well as the monkeys on Study 2, although the
monkeys outperformed the fish on the cup task in Study 3. This is quite different than what we observed for the monkeys, which performed better on some versions of the task than others. There are a couple of possible explanations for this. First, our results could be partly due to an artefact of the fishes’ apparent preference for some plates over others (which is potentially interesting in and of itself given that such a color preference has not been reported previously for any other population). Second, the fish may have generalized from their experience in natural contexts, using cues derived from their own ecology, priming them to succeed in all versions of the task. This would be an impressive feat, in particular, given the primates’ difficulty in learning the task originally (even with what we predicted to be more ecologically relevant cues, the monkeys still did not learn the color task in fewer trials than the fish used in this study or than the fish from the original study by Salwiczek et al., 2012). This possibility deserves further consideration because as of late, there is evidence for generalized rule learning in cleaners (Wismer, Grutter, & Bshary, 2016), and much evidence has emerged supporting greater cognitive ability in fish than was previously recognized (Agrillo et al., 2012; Brown et al., 2011; Bshary et al., 2014; Ferrari et al., 2005; Grosenick et al., 2007; Kendall et al., 2009; Piffer et al., 2013; Vail et al., 2013, 2014; for reviews, see: Brown, 2015; Bshary et al., 2002, 2014). These results might thus indicate that fish have some form of general intelligence that goes beyond the ability to readily solve problems only if presented within a precise ecological context.

Considering the second finding, we were very surprised by the Moorean fishes’ inability to solve the reversal task, particularly in comparison with the performance of the original Philippine fish. Because we wished to (1) give the fish comparable experience to the primates and (2) compare the fish to themselves in a within-subjects design, as we did with the primates, we initially tested the fish on the original version of the task (Study 1). This secondarily allowed
us to compare two different populations of cleaners. We found that, while their outcomes were at least not significantly different for the initial learning trials (64% of the Moorean fish learned the task compared with 100% of the Philippine fish), the Moorean fish were significantly worse at the reversal task than the original Philippine population. This may indicate differences in cognitive flexibility between these two populations, possibly due to differences in their interspecific social environment (such differences due to microecology have been described for cleaners caught from different microhabitats around Lizard Island, Great Barrier Reef; Wismer et al., 2014), experimental differences (such as the aforementioned differences in the length of the aquaria in the two new studies), or the unexpected plate preferences we found in our fish population. Unfortunately, we do not have access to the original Philippine population anymore, nor do we know exactly where they came from (they were obtained from a pet store), so it is difficult to test the first possibility directly. However, both the ecological and procedural possibilities can be tested by exploring population-level differences in cognition and behavior in both fish and other species. Although this was not the primary focus of our study, we find it a very intriguing avenue of research and are excited by the possibility of fully integrating ecology into studies of cognition.

Overall, we find that changing the cues given to subjects can dramatically influence their ability to learn a task. This has important ramifications for comparative work, where scientists (including us!) work very hard to equalize every possible aspect of the task across a species or population. However, these results indicate that this may backfire, with subjects not living up to their potential because they are struggling with the cue of the task rather than the task itself. While there are undoubtedly many ways to approach this issue, we have done so by first running studies that are, to the degree possible, identical, and then iterating the design for the lesser-
performing species to try to unpack what exactly caused the difficulty. This allows us to
triangulate in on the fairest comparison. We are excited to see an increasing volume of research
that integrates ecology, cognition and behavior to better understand the evolution of decision
making across species.
2.7 Tables

Table 1 Two-Choice Moorean Fish Plate Task

Number of trials needed for each subject to solve the plate task for the initial learning phase and the reversal phase in Study 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal</th>
<th>Ephemeral option</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>60</td>
<td>40</td>
<td>Plate 1</td>
</tr>
<tr>
<td>D2</td>
<td>70</td>
<td>80</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D3</td>
<td>20</td>
<td>&gt;100</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D4</td>
<td>20</td>
<td>&gt;100</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D5</td>
<td>20</td>
<td>&gt;100</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D6</td>
<td>30</td>
<td>&gt;100</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D7</td>
<td>60</td>
<td>&gt;100</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D8</td>
<td>100</td>
<td>&gt;100</td>
<td>Plate 2</td>
</tr>
<tr>
<td>D9</td>
<td>110</td>
<td>&gt;100</td>
<td>Plate 1</td>
</tr>
<tr>
<td>D10</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>D11</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>D12</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>D13</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>D14</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
</tbody>
</table>

Stimulus options (plate 1 vs. plate 2) are shown in Figure 2. Subjects that did not reach preference criterion in the initial learning phase were not further tested on the reversal phase.
**Table 2** Two-Choice Monkey-Fish Cup Task

Number of trials needed for each subject to reach preference criterion for the initial learning phase and the reversal phase in Study 2

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal</th>
<th>Ephemeral option</th>
<th>Initial preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monkeys</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wren$^1$</td>
<td>20</td>
<td>50</td>
<td>Colour 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Nala$^1$</td>
<td>20</td>
<td>60</td>
<td>Colour 2</td>
<td>Permanent</td>
</tr>
<tr>
<td>Lily</td>
<td>20</td>
<td>60</td>
<td>Colour 2</td>
<td>Permanent</td>
</tr>
<tr>
<td>Logan</td>
<td>30</td>
<td>80</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Nkima$^2$</td>
<td>40</td>
<td>40</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Liam</td>
<td>40</td>
<td>100</td>
<td>Colour 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Gambit$^{2, 3}$</td>
<td>50</td>
<td>&gt;100</td>
<td>Colour 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Griffin$^1$</td>
<td>70</td>
<td>70</td>
<td>Colour 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Gabe</td>
<td>100</td>
<td>80</td>
<td>Colour 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D10</td>
<td>20</td>
<td>20</td>
<td>Colour 1</td>
<td>Permanent</td>
</tr>
<tr>
<td>D1$^1$</td>
<td>50</td>
<td>&gt;100</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D11</td>
<td>50</td>
<td>&gt;100</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D14</td>
<td>50</td>
<td>&gt;100</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D5$^1$</td>
<td>60</td>
<td>&gt;100</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D7</td>
<td>70</td>
<td>&gt;100</td>
<td>Colour 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D3</td>
<td>100</td>
<td>&gt;100</td>
<td>Colour 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D8</td>
<td>&gt;100</td>
<td>—</td>
<td>Colour 1</td>
<td>—</td>
</tr>
</tbody>
</table>

Stimulus options (color 1 vs. color 2) are shown in Figure 2. Subjects that did not reach preference criterion in the initial learning phase were not further tested on the reversal phase.

$^1$ Subjects previously tested in Study 3.

$^2$ New subjects (not tested in Salwiczek et al., 2012)

$^3$ Subject was tested without the opaque face shield.
Table 3 Two-Choice Monkey-Fish Cup Task

Number of trials needed for each subject to reach preference criterion for the initial learning phase and the reversal phase in Study 3

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal</th>
<th>Ephemeral option</th>
<th>Initial preference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monkeys</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gabe(^1)</td>
<td>20</td>
<td>30</td>
<td>Cup 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Wren</td>
<td>30</td>
<td>70</td>
<td>Cup 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Gambit(^1, 2, 3)</td>
<td>30</td>
<td>100</td>
<td>Cup 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Liam(^1)</td>
<td>40</td>
<td>50</td>
<td>Cup 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Lily</td>
<td>40</td>
<td>90</td>
<td>Cup 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Nkima(^1, 2)</td>
<td>50</td>
<td>80</td>
<td>Cup 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Griffin</td>
<td>70</td>
<td>&gt;100</td>
<td>Cup 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Nala(^1)</td>
<td>100</td>
<td>100</td>
<td>Cup 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Logan(^1)</td>
<td>110</td>
<td>&gt;100</td>
<td>Cup 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td><strong>Fish</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>30</td>
<td>60</td>
<td>Plate 2</td>
<td>Permanent</td>
</tr>
<tr>
<td>D6</td>
<td>40</td>
<td>10</td>
<td>Plate 2</td>
<td>Permanent</td>
</tr>
<tr>
<td>D10(^1)</td>
<td>40</td>
<td>&gt;100</td>
<td>Plate 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D12(^1)</td>
<td>60</td>
<td>&gt;100</td>
<td>Plate 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D13</td>
<td>60</td>
<td>&gt;100</td>
<td>Plate 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>D4</td>
<td>70</td>
<td>&gt;100</td>
<td>Plate 2</td>
<td>Permanent</td>
</tr>
<tr>
<td>D8(^1, 4)</td>
<td>110</td>
<td>—</td>
<td>Plate 2</td>
<td>Permanent</td>
</tr>
<tr>
<td>D5</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
<td>—</td>
</tr>
</tbody>
</table>

Stimulus options (cup 1 vs. cup 2) are shown in Figure 2. Subjects that did not reach preference criterion in the initial learning phase were not further tested on the reversal phase.

\(^1\) Subjects previously tested in Study 2.

\(^2\) New subjects (not tested in Salwiczek et al., 2012).

\(^3\) Subject was tested without the face shield.

\(^4\) Subject was not tested on reversal test due to time constraints.
2.8 Figures

<table>
<thead>
<tr>
<th>Fish</th>
<th>Monkeys</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Fish Experimental Set-Up" /></td>
<td><img src="image2.png" alt="Monkeys Experimental Set-Up" /></td>
</tr>
</tbody>
</table>

**Figure 1** Monkey-Fish Experimental Set-Up

Subjects used a door system (fish: sliding door; monkeys: “Velcro” doors) before they could access the plates and food (the dimensions of the aquarium and enclosure are given in centimeters).
### Figure 2 Monkey-Fish Sets of Stimuli

Pairs of stimuli (named 1 and 2) used in the plate, color and cup tasks.
Figure 3 Monkey-Fish Initial Learning Phase

The number of trials required for Moorea fish and capuchin monkeys to develop a preference for the ephemeral stimulus in the plate task (capuchins data based on Salwiczek et al., 2012; note that one subject previously tested in the plate task was no longer available for the current study), and either stimulus in the color and cup tasks. Each dot represents one individual; black dots indicate a preference for the ephemeral stimulus, while white dots indicate a preference for the permanent stimulus. For monkeys, the grey dots designate new subjects not previously tested in the plate task. Fourteen fish were tested in the plate task, eight were tested only in the color or cup task, and four were tested in both the color and cup tasks. As in our previous work (Prêtôt et al., 2016a), while criterion was set at 10 sessions, if subjects were in the process of meeting criterion on the 10th session (e.g., preferred one stimulus on at least 7 out of 10 trials), they were allowed to continue until they either met criterion or failed to do so. The label “failed” corresponds to subjects who did not reach preference criterion.
3 STUDY 2: VARIATION IN PERFORMANCE OF ORANGUTANS, GORILLAS AND DRILLS IN TWO VERSIONS OF A CHOICE TASK DERIVED FROM THE CLEANER FISH ECOLOGY

3.1 Introduction

Species’ performance in cognitive tasks is influenced by a number of key factors. One recent hypothesis is that differences in cognitive performance are often due to differences in understanding species-specific cues needed to solve a task (Lotem & Halpern, 2012; for a review, see Rowe & Healy, 2014). For example, we recently showed that monkeys performed better in a dichotomous choice task derived from the cleaner fish ecology when the cues to solve the task were more ecologically relevant to primates (Prétôt et al., 2016b). In the task, if the subjects first chose one of the options (ephemeral), they received both food items, but if they chose the other (permanent) option first, the ephemeral option was no longer available. Here, we replicated the study with three additional primate species (Pongo spp., Gorilla gorilla gorilla and Mandrillus leucophaeus), to see if our results generalized across the primate taxon. The task was based on the cleaning interaction between the cleaner wrasse Labroides dimidiatus and its “client” fish species, from which the wrasse eats ectoparasites as its primary diet.

In the wild, cleaner fish remove parasites and other material from client reef fish, which visit them at their so-called cleaning stations. Clients have been categorized as either residents, which live in small territories and interact most frequently with the same cleaner fish (and are therefore a permanent food source for the cleaner), or as choosy clients, which have larger home ranges that cover several cleaning stations and can choose the cleaners with whom to interact (and are therefore an ephemeral food source for the cleaner). Cleaners typically compete among
each other over access to choosy clients, while they each have exclusive access to their resident clients; consequently, choosy clients are expected to use their choice options by visiting stations where the service is better. Field observations found that choosy clients have priority for cleaning access over the residents (Bshary, 2001). This priority of access is likely because visitors will leave if they are not cleaned rapidly, and they are less likely to return to the same cleaner if it has been ignored in the past or received a bad service access (Bshary & Schäffer, 2002).

To test this in the lab, Bshary and Grutter (2002) replaced client fish with plates; one plate mimicked the choosy client, while the other mimicked the resident client. Importantly, the two plates contained a same reward and only differed in size. In the task, if the subject picked the food from the choosy plate first, it could then take the food from the resident plate too (the same way as a client fish would prioritize cleaning access to a choosy client over a resident client). In contrast, if the resident plate was selected first, the choosy plate was withdrawn (the same way as a choosy client would leave and switch of cleaning partner if it is not inspected first). In this case, the subject obtained only the food from the resident plate. Therefore, the optimal foraging strategy was to take the food from the choosy plate first before the food from the guaranteed, resident plate. Within just a few trials, cleaner fish inspected the choosy plate first, supporting previous field observations of this behavior.

In a subsequent study, Salwiczek et al. (2012) tested new cleaner fish subjects, capuchin monkeys, chimpanzees, and orangutans, on the plate task. The goal of this study was to compare the performance of fish and two primate species, capuchins and chimpanzees, who converged on their tendency to cooperate with one another (e.g., capuchins: Brosnan, 2010; chimpanzees: Boesch & Boesch, 1989; Suchak et al., 2014) and their propensity to eat both mobile and
immobile food sources, which may roughly correlate with the stable resident and mobile choosy clients (e.g., plant materials vs. hunting for insects and smaller vertebrates; see Fragaszy et al., 2004; Goodall, 1986). In addition, their performance was contrasted with orangutans, which primarily eat fruits (Galdikas, 1988) and less frequently insects or other mobile animal protein sources (Rijksen, 1978). We hypothesized that if brain size is the key feature in determining whether species can solve this task, one would expect that primates, who have relatively large brain, outperform fish. However, if ecology is more important, then one would predict that fish outperform primates, or a subset of primate species, in this task derived from cleaner fish ecology. Supporting the latter prediction, the fish performed better than all primates in the task. However, to draw the general conclusion that it is the ecological relevance of a task that primarily influences species’ abilities to solve it, we still needed to demonstrate that primates outperformed fish in a similar task derived from their ecology.

To address these points, we tested capuchins and new fish subjects in versions of the plate task designed to be more relevant to primates (Prétôt et al., 2016b). In the original design based on cleaners, the plates differed but the food outcomes were identical, just as cleaners always consume the same foods, but acquire these foods from different clients. This may have given the cleaners an advantage simply because they were already predisposed to focus on the way the food was presented (i.e., the plate, which was indeed the relevant stimulus) rather than the food itself, which was uninformative as the foods were identical. In a first study, we changed the color of the food item rather than the design/color of the plates. Food color is potentially a more ecologically relevant cue to primates than fish. In a second study, the foods were hidden, under cups for the monkeys and behind plates for the fish, as primates’ performance in choice tasks is generally influenced by the presence of visible rewards (e.g., Boysen & Berntson, 1995;
Monkeys improved their performance in both tasks as compared to the original plate task, indicating that these more salient cues did support the primates’ ability to solve the task. However, they do not rule out the (very likely) possibility that other factors influence species’ performance in these tasks. Indeed, recent studies have found that parrots, whose ecology and cognition are presumably more like those of primates than fish, performed at levels equal to the fish in the plate task (Pepperberg & Hartsfield, 2014), while two other species, pigeons and rats, perform as poorly as the primates (Zentall et al., 2016, 2017).

The goal of the current study was to see if the results of the color and cup tasks were capuchin-specific, or if they generalized across the primate taxon. To do so, we tested three additional primate species on these tasks: orangutans, who failed to solve the plate task in a prior study (Salwiczek et al., 2012), and two new species, gorillas and drill monkeys. All three species have relatively comparable foraging habits, as they all consume mostly plant matter (predominantly fruits, but also leaves, seeds and bark). All differ, however, from chimpanzees and capuchins in that they eat ephemeral prey such as insects or other mobile animal protein sources less frequently (orangutans: Galdikas, 1988; Knott, 1998; Rijksen, 1978; gorillas: Doran & McNeilage, 2001; Doran et al., 2002; Tutin, 1996; drills: Astaras et al., 2008; Gonzalez-Kirchner & de la Maza, 1996), and do not cooperate to the same extent in the wild. As a result, we could see whether the results of our prior studies (Salwiczek et al., 2012; Prétôt et al., 2016b) were 1) due to some aspect of capuchins’ and chimpanzees’ specific ecologies (i.e., frequency of encounter with ephemeral prey), 2) specific to having a large brain-to-body ratio (true of the ape species and capuchins; e.g., Deaner et al., 2007), 3) related specifically to higher levels of cooperation in the wild (as in chimpanzees and capuchins), or 4) widespread across the primate
Our primary goal was to compare all three species’ responses across all three tasks (the original plate task and the two new studies conducted here). However, to do this, we also needed to test gorillas and drills on the original task (Salwiczek et al., 2012), in order to give them equivalent experience. Therefore, gorilla and drill subjects were first tested on the plate task (Study 1) before being tested in either the color task (Study 2) or the cup task (Study 3). Because some orangutan subjects had previous experience with the task from Salwiczek et al. (2012), we did not replicate the plate task with them. We predicted that gorillas and drills, like capuchins and orangutans tested previously, would fail in the plate task. However, we expected all three species, like capuchins, to solve the color and cup tasks, because the cues to solve the tasks were designed to be relevant to primates.

3.2 Methods

3.2.1 Subjects and Housing

We tested six orangutans, nine western lowland gorillas, and six drill monkeys, housed in social groups at Zoo Atlanta, Atlanta, GA, USA (for group compositions, see Table 4). All subjects were tested while off exhibit. Orangutan tests subjects were one Sumatran male (Dumadi), four Sumatran females (Madu, Biji and Blaze, who was tested with her 2-year-old infant male), one Bornean male (Satu) and one hybrid male (Chantek). Three of them (Satu, Madu and Chantek) have had previous experience with Salwiczek et al. (2012)’s original plate study. All tests subjects were also tested in a computerized-touchscreen task during the period of our study, and some had participated in other cognitive tests prior this one (e.g., Brosnan, Flemming, Talbot, et al., 2011; Diamond, Stoinski, Mickelberg, et al., 2016; Flemming, Jones,
Mayo, et al., 2012; Talbot, Mayo, Stoinski, & Brosnan, 2015). One female, Madu, was reared in a computer-enriched environment at the Georgia State University Language Research Center, where she participated in various cognitive tests (e.g., see Beran, 2002; Washburn, Gulledge, James, & Rumbaugh, 2007). In addition, one male, Chantek, had previous experience using sign language to communicate with humans (see Miles, 1990).

Gorilla tests subjects were six bachelor males (Jasiri, Kidogo, Kekla, Stadi, Charlie and Mbeli), two geriatric females (Choomba and Shamba) and one geriatric male (Ozzie). Although some subjects have been the focus of a large number of behavioral studies, their experience with cognitive tests prior ours was relatively limited (e.g., Anderson, Stoinski, Bloomsmith, et al., 2005; Drayton, Brosnan, Carrigan, & Stoinski, 2013; Stoinski, Wrate, Ure, & Whiten, 2001).

Drill tests subjects were one male (Bobby) and five females (Inge, Drew, Lucy, Achi and Amaka). Although they interact daily with their human keepers in the context of feeding and training, they have never received any cognitive test prior to ours.

All subjects had indoor/outdoor access and extensive material enrichment (climbing structures, ropes and swings, barrels, and other toys), and were tested off-exhibit. All subjects were fed their usual diet consisting of primate chow, fruits and vegetables prior or after testing. In addition, feeding enrichment was provided on a daily basis as part of the husbandry routine. At no time were the subjects ever food or water deprived. Studies involved a single subject at a time. All subjects participated voluntarily, being called in from their social groups and tested in one of the indoor dens of their living area. If possible subjects were separated from other individuals to limit distractions (however, unweaned infants always accompanied their mothers). All procedures used in this research were approved by the Scientific Review Committee of Zoo
Atlanta and met the ethical standards of the United States. Zoo Atlanta is fully accredited by the Association of Zoos and Aquariums (AZA).

### 3.2.2 General Procedure

The experimental design was based on the studies of Bshary and Grutter (2002), Salwiczek et al. (2012), and Prétôt et al. (2016b). Subjects had to choose between two stimuli, each assigned to one of two specific roles, permanent or ephemeral. Choosing the permanent stimulus always resulted in an immediate reward and this stimulus was available throughout the trial. The ephemeral stimulus offered the same immediate reward, but only if it was the first one the subjects chose. If subjects chose the permanent stimulus first, the ephemeral stimulus was withdrawn out of reach of the subject (for a total of only one reward). Thus, the optimal outcome was to pick the ephemeral plate first for a first reward, which allowed the subject to also obtain the permanent stimulus for an additional reward (for a total of two rewards). The side on which each stimulus was presented for each trial was randomly determined, but counterbalanced within a session so that they were on each side an equal number of times, but with no more than three trials in a row on the same side. Each subject received 10 sessions of 10 trials each (unless otherwise noted; see Learning Criterion section below). Subjects that reached preference criterion in the initial learning test were then tested on a reversal test. The procedure was identical to the initial test, but the role of each stimulus was reversed (i.e., the ephemeral stimulus became the permanent one, and vice-versa).

The presentation of the stimuli and the choices slightly differed between the ape and the monkey studies. In compliance with the safety procedures, orangutans and gorillas were presented the two stimuli out of reach (the distance varied between 40 and 60cm), and they
indicated their choice by gesturing to it with their hand (for similar procedures, see Brosnan et al., 2011; Flemming et al., 2012). They always received the choice that they indicated. Although this is different than in previous work, as the subjects were unable to grab the reward themselves, this was essential due to practical demands of the zoo environment and a procedure with which all of the apes were familiar. Occasionally they gestured at both stimuli (e.g., by stretching out both hands or moving one hand from one stimulus to the other), in which case the experimenter withdrew both stimuli and restarted the trial. As with other such studies, immediately after a choice, the selected option was moved forward and delivered to the subjects through an opening located below the front mesh.

Unlike orangutans and gorillas, drill monkeys made their decisions using a procedure more similar to the one of capuchin monkeys in prior studies (see Prétôt et al., 2016b; Salwiczek et al., 2012). However, as subjects were tested directly on their front enclosure, the two stimuli were presented such that they were both visible, but far enough away from the mesh that they could not be reached simultaneously. Despite the distance between the two stimuli, however, some subjects became very quick at selecting the foods, often trying to obtain both stimuli. As with the apes, if a subject attempted to select both stimuli simultaneously, the stimuli were quickly removed and the trial restarted.

For all species, the foods were presented on two plates that were located at equal distance from the subject and that the experimenter could move independently. The trial started once the subjects faced the two food options. If the subject chose the permanent stimulus first, the ephemeral stimulus was quickly withdrawn out of reach of the subject. Subjects were tested up to four sessions a week, but never in more than one session per day. The inter-trial interval (ITI) was 30 seconds, chosen because, in previous testing, it was the ITI at which the orangutans did
the best (Salwiczek et al., 2012). We used 750 mg banana flavored pellets for orangutans, 1g Piña Colada flavored pellets for gorillas (unless otherwise noted), and raw shelled peanuts for drills. These foods could be dyed different colors, as needed for Experiment 2.

**Learning Criterion and Statistics**

We used the same learning criterion as in earlier work (Prêtôt et al., 2016a & 2016b). This criterion was based on a two-tailed binomial/sign test, such that subjects were considered to have solved the test when they met the preference criterion for choosing the payoff-maximizing ephemeral option. Subjects met the preference criterion when they chose the stimulus (1) 10/10 trials on one session (but only if the subject selected each of the two stimuli in at least one trial of a previous session in the initial learning test, in order to ensure that they had experience with both outcomes), (2) 9/10 or 8/10 trial on two consecutive sessions or (3) 7/10 trials on three consecutive sessions. Although we planned to limit subjects to 10 sessions, if a subject chose either option 9/10, 8/10 or 7/10 in the last session (i.e., Session 10), it was given another session (or two, in the case of sessions with 7/10, followed by a second session with 7 or more out of 10) of 10 trials to maximize its chances of reaching criterion. They failed if they either developed a preference for the permanent option (using the same criterion as for the ephemeral preference) or did not develop any preference within 100 trials. Subjects who developed a preference for either option were then tested on a reversal test so that all subjects had similar experience prior to subsequent studies.

To compare subjects’ performance between tasks, we used a two-tailed Fisher’s exact test that allowed us to compare the number of individuals who met criterion for preferring the ephemeral option to those who did not; this way, we could include data from all subjects, even
those who developed a preference for the permanent option. To compare subjects’ performance across species, we used a two-tailed Kruskal-Wallis test and a Scheffe’s test (as a post-hoc test; see Salwiczek et al., 2012). Subjects that failed were assigned 110 trials for the purposes of analysis, which was one session greater than the minimum number of trials at which testing ceased (see Prétôt et al., 2016a). For this comparison, we used the data from subjects who either preferred the ephemeral option or did not develop any preference, but did not use the data from the subjects who significantly preferred the permanent stimulus (although we reported them), because they were difficult to interpret (see Prétôt et al., 2016b). Indeed, if subjects initially preferred the ephemeral stimulus and then reversed to prefer the newly ephemeral (formerly permanent) option, this would indicate that they were consistent in their preference. Similarly, reversing to maintain a preference for the newly permanent (formerly ephemeral) option would indicate a consistent preference for the permanent stimulus. However, maintaining a preference for the newly ephemeral (formerly permanent) option would be uninformative, because we would not know whether subjects preferred that particular stimulus (e.g., based on design or color), or whether they recognized the utility of choosing the ephemeral option because it became the food-maximizing option after the two options switched role. We used a within-subjects design for the comparisons between tasks, and a between-subjects design to compare performance between species.

3.3 Study 1: Plate Task in Gorillas and Drills

Procedure

As in Prétôt et al. (2016b), we replicated the original plate task to give gorillas and drills the same experience as three of the orangutans, and compare how subjects did on the subsequent
studies. In the task, the nine gorillas and the six drills had a choice between two food items placed on two plates that were similar in size and shape, but differed in color and pattern (for a better comparison, the design of the plates were kept similar across species; see Figure 4 for details of plates in all studies to date). Plate 1 had two yellow triangles and one green circle, while Plate 2 had two blue rectangles (one of them striped). The gorilla plates were approximately 60 x 12 x 2cm, whereas the drill plates were approximately 34 x 10 x 0.5cm (i.e., scaled for body size). The plates were equipped with a handle at the bottom to make the manipulations easier. Five of the nine gorillas and three of the six drills were tested with Plate 1 as the ephemeral choice, while the other four gorillas and three drills were tested with Plate 2 as the ephemeral choice. Note that one male gorilla (Stadi) was tested using purple grapes rather than pellets, because he would not work for pellets.

Results

Individual data for all subjects are in Table 5.

Initial learning phase

None of the gorilla and drill subjects tested on the initial learning phase solved the task ($N_{\text{gorillas}} = 9$, $N_{\text{drills}} = 6$). Given that no subject met criterion, no subject was tested on the reversal test and the study ended.

Cross-species comparison (between-subjects)

We compared the performance of gorillas and drills in the plate task with the performance of the capuchin monkeys and the orangutans in Salwiczek et al. (2012). There was
no significant difference in learning speed performance between the species (Kruskal-Wallis exact test: df = 3, H = 0, P = 1.00; see Figure 5).

Discussion

Consistent with our prior study (Salwiczek et al., 2012), all the primates in the current study failed to solve the plate task. One possible explanation for their struggle is that they had difficulty recognizing which cue was the important one for the task (i.e., that they should be distinguishing based on plate color/pattern). Unlike cleaner fish, which choose food based on which client fish species they are feeding on, primates grab their foods directly, so the substrate on which the food is presented may not be relevant to them. To test this hypothesis, in the subsequent studies, we retested subjects in two versions of the original plate task with cues that we hypothesized were more relevant to primates than fish.

3.4 Study 2: Color Task

Procedure

The basic procedure was identical to the plate task described in Study 1, except that the plates were similar in color and pattern, but the foods differed in color. We tested six orangutans, and the same gorilla and drill subjects as in Study 1. We used the same plates as in Study 1 (but eliminated colors and patterns) for both gorillas and drills. Orangutans were tested using the same plates as the gorillas. Rewards were colored pink (Color 1) or black (Color 2) with spray food coloring (Figure 4). Note that two gorillas were tested with different combinations of foods and colors. One female (Shamba) who was originally presented with pink pellets as ephemeral foods and black pellets as permanent foods, would not eat any of the black pellets in Session 1.
Consequently, we tested her with a new combination of colors, blue and red, in Session 2. In addition, one male (Stadi) was tested using a combination of green and black grapes. Although we strived to provide all tests subjects with foods that were equally attractive, we could not know whether this subject perceived any difference in flavor between the two grapes; however, we assumed that if he showed any, we would have seen it early in the test, which was not the case (see Table 6). Five of the nine gorillas, and three of the orangutans and drills were tested first in this study and subsequently on Study 3 (cup task), while the other four gorillas and three orangutans and drills were tested first on Study 3 and subsequently on this study.

Results

Individual data for all subjects are in Table 6.

Initial learning phase

Four of the six orangutans tested on the initial learning phase solved the task within 60 trials (range 20-60 trials, mean ± SD = 37.50 ± 17.08 trials). Three of the nine gorillas reached preference criterion within 120 trials; two individuals solved the task and preferred the ephemeral stimulus in 40 and 120 trials, while one developed a preference for the permanent stimulus in 110 trials. None of the drills tested on the initial learning phase solved the task ($N_{drills} = 6$).

Reversal learning phase

Three of the four orangutans that solved the initial learning task also solved the reversal task within 90 trials (range 50-90 trials, mean ± SD = 70.33 ± 20.01 trials), while the remaining
individual did not reach criterion. The two gorillas that reached preference criterion for the ephemeral stimulus in the initial learning task did not solve the reversal task, while the gorilla that preferred the permanent stimulus never developed any preference for either stimulus in the reversal task.

_Plate vs. color tasks (within-subjects)_

For the analysis, we only included subjects that completed both the plate and the color tasks (three of the six orangutans, the nine gorillas and six drills). Note that orangutans’ small sample size did not allow for a statistically significant analysis. The three primate species did not show any difference in performance between the color task (two of three orangutans and two of nine gorillas succeeded, while all six drills failed in the task) and the plate task (all subjects of all species failed in the task; Fisher’s exact test, orangutans: $P = 0.40$; gorillas: $P = 0.47$; drills: $P = 1.00$; Figure 5).

_Cross-species comparison (between-subjects)_

We compared the performance of orangutans, gorillas and drills in the color task with the performance of capuchin monkeys in Prêtôt et al. (2016b). There was a significant difference in learning speed between the species (Kruskal-Wallis exact test: $df = 3$, $H = 14.59$, $P < 0.001$; Figure 5). Post-hoc comparisons revealed that the capuchins performed better than the drills and the gorillas (Scheffe’s test: all $Ps \leq 0.01$), but not the orangutans ($P = 0.90$). Orangutans also outperformed the drills ($P = 0.04$) and showed a tendency to perform better than the gorillas ($P = 0.08$), while the drills were not distinguishable from the gorillas’ behavior ($P = 0.97$).
Discussion

Our results supported partially our hypothesis. As predicted, orangutans improved their performance in the task as compared to the plate version (although our sample size was too small for statistical significance), supporting our prior results with capuchin monkeys. Together, these species performance indicates that changing the task cues from the color of the plate to the color of the food helps them to solve the task, possibly because it is a more salient cue. Gorillas and drills, however, did not do better in the task, suggesting that this modification to the task does not work universally among primates, or that there another factor is influencing the responses of the gorillas and drills. We discuss why the task might still be challenging for some primates in the General Discussion section.

3.5 Study 3: Cup Task

Procedure

The procedure was identical to Study 1, except that the two foods were hidden under two different opaque cups (used in Prétôt et al., 2016b; see Figure 4) so that subjects could see the experimenter bait the cup, but could not see the food during the choice. One cup was yellow with one red dot on each side (Cup 1) and the other was purple with a blue triangle on each side (Cup 2). The two plates that carried the cups were the same as in Study 2 (unmarked and identical). At the beginning of each trial, the subjects saw the experimenter hiding a same food item under each cup. They were then presented the choice between the two cups. Once gorillas and orangutan subjects made a choice (by gesturing at the cup they wanted), the experimenter lifted the cup for them, before pushing the plate forward to give subjects access to the food underneath the front mesh. If the ephemeral cup was selected first, the experimenter removed the ephemeral
cup first, gave the subject access to the food, and only then did the same for the permanent cup. Again, drills were presented the cups so that they could make a choice either by lifting up the cup themselves and take the food through the mesh, or by touching the cup (which varied across subjects), in which case the experimenter lifted the cup for them and let them take the food.

Results

Individual data for all subjects are in Table 7.

Initial learning phase

One of the six orangutans tested on the initial learning phase solved the task within 30 trials, and one of the nine gorillas reached preference criterion for the permanent stimulus with 20 trials. As in Study 2, none of the drills tested on the initial learning phase solved the task ($N_{drills} = 6$).

Reversal learning phase

The orangutan who solved the initial task solved the reversal task within 111 trials, while the gorilla who reached preference criterion for the permanent stimulus never developed any preference for either stimulus in the reversal task.

Plate vs. cup tasks (within-subjects)

As in Study 2, we only included subjects that completed both the plate and the cup tasks (three of the six orangutans, the nine gorillas and six drills). Again, the three primate species did not show any difference in performance between the cup task (all three orangutans, nine gorillas
and six drills failed in the task) and the plate task (all subjects of all species failed in the task; Fisher’s exact test: all $P$s = 1.00; Figure 5).

**Color vs. cup tasks (within-subjects)**

For the analysis, we included subjects that completed both the cup and the color tasks (the nine gorillas and the six orangutans and drills). Subjects from the three primate species did not show any difference in the initial learning phase between the color task (four of six orangutans and two of nine gorillas succeeded, while all six drills failed in the task) and the cup task (one of six orangutans succeeded, while all nine gorillas and six drills failed in the task; Fisher’s exact test, orangutans: $P = 0.24$; gorillas: $P = 0.47$; drills: $P = 1.00$; Figure 5).

**Cross-species comparison (between-subjects)**

We compared the performance of orangutans, gorillas and drills in the cup task with the performance of the capuchin monkeys in Prêtôt et al. (2016b). There was a significant difference in learning speed between the species in (Kruskal-Wallis exact test: $df = 3$, $H = 18.54$, $P < 0.001$; Figure 5). Post-hoc comparisons revealed that the capuchins outperformed all three species (Scheffé’s test: all $P$s ≤ 0.02), which were not distinguishable from one another (all $P$s ≥ 0.77).

**Discussion**

Interestingly, none of the primate species improved their performance in the cup task, in which the foods were hidden, as compared to the plate task, in which the foods were visible. We had predicted that they would do better due to the presumably lessened prepotent response associated with visible rewards (since the food was hidden). The overall poor performance of
orangutans in the cup task was particularly surprising, given their well-known inhibitory skills (e.g., see Amici, Aureli, & Call, 2008; Parrish, Perdue, Stromberg, et al., 2014; Shumaker, Palkovitch, Beck, et al., 2002; Vlamings, Hare, & Call, 2010). Although we do not know why orangutans performed so poorly in the task, our results nonetheless support the hypothesis that changing the task cues influences performance differently across primates.

3.6 General Discussion

In a prior study, we found that capuchin monkeys performed better in a dichotomous choice task after the cues were changed from being primarily motivated by the cleaner fish ecology towards cues that we predicted were more relevant to primates’ behavior (Prêtôt et al., 2016b). Here, we tested whether these results generalized across the primate taxon by testing three additional species in the tasks. As predicted, gorillas and drills failed in the plate task, which was consistent with our prior findings on both capuchin monkeys and orangutans. Besides this, however, the choices of the orangutans, gorillas and drills largely deviated from the capuchins’ results, despite the fact that we expected all three species, like capuchins, to do better because these tasks were designed specifically to make use of cues that might be familiar to primates (the color task) or minimize cues that might be distracting (the cup task). As predicted, orangutans did perform overall better in the color task than in the original plate ask, which was also how the capuchins had behaved. However, unlike the capuchins, they did not improve relative to the plate task in the cup task, in which the food was hidden. Moreover, the gorillas and drills performed poorly on all tasks.

Although we can only speculate on why these primates in general did so poorly (excepting the orangutans in the color task), we see at least two possible explanations. First, it is
possible that the cues were (inadvertently) more salient to capuchins than to the other species, which would have meant that the capuchins were better than the other primates at finding the cues associated with success in the tasks. We cannot rule this out; however, we also cannot think of any a priori reason why it should be the case. All four species are diurnal primates with good color vision, and all eat fruits (which has been hypothesized as one driver of primates’ reliance color vision; e.g., see Osorio & Vorobyev, 1996; Osorio, Smith, Vorobyev, & Buchanan-Smith, 2004; Regan, Julliot, Simmen, et al., 2001). Indeed, color is presumably even more important in the drills and great apes than in the capuchins, in which females have concealed ovulation and males lack conspicuous sexual coloration (e.g., see Carnegie, Fedigan, & Ziegler, 2005; Carosi, Heistermann, & Visalberghi, 1999).

Second, and, we believe, most likely, the primates might have done poorly in the tasks for reasons that were external to the task. In particular, none of the subjects that we tested in the current study have a lot of experience with cognitive testing, especially relative to the capuchins. Perhaps more tellingly, the gorillas and drills have even less experience than the orangutans, and also did relatively worse than the orangutans. Indeed, while some of our gorilla subjects have had participated in studies prior ours, most of these studies were behavioral, and only a few have used experimental procedures. This was, to our knowledge, the first cognitive study ever performed on these drill subjects. The orangutans’ increased experience with testing may explain why they did better than the gorillas and drills on the color task.

We do not know why the orangutans performed poorly in the cup task, in which foods were hidden, while they did relatively well on the color task, in which the foods were colored rather than the plates. On one hand, given evidence that orangutans generally possess advanced inhibitory skills as compared to other primates (e.g., see Amici et al., 2008; Shumaker et al.,
2002; Vlamings et al., 2010), we had expected them to do well in both tasks. On the other hand, performing poorly in the cup task despite high inhibitory control may have indicated that inhibition was not the key issue that affects subjects’ responses in the task (at least in orangutans). If it was, then the orangutans may not have needed the extra boost from hiding the foods to solve it.

Overall, what do our results tell us about species’ performance in the task? Primarily, our findings indicate that performance in the task vary greatly across species of the same taxon, which indicates, not surprisingly, that these responses are driven by very specific features of individuals species’ ecologies rather than broad generalities that apply to an entire taxon. Unfortunately, the (very likely) influence of external factors, including subjects’ levels of experience with testing procedures, did not allow us to determine with much certainty what was the underlying cause of such variability (although these results may indicate that difficulty with inhibition is not a key aspect that drives subjects’ responses, at least in orangutans). Most importantly, these results (re-) emphasize the importance of testing multiple species in similar tasks whenever possible (as we did here and in prior studies), which should facilitate identifying the factors that may influence species’ performance in the tasks. Indeed, testing any one of these species separately would have led to a very different set of conclusions than we can now reach.
3.7 Tables

Table 4 Orangutan-Gorilla-Drill Subjects General Information

Social group, age, gender and studies conducted

<table>
<thead>
<tr>
<th>Subject</th>
<th>Group</th>
<th>Age</th>
<th>Gender</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orangutans</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satu$^1$</td>
<td>1</td>
<td>12</td>
<td>M</td>
<td>2-3</td>
</tr>
<tr>
<td>Madu$^1$</td>
<td>2</td>
<td>32</td>
<td>F</td>
<td>2-3</td>
</tr>
<tr>
<td>Chantek$^1$</td>
<td>3</td>
<td>38</td>
<td>M</td>
<td>2-3</td>
</tr>
<tr>
<td>Dumadi$^2$</td>
<td>3</td>
<td>9</td>
<td>M</td>
<td>2-3</td>
</tr>
<tr>
<td>Blaze$^2$</td>
<td>4</td>
<td>19</td>
<td>F</td>
<td>2-3</td>
</tr>
<tr>
<td>Biji$^2$</td>
<td>5</td>
<td>45</td>
<td>F</td>
<td>2-3</td>
</tr>
<tr>
<td><strong>Gorillas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jasiri</td>
<td>1</td>
<td>17</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Kidogo</td>
<td>1</td>
<td>17</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Kekla</td>
<td>2</td>
<td>26</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Stadi</td>
<td>2</td>
<td>24</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Charlie</td>
<td>2</td>
<td>19</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Mbeli</td>
<td>3</td>
<td>13</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Shamba</td>
<td>4</td>
<td>56</td>
<td>F</td>
<td>1-3</td>
</tr>
<tr>
<td>Ozzie</td>
<td>4</td>
<td>54</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Choomba</td>
<td>4</td>
<td>52</td>
<td>F</td>
<td>1-3</td>
</tr>
<tr>
<td><strong>Drills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inge</td>
<td>1</td>
<td>30</td>
<td>F</td>
<td>1-3</td>
</tr>
<tr>
<td>Bobby</td>
<td>1</td>
<td>25</td>
<td>M</td>
<td>1-3</td>
</tr>
<tr>
<td>Drew</td>
<td>2</td>
<td>19</td>
<td>F</td>
<td>1-3</td>
</tr>
<tr>
<td>Lucy</td>
<td>2</td>
<td>18</td>
<td>F</td>
<td>1-3</td>
</tr>
<tr>
<td>Achi</td>
<td>2</td>
<td>16</td>
<td>F</td>
<td>1-3</td>
</tr>
<tr>
<td>Amaka</td>
<td>2</td>
<td>14</td>
<td>F</td>
<td>1-3</td>
</tr>
</tbody>
</table>

$^1$ Subjects previously tested in the plate task (Salwiczek et al., 2012).

$^2$ Subjects only tested in Studies 2 and 3.
**Table 5** Two-Choice Gorilla-Drill Plate Task

Number of trials needed for each subject to solve the plate task for the initial learning phase and the reversal phase in Study 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal</th>
<th>Ephemeral option</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gorillas</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidogo</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Shamba</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Charlie</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
<tr>
<td>Choomba</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
<tr>
<td>Jasiri</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
<tr>
<td>Ozzie</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
<tr>
<td>Stadi¹</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Kekla</td>
<td>&gt;97²</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Mbeli</td>
<td>&gt;93²</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td><strong>Drills</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achi</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Inge</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Drew</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 1</td>
</tr>
<tr>
<td>Amaka</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
<tr>
<td>Bobby</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
<tr>
<td>Lucy</td>
<td>&gt;100</td>
<td>—</td>
<td>Plate 2</td>
</tr>
</tbody>
</table>

Stimulus options (Plate 1 vs. Plate 2) are shown in Figure 4. Subjects that did not reach preference criterion in the initial learning phase were not further tested on the reversal phase.

¹ Subject was tested with grapes.

² Subjects did not complete one session.
Table 6 Two-Choice Orangutan-Gorilla-Drill Color Task

Number of trials needed for each subject to reach preference criterion for the initial learning phase and the reversal phase in Study 2

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal</th>
<th>Ephemeral option</th>
<th>Initial preference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orangutans</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chantek¹,²</td>
<td>20</td>
<td>50</td>
<td>Color 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Satu¹,²</td>
<td>40</td>
<td>90</td>
<td>Color 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Dumadi</td>
<td>60</td>
<td>71³</td>
<td>Color 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Biji</td>
<td>30</td>
<td>&gt;100</td>
<td>Color 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Madu¹</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Blaze²</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
<tr>
<td><strong>Gorillas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mbeli²</td>
<td>40</td>
<td>&gt;100</td>
<td>Color 1</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Stadi⁴</td>
<td>120</td>
<td>&gt;100</td>
<td>Green</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Shamba⁵</td>
<td>110</td>
<td>&gt;100</td>
<td>Blue</td>
<td>Permanent</td>
</tr>
<tr>
<td>Charlie</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Choomba²</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Jasiri²</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Kidogo</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
<tr>
<td>Ozzie</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
<tr>
<td>Kekla²</td>
<td>&gt;99³</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
<tr>
<td><strong>Drills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lucy</td>
<td>&gt;120</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
<tr>
<td>Inge²</td>
<td>&gt;104³</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
<tr>
<td>Amaka</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Bobby</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Drew²</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 1</td>
<td>—</td>
</tr>
<tr>
<td>Achi²</td>
<td>&gt;100</td>
<td>—</td>
<td>Color 2</td>
<td>—</td>
</tr>
</tbody>
</table>

Stimulus options (Color 1 vs. Color 2) are shown in Figure 4. Subjects that did not reach preference criterion in the initial learning phase were not further tested on the reversal phase.

¹ Subjects previously tested in the plate task (Salwiczek et al., 2012).
² Subjects previously tested in Study 3.
³ Subject did not complete one or more sessions.
⁴ Subject was tested with grapes.
⁵ Subject was tested with different color food.
Table 7 Two-Choice Orangutan-Gorilla-Drill Cup Task

Number of trials needed for each subject to reach preference criterion for the initial learning phase and the reversal phase in Study 3

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal</th>
<th>Ephemeral option</th>
<th>Initial preference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orangutans</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dumadi&lt;sup&gt;1&lt;/sup&gt;</td>
<td>30</td>
<td>111&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Cup 2</td>
<td>Ephemeral</td>
</tr>
<tr>
<td>Blaze</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Madu&lt;sup&gt;1,3&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Satu&lt;sup&gt;3&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Chantek&lt;sup&gt;3&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Biji&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td><strong>Gorillas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shamba&lt;sup&gt;1&lt;/sup&gt;</td>
<td>20</td>
<td>&gt;100</td>
<td>Cup 1</td>
<td>Permanent</td>
</tr>
<tr>
<td>Choomba</td>
<td>&gt;110</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Stadi&lt;sup&gt;1,4&lt;/sup&gt;</td>
<td>&gt;110</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Jasiri</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Kidogo&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Ozzie&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Charlie&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Mbeli</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Kekla</td>
<td>&gt;98&lt;sup&gt;2&lt;/sup&gt;</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td><strong>Drills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inge</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Amaka&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Bobby&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 1</td>
<td>—</td>
</tr>
<tr>
<td>Achi</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Drew</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
<tr>
<td>Lucy&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&gt;100</td>
<td>—</td>
<td>Cup 2</td>
<td>—</td>
</tr>
</tbody>
</table>

Stimulus options (Cup 1 vs. Cup 2) are shown in Figure 4. Subjects that did not reach preference criterion in the initial learning phase were not further tested on the reversal phase.

<sup>1</sup> Subjects previously tested in Study 2.

<sup>2</sup> Subject did not complete one session.

<sup>3</sup> Subjects previously tested in the plate task (Salwiczek et al., 2012).

<sup>4</sup> Subject was tested with grapes.
3.8 Figures

Figure 4 Orangutan-Gorilla-Drill Sets of Stimuli

Pairs of stimuli (named 1 and 2) used in the plate, color and cup tasks.
Figure 5 Orangutan-Gorilla-Drill Initial Learning Phase

The number of trials required for capuchins, drills, gorillas and orangutans to develop a preference for the ephemeral stimulus in the plate task (capuchins’ data based on Salwiczek et al., 2012) and to develop a preference for either stimulus in the color and cup tasks (capuchins’ data based on Prétôt et al., 2016b). Each dot represents one individual; black dots indicate a preference for the ephemeral stimulus, while white dots indicate a preference for the permanent stimulus. For capuchins and orangutans, the grey dots designate new subject not previously tested in the plate task. As in our previous work (Prétôt et al., 2016a & 2016b), criterion was set at 10 sessions, but if subjects chose one stimulus at least 7 out of 10 times on the 10th session, they were allowed to continue until they either met criterion or failed to do so; “failed” indicates subjects that did not reach preference criterion.
4 STUDY 3: COMPARING SPECIES DECISIONS IN A DICHOTOMOUS CHOICE TASK: ADJUSTING TASK PARAMETERS IMPROVES PERFORMANCE IN MONKEYS


4.1 Introduction

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

Recently, Salwiczek et al. (2012) conducted a cross-taxon study designed to compare the performance of cleaner fish with several nonhuman primate species (capuchin monkeys, chimpanzees and orangutans) on a seemingly simple two-choice discrimination task derived from the cleaner’s natural habitat (Bshary & Grutter, 2002). The goal of this study was to compare two distantly related taxa that have both evolved complex social and foraging behaviors
(especially those related to cooperation between individuals) to see if the primates could solve it as well as the fish despite the lack of ecological relevance. If this had been the case, it would have indicated that the primates were able to use their relatively more developed cognitive ability to overcome their lack of experience. However, it was not; the adult fish outperformed all of the primate species (and juvenile cleaner fish). This was surprising, and indicated that the task might not be as easy as it seemed. For example, the primates might have failed at the task because of the general tendency animals show to prefer the first of two rewards that they encounter (e.g., Davison & Baum, 2000; Timberlake, Gawley, & Lucas, 1987; also see Salwiczek et al., 2012). In light of this literature, it is also surprising that the cleaner fish performed as well as they did on the task. Unfortunately, the study design did not allow us to further explore which factors made the task difficult for the primates. Thus, the goal of the current study was to test two potential explanations, first, that the primates would do better in a computerized task that minimized extraneous cues and the presence of the experimenter, and second, that they would do better in a task designed to more closely mimic their ecology.

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

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

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

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

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

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

4.2.1 Subjects and Housing

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

Capuchins

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

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

Macaques

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

4.2.2 Apparatus

Subjects were tested using the Language Research Center’s Computerized Test System comprising a personal computer, digital joystick, color monitor, and pellet dispenser (Evans, Beran, Chan, et al., 2008; Richardson, Washburn, Hopkins, et al., 1990; Washburn & Rumbaugh, 1992). Monkeys manipulated the joystick to produce isomorphic movements of a computer-graphic cursor on the screen. Contacting appropriate computer-generated stimuli with the cursor provided them a 45-mg (capuchins) or 94-mg (macaques) banana-flavored chow pellet (Bio-Serv, Frechtown, NJ) using a pellet dispenser that interfaced to the computer through a digital I/O board (PDISO8A; Keithley Instruments, Cleveland, OH). All monkeys had previously participated in multiple cognitive and behavioral experiments involving this computerized test system (e.g., Beran, 2007, 2008; Beran & Parrish, 2012; Beran, Evans, Klein, & Einstein, 2012; Beran, Harris, Evans, et al., 2008a; Brosnan, Wilson, & Beran, 2012; Evans & Beran, 2012). Monkeys were tested on the computerized apparatus while in their testing enclosures.
4.2.3 General Procedure

The experimental design was based on the studies of Bshary & Grutter (2002) and Salwiczek et al. (2012). Two icons, one representing an ephemeral choice and one a permanent choice (see Table 8), were presented on a computer screen and subjects had to choose between them. If the permanent stimulus was chosen first, the subject could not make another choice, but if the subject chose the ephemeral stimulus first, they could then choose the permanent one as well, resulting in one reward for each choice. The optimal outcome was always to select the ephemeral stimulus first, which ultimately gave the subject two rewards. The location of each stimulus was randomly determined, but in conditions in which only two locations were used (e.g., left/right), the stimulus position was counterbalanced within a session so that they were presented on each side an equal number of times, with the constraint that there were no more than three trials in a row on the same side (as in Salwiczek et al., 2012; see details in the procedures, below). Each trial ended once the monkeys made a choice between the options. Each subject received 10 sessions of 10 trials each, unless otherwise noted. The inter-trial time interval varied depending on the task (see detail in the procedures, below).

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

Learning Criterion and Statistics

In Study 1, we initially used the same criterion as was used in Salwiczek et al. (2012). This criterion was based on a two-tailed signs test, such that significance was reached when a subject developed a preference for the ephemeral source in 10/10 or 9/10 trials on one session, in \( \geq 8/10 \) on two successive sessions, or in \( \geq 7/10 \) trials on three consecutive sessions; however, unlike primates in the previous study, some monkeys developed a preference for the permanent over the ephemeral stimulus, so for Study 2, developing a preference for either stimulus met criterion. Two additional changes to Study 2 were that 1) we still used the 10/10 criterion on one unique session, but only if the subject selected each of the two stimuli in at least one trial of a previous session to ensure that they had experience with both outcomes (note that this did not apply to the reversal phase, because a subject had experienced both options in the initial phase), and 2) we required two sessions of 9/10 because it came to our attention that, while a single 9/10
session is used frequently as a criterion in cognitive and behavioral testing, Monte Carlo simulations show that this criterion is not sufficiently strict (see Grant, 1946).

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

Due to our small sample sizes, we also provided a measure of the effect size (Pearson’s correlation coefficient, $r$; see Rosenthal, 1991). Cohen (1988)’s classification of effect size magnitude was used, whereby $r \leq 0.09$ mean that there is “no effect” of the treatment, $r = 0.10-0.29$ is considered a “small effect”, $r = 0.30-0.49$ is a “moderate effect”, and $r \geq 0.50$ is a “large effect”.
4.3 Study 1

4.3.1 Procedure

Study 1 was carried out between July and December 2012 with the same 8 capuchin subjects tested in the original plate study (Salwiczek et al., 2012) and two additional subjects who had no experience with the original paradigm, and between August and November 2012 with macaques. At the beginning of each trial, two stimuli that differed in shape and color appeared simultaneously on the computer screen (the orientation of the stimuli differed across the versions of the task; see Table 8). The monkeys used a cursor (a red dot on the screen) to select one of the stimuli.

Study 1A

Subjects were tested on four versions of the task that varied on inter-trial time interval (ITI) and stimulus location until they met criterion (see Table 8 for a summary of procedures in the order in which the subjects experienced them). Once they met criterion, they were not tested on subsequent versions. Additionally, once they met criterion, they were given a reversal test using the same version of the task on which they had just passed criterion. We used a green pentagon and a purple cross stimuli in all the versions (see Table 8). Note that the ITI was the interval between stimulus presentations, thus the subject could have taken more time to actually make a choice. Rhesus monkeys sometimes failed to re-engage in the task after an ITI (this is why trial counts were not always multiples of 10), thus we added a tone concurrent with the presentation of a new stimulus to attract their attention in an attempt to ameliorate this problem. Capuchin monkeys typically completed their sessions and did not have a tone in any condition.
Version 1 had an ITI of five minutes and a left-right presentation of the stimuli (e.g., one stimulus was on each side of the screen); if they failed to meet criterion, subjects were tested in subsequent versions that were altered in ways that we hypothesized would improve performance. Version 2 had an ITI of only 1 minute, because in the previous study a decreased ITI increased performance on the task (see Salwiczek et al., 2012). For this version, we used a maximum of 50 trials because the three capuchin subjects tested on that version developed a side bias, which did not occur in other conditions. To avoid the subjects developing side biases, Version 3 was the same as Version 2, except that the stimuli were randomized across four positions that were either lined up on a left-right or up-down axis. Version 4 replicated Version 3, except with a shorter ITI of 30 seconds.

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

*Study 1B*
Each capuchin subject who passed both the initial and reversal learning phases of Study 1A was subsequently tested in a generalization task in which both ephemeral and permanent stimuli were replaced with new ones to see how quickly the subjects were able to apply the same rules to a new situation (Harlow, 1949; Schrier, 1984; Barros, Galvão, & Mcilvane, 2002). For this task, subjects were tested in the same version of Study 1A on which they had met criterion, with the exception that we did not use Version 1, with the ITI of 300 seconds, to maintain subjects’ motivation to participate. Similarly to Study 1A, subjects who did not succeed in the first version on which they were tested progressed to the next version in the order indicated in Table 8. To be consistent with Study 1A, we used a maximum of 50 trials in Version 2 of Study 1B. We used a blue circle-top rectangular and a brown L-shape stimuli in all versions (see Table 8). Note that one subject (Nkima) always chose the permanent icon after 60 trials on the reversal phase of Version 3, at which point we decided to terminate the study. Rhesus monkeys were not run on Study 1B.

4.3.2 Results

Individual data for all subjects are in Tables 9 (capuchins) and 10 (macaques). Subjects’ performance across the conditions is shown in Figures 6 (initial phase) and 7 (reversal phase).

Capuchin Study 1A

Initial Learning Phase. Six out of the nine subjects tested in the initial learning phase of Version 1 reached criterion within 100 trials (or 10 sessions; range: 20-100 trials, \(X = 45\) trials, SD ± 32.71). The three subjects who did not meet criterion in Version 1 also did not meet criterion in Version 2, but all did so in Version 3 (range: 30-96 trials, \(X = 62\) trials, SD ± 33.05).
Reversal Learning Phase. Two of the six subjects who had passed the initial learning phase of Version 1 reached criterion on the reversal phase within 110 trials (range: 60-110 trials, \( X = 85 \) trials, SD ± 35.36). The other four subjects met criterion in Version 2 (range: 10-40 trials, \( X = 25 \) trials, SD ± 12.91). One of the three subjects who reached criterion on the initial learning phase of Version 3 reached criterion on the Version 3 of the reversal phase (110 trials). Of the other two subjects, one met criterion in Version 4 (30 trials) and one never reached criterion.

Initial vs. Reversal Learning Phases. Subjects who met criterion in Version 1, and therefore were the quickest in the initial learning phase, performed better on the initial learning phase than on the reversal phase (Wilcoxon signed-ranks test: \( W = 0, Z = -2.20, N = 6, P = 0.03, r = -0.64 \)). This was different than in the original plate task, where capuchins met criterion faster on the reversal phase than the initial phase (see Salwiczek et al., 2012). Considering all subjects who passed in all versions, however, subjects showed no significant difference in speed of learning between the two phases (Wilcoxon signed-ranks test: \( W = 11, Z = -0.98, N = 8, P = 0.33, r = -0.25 \)).

Computer vs. Plate Tasks (within-subjects). We compared the performance of the eight capuchin subjects who participated in the initial learning phase of the original plate task (Salwiczek et al., 2012) with their own data on the computerized task (note that there is an order confound as all plate testing was completed prior to computerized testing, although three years separated the two experiments). Subjects performed only marginally better on the computerized learning than the previous plate task, although the effect size is at the high end of medium, indicating that this may be a real effect (Wilcoxon signed-ranks test: \( W = 0, Z = -1.84, N = 7, P = 0.13, r = -0.49 \); Figure 6). Interestingly, the monkeys’ performance on the reversal phase of the
computerized task was marginally worse than the reversal phase of the previous plate task (Wilcoxon signed-ranks test: $W = 2.50, Z = -1.70, N = 7, P = 0.13, r = -0.45$; Figure 7).

**Macaque Study 1A**

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

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

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

*Ephemeral vs. Permanent Choice.* There was no significant difference in learning speed between subjects who preferred the ephemeral icon and those who preferred the permanent one.
(Mann-Whitney $U$ test: $U = 4, Z = -1.16, N_{\text{ephemeral}} = 4, N_{\text{permanent}} = 4, P = 0.29, r = -0.41$; see Table 10).

**Capuchins vs. Fish**

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

**Capuchins vs. Macaques**

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

**Capuchin Study 1B**

*Initial Learning Phase.* We tested the subjects who passed both the initial and the reversal learning phases in any version of Study 1A on a generalization task designed to determine if they would extrapolate to a novel situation. Five out of six subjects who met criterion on Version 1 of Study 1A met criterion within 30 trials in Version 2 of Study 1B (we did not use Version 1, so subjects started on Version 2; range: 10-30 trials, $X = 22$ trials, SD ± 10.95). The only subject who did not reach criterion in Version 2 did so in Version 3 (30 trials).
The two subjects who did not reach criterion on the learning phase until Versions 3 and 4 met criterion on this generalization task in 30 trials (subject who met criterion on Version 3) and 80 trials (subject who met criterion on Version 4), respectively.

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

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

### 4.3.3 Study 1 Discussion

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

These results are very important as they again emphasize the importance of how a task is presented in the subjects’ ability to learn it. On the one hand, these results negate those from Salwiczek et al. (2012), as the monkeys were able to learn the task, but on the other hand, they emphasize that the monkeys could not learn using the same task parameters as the fish, but required alternate ones. This complicates comparative research as it requires studies to test both the different species on identical parameters (for a direct comparison, as we did in Salwiczek et al. (2012)) and also to test species using parameters that are adapted to that species’ preferences and abilities (as we did here) to rule out the possibility that the difference is not in the task itself, but in how the task is presented. Of course, variations that allow a species to learn to the best of their ability could also derive from differences in cognitive architecture, ecology, or how they interact with the task, and in most cases will likely differ on more than one dimension. Thus, it is important to test several of these possibilities in an attempt to pin down which feature(s) led to the species difference, which is what we do in the current paper.

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

In an attempt to address this confound, we also tested rhesus monkeys, who had not previously been tested. All macaques reached preference criterion, however while the capuchins were more likely to prefer the payoff maximizing choice, the macaques were equally likely to prefer to maximize payoffs or to avoid what likely appeared to be an unreliable stimulus (e.g., the ephemeral was often not available for them to choose). This was a surprising finding, given these rhesus macaques’ higher performance in cognitive tasks as compared to capuchin monkeys (e.g., Beran & Smith 2011; Beran, Perdue, & Smith, 2014; Beran, Smith, Coutinho, et al., 2009).

It is difficult to know exactly what subjects based their decision on, but we discuss this further in the General Discussion, including data from both Studies 1 and 2.

4.4 Study 2

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

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

4.4.1 Procedure

Study 2A was carried out between September and October 2014 with capuchins, and between February and March 2015 with macaques, while Study 2B was carried out between November and December 2015 with both monkey species. We tested eight of the nine capuchins used in Study 1, plus an additional subject with no previous experience with any version of the paradigm, and seven of the eight macaques used in Study 1. In order to compare the performance of monkeys in both Study 1 and Study 2, subjects were again tested in Version 1 of Study 1, using novel stimuli, except that when the subject touched the start button to begin the trial, both stimuli appeared and the ephemeral stimulus began moving from its initial position towards a lower corner of the screen. In Study 2B, both icons vibrated on the spot for 200 milliseconds before the ephemeral icon (still vibrating) began moving. The subjects had to track and touch the ephemeral stimulus first, before it moved off the screen and could not be “caught” (the cursor moved only slightly faster than the ephemeral stimulus). We used a yellow heart and a purple-pink ring stimuli for Study 2A, and a blue waive and an orange moon for Study 2B (see Table 8).
Note that two macaques (Obi and Luke) received two reversal sessions on one day (respectively in Study 2A and Study 2B). Also, one macaque (Luke) received an additional session because the data of the second session were lost.

4.4.2 Results

Individual data for all subjects tested on Study 2A and Study 2B are in Table 11. Subjects’ performance across the conditions is shown in Figures 6 and 7.

Capuchin Study 2A

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

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

Initial vs. Reversal Learning Phase. As in Study 1, there was a trend for capuchins to learn the initial learning phase in fewer trials than the reversal phase (Wilcoxon signed-ranks test: \( W = 2, Z = -2.05, N = 8, P = 0.06, r = -0.51 \)).
Within-subjects comparison with previous studies. There was no significant difference between Study 1A and Study 2A in monkeys’ learning performance in either the initial (Wilcoxon signed-ranks test: \( W = 2, Z = -1.48, N = 6, P = 0.19, r = -0.43 \)) or the reversal phase (Wilcoxon signed-ranks test: \( W = 0, Z = -1.34, N = 6, P = 0.50; r = -0.39; \) Figures 6 & 7). Unlike in Study 1A, however, subjects were faster to learn the initial phase of Study 2A than in the previous plate task (Wilcoxon signed-ranks test: \( W = 0, Z = -2.23, N = 6, P = 0.03, r = -0.64 \)).

Capuchin Study 2B

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

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

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

Within-subjects comparison with previous studies. Although nonsignificant, subjects showed a trend towards being faster on the initial phase of Study 2A than Study 2B (Wilcoxon
signed-ranks test: $W = 3, Z = -1.87, N = 8, P = 0.08, r = -0.47$; Figures 6 & 7). Like in Study 2A, subjects did not differ in performance between Study 2B and Study 1A (Wilcoxon signed-ranks test: $W = 7, Z = -0.74, N = 8, P = 0.59, r = -0.19$); however, unlike in Study 2A, they did not differ in performance between Study 2B and the previous plate task (Wilcoxon signed-ranks test: $W = 1, Z = -0.45, N = 7, P = 1.00, r = -0.12$).

**Macaque Study 2A**

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

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

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

*Study 1A vs. Study 2A (within-subjects).* Although all seven subjects performed better in Study 2A than in Study 1A, there were too few subjects who initially preferred the ephemeral icon in both studies for statistical significance; we note, however, that all three subjects who had an initial preference for the permanent icon in Study 1A preferred the ephemeral icon in Study 2A, and three of the four subjects who developed a preference for the ephemeral icon in both tasks learned the initial task more rapidly in Study 2A than in Study 1A (Wilcoxon signed-ranks test: $W = 3, Z = -1.87, N = 8, P = 0.08, r = -0.47$; Figures 6 & 7).
test: \( W = 2, Z = -1.10, N = 4, P = 0.38; r = -0.39 \); Figures 6 & 7). Three of these four subjects also learned the reversal phase of Study 2A more rapidly than in Study 1A (Wilcoxon signed-ranks test: \( W = 4, Z = -0.37, N = 4, P = 0.88; r = -0.13 \)).

**Macaque Study 2B**

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

*Reversal Learning Phase.* None of the four subjects who preferred the ephemeral icon in the initial phase reversed their preference within 100 trials. The two subjects who initially preferred the permanent icon reached a significant preference for their previously preferred icon, which now signaled the ephemeral reward, in 20 and 60 trials (although again we cannot interpret this result).

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

*Within-subjects comparison with previous studies.* Too few subjects initially preferred the ephemeral icon in the studies for statistical significance; however, we note that three of the four subjects who learned the initial task in both studies learned it more rapidly in Study 2A than in Study 2B, while the last one performed equally in both studies (Wilcoxon signed-ranks test: \( W = 4, Z = -0.37, N = 4, P = 0.88; r = -0.13 \)).
0, $Z = -1.60$, $P = 0.25$; $r = -0.57$; Figures 6 & 7). Three of the four subjects who performed the reversal task in both studies learned it more rapidly in Study 2A than in Study 2B, while the last one failed to learn it in Study 2B (Wilcoxon signed-ranks test: $W = 1$, $Z = -1.46$, $P = 0.25$; $r = -0.52$). Two of the three subjects who learned the initial task in both studies learned it more rapidly in Study 1A than in Study 2B, while the last one showed the opposite (Wilcoxon signed-ranks test: $W = 1$, $Z = -1.07$, $P = 0.50$; $r = -0.44$). The three subjects learned the reversal task more rapidly in Study 1A than in Study 2B (Wilcoxon signed-ranks test: $W = 0$, $Z = -1.60$, $P = 0.25$; $r = -0.65$).

**Capuchins vs. Fish**

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

**Capuchins vs. Macaques**

As in Study 1A, there was no significant difference in performance between monkey species in either Study 2A (Mann-Whitney $U$ test: $U = 15$, $Z = -1.53$, $N_{\text{capuchins}} = 8$, $N_{\text{macaques}} = 7$, $P = 0.16$; $r = -0.40$) or Study 2B (Mann-Whitney $U$ test: $U = 21.50$, $Z = -0.44$, $N_{\text{capuchins}} = 10$, $N_{\text{macaques}} = 5$, $P = 0.69$; $r = -0.11$). Also, although the two species did not differ on the reversal phase in Study 2A (Mann-Whitney $U$ test: $U = 20.50$, $Z = -0.88$, $N_{\text{capuchins}} = 8$, $N_{\text{macaques}} = 7$, $P = 0.40$, $r = -0.23$), capuchins showed a trend towards being faster than macaques in Study 2B (Mann-Whitney $U$ test: $U = 1.50$, $Z = -2.00$, $N_{\text{capuchins}} = 4$, $N_{\text{macaques}} = 4$, $P = 0.11$; $r = -0.71$).
4.4.3 Study 2 Discussion

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

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

4.5 General Discussion

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

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

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

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

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

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

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

Table 8 Computer Tasks

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

<table>
<thead>
<tr>
<th>Version</th>
<th>ITI (s)</th>
<th>Location</th>
<th>Study 1A</th>
<th>Study 1B</th>
<th>Study 2A</th>
<th>Study 2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1</td>
<td>300</td>
<td>Left/right</td>
<td>[Image]</td>
<td>No version 1</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
<tr>
<td>Version 2</td>
<td>60</td>
<td>Left/right</td>
<td>[Image]</td>
<td>[Image]</td>
<td>No version 2</td>
<td>No version 2</td>
</tr>
<tr>
<td>Version 3</td>
<td>60</td>
<td>Random</td>
<td>[Image]</td>
<td>[Image]</td>
<td>No version 3</td>
<td>No version 3</td>
</tr>
<tr>
<td>Version 4</td>
<td>30</td>
<td>Random</td>
<td>[Image]</td>
<td>[Image]</td>
<td>No version 4</td>
<td>No version 4</td>
</tr>
</tbody>
</table>
Table 9 Capuchin Computer Study 1

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

<table>
<thead>
<tr>
<th>Subject</th>
<th>Capuchin Study 1A</th>
<th>Capuchin Study 1B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
<td>Reversal</td>
</tr>
<tr>
<td>Nala</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>Logan</td>
<td>30</td>
<td>110</td>
</tr>
<tr>
<td>Liam</td>
<td>20</td>
<td>&gt;110</td>
</tr>
<tr>
<td>Nkima</td>
<td>30</td>
<td>&gt;120</td>
</tr>
<tr>
<td>Gambit</td>
<td>100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Wren</td>
<td>70</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Griffin</td>
<td>&gt;120</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Drella</td>
<td>&gt;130</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Gübe</td>
<td>&gt;120</td>
<td>&gt;51</td>
</tr>
</tbody>
</table>
Table 10 Macaque Computer Study 1A

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

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial</th>
<th>Reversal 1</th>
<th>Reversal 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macaque Study 1A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chewie</td>
<td>10</td>
<td>40&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Han</td>
<td>41&lt;sup&gt;a&lt;/sup&gt;</td>
<td>62&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Murphy</td>
<td>45&lt;sup&gt;a&lt;/sup&gt;</td>
<td>88&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Obi</td>
<td>50</td>
<td>73&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Luke</td>
<td>20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Gale</td>
<td>20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>20</td>
<td>133&lt;sup&gt;a,d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Lou</td>
<td>29&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Hank</td>
<td>40&lt;sup&gt;b&lt;/sup&gt;</td>
<td>10</td>
<td>86&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
Table 11 Capuchin-Macaque Computer Study 2

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

<table>
<thead>
<tr>
<th>Subject</th>
<th>Study 2A</th>
<th>Study 2B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
<td>Reversal</td>
</tr>
<tr>
<td>Capuchin Study 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nala</td>
<td>50</td>
<td>60</td>
</tr>
<tr>
<td>Logan</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>Liam</td>
<td>20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>50</td>
</tr>
<tr>
<td>Nkima</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>Gambit</td>
<td>20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>20</td>
</tr>
<tr>
<td>Wren</td>
<td>30</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Griffin</td>
<td>30</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Gabe</td>
<td>50</td>
<td>&gt;100</td>
</tr>
<tr>
<td>Lily</td>
<td>40</td>
<td>30</td>
</tr>
<tr>
<td>Widget&lt;sup&gt;a&lt;/sup&gt;</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Macaque Study 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chewie</td>
<td>20</td>
<td>&gt;109&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Han</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Murphy</td>
<td>14&lt;sup&gt;c&lt;/sup&gt;</td>
<td>33&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Obi</td>
<td>28&lt;sup&gt;c&lt;/sup&gt;</td>
<td>55&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Luke</td>
<td>20</td>
<td>76&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Lou</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Hank</td>
<td>107&lt;sup&gt;c&lt;/sup&gt;</td>
<td>20</td>
</tr>
</tbody>
</table>
4.7 Figures

Figure 6 Capuchin-Macaque Initial Learning Phase

The number of trials required for individuals to select either the ephemeral or the permanent stimulus on the original manual plate task (data from Salwiczek et al., 2012), computer Study 1A and Study 2, as compared to the fish. The black dots designate subjects that developed a preference for the ephemeral stimulus, whereas the white dots designate those who preferred the permanent stimulus. The grey dots designate new capuchin subjects who were not tested in the original plate task (Salwiczek et al., 2012): the dark grey dots are for those who developed an initial preference for the ephemeral icon, whereas the light grey dot is for the subject who developed an initial preference for the permanent icon. Note that one capuchin subject was dropped from Study 1A, and one capuchin and one macaque subjects were not tested subsequent to Study 1A.
The number of trials required for individuals to reverse their preference on the original manual plate task (data from Salwiczek et al., 2012), computer Study 1A and Study 2, as compared to the fish. We only included data from those subjects who preferred the ephemeral stimulus on the initial learning phase (see Figure 6). The black dots designate subjects who were tested in all possible conditions (capuchins: manual + computer tasks; macaques: computer tasks), whereas the grey dots designate new capuchin monkeys who were not tested in the original plate task (Salwiczek et al., 2012).
5 STUDY 4: MONKEYS SHOW BACKWARDS INDUCTION IN A MANUAL MAZE TASK

Currently in review as Prétôt, L., & Brosnan, S. F. Monkeys show backwards induction in a manual maze task. *Journal of Comparative Psychology.*

5.1 Introduction

The ease with which species can solve apparently similar problems can be due to a variety of features, but one that has gained attention lately is the degree to which different species are selected to focus on cues that are relevant to them (Lotem & Halpern, 2012; Prétôt et al., 2016a & 2016b). How species perceive relevant cues may influence their ability to solve a task, irrespective of their cognitive abilities. For example, previous research found that cleaner fish outperformed monkeys and apes in a task derived from fish ecology, suggesting that the ecological relevance of the task was likely to determine species’ performance (Salwiczek et al., 2012). To draw the general conclusion that ecology primarily influences species’ ability to solve a task, however, we must in concert consider alternative hypotheses to explain these findings. In the current study, we followed up on previous work that compared fish and primates to explore one possible cognition limitation in the primates, their ability at backwards induction.

This earlier work was based on a task that simulated the cleaner fish’s natural ecology (Salwiczek et al., 2012). In the wild, cleaner fish remove parasites and other material from client reef fish (this is the cleaners’ sole source of food), which actively visit cleaners at their so-called cleaning stations. There are two categories of clients, visitors and residents; visitors live in large territories that cover many cleaning stations and can choose the cleaners with whom to interact,
and are therefore an ephemeral food source for the cleaners. In contrast, residents live in smaller territories and interact most frequently with the same cleaner, making them a permanent food source. Field observations found that visitors have priority of cleaning access over the residents (Bshary, 2001). This priority of access is likely because visitors will leave if they are not cleaned rapidly, and they are less likely to return to the same cleaner if it has been ignored in the past or has received bad service (i.e., been bitten; see Bshary & Schäffer, 2002).

Following the latter observation, researchers designed a lab experiment that simulated this interaction (Bshary & Grutter, 2002) with two different plates of food (each containing the same food, and thus were equally attractive). In the task, if the cleaner picked the food from the permanent plate first, the ephemeral plate was immediately removed, just as visitor clients would leave if not inspected rapidly. The permanent plate, however, always stayed until the fish had stopped feeding on it, just as resident clients often queue for service if the cleaner inspects another client. Within just a few trials, cleaner fish inspected the ephemeral plate first, supporting previous field observations of this behavior.

In a subsequent study, Salwiczek, Prétôt et al. (2012), tested cleaner fish and three nonhuman primate species in the same plate task. The goal of this study was to compare the performance of two animal taxa, cleaner fish and primates, which evolved complex social and foraging behaviors but differ substantially in brain size. If brain size (as a proxy for cognitive ability) is the key feature in determining whether they can solve this task (see Deaner et al., 2006, 2007; Reader et al., 2011; but see Healy & Rowe, 2013), one would expect that primates would outperform fish in the task. However, if natural ecology is more important, then one would predict that fish would outperform primates in this task, which derived from cleaner fish ecology. Although chimpanzees and capuchins hunt for meat and catch mobile insects and small
vertebrates, and in doing so, encounter ephemeral food sources (Boesch, 1994; Fragaszy et al., 2004; Rose, 1997), they generally encounter ephemeral food sources unpredictably and opportunistically, and thus the ecological constraints are quite different from those of the fish, for whom the interaction with ephemeral sources is predictable. Supporting the latter prediction, fish outperformed all of the primate species. However, to draw the general conclusion that it is the ecological relevance of a task that primarily influences species’ abilities to solve it, we needed to both demonstrate that primates outperformed fish in a similar task derived from primate ecology, and also rule out possible cognitive explanations for the primates’ poor performance.

To address the first point, in follow-up studies, fish and monkeys were tested on a series of tasks designed to be more relevant to primates (Prétôt et al., 2016b; also see Prétôt et al., 2016a). In the original design based on cleaners, the plates differed but the food outcomes were identical, just as cleaners always consume the same foods, but acquire these foods from different clients during social interactions. This may have given the cleaners an advantage simply because they were already predisposed to focus on the way the food was presented (i.e., the plate, which was indeed the relevant stimulus) rather than the food itself, which was uninformative as the foods were identical. To determine the degree to which this cue influenced the primates’ ability to learn the task, we tested fish and monkeys using two variations designed to be more salient to monkeys (Prétôt et al., 2016b).

In the first study, the plates were identical but the food items differed (only in color), which is a cue more ecologically relevant to the monkeys. In the wild, cleaner fish consume small invertebrates on the surface of the client reef fish (Côté, 2000; Randall, 1958), which only become visible at short range (i.e., that need to be searched for and found). Parasite abundance varies between species, partly as a correlate of client body size (Grutter, 1995); therefore,
cleaners should prefer certain clients over others because of their quality as a food patch (Grutter et al., 2005). In other words, cleaners should focus on the way the food is presented, rather than on the food itself. This was reflected in the original plate task adaptation, where the plate color and pattern were the relevant stimuli, rather than the foods, which were identical and uninformative (Salwiczek et al., 2012). For the primates, however, what is important is the food itself, not the food patch. Although foods may be associated with specific surroundings (e.g., a species of tree may provide hidden fruits), the general details of the source do not change (e.g., the fruits will not suddenly be found in a different species of tree) and the patch may not be informative about the quality of food (e.g., the position of the leaves will not tell whether the fruits are ripe; the fruits themselves must be inspected). Therefore, as food color is potentially a more ecologically relevant cue to primates than fish, we changed the color of the food item rather than the design/color of the plates.

In the second study, the food was hidden (under cups in monkeys and behind plates in fish), which should help with the general tendency of primates (and other species) to be distracted by seeing food immediately available during choice tasks (see Prétôt et al., 2016a). Monkeys improved their performance in both tasks as compared to the original plate task, indicating that these more salient cues did support the primates’ ability to solve the task. Contrary to our predictions, however, the fish kept performing at the same level in all tasks, which may indicate that they were able to generalize to novel cues (see other evidence for fishes’ impressive cognitive ability: e.g., Agrillo et al., 2012; Brown et al., 2011; Bshary et al., 2014; Ferrari et al., 2005; Grosenick et al., 2007; Kendal et al., 2009; Miletto Petrazzini et al., 2016; Piffer et al., 2013; Vail et al., 2013, 2014; Wismer et al., 2016; for reviews, see: Brown, 2015; Bshary et al., 2002, 2014).
Despite the monkeys’ success, these results suggest that other factors are affecting species’ response in the task. For example, Pepperberg and Hartsfield (2014) found that parrots performed equally to the fish in the original plate task. They argued that both species may have excelled in the task because they are naturally constrained to making a single choice at a time (with the mouth or bill), whereas primates typically use both hands to reach multiple foods or objects simultaneously, but were forced by the procedure to make only one choice. Indeed, capuchins did better on a computerized version in which they used a joystick to make their choice, potentially because they are accustomed to being able to make only a single choice at a time in this modality (Prétôt et al., 2016a). In studies using an analogue of the procedure, however, two other species that are naturally constrained to single choices, pigeons and rats, failed in learning the task (Zentall et al., 2016, 2017), emphasizing the importance of considering alternative explanations to species’ difference in performance in the task.

Salwiczek et al. (2012) proposed that part of the reason the primates performed so poorly in the task might have been a failure to use a higher cognitive mechanism, like backwards induction, to solve this task in lieu to the evolved predispositions seen in the fish. Backwards induction (first introduced by von Neumann & Morgenstern, 1944) is a problem-solving strategy, which consists of analyzing a problem from back to front in order to determine what sequence of optimal actions is required to solve it. Evidence of this capacity in animals was first demonstrated in a chimpanzee, Julia, using a tool-use procedure (Döhl, 1968). In this study, Julia learned to open 10 boxes in the appropriate sequence, using specific tools that she could find inside each box, to retrieve a food at the end. In our plate task, although the two foods were presented simultaneously, they were always offered (or removed) sequentially. Therefore, after subjects experienced and learned the outcome associated with each option, they could have used
backwards induction to determine what option they would have to choose next in order to obtain the additional reward from the subsequent (permanent) option.

Although backwards induction is a term that has been used mostly in human game theory, it is unclear how it differs from the planning capacity explored in the animal literature. Beran et al. (2015) broadly define planning as “requiring the organization of behavior in the present to obtain a future goal” (where the time frame is not specific), and emphasize that “although the cognitive processes involved in different types of planning might not necessarily be identical, they are likely related, and all have some degree of future-orientation.” According to this, our definition of backwards induction falls into one of these categories.

In Beran et al. (2015), four primate species (human children, chimpanzees, rhesus and capuchin monkeys) were tested in a computerized maze task. Subjects had to move a cursor through mazes presented on a computer screen, and avoid obstacles to reach a target located at the bottom of the screen. One important caveat is needed; although such procedure requires subjects to integrate somehow future consequences into their immediate decision to succeed in the task, it does not allow us to determine whether subjects solved the mazes by looking ahead, that is they started with the cursor location (at the top) and figured out which way might lead the cursor to the target (at the bottom), or by looking backwards, that is they first located the endpoint of the problem (the target) and only then found their way to the start point (the cursor). Because there is no evidence that these two strategies truly differ, nor do we know how significant the difference (if any) between the two is, we use the term “backwards induction” throughout this paper, because we believe this is specifically the strategy that could help primates to learn the plate task.
Although there is evidence that apes (and other animals) show, under our definition, some degree of backwards induction in various contexts and modalities (e.g., joystick-based computerized maze tasks: Beran et al., 2015; Fragaszy et al., 2003, 2009; touchscreen-based computerized maze task: Iversen & Matsuzawa, 2001; manual “paddle-box” task: Tecwyn et al., 2013; manual finger-maze task: Völter & Call, 2014a; for a tool-use version, see Völter & Call, 2014b), few studies have investigated this ability in monkeys, and most of these exceptions have used computerized procedures (e.g., Beran & Parrish, 2012; Beran et al., 2015; Fragaszy et al., 2003, 2009; Pan et al., 2011). In Beran et al. (2015), for example, although all species succeeded in the maze task, the chimpanzees were overall more proficient than the monkeys in solving it (see also Fragaszy et al., 2003, 2009), and monkeys particularly struggled when required to move away from the target.

One particularly clever design, by Tecwyn et al. (2013), tested backwards induction using a “paddle-box” apparatus, which consists of a clear box with three levels of paddles that subjects had to rotate to drop a food reward into one of four openings located at the bottom of the box (and retrieve it from there). Tecwyn and colleagues tested orangutans and bonobos and found that most subjects succeeded in the task when the paddles were set up in a flat orientation (although they failed when the paddles were initially positioned in diagonal orientations). Such a procedure presented two main advantages as compared to the tool-use and computer-based tests of planning. First, unlike tool-use tasks, the paddle-box task is quite intuitive for primates, as it does not involve any complex tool-use skill, nor does it depend on species-specific behaviors or competencies (although capuchin monkeys are usually good at tool-use tasks, in general primates have difficulty solving them; see Tomasello & Call, 1997). Second, unlike computerized mazes, it is a very natural food acquisition context, in which individuals have to go through multiple
steps to acquire the food. This may be particularly relevant for extractive foragers, like capuchin monkeys, making this a good task to further explore monkeys’ ability to use backwards induction.

In the current study, capuchin monkey (Cebus [Sapajus] apella) subjects were tested for backwards induction in a “paddle-box” task based on the design by Tecwyn et al. (2013). There were three levels of paddles and, accordingly, three degrees of difficulty to retrieve the food. If the monkeys show backwards induction, they should anticipate which way to turn the paddles to release the reward. If they fail to follow the correct path, the food would be lost. Based on capuchins’ performance in previous studies (i.e., Beran et al., 2015), we predicted that monkeys would solve our manual maze task in all three degrees of difficulty. We also predicted that they would understand the task, leading to them passing the final generalization phase of testing. These results would provide evidence that capuchin monkeys show backwards induction in a food acquisition context, and that they can show this ability in different contexts and modalities (i.e., in previous work and in the current project), ultimately supporting the hypothesis that capuchins have at least this cognitive skill that could help to solve the two-choice task of Salwiczek et al. (2012).

5.2 Methods

5.2.1 Subjects and Housing

We tested 10 captive-born brown capuchin monkeys (5 males, average age: 14 years, range 9–19 years; 5 females, average age: 15 years, range 8–20 years) from two long-term, stable, mixed-sex social groups housed at the Language Research Center of Georgia State University, Atlanta, Georgia, U.S.A. One subject was discontinued in the study because she
often chose not to participate, resulting in a total sample size of nine individuals. Subjects were always housed with their social groups except when they separated voluntarily for behavioral and cognitive testing. They were fed a diet according to their species-specific needs that included primate chow and fresh fruits and vegetables. They also received enrichment foods several times per day. Animals were never deprived of food or water for testing purposes. Running water was available ad libitum, including during testing. All of the capuchin monkey experiments were approved by the Georgia State University IACUC (A16031) and met the ethical standards of the United States and the American Society of Primatologists. Georgia State University is fully accredited by AAALAC.

The monkeys lived in two large indoor/outdoor enclosures. Each enclosure contained ample three-dimensional climbing space as well as trapezes, perches and enrichment items. The enclosure for each social group was divided into an indoor area and a larger outdoor area (approximately one-half to two-thirds of their total space). The subjects had previously been trained to voluntarily enter testing boxes attached to their indoor area, which allowed us to separate individuals from their group for testing. They could choose not to participate at any time by walking away from the experimenter, and there were no consequences for the monkeys if they decided not to participate. No subject was ever involved in more than one session of this test on any given day.

5.2.2 General Procedure

Subjects were tested using a paddle-box apparatus (based on Tecwyn et al., 2013; Figure 8), which consisted of a clear acrylic box (approximately 60 x 45 cm) containing paddles that subjects rotated using handles to drop one food reward into a funnel connected to one of four
openings (A-D; Figure 8). Rewards were 1g banana-flavor precision pellets (Bio-Serv, Flemington, NJ, U.S.A.). Subjects could only obtain the food from the opening connected to the funnel, which was made large enough for the pellet to go through but too small for monkeys’ hands to reach through. The whole system was mounted vertically on a computer cart. Subjects were tested individually in a large testing cage (70 x 70 x 60 cm) and would activate the handles through five large horizontal slits (55 x 5 cm) cut in a clear polycarbonate front to the testing cage.

The paddle-box apparatus had three degrees of difficulty depending on the level of the paddle at which the pellet was initially placed. The first level consisted of three paddles that could be rotated to the left or the right towards the funnel, into which the pellet dropped directly. The second level consisted of two paddles that, once turned, dropped the food into a single paddle of the first level; only then could the subject rotate that paddle (on level 1) to drop the food into the funnel (if the food was in the appropriate paddle) to retrieve it. Finally, the third level consisted of three paddles that, once turned, dropped the food into one paddle on the second level, which the subject had to rotate to drop the food into the first level, from which it could be retrieved via the funnel. Note that at any point subjects could send the food in the wrong direction, which was irreversible. In the case of levels 2 and 3, failing to properly plan ahead meant that the food could fail to end up in the paddle(s) on level 1 that allowed it to be retrieved. It is this aspect that required backwards induction (especially on the novel generalization trials, which they could not solve based on prior experience).

Prior to a trial, the experimenter placed the reward on a predetermined paddle (see schedule below). A trial began once the subject was presented the apparatus and rotated the paddle containing the food item, at which point the experimenter stepped back (in order to
minimize experimenter cueing). Each trial ended once the food was retrieved through the funnel (correct trials), was lost by dropping on the bottom of the cart or the room floor (incorrect trials), or was still located on any paddle after 60 seconds (this occurred only in a few trials at the beginning of the study, likely due to subjects’ lack of experience with the apparatus). Each trial was video-recorded and checked for accuracy after each session.

To evaluate subjects’ degree of skill in the task, the locations of both food and funnel were assigned according to the probability of retrieving the reward, starting with the locations that gave the highest probability, and moving sequentially to those that gave the lowest probability. Only once the subject solved the task with the food located on the lowest level (i.e., paddles 1 to 3) was it tested with the food placed on higher levels (i.e., paddles 4 to 5 and, if they met criterion on the second level, paddles 6 to 8). The locations of both the paddle and the funnel were chosen so that they gave subjects 50%, 25%, and 12.5% of probability to retrieve the food (see Table 1). This excluded 14 combinations; 10 combinations that had 0% of success (1-C, 1-D, 2-A, 2-D, 3-A, 3-B, 4-D, 5-A, 6-D, 8-A), the two combinations that had 37.5% of success because there was only one starting paddle location for which this probability held (see selection criteria, below), and two combinations that were too ambiguous because they lead to unpredictable outcomes (6-A, 8-D; see “?” Table 12). For example, with an 8-D combination, if the subject rotated the paddle clockwise, we could not predict (nor could the monkeys) whether the food would drop directly into the funnel, or skip paddle 5 and fall onto paddle 3 and be retrieved, or roll too far on the right and be lost by dropping on the floor of the cart.

Subjects were tested in up to five phases, starting with phase 1, which had the lowest degree of difficulty (50% chance of receiving the reward) and the lowest level of paddles (one level), and working up to increased difficulty (up to 12.5% in phase 5; Table 13). Within each
phase, only the paddles that were from a same level and that had a same probability of releasing the food were used. Subjects were randomly assigned to one of three groups so that all subjects within the same group received the same trials, but the different groups received different trials. For each group, we selected trials following three criteria that maximized subjects’ exposure to different trials while saving new ones for the generalization test (see below). First, there were only two different trials in each phase; for example, phase 1 had six possible combinations (1A, 1B, 2B, 2C, 3C and 3D; Table 13), so each group received two different trials three times for a total of six trials; phase 2, however, had only two possible combinations (4B and 5C), so subjects received each combination three times for the same total number of trials. Second, the two trials required subjects to rotate the paddles evenly to the left and the right; for example, for one monkey group, phase 1 had three 1A trials, which required a rotation to the left side and three 2C trials, which required a rotation to the right side. Third, the two trials always involved different paddles holding the food (which excluded the trials that gave 37.5% of probability to succeed); for example, phase 1 had three trials with the food located on paddle 1 and three trials with the food located on paddle 2.

Subjects received one six-trial session per day and were tested in up to 15 sessions in each phase (unless otherwise specified). Subjects who succeeded in the task were then given a generalization test, consisting of three six-trial sessions of a combination of familiar and unfamiliar trials (i.e., novel trials that they would not have experienced in any previous phase). There were eight unfamiliar trials and 10 familiar trials per subject (for a total of 18 trials). The first and second sessions had three pairs of trials from all three degrees of difficulty (and all three levels of paddles), each pair made of one familiar and one unfamiliar trial. Because there were not enough new trials remaining after the second session for three unfamiliar trials, the final
session was made of only two unfamiliar trials from the first level of paddles, and two pairs of familiar trials from the two other levels (levels 2 and 3).

As in Tecwyn et al. (2013), the reward occasionally ended up in an unexpected location (e.g., on a different paddle, stuck on a paddle, or accidentally lost by jumping out of the apparatus). This mostly occurred when a subject rotated the paddle too quickly or slowly, or when a paddle did not function properly (because it was too tight or loose). When this happened, the experimenter withdrew the apparatus as quickly as possible before the trial ended and restarted the trial (although in a few occasions, the subject was fast enough to grab the food before the experimenter could take the apparatus away). This accounted for fewer than 2% of the total trials of the study.

Learning Trials

Four of the nine subjects tested in phase 1 reached the learning criterion within 10 sessions (see section below). The other five subjects that did not solve the task showed a robust side bias in sessions 6-10 (binomial/sign test: all $P$s < 0.001). In an attempt to eliminate the side bias, subjects received five learning sessions of six forced trials each, directly following session 10. In those trials, a T-shape pipe that functioned as a blocker was inserted between paddles, requiring the subject to rotate the paddle towards the less-preferred side to retrieve the food. Because each learning trial was designed to always lead to a correct outcome, subjects always retrieved the food reward. Following the learning phase, subjects were tested again in 10 sessions of phase 1 (without the learning trials). No subsequent learning trials were used.

Criterion and Statistics
The learning criterion for all tests was based on a two-tailed exact binomial test. In phases 1-5, for those locations that gave 50% of probability of success, subjects needed to get at least 5/6 correct trials on two consecutive sessions to reach criterion of significance. For the locations that gave 25% and 12.5% of chance to succeed, subjects needed respectively at least 4/6 and 3/6 correct trials on two consecutive sessions to reach criterion. Note that subjects always completed the six trials, whether or not they reached criterion before the six trials ended, to ensure all subjects had the same exposure to the task. Although our criterion was to limit subjects to 15 sessions, if a subject did 5/6 correct trials in the last session (i.e., session 15), it was given another set of six trials to maximize its chances of reaching criterion. This happened to one subject (Wren) in phase 1 (Table 14).

Because there were too few unfamiliar trials at the individual level for statistical significance, we reported the percentage of correct unfamiliar trials (i.e., out of eight trials) for each subject in the generalization test. At the group level, we used the total number of unfamiliar trials for each of the three degrees of difficulty; that is, 28 trials that had 50% of probability to succeed, and 14 that had 25% and 12.5% of probability to succeed. For those trials that had 50% of probability to succeed, subjects as a group reached criterion if they got at least 20/28 correct trials. For the locations that gave 25% and 12.5% of chance to succeed, all subjects together needed respectively at least 8/14 and 5/14 correct trials to succeed. As with the individual analyses, our criteria are based on two-tailed probabilities.

5.3 Results

Phases 1-5
Four of the nine subjects used for the analysis reached criterion of phase 1 within 60 trials (range 18–60 trials, mean ±SD = 34.50 ± 18.57 trials; Table 14 & Figure 9), while three of the other five did so after they received the learning trials (range 72–78 trials, mean ±SD = 74 ± 3.46 trials; note that the range includes the initial 60 trials and the trials subsequent to the learning trials, but not the 30 learning trials, which were forced choice). The last two subjects never reached criterion in phase 1, after 150 and 156 trials respectively (again not counting the 30 learning trials).

All seven subjects who passed phase 1 reached criterion in all subsequent phases in 72 or fewer trials on each phase (there were no additional learning trials in any phase; phase 2: range 12–72 trials, mean ±SD = 36 ± 22.72 trials; phase 3: range 18–72 trials, mean ±SD = 30.86 ± 19.42 trials; phase 4: range 24–66 trials, mean ±SD = 40.29 ± 13.73 trials; phase 5: range 12–60 trials, mean ±SD = 24 ± 17.32 trials).

**Generalization Test**

At the individual level, one subject was correct on 8/8 unfamiliar trials (100%; Table 15), two subjects were correct on 7/8 trials (87.5%), two were correct on 6/8 trials (75%), one was correct on 5/8 trials (62.5%), and one was correct on 4/8 trials (50%).

Taken as a group, monkeys reached learning criterion in all three degrees of difficulty of the test; they were correct on 25/28 trials in the lower degree of difficulty (50%), 12/14 in the intermediate degree (25%), and 8/14 in the higher degree (12.5%).
5.4 Discussion

Although previous studies have shown evidence of backwards induction in monkeys, most have used computerized procedures. Therefore, an obvious next step was to test animals in different contexts and modalities to probe the limits of their understanding. In the present study, we tested whether capuchin monkeys show backwards induction in a manual procedure. Subjects were tested in an adapted version of the “paddle-box” task, using an apparatus originally designed for apes (Tecwyn et al., 2013). The system, which allowed for various degrees of difficulty, consisted of paddles that subjects had to rotate in the correct pattern to drop a food reward through a funnel, and thereby gain access to it. This procedure was particularly appropriate for capuchins, who are very manipulative extractive foragers, therefore manipulating substrates to get at an otherwise unobtainable food should be ecologically relevant for them.

In our study, all seven subjects who passed phase 1 (the easiest phase, with a single level of paddles and no backwards induction) ultimately passed phase 5 (the most difficult phase, with three levels of paddles, and therefore three decisions to make). Thus, at bare minimum, subjects could learn to make the required series of choices to obtain the food. Subjects showed different levels of performance in the generalization test, but overall, despite the small sample size, they showed evidence of using their backward induction skills to solve novel trials. These results demonstrate that monkeys possess, at least to some degree, the capacity for backwards induction in a manual task, which adds to the corpus of studies showing that this trait is shared quite widely across the primate taxon.

What do these findings tell us about the poor performance of monkeys relative to cleaner fish in the previous choice-task study? Salwiczek et al. (2012) suggested that one way to solve the problem was to learn the outcome of each option and evaluate subsequently which one to
select first to maximize rewards. That means subjects should learn first to choose the ephemeral option over the permanent one because it leads to as twice as many rewards. Our findings support the view that capuchin monkeys have the capacity to show backwards induction. Therefore, the next question is why did they fail to use it? We see (at least) three possible reasons for this failure.

First, even if the monkeys show backwards induction, they might have failed to use this skill because the intermediate reward from the permanent plate may have interfered with learning mechanisms, in that it lowered the incentive value of the reward from the subsequent, ephemeral plate. That is, because the reward from the suboptimal, permanent option was received prior to the one from the optimal ephemeral option, blocking or overshadowing mechanisms may have lowered the incentive value of the ephemeral option. This intermediate reward contrasts with studies on backwards induction, including the current study, in which subjects typically receive the reward only if they succeed in completing the entire task.

Second, they might have failed to use backwards induction because of the lack of motivation to solve the task. Schubiger, Kissling, and Burkart (2016) recently found that two other species of monkeys performed better in a memory task after the task was made more cognitively challenging to them, that is, when the number of options presented increased and therefore the probability to obtain the reward concomitantly decreased (from 50% in a two-choice task to 11% in a nine-choice task). In Salwiczek et al. (2012), the primates’ chance of receiving the better reward (two pieces of food) was high, and moreover, they had a 100% chance to obtain at least one reward on every trial. Therefore, the primates’ lack of motivation to pay attention and make careful decisions in the task (rather than their inability to solve it) might explain their poor performance.
Finally, and most importantly for our cross-taxon comparative research program, the monkeys may have failed to use backwards induction in the previous two-choice task because they did not recognize the relevant cue associated with it in the task. That is, even if a species possesses a particular cognitive skill, the ability might never show in a context that is irrelevant (e.g., see Salwiczek et al., 2012; Prétôt et al., 2016a & 2016b, for the importance of considering a species natural ecology). Although we could not test any of these hypotheses using the current procedure, we think that it is important to take such factors into account when investigating species’ variation in performance in a task. This may be particularly important when looking at species with lesser cognitive abilities; although we assumed that monkeys should be able to learn a task that a fish can learn, which led us to explore further, if the reverse had happened it is likely that the conclusion would have been that the fish simply lacked the cognitive skill of the primates, and no further exploration would have been done to figure out what factor was limiting the fish.

Taken together, our findings emphasize the importance of testing species in what we (researchers) perceive to be the same question in multiple contexts and modalities to ensure that we have a good understanding of the real limits of their abilities. This way, we can present species with tasks that are “fair”, which means that take full account of the species-specific needs and avoid a biased interpretation of species skills. Again, this may be particularly important when testing species that are assumed to have less cognitive ability. In conclusion, although it is still unclear why the monkeys failed to use backwards induction in the previous task (Salwiczek et al., 2012), the current findings demonstrate that they show the skill under at least some contexts and emphasizes that subjects may not always use all of their cognitive abilities in every situation.
5.5 Tables

Table 12 Paddle-Box Task: Probability to Retrieve Food

Probability in percent to retrieve the food for each paddle given the funnel location if subjects’ choices were purely random. Only the combinations with 50%, 25% and 12.5% of probability to retrieve the food, were used in the current study.

<table>
<thead>
<tr>
<th>Paddle</th>
<th>Location A</th>
<th>Location B</th>
<th>Location C</th>
<th>Location D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>25</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>?</td>
<td>25</td>
<td>12.5</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>12.5</td>
<td>37.5</td>
<td>37.5</td>
<td>12.5</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>12.5</td>
<td>25</td>
<td>?</td>
</tr>
</tbody>
</table>
Table 13 Paddle-Box Task: Summary of the Trials

Summary of the trials in each phase (1-5) and in the generalization test (G), including the subject’s group (each group received a different set of trials), the probability to retrieve the food (in percent), the level at which the paddle holding the food was located, and the combinations of paddle and funnel location (see Figure 8).

<table>
<thead>
<tr>
<th>Phase</th>
<th>Group</th>
<th>Probability</th>
<th>Level</th>
<th>Combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>50</td>
<td>1</td>
<td>1A, 2C</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>50</td>
<td>1</td>
<td>1B, 3C</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>50</td>
<td>1</td>
<td>2B, 3D</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>50</td>
<td>2</td>
<td>4B, 5C</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>50</td>
<td>2</td>
<td>4B, 5C</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>50</td>
<td>2</td>
<td>4B, 5C</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>25</td>
<td>2</td>
<td>4C, 5D</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>25</td>
<td>2</td>
<td>4A, 5B</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>25</td>
<td>2</td>
<td>4C, 5B</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>25</td>
<td>3</td>
<td>6B, 8C</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>25</td>
<td>3</td>
<td>6B, 8C</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>25</td>
<td>3</td>
<td>6B, 8C</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>12.5</td>
<td>3</td>
<td>7D, 8B</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>12.5</td>
<td>3</td>
<td>6C, 7A</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>12.5</td>
<td>3</td>
<td>6C, 8B</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>50-12.5</td>
<td>1-3</td>
<td>1B, 2B, 3C, 3D, 4A, 5B, 6C, 7A</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>50-12.5</td>
<td>1-3</td>
<td>1A, 2B, 2C, 3D, 4C, 5D, 7D, 8B</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>50-12.5</td>
<td>1-3</td>
<td>1A, 1B, 2C, 3C, 4A, 5D, 7A, 7D</td>
</tr>
</tbody>
</table>
Table 14 Paddle-Box Task: Trials to Reach Criterion

Number of trials needed to reach criterion in each phase. Note that a minimum of 12 trials was required to meet criterion in all phases.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>Phase 4</th>
<th>Phase 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liam</td>
<td>18</td>
<td>30</td>
<td>24</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>Gabe</td>
<td>24</td>
<td>72</td>
<td>18</td>
<td>66</td>
<td>60</td>
</tr>
<tr>
<td>Logan</td>
<td>36</td>
<td>30</td>
<td>36</td>
<td>48</td>
<td>30</td>
</tr>
<tr>
<td>Nala</td>
<td>60</td>
<td>12</td>
<td>18</td>
<td>36</td>
<td>12</td>
</tr>
<tr>
<td>Nkima</td>
<td>72&lt;sup&gt;1&lt;/sup&gt;</td>
<td>36</td>
<td>30</td>
<td>36</td>
<td>18</td>
</tr>
<tr>
<td>Griffin</td>
<td>72&lt;sup&gt;1&lt;/sup&gt;</td>
<td>60</td>
<td>72</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Lily</td>
<td>78&lt;sup&gt;1&lt;/sup&gt;</td>
<td>12</td>
<td>18</td>
<td>42</td>
<td>12</td>
</tr>
<tr>
<td>Gambit</td>
<td>&gt;150&lt;sup&gt;1&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Wren</td>
<td>&gt;156&lt;sup&gt;1,2&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>1</sup> Subjects received five learning sessions of six forced trials before being tested again in phase 1. This number includes the original 60 trials and those subsequent to the learning trials, but not the learning trials.

<sup>2</sup> Subject did 5/6 correct trials in the last session, thus was given an additional set of six trials.
Table 15 Paddle-Box Task: Correct Unfamiliar Trials

Number of correct unfamiliar trials in each session of the generalization test. Note that there were three unfamiliar trials in session 1-2, and two in session 3.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Griffin</td>
<td>3/3</td>
<td>3/3</td>
<td>2/2</td>
<td>8/8</td>
</tr>
<tr>
<td>Nala</td>
<td>2/3</td>
<td>3/3</td>
<td>2/2</td>
<td>7/8</td>
</tr>
<tr>
<td>Lily</td>
<td>2/3*</td>
<td>3/3</td>
<td>2/2</td>
<td>7/8</td>
</tr>
<tr>
<td>Liam</td>
<td>2/3</td>
<td>2/3</td>
<td>2/2</td>
<td>6/8</td>
</tr>
<tr>
<td>Logan</td>
<td>2/3</td>
<td>2/3</td>
<td>2/2</td>
<td>6/8</td>
</tr>
<tr>
<td>Gabe</td>
<td>2/3</td>
<td>2/3</td>
<td>1/2</td>
<td>5/8</td>
</tr>
<tr>
<td>Nkima</td>
<td>2/3</td>
<td>1/3</td>
<td>1/2</td>
<td>4/8</td>
</tr>
</tbody>
</table>

*Because of an issue with the apparatus, the subject was given one unfamiliar trial again at the end of the session.
5.6 Figures

**Figure 8** Paddle-Box Apparatus

Each paddle (1-8) was controlled by a handle that subjects rotated to drop the food into a funnel connected to one of four openings (A-D; based on Tecwyn et al., 2013). Here, the food could only be retrieved through opening B.
Figure 9 Paddle-Box Task: Phase Results

Number of trials required for subjects to complete each phase. Each dot represents one individual. Note that a minimum of 12 trials was required to meet criterion in all phases.
6 CONCLUSIONS

6.1 Summary of Results

The overarching goal of my research is to understand the underlying causes of the diversity of behaviors across species. For my dissertation project, I focused primarily on the relative role of ecology in shaping species’ decision-making, a perspective that is often neglected in studies of cognition. To do so, I compared the performance of several nonhuman primate species with cleaner fish in tasks derived from the cleaner fish ecology.

In a prior study, my colleagues and I compared the performance of capuchin monkeys, orangutans and chimpanzees, with the bluestreak cleaner wrasse on a dichotomous choice task derived from the natural cooperative interaction between the cleaner fish and its client fish species (Salwiczek et al., 2012). We found that fish were better than primates at solving the task, indicating that ecology prepares species to be sensitive to particular cues when solving ecological problems. Species tested using cues not derived from their natural ecology typically required longer to solve the task, probably because they first had to determine which of the available cues were the most relevant to its solution. Consistent with this idea, when for my dissertation I re-designed the original fish task using cues that I hypothesized were more relevant to primates than fish, capuchin monkeys improved their performance (Study 1; see Prétôt et al., 2016b). These results indicated that, not surprisingly, many factors play into determining what species may learn with more or less ease, but that cues are an important part of decision-making.

Interestingly, the fish continued to perform at the same level in all tasks, which differed from the monkeys, who performed better on some versions of the task than others. One explanation for these results is that cleaner fish were able to generalize to novel cues (although I
could not rule out the possibility that they developed an apparent preference for some plates over others). This would add to the corpus of evidence showing that fish, as a group, possess greater cognitive ability than was previously recognized (e.g., see Wismer et al., 2016).

In a follow-up study, however, I found that the choices of three other primate species (orangutans, gorillas and drills) in these subsequent tasks largely deviated from the capuchins’ results, with all three species performing poorly (Study 2; Prétôt et al., in prep). Although one possibility is that the cues involved in the tasks may have been unintentionally capuchin-specific, which may have explained species’ variation, it was more likely that the tasks were challenging for the other species because my subjects lacked of experience with cognitive testing, especially relative to the capuchins. This latter possibility is supported by the orangutans’ performance; orangutans performed at the same level as the capuchins on one task, but not the other. Orangutans also had substantially more testing experience than the gorillas or drills, albeit quite a bit less than the capuchins, supporting the possibility that these results were heavily influenced by the subjects’ levels of experience.

In a third study, I tested monkeys in computerized adaptations of the plate task, which better control for unintentional extraneous cues associated with manual testing, like the presence of an experimenter and, again, visible foods during the subjects’ choice (Study 3; Prétôt, Bshary, & Brosnan, 2016a). Monkeys’ overall performance in these tasks matched the performance of the fish in the original study, which provided evidence that with appropriate changes in the procedure, the monkeys could eventually learn the task.

In a final study, I investigated one possible cognitive limitation to the primates in the task, the failure to use backwards induction to solve it (Study 4; Prétôt et al., in review). I tested capuchin monkeys using the manual “paddle-box” task, originally designed for testing planning
skills in apes. Most subjects solved all levels of difficulty in the task, and monkeys succeeded as a group in a generalization test. Although it is still unclear why, if they possess backwards induction, the monkeys failed to use it in the original plate task, these findings demonstrate that species may not always use all of their cognitive abilities in every situation.

6.2 Major Take-Home Messages

6.2.1 Importance of Cues

The primate results from Study 1 and Study 2 provide empirical evidence that species are sensitive to cues when solving problems, as predicted by Lotem and Halpern (2012), and that they respond differently to different cues in the task. Capuchin monkeys were originally presented the problem using cues derived from the cleaner fish ecology, which they were apparently unable to use (Salwiczek et al., 2012). In Study 1, the monkeys were presented the same problem but with cues that, I hypothesized, were more relevant to them than to the fish. Under this condition, monkeys improved their performance, which provided evidence for the cue hypothesis.

In Study 2, I found that orangutans, gorillas and drills, like capuchins, failed in learning the task with the fish cues; however, their choices largely deviated from the capuchins’ results in the other tasks, with all three species performing overall poorly. One possibility is that the cues I chose were more relevant for capuchins, for whom they were initially designed. Alternatively, their poor performance may have been due to their level of experience (a hypothesis supported by the performance of orangutans).

Although I can only speculate on this, the results of Study 4 might provide additional support for the cue hypothesis. I demonstrated that capuchin monkeys can show backwards
induction, a capacity they failed to use to solve the original plate task, in a very natural food acquisition context. It is possible that subjects failed to use backwards induction in the plate task because they did not recognize that it was relevant for solving the task. Maybe they use specific cognitive abilities in specific tasks, and so without the cues for those tasks, they do not trigger the right cognitive ability. That is, even if a species might possess the cognitive skill necessary to solve a problem, they may still fail to use it, because either they do not recognize it, or the context is so new that they have yet to associate it with a given cognitive ability. Alternatively, they may not have understood the task the way I expected them to, and so they may not have triggered the appropriate cognitive ability (backwards induction) because they did not realize what they were actually solving.

6.2.2 Importance of Experience

My results from the three zoo species revealed that experience with testing might also have heavily influenced species’ performance in the task. One very likely explanation is that the species performed overall poorly in the tasks because none had a lot of experience with cognitive testing (especially relative to the capuchins). In particular, orangutans, which had less experience with testing than capuchins, but more than gorillas and drills, performed worse than capuchins but better than the two other species (in at least one of the two versions of the task). These results emphasize the importance for researchers to accurately report and keep track of subjects’ testing history and take this into account when interpreting their data. This is also a way for future investigators to more easily identify whether experience can explain variation in subjects’ performance across studies and conditions.
In fact, the effect of experience on performance is a common issue in cognitive psychology, which is why it is important to know what role it plays in the general context of species’ learning and performance. For example, in another study conducted at Zoo Atlanta, Anderson et al. (2005) tested gorillas’ numerical competence and compared their performance to those of chimpanzees and orangutans previously tested on similar tasks (e.g., see Beran, 2001; Call, 2000; Rumbaugh, Savage-Rumbaugh, & Hegel, 1987). The gorillas performed relatively poorly in the learning task, and worse than chimpanzees and orangutans. However, they performed at levels comparable to the other species after they received a specific training procedure. The authors suggested that the gorillas’ poor performance was likely due to their more limited experience with cognitive testing. Although subjects have had participated in some cognitive tasks prior to this study, their level of experience never reached the one of chimpanzees and orangutans, who also had prior experience with other numerical competence experiments. Note that this phenomenon is widespread across primates (and presumably other taxa as well). For example, Baenninger and Newcomb (1989)’s meta-analysis on human spatial cognition revealed that the performance of adults in spatial ability tests could be improved with training procedures. In a more recent study, humans and monkeys learned to choose the more optimal strategy as a result of increased experience with a probability-based task (Klein, Evans, Schultz, & Beran, 2013).

6.2.3 Importance of Designing Careful Procedures

Despite my best efforts, there were some practical concerns that may have limited my results. In particular, I could not rule out the (very likely) possibility that species’ performance in the task was heavily influenced by external factors, in particular the choice of testing subjects
(and environment), which is an issue present in many comparative studies and thus an important one to be addressed. In my research, fish and primates differed substantially in raising/living conditions and experience with cognitive testing. At the taxon level, cleaner fish subjects were all wild-caught, thus never had any experience with testing prior my work, and must actively and constantly search for food in the wild, which is a matter of survival for them. In contrast, most of the primates were born in captivity, and many have participated in various studies in the past. Importantly, the primate subjects receive food sufficiently multiple times per day, for which they do not have to compete at the same levels as the fish (in particular, food is not a limited resource). It is thus possible that differences in environmental pressure and/or testing experience may have influenced taxa differently in the tasks.

At the species level, it is important to note that the capuchin monkey subjects live in a different setting than the other primates tested. All four primates (capuchins, gorillas, orangutans and drills) live in large social groups with species-typical demographics. In addition, all four species enclosures consist predominantly of natural outdoor space, and all subjects have extensive enrichment (for climbing, swinging, hiding, etc.). Finally, no individuals for any species are ever deprived of food, water, treats, outdoor access, or access to social group peers, except for veterinary procedures, and certainly never to motivate testing. All subjects at both facilities received food multiple times per day regardless of testing. However, the capuchins live in a research setting, and participate daily in cognitive and behavioral tasks (no invasive work is or has even been done with the capuchins). As a result, they have more experience than the orangutans, which are tested several times per week, and fundamentally more than the other species, who are tested more rarely or not at all. Supporting this, prior studies have shown that rearing/housing conditions influenced primates’ performance in a cognitive task. For example,
Vlamings et al. (2010) found that, unlike orangutans, chimpanzees and bonobos housed in sanctuaries outperformed their counterparts housed in zoos in an inhibitory task. Although I do not know why the two populations behaved differently, such results emphasize the importance of considering such factors when conducting comparative research (Boesch, 2007).

Additionally, the capuchins are tested in individual test boxes that they have been trained to enter when they are willing to test (again, there are no negative repercussions of any sort for failing to enter the boxes). Although they can still hear and see the other members of their social group, this minimal separation likely minimizes distraction (e.g., from the group), which presumably facilitates learning. In contrast, orangutans, gorillas and drills were all tested at the zoo, where they were tested directly in their home enclosure. Both they and their social group members were free to enter and leave the testing area at any time, which might have been a distraction and made the task more difficult for them to learn.

Another important note is on the importance of considering differences in natural environments (population-level) as well as captive ones. For example, I found difference in performance in the tasks between Moorean and Philippine cleaner fish, with the former performing overall lower than the latter (at least on the reversal tests). These results may indicate differences in cognitive flexibility between the two populations, possibly due to differences in their interspecific social environment (see Wismer et al., 2014); in particular, the Philippines have as twice as many species of reef fish than French Polynesia (which includes Moorea; see Carpenter & Springer, 2005). Therefore, Philippine fish have increased exposure to different species and therefore may have more complex interactions with client fish than Moorean fish.

There is also evidence that captive and wild animals often behave differently in different contexts. In particular, captive individuals have been shown to outperform their wild
counterparts in various cognitive tasks, possibly because they are less neophobic and more exploratory (e.g., Benson-Amram, Weldele, & Holekamp, 2012; Damerius, Forss, Kosonen, et al., 2017; Laidre, 2008; Reader & Laland, 2003; Visalberghi, Janson, & Agostini, 2003). Finally, it may not just be the difference between captive and wild animals that impacts performance. For example, Damerius et al. (2017) found that captive orangutan subjects that had experienced different rearing and housing conditions during ontogeny performed differently in cognitive tasks. In particular, those that had increased exposure to humans typically performed better as compared to those that did not.

### 6.2.4 Unexpectedly Good Performance of Fish

Although I had no prediction for the effect of the cup task on fish performance in Study 1, I expected primates to outperform fish at least in the color task, because I hypothesized that the fact that the cue was the color of the food rather than some aspect of the plate would be more ecologically relevant for primates than fish. However, this was not the case; fish performed surprisingly well in all tasks. One possible explanation for these results is that the fish might have developed the decision rule to preferentially approach the ephemeral option under natural conditions and then applied the same rule to all experiments (i.e., they generalized across contexts), even if they didn’t necessarily recognize the task as the same one (see Wismer et al., 2016). Monkeys, however, rarely face a natural situation that calls for the discrimination between ephemeral and permanent sources of food and so may have struggled when the cue was the plate color compared to the other conditions. Supporting this hypothesis, prior studies have found that adult cleaner fish outperformed juvenile fish in the plate task, presumably because adults have
more encounters with clients in the wild and have experienced the interaction for longer than the juveniles (Salwiczek et al., 2012).

Finally, I could not rule the possibility that Moorean fish developed an apparent preference for one plate over the other in the tasks. This was somehow surprising, because none of the Philippine adult fish showed such biases in the previous study (Salwiczek et al., 2012). However, I find it very interesting that one population should show a much stronger color preference than the other, and I hope to explore this topic further in the future (see Importance of Designing Careful Procedures section above).

Taken together, the fish results raise new important questions about the influence of experience and the use of presumably general cognitive capacities in the task. Although both are very likely involved, additional studies are necessary to further distinguish the relative impact of each of these factors separately. For example, Salwiczek et al. (2012) suggested that a logical follow-up study to investigate the role of experience in the task could consist in replicating this study with adult subjects that were kept in captivity without simultaneous exposure to both visitor and resident clients.

6.3 Implications

6.3.1 Scientific Contribution

Taken together, the findings of my dissertation project represent an important contribution to the fields of psychology and evolutionary biology, in two key ways. First, they emphasize the importance of understanding a species’ ecology to better understand its cognition. Second, they represent a nice example of the template promoted by Kamil (1998) for conducting comparative research. Below, I expand on these two issues.
First, my findings show that ecology influences cognition. They emphasize the importance of a species’ natural ecology for understanding its psychology, a perspective that is often neglected in studies of cognition. Indeed, many studies on cognition and decision-making assume that the critical factor – or only – of success in any problem-solving tasks resides in the range of the species’ cognitive skills or cognitive flexibility (e.g., for an approach using the serial reversal task, see Bitterman, 1965, 1975). Based on my results, I argue that some aspects of a species’ ecology, which include species-relevant cues and experience, also directly influence species’ decision-making. Importantly, one must keep in mind that both ecology and cognition work together rather than in competition, and that neither should be treated as if it is the only explanation for a species’ variance in a task.

In fact, both aspects answer completely different questions of decision-making, with the former focusing more on ultimate questions, and the latter about its proximate explanations (Mayr, 1961; also see Laland, Sterelny, Olding-Smee, et al., 2012). Ultimate explanations focus on the historical reasons of a behavior, whereas proximate explanations focus on the immediate causes and/or mechanisms underlying a behavior. For example, to understand the phenomenon of migration in birds, one must understand why birds migrate (ultimate) and how they migrate (proximate). Birds migrate south because their food supply becomes scarce during the cold season (ultimate), and it is the changing length of the days near winter that stimulates hormonal and behavioral changes that result in migration (proximate). Tinbergen (1963) proposed a refinement of Mayr (1961)’s dichotomous view, breaking down the two-question classification into a four-question classification, which included two proximate causes of behavior – the mechanism (causation) and ontogeny (development) – and two ultimate ones – function (adaptation) and phylogeny (evolution). Importantly, Tinbergen viewed all four levels of analysis
as equally important, nonmutually exclusive, and complementary. In other words, to fully understand a behavior, one must answer not just one, but all four questions.

The view that I promote in this dissertation is in accordance with Kamil (1998)’ synthetic approach to the study of animal behavior, which integrates both cognitive and ecological research programs by including all four levels of Tinbergen (1963)’s analysis. The use of an integrative multilevel approach eliminates the false dichotomies between cognitive and ecological approaches. Again, both approaches must be viewed as nonmutually exclusive, and complementary.

The second major contribution of my dissertation project is that it represents a nice example of the template promoted by Kamil and colleagues for researchers interested in developing comparative research programs that investigate specific questions on decision-making (e.g., see Balda & Kamil, 1989; Kamil et al., 1994; Olson et al., 1995). My findings, and the difficulties I encountered, emphasize the importance of designing careful procedures when conducting cross-species comparisons; in particular, care must be taken when testing species that differ substantially in various aspects of behavior, such as body plan, ecology, and cognitive abilities. To address this issue, I encourage researchers to use a two-step analysis, as I did here, to compare species in a task, first by testing them in the same basic procedure (i.e. in which all parameters are kept identical), which allows to establish a baseline for comparing further results (see Salwiczek et al., 2012), and only then by altering the procedure in ways that address species’ specific strengths in the task (as in Prêtôt et al., 2016a & 2016b). Neither approach is ideal in isolation, but combining the two across a long-running research program allows for a very strong comparative test.
For my dissertation, I used a cross-taxon comparative approach to investigate the impact of ecology on decision-making. I had the privilege to test several primate species, including capuchin and rhesus monkeys at the Language Research Center of Georgia State University, orangutans, gorillas and drill monkeys at Zoo Atlanta, and cleaner fish in Moorea, French Polynesia, in collaboration with Prof. Redouan Bshary from the University of Neuchâtel, Switzerland. These collaborations across disciplinary, institutional and laboratory boundaries, in which my colleagues and I participate enthusiastically, represent one powerful way (among others) to provide full answers and solutions to different problems relevant to psychology. Most importantly, I encourage researchers to 1) consider alternate hypotheses when exploring questions of psychology, 2) address relevant ecological constraints that others might be willing to test, and 3) integrate their research findings into the broader scheme of the existing literature.

6.3.2 Interface Between Ecology and Cognition

The findings of my dissertation indicate that species’ natural ecologies likely influence individuals’ capacities to solve problems; different species responded differently when presented the same task, presumably because they rely on different specific cues to solve it. In my studies, cleaner fish outperformed all primate species in the version of the task based on the fish ecology, whereas capuchin monkeys, and orangutans to some extent, performed equally to the fish in a version of the task designed to be more ecologically relevant to primates (Studies 1 and 2).

In addition, individuals differed substantially in the tasks, both in terms of learning speed and choice pattern (i.e., preference for either the ephemeral or permanent option). These results indicate that the change of cues impacted individuals differently; that is, the new cues helped some individuals more than others to solve the task. Although I do not know precisely what
caused such variation, it is likely that individual life history (among other factors) played a major role in the way subjects made decisions. To begin to address the issue on variation in performance, I first define cognition and decision-making, and then discuss the possible mechanisms involved in the acquisition of the task.

Cognition is broadly defined as the action or faculty of thinking (i.e., processing information mentally), which encompasses various concepts of psychology, such as knowledge, consciousness, insight, reasoning, and decision-making (see Benjafield, Smilek, & Kingston, 2010). Decision-making, which is the major focus of my dissertation, is the cognitive process in which individuals choose (whether consciously or not) between two or more options available to them. Choosing often depends on numerous factors, but one that is likely to impact subjects’ performance is the environmental aspect, in particular the information from meaningful objects and events present in the real world (which is also the major focus of the ecological approach; see Gibson, 1950, 1966). It is important that individuals rely on a schema, which represents what is likely to be found in the environment (Neisser, 1976); then, the frequency at which individuals encounter expected and unexpected information helps them become more and more sophisticated with how they deal with the world.

Because animals are constantly exposed to a greater amount and variety of sensory data than they can absorb, they must focus on information that is ecologically relevant, while ignoring that which is irrelevant, as predicted by Lotem and Halpern (2012)’s model. This salience can be either innate or acquired through experience. For example, when an animal recognizes food in a natural context, this reinforcer becomes associated with objects or features close (in time and space) to the food. Even if they do not act directly as reinforcers, these data include useful information because they help an individual to locate food more efficiently. Following this logic,
based on their previous experience, the cleaner fish in the previous study (Salwiczek et al., 2012) might have immediately seen the design of the plates (color and shape) as the salient information they required to make a decision, whereas the primates did not.

As mentioned earlier, the ecological approach is highly functional; that is, species are good at solving problems that are important for their fitness. The approach is, however, neutral with respect to the underlying mechanisms involved. Consequently, greater skill at a task may not indicate greater intelligence per se, but rather that the species use different learning mechanisms or focus on different aspects of the task. Therefore, one cannot assume that because different species show similar behaviors, the learning mechanisms used must be the same, or that learning, faster or not, implies the use of a more sophisticated mechanism (for a discussion on the topic, see Chittka & Jensen, 2011; Savage-Rumbaugh, Rumbaugh, Smith, & Lawson, 1980). In fact, it is likely that cleaner fish solved the task using some type of associative learning (for reviews, see Dickinson, 2012; Heyes, 2012), like operant conditioning, although this remains to be tested in the future. In the wild, cleaner fish may have learned to prioritize visitor over resident clients without an understanding of the future consequences of failing to do so; instead, they may have learned a more simple rule, by which feeding first on clients that they encounter occasionally (i.e., visitor clients) provides them with more food than those they see more often (i.e., resident clients). Alternatively, they may have perceived the leaving of the food source as a negative reinforcer, and therefore chose the ephemeral food source first before approaching the permanent one (Salwiczek et al., 2012). In either case, fish may have succeeded in the task, because they were better conditioned than primates, prior to the onset of these studies, at identifying and discriminating relevant stimuli.
Primates may have failed in the task for two reasons. First, experience with cognitive tests (including dichotomous choice tasks) often requires discriminating between food quantities or qualities, rather than between identical and immediate rewards (as in the current tasks); consequently, subjects may have been looking for other cues or some more complex rule than associative learning to solve the task, while completely missing the discriminative stimulus. Second, some inhibitory phenomena, like blocking (and possibly overshadowing) may have interfered with learning mechanisms in the task (see Kamin, 1968; Salwiczek et al., 2012). Specifically, the food may have become a conditioned stimulus that was stronger than the plates, and which then blocked conditioning to the plates. By altering the task in ways that were designed to increase the salience of one or more aspects of the stimuli to the primates (e.g., making the food itself discriminative, or removing the food so as to eliminate those prepotent cues), however, capuchin monkeys and orangutans overcame this effect and improved their performance. This provided evidence as to which features of the task make it easier to attend to by the fish and the primates.

If one assumes that simple learning rules are the key mechanism to solve the task, then having evolved a small brain and relying on basic learning mechanisms may be sufficient to learn problems, perhaps even to outperform more complex organisms in similar tasks. For example, prior studies have shown that even invertebrates, like bees, outperform many vertebrate species, including monkeys and human infants, when placed into a simple key stimulus-response context (Chittka & Jensen, 2011; also see Pearce, 2008). Similar results have been reported in pigeons and rats, which outperform humans in various tasks (e.g., Herbranson & Schroeder, 2010; Vermaercke, Cop, Willems, et al., 2014).
Although Lotem and Halpern’s (2012) model predicts that the environment favors the use of simple learning mechanisms to solve ecological problems, it does not exclude the possibility that more advanced mechanisms may be in play too. In fact, the model suggests that advanced cognitive abilities may have evolved from more simple associative principles. Therefore, it is possible that interspecies (and possibly inter-individual) variation in the task may be due to subjects using a variety of different learning strategies, from simpler to more complex cognitive abilities. Although it is impossible to know exactly which particular mechanism was used by which subject in the task (see Chittka & Jensen, 2011), I see at least three cognitive processes that may have accounted for variation in the task: attention (e.g., attention biases), perception (e.g., object and pattern recognition), and executive functions (e.g., inhibitory control, cognitive flexibility and working memory).

Regarding attention, species and individuals pay particular attention (whether consciously or not) to salient stimuli that are present in their external environment. For example, humans (Navon, 1977) and chimpanzees (Hopkins & Washburn, 2002) attend to stimulus global patterns (i.e., shape and overall contours), whereas rhesus macaques (Hopkins & Washburn, 2002) and capuchin monkeys (Spinozzi, De Lillo, & Truppa, 2003; Truppa, Carducci, De Simone, et al., 2017) attend to more local cues. Each processing presents its own advantages, but the kind of bias each species shows reflects how they naturally perceive and treat stimuli in the real world (see Neiworth, Gleichman, Olinick, & Lamp, 2006). To my knowledge, no existing studies have investigated attention biases in cleaner fish; however, the results of the plate task might indicate that, unlike capuchins, fish (and possibly chimpanzees; see Salwiczek et al., 2012) used a global-processing strategy to solve the task; that is, they focused on the overall design of the plates.
Regarding perception, prior studies have shown that primates have the visual capacity to recognize familiar and unfamiliar individuals, a capacity widely present across highly social species (for a review, see Leopold & Rhodes, 2010; for primates, see Parr, Winslow, Hopkins, & de Waal, 2000; Talbot, Leverett, & Brosnan, 2016; Talbot, Mayo, Stoinski, & Brosnan, 2015). There is also experimental evidence that cleaner fish can discriminate between familiar and unfamiliar clients in the absence of cues such as territory, behavior and location (see Tebbich, Bshary, & Grutter, 2002). Although it is still unclear whether cleaners recognize all their clients individually or just some of them (or what cues are key for the discrimination), given that cleaner fish have more than 2000 cleaning interactions per day (Grutter, 1995), and that these encounters often take place in the presence of potential clients (see Bshary & D’Souza, 2005; Pinto et al., 2011), the results of the plate task suggest that cleaners may be even better than primates at recognizing individuals, which in turn may have made the fish better than the primates at recognizing individual plates.

Regarding executive function, prior studies have shown that orangutans outperform some other primate species, including gorillas and capuchins, in various problems that require the capacity to inhibit certain prepotent responses associated with the visibility of rewards (e.g., see Amici et al., 2008; Parrish et al., 2014; Shumaker et al., 2002; Vlamings et al., 2010). To support this, I found that orangutans also outperformed gorillas (but not capuchins) in the color task, where the food was visible, possibly because of their superior inhibitory skills. In a more recent study, Zentall and colleagues proposed that the mechanism responsible for acquisition of the fish task is likely self-control. According to this hypothesis, fish outperformed primates in the task because they were less impulsive (Zentall et al., 2016). Unlike primate subjects, for whom the risk of making an incorrect trial may not have direct repercussion on their fitness, cleaner fish
must be relatively cautious when interacting with their client fish, because cheating, such as
eating the client fish’s mucus rather than parasites, might have serious consequences (e.g.,
terminate the interaction, risk to be eaten or chased by the client, etc.).

To test for this inhibition hypothesis, rats were subjected to a version of the task where
the food reward was released after 20 seconds following the first choice, which the authors
hypothesized, would improve performance (Zentall et al., 2017). Under these conditions,
subjects eventually learned to solve the task. The authors suggested that delays to reinforcement
facilitate the acquisition of the task by minimizing the impulsive response associated with the
immediacy of the first reward. Whereas the results of the capuchin monkeys in the plate task, the
computerized-based tasks, and the cup task may support the impulsivity hypothesis (in that they
improved their performance when the food became not immediately available to them), the
performance of capuchins and orangutans in the color task (where the food is immediately
available) seem to argue against it. These results indicate that impulsivity may be influencing
performance, but cannot alone explain all species’ variation in the task, and that other cognitive
mechanisms are likely explaining success in the task.

What mechanisms species used to solve the task, and the inter-individual variation in
performance found in the tasks, remain elusive. If my subjects (within the same species or
between populations) had both the same evolutionary history and same experience with the
experimental procedure, why did they still show variation in performance in the task? Why did
some subjects learn the task quickly, whereas others never did so? Why did some subjects
behave rather counter-intuitively by choosing the suboptimal permanent option? When
comparing the performance of species in a similar task, it is hard to determine the exact strategy
by which subjects solve it. That is, even if they perform similarly, some individuals might use
relatively simple heuristics, whereas others might use a more abstract reasoning (see Chittka & Jensen, 2011). One way to address this issue, however, is to test multiple individuals (and species) on multiple tasks and measures; then, if they show similarities in performance in various contexts, it would provide converging evidence for similarities in processing as well (see Kamil, 1998). This was the approach for my dissertation project, and whereas a single dissertation project does not allow sufficient time to address this question in full, it is my hope that my continued work and others’ will eventually answer this question.

In conclusion, my results indicate that ecology, coded in the genetic and neural systems, appears to predispose animals to attend to particular types of stimuli, to respond to particular contingencies, and even to inhibit responses under particular conditions. My research emphasizes the importance and power of comparative research to elucidate these differences.

6.4 Future Directions

6.4.1 Refining Methods to Test the Ecological Hypothesis

In a series of studies, my colleagues and I found that monkeys did better in the task derived from fish ecology (Salwiczek et al., 2012) when the task cues were changed to be more relevant to primates (Prétôt et al., 2016b). Nevertheless, the primates did not outperform the fish. There are at least two possible explanations for this. First, it may be that the task was still difficult for the monkeys, presumably because it was fundamentally motivated by fish, rather than monkeys, behavior. Second, it is possible that the fish performed unexpectedly well in all tasks because they were able to generalize. Therefore, an important next step is to test fish and monkeys in an adaptation of the fish procedure entirely derived from primates’ natural behavior. I would choose prosociality as primate behavior of interest for this study.
Prosociality is defined as an altruistic behavior in which conspecifics are sensitive to the needs of others. Several studies have shown that this trait may not be unique to the human species and that nonhuman primates also show some degree of prosocial behavior (e.g., Horner, Carter, Suchak, & de Waal, 2011; Lakshminarayan & Santos, 2008; de Waal et al., 2008). In a future study, I will thus adapt a paradigm typically used to test prosocial behavior in primates to this procedure (derived from fish ecology), in order to test the degree to which primates and fish do better in the basic task when the underlying structure is directly applicable to a natural situation for monkeys, as opposed to fish (Prétôt et al., 2016b).

Prosocial behavior represents a promising model to test my hypothesis for several reasons. First, the intuitive nature of most tasks designed to test prosocial behavior makes it relatively convenient to test in a variety of species. In my case, the experimental design is based on a simple choice between two options, which is easily done by fish and primates as well as other species. Second, the procedure used for testing prosocial behavior will be quite similar to the one used in Salwiczek et al. (2012), a prerequisite for my comparison. Third, prosocial behavior has been reported in primates in both experimental studies (e.g., chimpanzees: Horner et al., 2011; Melis, Warneken, Jensen, et al., 2011; bonobos: Hare & Kwetuenda, 2010; Hare, Melis, Hastings, et al., 2007; capuchins: Lakshminarayanan & Santos, 2008; de Waal et al., 2008; marmosets and tamarins: Burkart, Fehr, Efferson, & van Schaik, 2007; Cronin, Schroeder, & Snowdon, 2010) and in the field (chimpanzees: Boesch & Boesch-Achermann, 2010; Boesch, Bolé, Eckhardt, & Boesch, 2010; capuchins: Crofoot, Rubenstein, Maiya, & Berger-Wolf, 2011; Perry & Rose, 1994; Rose, 1997; Sargeant, Wikberg, Kawamura, & Fedigan, 2015; for reviews, see de Waal, 2009; de Waal & Suchak, 2010). This ecological relevance for primates is an important aspect of this particular component of my comparative program.
6.4.2 Children’s Cooperative Decision-Making in Socially Valid Context

The findings of my dissertation project give rise to an important and untested question: Is it possible that the species differed in performance because they perceived the task differently? Specifically, the fish may have performed relatively well in all tasks because they perceived it as social and cooperative (it simulates the cleaning interactions with their client fish), whereas primates, who never experience this situation, may have perceived it as a nonsocial choice task (Salwiczek et al., 2012). If so, a change in protocol for the primates that lets them perceive the task as more social and cooperative should lead to faster learning.

In a recent work, I tested this hypothesis by presenting capuchin monkeys with a version of the task in which they chose the food from between two experimenters rather than two plates (as in the original study; Prétôt et al., in prep). I predicted that the presence of a pair of real partners (rather than plates) would increase subjects’ overall performance in the task as compared to the original task. Contrary to my hypothesis, however, the monkeys did not improve their performance in the adapted task, indicating that making the task more social did not seem to affect their behavior. To further test the possibility that the presence of a conspecific is a key component of success in the task, I am currently replicating this study with children. My preliminary data seem to indicate that both the context and the social component may influence decisions in the task. Indeed, children tested at a local daycare, where I had the option to test them over multiple days, seem to perform overall better than those tested in the on-campus lab or at a local science museum, who were tested only once. In addition, children at the daycare perform apparently better in the social than the nonsocial task, which demonstrates that working with a conspecific likely improves performance in humans.
Taken together, my research findings emphasize the importance of testing species in ecologically and socially valid contexts in order to understand the full scope of their cognitive abilities. In future work, I hope to employ a similar approach to investigate cooperation in humans. Specifically, my proposed projects for my postdoctoral fellowship are united by a strong commitment to studying children’s cooperative decision-making in a socially valid context, one in which children engage in actual cooperation with partners to acquire and divide collective resources. This study will be among the first to investigate how children solve cooperative dilemmas, and will extend existing work by increasing the social validity of established paradigms.
REFERENCES


adult cotton top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 120*, 323–330.


Prétôt, L., & Brosnan, S. F. (*in review*). Monkeys show backwards induction in manual maze task. *Journal of Comparative Psychology*.


Prétôt, L., Mickelberg, J., Carrigan, J., Stoinski, T., Bshary, R., & Brosnan, S. F. *(in prep).*

Variation in performance of orangutans, gorillas and drills in two versions of a choice task derived from the cleaner fish ecology.


