



# Animal cognition in the field: performance of wild vervet monkeys (*Chlorocebus pygerythrus*) on a reversal learning task

L. Tamara Kumpan<sup>1</sup> · Eve A. Smeltzer<sup>1</sup> · Julie A. Teichroeb<sup>1</sup>

Received: 10 July 2019 / Revised: 21 January 2020 / Accepted: 27 January 2020 / Published online: 5 February 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

Increasingly, researchers are moving animal cognitive research into wild field settings. A field-based approach offers a valuable complement to laboratory-based studies, as it enables researchers to work with animals in their natural environments and indicates whether cognitive abilities found in captive subjects are generalizable to wild animals. It is thus important to field-based research to clarify which cognitive tasks can be replicated in wild settings, which species are suitable for testing in the wild, and whether replication produces similar results in wild animals. To address these issues, we modified a well-known lab test for field applications. The transfer index (TI) is a reversal learning task that tests whether animals rely on more associative or rule-based learning strategies (Rumbaugh in *Primate behavior: developments in field and laboratory research*. Academic Press, Inc., New York, pp. 2–66, 1970). In this paper, we detail changes needed to use a TI-like task in the field, here referred to as the Field Reversal Index (FRI). We tested a sample of nine wild vervet monkeys (*Chlorocebus pygerythrus*) on the FRI task at Lake Nabugabo, Uganda. We show that wild primates can successfully be tested on reversal learning paradigms, and present findings that reinforce previous conclusions from captive experiments. Our results indicate that vervets, like other cercopithecoids, rely on associative learning rather than rule-based learning. Further, our results are consistent with previous research that reports improved performance post-reversal in younger individuals relative to older individuals. The FRI enables researchers to test animals both in the wild and in captivity to facilitate direct comparisons between the learning abilities of captive and wild animals.

**Keywords** Reversal learning · Associative learning · Field research · Cognitive evolution · Non-human primates · Cognitive ecology

## Introduction

Much of our present understanding of primate cognition originates from research conducted in captive settings, primarily laboratories and zoos. In captive settings, researchers can control for many extrinsic and intrinsic factors that may influence animal behaviour, and this degree of environmental control has provided valuable insights into animal cognition and behaviour. However, captivity removes animals

from their natural settings and may create disturbances in normal behavioural patterns and ecological conditions that can influence performance, for example, stress resulting from altered grouping sizes or atypically small territories. Although field studies can compensate to some extent and complement laboratory-based work by testing animals in their natural settings, it introduces a unique set of challenges. For example, sample sizes by necessity are typically smaller than those possible in captivity (MacDonald and Ritvo 2006), and the external (e.g., food abundance) as well as internal (e.g., motivation) environment of the animal is impossible to control (Cauchoix et al. 2017). Further, it is often difficult to account for the effects of other extraneous variables such as observational learning, which can occur quite easily especially in group-living species. However, field-based cognitive research is possible if these challenges are met with innovative solutions.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10071-020-01356-5>) contains supplementary material, which is available to authorized users.

---

✉ L. Tamara Kumpan  
tamara.kumpan@mail.utoronto.ca

<sup>1</sup> Anthropology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 1A4, Canada

It is imperative to adapt captive paradigms for wild implementation, as these comparisons can determine whether the cognitive abilities found in captive animals are generalizable to their wild counterparts. The ecological and evolutionary processes that govern behaviour and cognition can be disentangled by increasing the scope of cognitive research conducted in the field (Janson 2012). It is therefore important to clarify which types of tasks can successfully be replicated in wild settings, which species are suitable for testing in the wild, and whether this reproduction of lab-based tasks produces similar results in wild animals. In this paper, we address this issue by modifying a reversal learning task previously conducted with captive primates, the transfer index (TI), and present the first results for wild primates tested using a reversal learning paradigm.

Rumbaugh (1969, 1970) introduced the TI task to provide comparative information on cognitive evolution across a variety of taxa. To date, the TI has been used especially with captive primates including cercopithecoids (Rumbaugh and Pate 1984; Washburn et al. 1989; Kinoshita et al. 1997; Bonte et al. 2014), stepsirrhines (Rumbaugh and Pate 1984), cebids (De Lillo and Visalberghi 1994), and hominoids (Rumbaugh and Gill 1971; Rumbaugh and Pate 1984), but to our knowledge, it has not yet been conducted with wild primates. The task consists of an initial set of two-choice discrimination trials (i.e., the prereversal trials) in which participants must choose one of two visual stimuli associated with either a food reward (the S+) or no food reward (the S-). Individuals are trained to an accuracy criterion of either 67% (weak association) or 84% (strong association). The choice task is then repeated using different sets of stimuli.

In a typical TI, novel pairs of computer-generated stimuli (e.g., varying colourful shapes) are used in every new trial, and if the study subject selects the incorrect stimulus (the S-), it results in termination of the trial with no food reward. Once participants can discriminate the stimuli to a predetermined accuracy criterion (67% or 84%), the reward contingencies associated with the stimuli are immediately switched in a set of reversal trials (i.e. S+ becomes S-, and vice versa) immediately following the discrimination session. Standard procedure for the TI dictates that the stimuli stay reversed for the duration of 11 trials following the reversal, and then TI scores are calculated based on correct responses in this set of reversal trials (Rumbaugh and Pate 1984). Specifically, the TI score is the ratio between the percentage of correct responses in the reversal trials and the percentage correct in the prereversal trials. This ratio provides an informative measure of performance that captures changes in accuracy from the prereversal to the reversal phase. Participants normally experience several reversals across numerous completed TI tasks to determine if they have learned the concept of reversal.

The TI is a reversal learning task, and subjects complete the task either through simpler associative learning processes or more complex rule-based learning processes. If the subject achieves a low proportion of correct selections in the post-reversal trials, this indicates that they learn largely through associative processes, as the individual's response to the original S+ built up in the prereversal trials must be extinguished before they can learn the new association in the post-reversal trials (Tomasello and Call 1997). An animal that relies on associative learning would thus show greater difficulty in the post-reversal trials when the prereversal trials involved a stronger association of the original S+ (i.e., 84% criterion) than a weaker association (i.e., 67% criterion). Conversely, if the subject attains a high proportion of correct selections in the post-reversal trials, the subject is presumed to learn through rule-based cognitive mechanisms (e.g., a rule such as "pick the opposite", or the win-stay lose-shift rule). Rule-based learners would thus perform more accurately in the post-reversal trials after more strongly establishing an association in the prereversal trials that indicated the S+ as a "win" (i.e., 84% criterion) compared to more weakly establishing that association (i.e., 67% criterion). This application of rule-based mechanisms represents "transfer" of learning across the discrimination trials (Rumbaugh and Pate 1984). Here, transfer of learning specifically captures the change in performance from the 67% criterion to the 84% criterion. Thus, the task enables a comparative perspective of cognitive processes as the TI measures relative performance across both accuracy criteria after all individuals are brought to the same predetermined accuracy levels.

At present, the TI task is computer automated and participants interact with the program via a touchscreen or a joystick (Beran et al. 2008; Bonte et al. 2014; O'Hara et al. 2015). To extend previous findings of the TI to wild populations, we needed to adapt the task for use in the field without complex technological equipment. Due to several key alterations made to the standard TI procedure, which we discuss in detail in "Procedure", we refer to our modified version of the task as the field reversal index (FRI). Using our modified reversal learning test, we assessed reversal abilities in wild vervet monkeys (*Chlorocebus pygerythrus*) and discuss their performance relative to captive cercopithecoids and other primates. As the TI allows one to compare associative and rule-based learning, we thus present data for the first wild primates to be explicitly analyzed for their use of either associative or rule-based learning mechanisms. We hypothesized that wild vervets would perform comparably to their captive counterparts. That is, we expected that wild vervets would demonstrate associative learning, consistent with significantly higher scores in the 67% accuracy condition, as has been found in previous studies that test the reversal abilities of captive cercopithecoids (*Macaca mulatta*, *Cercopithecus*

*aethiops*, Rumbaugh and Pate 1984; *M. mulatta*, Washburn et al. 1989; *M. fuscata*, Kinoshita et al. 1997; *Papio papio*, Bonte et al. 2014). Several lines of cognitive research suggest that wild species should perform similarly to captive species. First, semi-free ranging macaques (*M. mulatta*) tested in field and laboratory conditions performed similarly in both environments (Gazes et al. 2013). Second, free-ranging great tits (*Parus major*), the only animals directly compared in wild and captive conditions for reversal learning abilities, also performed similarly in both conditions (Cauchoix et al. 2017). These findings suggest that cognitive experimentation in naturalistic environments does not significantly alter results. Therefore, we did not expect the learning strategy used by wild vervets on our reversal learning task to differ from that of captive cercopithecoids.

Additionally, we hypothesized that younger individuals would achieve significantly higher scores than older individuals, similar to previous testing in captive primates. Young baboons (*Papio papio*) performed better on a TI task than adults (Bonte et al. 2014), with older baboons perseverating more than younger baboons. Similarly, Japanese macaques (*M. fuscata*) declined in performance on the TI as a function of age beginning around 3–5 years old (Kinoshita et al. 1997). Other reversal learning tasks have reported similar impairments in older individuals, including a greater degree of perseverative responding in aged dogs (Tapp et al. 2003) and rhesus macaques (Voytko 1999), an increased number of trials required to reach criterion in older rats (Schoenbaum et al. 2002), as well as a tendency for juveniles to learn faster than older individuals in tits (Morand-Ferron et al. 2015). Given this pervasive tendency for juveniles to outperform adults in reversal learning tasks, we expected to find a similar effect of age in wild vervets with juveniles and sub-adults achieving the highest reversal scores and adults achieving lower reversal scores.

## Methods

### Study site and subjects

Data were collected from July 22nd, 2017 until August 22nd, 2017 at the Lake Nabugabo Research Centre, located in central Uganda (0° 22'–12° S and 31° 54' E). Lake Nabugabo is a small satellite lake of Lake Victoria (approximately 8.2 × 5 km) lying at an elevation of 1136 m. Our research site lies on the western end of Lake Nabugabo, in an area comprised of agricultural fields, mixed forest fragments, degraded forest, and tourist camps. The north-western region of the lake is bordered by forests, and the remainder of the lake is surrounded by dense wetlands, areas of natural regenerating vegetation, and grasslands (Hanna et al. 2016).

Our research subjects were a group of vervet monkeys referred to as KS group, which was habituated for 1 year prior to data collection. Just prior to this study, we conducted a nutrition-based foraging experiment with KS group (Kumpan et al. 2019) that involved daily food supplementation with popcorn for a period of 2 months but this group was otherwise experimentally naïve. KS group contained 39–40 members at the time of data collection (5 adult males, 11 adult females, 3 sub-adult males, 5 sub-adult females, 15–16 juveniles and infants), and maintained predictable ranging habits. As such, we placed the FRI setup in a strategic area that was visited at least once a day by KS group. We tested a total of 9 monkeys across 3 age groups: Soya (SY, adult female), Carrot (CT, adult female), Tomato (TM, adult female), Jam (JM, adult male), Mint (MT, sub-adult female), Pomelo (PM, sub-adult male), Vanilla (VN, sub-adult male), Leek (LK, sub-adult male), and Asparagus (AS, juvenile female).

### Age and sex determinations

Exact birthdates were unknown for most individuals, hence we culminated criteria from both observational and morphometric studies of vervet development to estimate age-sex classes (Table 1). We relied on visible physical features that correspond to behavioural changes. The transition from infant to juvenile at twelve months is consistent with other observational studies (Seyfarth and Cheney 1986) and the relative timing of the first adult teeth (Bolter and Zihlman 2003). After twelve months, individuals are rarely carried or in nipple contact with their mother. Individuals transition from juveniles to sub-adults at sexual maturity. Vervets are sexually dimorphic, thus females and males transition into sexual maturity at different times. Many studies agree that females reach sexual maturity at around three years old (Bramblett 1980; Horrocks 1985; Struhsaker 1967a, b), which is approximately when adult canines erupt in vervets (Bolter and Zihlman 2003; Turner et al. 1997). Thus, we used the presence of adult canines (visible when an individual yawns, eats, and plays) as an indicator of sub-adulthood for females. However, since males grow over a longer period than females and reach sexual maturity later (Bolter and Zihlman 2003; Bramblett 1980; Turner et al. 1997), it is more difficult to reliably determine their age class. Bolter and Zihlman (2003) found that males who had adult canines but had not yet achieved full adult dentition (36–42 months) had small, undescended testicles. Coincidentally, the transition between these dental age groups occurs simultaneously with an increase in male body mass and trunk length beyond that of a fully adult female (Bolter and Zihlman 2003; Horrocks 1985; Turner et al. 1997). Therefore, to determine age class in male vervets, we used the presence of adult canines in addition to body size exceeding that of an adult female.

**Table 1** Age-sex determinations for wild vervet monkeys

	Males		Females	
	Age (months)	Physical/behavioural characteristics	Age (months)	Physical/behavioural characteristics
Infant	0–12	Still nurse or have nipple contact with mother Carried by mother (> 6 months usually only carried when in danger)	0–12	Still nurse or have nipple contact with mother Carried by mother (> 6 months usually only carried when in danger)
Juvenile	12–48	Independent from mother Primarily socializes with playmates	12–36	Independent from mother Primarily socializes with playmates
Sub-adult	48–1st Transfer	Has adult canines Body size is larger than an adult female	36–1st infant	Has adult canines Sexually active (gives and receives sex solicitation)
Adult	Non-natal	Descended testes Actively participate in male dominance hierarchy	Premi- or multi-parous	Nipple elongation

**Fig. 1** The test apparatus with tarp screen attached to the back of the entire length of the platforms

Both females and males reach full adulthood when they invest socially in their offspring. For females, adulthood is marked by the birth of their first infant and can thus be identified by nipple elongation (Turner et al. 1997). Males reach adulthood around the same time as when they transfer from their natal group (Cheney and Seyfarth 1983).

## Apparatus

The test apparatus consisted of two wooden platforms placed directly next to each other with a protective tarp “screen” attached to the back of the entire length of the platforms to prevent observational learning (Fig. 1). Platforms were wooden tables 0.75 m high, with a square flat top 0.75 m × 0.75 m in size. The tarp screen was 0.5 m tall and 1.5 m in width. Platforms had “x’s” drawn 20 cm apart and

10 cm in from the edges of the table top, to ensure that cups (under which food rewards were later hidden) were always placed an equal distance apart. Individuals approached the apparatus from different angles; therefore “x’s” were drawn in 3 different areas: the right side of the platforms, the left side, and in the back-center. The “x’s” ensured that individuals were presented with cups directly next to each other regardless of the direction of approach (Fig. 2).

## Stimuli

The stimuli used in this experiment were two sets of brightly coloured plastic drinking cups (Fig. 3), a blue and pink set (8 cm in width and 12.6 cm in height), and an orange and green set (10.2 cm in width and 11 cm in height). We placed the cups on the FRI platform setup two at a time, onto the pre-measured “x’s” drawn into the platforms before the start of data collection and placed a slice of banana reward under the S+ cup.

To account for the possibility of colour preferences skewing performance, we conducted a set of control trials for each monkey who had participated frequently in a previous foraging experiment with a similar baited platform setup (Kumpan et al. 2019). In colour preference trials, we baited both cups in each colour set with a piece of banana and monkeys were permitted to take both rewards from under each cup (every individual took both rewards in succession). We randomized cup position for each trial using a mobile phone random-number generating application. We determined colour choice by recording the first cup removed by the subject to obtain a reward. We then conducted binomial tests for each individual with both the pink-blue cup set and the orange-green cup set. Based on the results of the binomial tests, the colour preference trials revealed no significant colour preferences for any monkey included in the analysis.



Fig. 2 Apparatus showing cup placement from all potential angles of approach

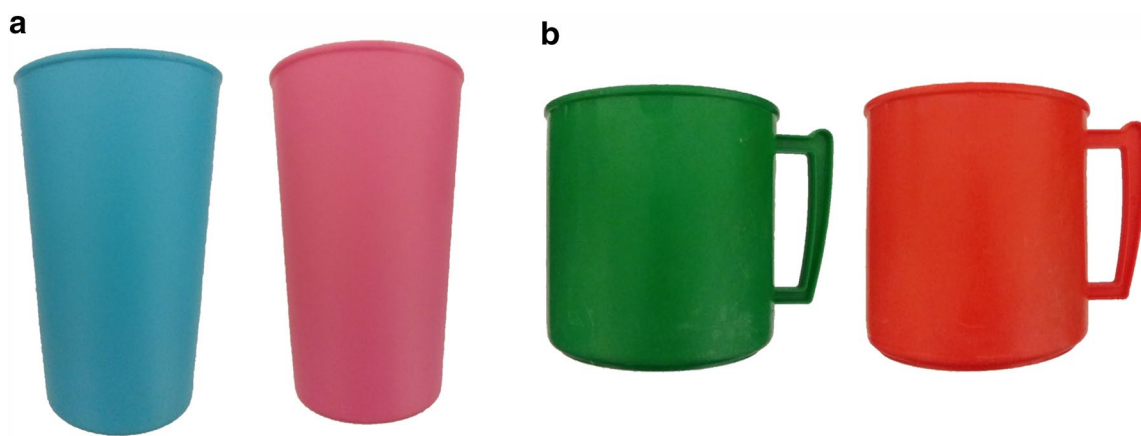


Fig. 3 The two stimulus cup sets: blue and pink (a) and green and orange (b)

## Procedure

The FRI takes advantage of the previous standards set by Rumbaugh (1969, 1970) and also includes a 67% and 84% accuracy criterion in the prereversal trials, followed by 11 post-reversal trials. Here, we refer to a “trial” as a single cup selection within either the prereversal or post-reversal portion of the task. We refer to the “prereversal” trials as the starting component of the FRI task in which we presented two cups, one with a food reward concealed underneath (the S+) and the other without a reward (the S−). We refer to the “post-reversal” trials as the second component of the FRI task in which the contingencies of the cups established in the prereversal trials were switched, such that the previously rewarding cup was no longer rewarding and became the S−, and the previously unrewarding cup now provided a reward and became the S+. Cup sets were alternated for each individual upon completion of both the prereversal and post-reversal phases of the FRI task, such that the next task began with the other set of cups. We refer to the “FRI task” as a completed set of both prereversal trials and 11 post-reversal trials.

This modified version of the TI is unique in two key ways. First, we allow participants to continue the task after an incorrect selection. Withholding a reward from the monkeys may have upset them, causing them to become skittish and putting the experimenter at risk of injury, so we simply allowed the individual to receive a reward after an incorrect selection. Second, we used physical items rather than computerized items, and used only 2 stimuli sets rotated throughout the entirety of the experiment such that no individual received the same set twice in a row. To run a standard TI without computerization we would have required 246 unique stimuli item sets which was not feasible at our remote field site (and may not be feasible for other field researchers without access to technological equipment and limitations in item availability). To address this issue, we utilized 2 stimuli sets (i.e., coloured cups) throughout the entire experiment and consider the potential issues associated with these changes in our discussion.

## Motor learning and cup habituation

Vervet participants did not readily attempt to move or lift the baited cups when first presented with them in the colour preference control. As such, we trained monkeys to lift cups to access the piece of banana underneath. To do this, we baited the platforms with a piece of banana (without cups) to attract a monkey for habituation. The single piece of banana was placed directly onto the midpoint of the line where the two wooden platforms were joined. After a

monkey obtained the initial reward, the experimenter presented two more banana pieces to the monkey and covered them with a set of cups (i.e., pink-blue or orange-green). We placed the cups directly onto the pre-measured “x’s” drawn into the platforms, either on the left end, right end, or center, depending on from which direction the monkey approached. After monkeys observed that cups concealed a banana reward, they usually lifted or knocked over a cup to access the reward. However, some monkeys needed to observe how to do this several times before attempting to move a cup. As it was impossible to predict all individuals who participated in the experiment, counterbalancing the initial cup colours experienced by each individual in the colour preference trials was not possible. To account for this, the first set of cups presented to each monkey in the preference control trials was determined by a random-number generating app, and then followed with the alternate set of cups to ensure that individuals were tested for colour preferences on both sets of colour stimuli. Data collection for the colour preference trials began after a total of 10 monkeys demonstrated proficiency with the task of lifting the cups, which we determined by 5 trials in a row of cup lifting. After this point, all tested monkeys regularly lifted the cups. The minimum of 10 monkeys was initially set as it corresponded with the individuals in KS group known to participate in experiments based on a previous study (Kumpan et al. 2019). We then recorded the colour of each initial cup selection for 15 consecutive trials per individual, and then ran binomial tests for each set of coloured cups per individual. Individuals completed all 15 colour preference trials on the same day.

## Acquisition and reversal

After completing the colour preference control trials, we presented each monkey with a randomized set of stimuli



**Fig. 4** Photo showing a sub-adult male (Leek) participating in a FRI task

to begin the experiment, either pink-blue or orange-green cups. The reward (a slice of banana) was placed under the S+ cup and to access it the monkey had to lift the correct cup (Fig. 4). The position of the S+ cup (right or left) was semi-randomized, but the colour remained the same, with the S+ placed on the same side up to a maximum of 3 trials in a row. After each trial, the cups were rebaited only when a monkey left the platforms and was at least 10 m away, to ensure that cups could be rebaited without monkeys watching. Similar distance-based re-baiting procedures have been used by researchers analyzing spatial heuristic use in vervets (Teichroeb 2015) and spatial recall in baboons (Vauclair 1990). In addition, we baited cups behind a large removable opaque barrier that blocked the entire cup setup, to assure that individuals could not see under which cup the reward was hidden. The experimenter avoided eye contact with any individuals while baiting the setup, and quietly placed the cups onto the platform to avoid providing any sound cues as to the location of the food reward. All individuals included in the analysis learned to vacate the platforms almost immediately after collecting their reward, and typically reverted to watching the group or foraging while the cups were rebaited (see Videos S1 and S2 in the Supplemental Material). After re-baiting, all individuals returned when called with a sound cue or when the barrier was removed and the cups were once again visible. The sound cue was a tut produced by pursing the lips and repeatedly sucking in air.

In the two-choice visual discrimination procedure, we presented a set of coloured plastic cups to tested individuals, where one cup was empty and the other contained the S+ reward (a slice of banana hidden underneath). These pre-reversal trials continued until the individual reached an accuracy criterion of either 67% or 84%, as determined by the standard procedure of Rumbaugh (1970) (Table 2). If a monkey selected the incorrect cup, we let them continue until they selected the correct cup, however this was recorded as an error. Once the individual achieved the number of correct

responses for 67% or 84% accuracy, we immediately began the post-reversal trials. We administered the two conditions opportunistically in that we could not control the order in which an individual completed a 67% or 84% task. As such, we recorded 67% and 84% accuracy trials as they occurred, rather than truly randomizing when an individual completed a 67% or 84% accuracy trial. For example, one individual may have completed two 67% FRI tasks in a row before completing an 84% FRI task.

In the post-reversal trials, we switched the cup contingencies such that the rewarding S+ cup was no longer rewarding, and the previously un-baited S– cup became rewarding. The cups remained reversed as described for the next 11 trials, according to standard TI procedure (Rumbaugh 1970). Following Rumbaugh (1970), we omitted the first reversal trial from our calculation of FRI scores, as the first post-reversal trial serves only to signal a reversal of cues and is not included in the score calculation (Rumbaugh 1970). Scores were therefore calculated based on the 10 remaining post-reversal trials. Thus, as in the TI task, FRI scores were the percent of correct responses from the post-reversal trials (excluding the first trial) divided by the percentage for the pre-reversal trials (the accuracy criterion) (Rumbaugh and Pate 1984).

During all trials, we recorded the date, the side of the selected cup, the number of errors, and the time the participant selected a cup. Trials were aborted ( $n = 12$ ) when the tested individual: (1) was interrupted or displaced by another individual, (2) was within 10 m of another test subject who could observe the trial outcome, (3) was distracted by an alarm call, or (4) left the platforms for longer than 5 min before returning to continue their session. The most common cause of abortion was another individual coming within 10 m of the platforms. In addition, we discarded trials when the tested individual had exceeded the performance expectation and selected the correct cup at an accuracy greater than 84%. We did this to ensure all participants were tested at the same standard levels of pre-reversal accuracy. In these cases, the subject received a 2-min “cool-down” period where cups were not set up for a new trial, after which the individual was then presented with the alternate set of cups in a new FRI task.

## Data analyses

A total of 123 FRI tasks were included in the analysis ( $n = 9$  individuals). To analyze vervet performance in the FRI task, we used five linear mixed-effects models run with the lme4 package in R (Bates et al. 2015). In each model, the independent variable was the FRI score and we included all trials where individuals reached criterion. We first examined the effect of trial type (67% or 84%) on scores using “type” as a fixed factor and controlling for

**Table 2** Number of correct responses required depending on trial number to achieve the 67% and the 84% accuracy criterions, following standard TI testing procedure from Rumbaugh and Pate (1984)

Condition (%)	Number of correct responses	Trial number
67	7 or 8	11
	9	14
	10	16
	12	19
	14 within last 21 trials	22–60
84	9	11
	14	17
	17 or 18	21
	17 or 18 within last 21	22–60

animal ID by including it as a random factor. Then, we assessed any effect of individual ID using it as the fixed factor and controlling for the trial type as a random factor. We also tested for the effects of sex and age by including them as fixed factors in their own models, controlling for trial type as random factors. For these analyses, we scored ages in three categories: juvenile, sub-adult, or adult. Finally, to determine if the effect of age that we identified was present in both types of trials (i.e., 67% and 84%), we ran a model with age and trial type included as fixed factors and individual ID included as a random factor. To determine the significance of these models, we used Likelihood Ratio Tests to compare them to null models that included only the random factors. Statistics and models were analyzed in R version 3.5.1 (R core team, 2018), with an alpha level of 0.05 set for significance.

## Results

The averaged FRI values for all nine individuals were 0.761 for the 67% accuracy condition and 0.535 for the 84% accuracy condition (values and ranges shown in Table 3). Overall, vervets scored significantly higher in the 67% accuracy condition than they did in the 84% accuracy condition (Linear mixed-effects model: Estimate =  $-0.228$ , SE =  $0.031$ ,  $t = -7.41$ ,  $P < 0.00001$ ; Table 4). Individual ID was not associated with a better or worse FRI score overall (Estimate =  $0.0009$ , SE =  $0.008$ ,  $t = 0.116$ ,  $P = 0.908$ ) and neither was the sex of the tested animal (Estimate =  $0.0406$ , SE =  $0.032$ ,  $t = 1.264$ ,  $P = 0.207$ ). However, we did find that younger individuals performed better overall (Estimate =  $-0.063$ , SE =  $0.023$ ,  $t = -2.805$ ,  $P = 0.006$ ) and this result was consistent within trial type (Estimate =  $-0.233$ , SE =  $0.030$ ,  $t = -7.75$ ,  $P < 0.0001$ ), meaning that age

**Table 3** Averaged transfer index (TI) values for both accuracy conditions and their ranges for all Nabugabo vervets tested ( $n = 9$ )

Vervet	Age	Sex	$\bar{x}$ TI at 67%	TI range 67%	$\bar{x}$ TI at 84%	TI range 84%
SY	Adult	F	0.745	0.298–0.895	0.476	0.238–0.595
CT	Adult	F	0.647	0.597–0.746	0.476	0.357–0.595
TM	Adult	F	0.646	0.477–0.746	0.357	0.357
JM	Adult	M	0.895	0.895	0.476	0.476
MT	Sub-adult	F	0.746	0.447–1.044	0.488	0.119–0.714
PM	Sub-adult	M	0.716	0.447–0.895	0.595	0.357–0.833
VN	Sub-adult	M	0.746	0.597–0.895	0.607	0.357–0.833
LK	Sub-adult	M	0.970	0.895–1.044	0.536	0.476–0.595
AS	Juvenile	F	0.880	0.447–1.044	0.571	0.119–0.833

**Table 4** Results of the linear mixed-effects models on FRI task performance in nine wild vervet monkeys

Independent variable	Fixed effects	Random effects	Estimate	SE	$t$	$p$ value <sup>†</sup>
FRI score	Trial type	ID	$-0.228$	$0.031$	$-7.41$	$<0.0001^*$
FRI score	ID	Trial type	$0.0009$	$0.008$	$0.116$	$0.908$
FRI score	Sex	Trial type	$0.0406$	$0.032$	$1.264$	$0.207$
FRI score	Age	Trial type	$-0.063$	$0.023$	$-2.805$	$0.0058^*$
FRI score	Age, and Trial type	ID	$-0.064$	$0.022$	$-2.847$	
			$-0.233$	$0.030$	$-7.705$	$<0.0001^*$

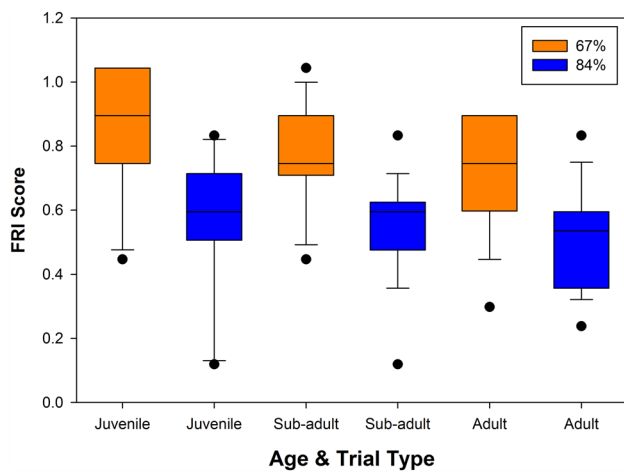
<sup>†</sup> $p$  value attained with Likelihood Ratio Tests comparing the model to a null model containing only the random factors

\*Significant results

**Table 5** Participant age and performance data

Age group	$n$	$\bar{x}$ TI at 67%	TI range 67%	$\bar{x}$ TI at 84%	TI range 84%	Transfer of learning
Juvenile	1	0.880	0.477–1.044	0.571	0.119–0.833	$-0.31$
Sub-adult	4	0.766	0.447–1.044	0.546	0.119–0.833	$-0.22$
Adult	4	0.723	0.298–0.895	0.512	0.238–0.595	$-0.21$
Overall	9	0.761	0.298–1.044	0.535	0.119–0.833	$-0.23$





**Fig. 5** Box plots showing vervet performance on the FRI task in the 67% criterion and the 84% criterion relative to age. Boxes show the upper and lower quartile, the line is the median, and the whiskers display the highest and lowest values excluding outliers, which are represented by dots

affected performance in both the 67% and 84% criterions (Table 5; Fig. 5).

## Discussion

The increasing interest in field-based comparative cognitive testing highlights the importance of identifying which types of tasks are conducive to field testing and the species that are responsive to testing in the wild. Our results show that we can successfully test wild primates using reversal learning paradigms, and that wild cercopithecoids perform similarly to their captive counterparts. Vervet reversal performance declined in the 84% accuracy condition in relation to the 67% accuracy condition ( $-0.23$ ), thus our results support previous conclusions that vervets, like other cercopithecoids, rely on associative learning strategies rather than rule-based strategies (Rumbaugh and Pate 1984). This decline represents a negative transfer of learning in that vervets appeared to have learned an association that they had difficulty reversing in the 84% accuracy condition but were able to reverse more easily in the 67% accuracy condition. Our data are in line with previous results which suggest that vervets perform at intermediate levels relative to other primates, better than strepsirrhines but worse than apes (Rumbaugh and Pate 1984). The mean scores and negative transfer we reported for wild vervets are also on par with results that have been reported for other cercopithecoids including captive rhesus macaques (*Macaca mulatta*, Washburn et al. 1989) and baboons (*Papio papio*, Bonte et al. 2014). Although baboons achieved higher scores in both accuracy criterions,

they showed a similar degree of negative transfer ( $-0.25$ , Bonte et al. 2014) as we found for wild vervets ( $-0.23$ ).

While we cannot directly compare the values we obtained for wild vervets with those in captivity (Rumbaugh and Gill 1971) due to differences in the testing procedure, we show that both wild and captive vervets appear to rely on associative learning strategies. Our successful reversal learning procedure with wild primates suggests that direct comparisons between captive and wild animals could be facilitated through implementation of the FRI. Cognitive experiments that compare wild and captive animal cognition have produced somewhat variable results. Wild spotted hyenas (*Crocuta crocuta*) were less successful than captive hyenas at solving a novel task administered using a puzzle box apparatus (Benson-Amram et al. 2013). However, a growing amount of literature suggests that captive and wild species display fundamental similarities in cognitive skills; for example, in the only direct comparison of cognitive ability between captive and wild animals, free-ranging great tits (*Parus major*) showed no differences in performance relative to captive tits on a spatial reversal learning task (Cauchoix et al. 2017). Our results add to this growing body of work. However, it is important to recognize unforeseen factors that may have contributed to our results, such as a personality bias in participants or our modification of the TI task to the FRI task.

In this study, our small sample size may not accurately represent the study population and might instead indicate a bias in our sample introduced by underlying differences in risk-taking. In our study, participants were required to approach a human observer and manipulate a novel object to obtain the reward. Of the 40 study vervets, 12 interacted with the task setup and only 9 were successful in completing a task. Thus, our sample may have been composed of “bold-type” individuals who were unafraid of approaching humans and undertaking a novel task. If boldness improves performance on the task by influencing cognitive flexibility or inhibitory control, it is possible that our FRI values are skewed to reflect this. Boldness has been linked with learning ability in associative tasks in some species, including guppies (*Poecilia reticulata*, Dugatkin and Alfieri 2003) and wild cavies (*Cavia aperea*, Guenther et al. 2013). On the contrary, it has also been shown that timid and reactive black-capped chickadees (*P. atricapillus*) outperformed more proactive and bold individuals on a reversal learning task (Guillette et al. 2011). To further complicate things, in wild tits, Morand-Ferron et al. (2015) found that participation and learning rates in a reversal task were not significantly influenced by personality differences. Similarly, in this study, we did not find individual differences in performance corresponding to variability in the number of completed FRI tasks per individual in our analysis (range for FRI tasks completed by an individual: 2–27), indicating that

timid individuals with a lower rate of participation did not differ strongly in performance from those who participated extensively. This may suggest that any individual variability in human tolerance or risk-taking present in our sample is not likely to have influenced accuracy in the post-reversal condition of the reversal learning experiment.

Our results confirmed our expectations that younger individuals would achieve higher scores than older individuals, as we found that juveniles and sub-adults outperformed adults in both accuracy criteria (Fig. 5). This age effect on performance echoes what has been reported in captivity for other cercopithecoids (Kinoshita et al. 1997; Bonte et al. 2014) and various other taxa (Voytko 1999; Schoenbaum et al. 2002; Tapp et al. 2003; Morand-Ferron et al. 2015). Kinoshita et al. (1997) reported a positive transfer of learning in 2- and 3-year old Japanese macaques using the TI, but a negative transfer of learning beginning at 5-years old and continuing into adulthood. Further, Bonte et al. (2014) found that the percentage correct in the post-reversal trials showed a negative relation to age in baboons, with younger individuals outperforming older individuals especially in the 84% accuracy criterion. Our FRI results suggest that, like previous TI results for captive primates, there is a decline in cognitive flexibility associated with age in vervets in the broad age categories of juvenile, sub-adult, and adult, although it is unfortunately unclear when this decline begins as the precise age in years of each individual included in our dataset was unknown.

The modifications we made to the standard TI may have influenced our results. First, we allowed participants to continue the task and receive a reward after an incorrect response rather than terminating the trial and withholding a reward after an incorrect response. This may create concern that individuals would learn more slowly without negative consequences for incorrect responses. However, previous work incorporating a non-binary reward scheme suggested that eliminating negative consequences may instead lead to faster learning. Macaques (*M. mulatta*) offered a smaller reward rather than no reward after an incorrect response showed quicker learning than those offered no reward after an incorrect response (Fischer and Wegener 2018). Similarly, by allowing vervets to retrieve the S+ reward after an incorrect response (S−) in our experiment, we increased the number of trials on which the monkeys could learn due to the lack of terminated trials. The reward provided feedback that reinforced the location of the S+ reward as individuals were given an opportunity to immediately redirect their response toward the correct option after an error (Fischer and Wegener 2018). This alternative reward scheme may also reduce frustration and stress in the animal during the task by keeping error rates low. Second, we used 2 stimuli sets rather than the 246 sets we would have required for a standard TI, thus it is possible that our altered

TI methodology (the FRI) introduced an associative bias. Thus, regardless of the difficulties, future work analyzing reversal learning in wild animals would benefit from including a novel stimulus set in every new task. Despite these key differences, our findings contribute to the growing consensus in comparative cognition, that animals in captivity do not experience impaired cognitive ability. Despite its limitations, the FRI is an effective tool to understand cognitive evolution in wild primates. Further, this study may help future researchers estimate the percentage of subject participation they can expect when administering challenging interactive tasks with wild primates.

Participation in an interactive cognitive task by wild animals may be limited by multiple factors, and success may likewise be limited by task complexity (van Horik et al. 2017). The low proportion of individuals that participated in (30%) and successfully completed our FRI task (23%) is similar to rates of participation and success achieved by other studies implementing comparable interactive problem-solving tasks with wild animals. For example, Benson-Amram et al. found a success rate of 15% of 62 wild hyenas when given a puzzle box (*Crocuta crocuta*, 2012), and Morand-Ferron et al. (2011) indicated a success rate of 14% among great tits (*Parus major*) and blue tits (*Cyanistes caeruleus*) when analyzing innovation using a lever-pulling device. Likewise, wild meerkats (*Suricata suricatta*) given interactive foraging tasks showed a participation rate of 47% of 135 individuals, and of these participating individuals only 8% were successful (Thornton and Samson 2012). Wild vervet groups showed a success rate of 32% with a baited box, and those groups with an especially low level of contact with humans showed a success rate of 7% with the baited box (van de Waal and Bshary 2010). These findings, as well as our own results, suggest that universal participation in comparative cognition research with wild animals is rarely achieved. Future field-based work should attempt to maximize sample sizes by considering past studies.

We conclude with suggestions to achieve larger sample sizes through habituation, experimental placement, and minimizing human presence. Larger sample sizes could be accomplished by working with wild groups who are more habituated to the presence of humans. Before this study, the vervets were followed for week-long focal periods that occurred once a month over the span of a year. While a portion of group members were comfortable with the presence of humans and were willing to participate in the task, a small sub-set of individuals were still timid and frequently lingered out of sight. Thus, a more intensive habituation period prior to testing may help increase the number of participants. It is also important to consider the placement of an experimental setup with the study species in mind. For example, vervets preferred an experimental setup located near an accessible “escape route”, such as a

tree they could climb when they felt unsafe (e.g., hearing an alarm call during the task). Future researchers should attempt to allow animal behaviour to dictate changes to experimental setup such as placement, as this may help to increase sample size. Further, allowing individuals to explore the experimental setup prior to testing without humans present might also help increase sample size. If possible, recording selections using a video camera, without a human present, may increase the participation of timid individuals. While field studies cannot always include the extensive experimental controls of a lab setting, testing animals under natural conditions (i.e., variable resources and predation pressure) presents a valuable opportunity to gain increased insights into the cognitive plasticity of a species.

**Acknowledgements** We are grateful for the help of Dr. Sara Shettleworth, Dr. Lauren Schroeder, Keegan Selig, Tiffany Atkinson, Richard Oldham, Samantha Stead, Dr. Dennis Twinomugisha, and Dr. Lauren Chapman, as well as Motovu Ponsiano, Katwere Livingstone, and Hillary Oscar. We also thank our reviewers for their insightful and helpful comments.

**Funding** This study was funded by the Natural Science and Engineering Research Council of Canada (Grant number: RGPIN-2016-0632).

### Compliance with ethical standards

**Ethical approval** All applicable international, national, and institutional guidelines for the care and use of animals were followed. The methods used in this cognitive experiment were approved by the Uganda Wildlife Authority (UWA), the Uganda National Council for Science and Technology (UNCST), and the University of Toronto Animal Care Committee (UACC). Individuals were not captured, handled, or restrained for the duration of data collection. Participation was entirely voluntary and opportunistic, and individuals were free to leave the platform setup at any time. The authors declare no conflicts of interest.

**Conflict of interest** Kumpan, Smeltzer, and Teichroeb declare that they have no conflicts of interest.

### References

- Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 64:1–48
- Benson-Amram S, Holekamp KE (2012) Innovative problem solving by wild spotted hyenas. *Proc R Soc B* 279:4087–4095
- Benson-Amram S, Weldele ML, Holekamp KE (2013) A comparison of innovative problem-solving abilities between wild and captive spotted hyenas, *Crocuta crocuta*. *Anim Behav* 85:349–356
- Beran MJ et al (2008) Discrimination reversal learning in capuchin monkeys (*Cebus apella*). *Psychol Rec* 58:3–14
- Bergman TJ, Kitchen DM (2008) Comparing responses to novel objects in wild baboons (*Papio ursinus*) and geladas (*Theropithecus gelada*). *Anim Cogn* 12:63–73
- Bolter DR, Zihlman AL (2003) Morphometric analysis of growth and development in wild-collected vervet monkeys (*Cercopithecus aethiops*), with implications for growth patterns in Old World monkeys, apes, and humans. *J Zool* 260:99–110
- Bonte E, Kemp C, Fagot J (2014) Age effects on transfer index performance and executive control in baboons (*Papio papio*). *Front Psychol* 5:1–7
- Boogert NJ, Reader SM, Laland KN (2006) The relation between social rank, neophobia and individual learning in starlings. *Anim Behav* 72:1229–1239
- Bramblett CA (1980) A model for development of social behavior in vervet monkeys. *Dev Psychobiol* 13:205–223
- Carere C, Locurto C (2011) Interaction between animal personality and animal cognition. *Curr Zoo* 57:491–498
- Carter GG, Forss S, Page RA, Ratcliffe JM (2018) Younger vampire bats (*Desmodus rotundus*) are more likely than adults to explore novel objects. *PLoS ONE* 13:e0196889
- Cauchoix M, Hermer E, Chaine AS, Morand-Ferron J (2017) Cognition in the field: comparison of reversal learning performance in captive and wild passerines. *Sci Rep* 7:12945
- Cheney DL, Seyfarth RM (1983) Dispersal in free-ranging vervet monkeys: social and genetic consequences. *Am Nat* 122:392–412
- De Lillo C, Visalberghi E (1994) Transfer index and mediational learning in tufted capuchins (*Cebus apella*). *Int J Primatol* 15:275–287
- Dugatkin LA, Alfieri MS (2003) Boldness, behavioral inhibition and learning. *Ethol Ecol Evol* 15:43–49
- Fischer B, Wegener D (2018) Emphasizing the “positive” in positive reinforcement: using nonbinary rewarding for training monkeys on cognitive tasks. *J Neurophysiol* 120:115–128
- Fragaszy DM, Adams-Curtis LE (1991) Environmental challenges in groups of capuchins. In: Box H (ed) *Primate responses to environmental change*. Chapman and Hall, London, pp 239–264
- Gazes RP, Brown EK, Basile BM, Hampton RR (2013) Automated cognitive testing of monkeys in social groups yields results comparable to individual laboratory-based testing. *Anim Cogn* 16:445–458
- Greenberg R (2003) The role of neophobia and neophilia in the development of innovative behaviour of birds. In: Reader SM, Laland KN (eds) *Animal Innovation*. Oxford University Press, Oxford, pp 175–196
- Guenther A, Brust V, Derse M, Trillmich F (2013) Learning and personality types are related in Cavies (*Cavia aperea*). *J Comp Psychol*. <https://doi.org/10.1037/a0033678>
- Guillette LM, Reddon AR, Hoeschele M, Sturdy CB (2011) Sometimes slower is better: slow-exploring birds are more sensitive to changes in a vocal discrimination task. *Proc R Soc B* 278:767–773
- Hanna DEL, Buck DG, Chapman LJ (2016) Effects of habitat on mercury concentrations in fish: a case study of Nile perch (*Lates niloticus*) in Lake Nabugabo, Uganda. *Ecotoxicology* 25:178–191
- Horrocks JA (1985) Life-history characteristics of a wild population of vervets (*Cercopithecus aethiops sabaeus*) in Barbados, West Indies. *Int J Primatol* 7:31–47
- Isden J, Panayi C, Dingle C, Madden J (2013) Performance in cognitive and problem-solving tasks in male spotted bowerbirds does not correlate with mating success. *Anim Behav* 86:829–838
- Janson CH (2012) Reconciling rigor and range: observations, experiments, and quasi-experiments in field primatology. *Int J Primatol* 33:520–541
- Judge PG, Evans DW, Schroepfer KK, Gross AC (2011) Perseveration on a reversal-learning task correlates with rates of self-directed behavior in nonhuman primates. *Behav Brain Res* 222:57–65
- Kinoshita M, Ohta H, Matano S (1997) Age change in learning ability of Japanese monkeys (*Macaca fuscata*) measured by the transfer index. *Psychol Rep* 80:467–473
- Kumpan LT, Rothman JM, Chapman CA, Teichroeb JA (2019) Playing it safe? Solitary vervet monkeys (*Chlorocebus pygerythrus*) choose high-quality foods more than those in competition. *Am J Primatol* e23002

- Lupien SJ, McEwen BS, Gunnar MR, Heim C (2009) Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat Rev Neurosci* 10:434–445
- MacDonald SE, Ritvo S (2016) Comparative cognition outside the laboratory. *Compar Cogn Behav Rev* 11:49–61
- Morand-Ferron J, Cole EF, Rawles JE, Quinn JL (2011) Who are the innovators? A field experiment with 2 passerine species. *Behav Ecol* 22:1241–1248
- Morand-Ferron J, Hamblin S, Cole EF, Aplin LM, Quinn JL (2015) Taking the operant paradigm into the field: associative learning in wild great tits. *PLoS ONE* 12:e0133821
- O'Hara M, Huber L, Gajdon GK (2015) The advantage of objects over images in discrimination and reversal learning by kea, *Nestor notabilis*. *Anim Behav* 101:51–60
- Rangel-Negrin A, Alfaro JL, Valdez RA, Romano MC, Serio-Silva JC (2009) Stress in Yucatan spider monkeys: effects of environmental conditions on fecal cortisol levels in wild and captive populations. *Anim Conserv* 12:496–502
- Rowe C, Healy SD (2014) Measuring variation in cognition. *Behav Ecol* 25:1287–1292
- Rumbaugh DM (1969) The transfer index: an alternative measure of learning set. In: Proceedings of the 2nd International Primatological Society Congress, Atlanta
- Rumbaugh DM (1970) Learning skills of anthropoids. In: Rosenblum L (ed) *Primate behavior: developments in field and laboratory research*. Academic Press Inc, New York, pp 2–66
- Rumbaugh DM (1971) Evidence of qualitative differences in learning processes among primates. *J Comp Physiol Psychol* 76:250–255
- Rumbaugh DM, Arnold RC (1971) Learning: a comparative study of Lemur and *Cercopithecus*. *Folia Primatol* 14:154–160
- Rumbaugh DM, Gill TV (1971) The learning skills of great apes. *J Hum Evol* 2:171–179
- Rumbaugh DM, Pate JL (1984) The evolution of cognition in primates: a comparative perspective. In: Roitblat HL, Bever TG, Terrace HS (eds) *Animal cognition*. Psychology Press, New York, pp 565–585
- Schoenbaum G, Nugent S, Saddoris MP, Gallagher M (2002) Teaching old rats new tricks: age-related impairments in olfactory reversal learning. *Neurobiol Aging* 23:555–564
- Seyfarth RM, Cheney DL (1986) Vocal development in vervet monkeys. *Anim Behav* 34:1640–1658
- Sih A, Del Giudice M (2012) Linking behavioural syndromes and cognition: a behavioural ecology perspective. *Phil Trans R Soc B Biol Sci* 367:2762–2772
- Struhsaker TT (1967a) Behavior of vervet monkeys (*Cercopithecus aethiops*). *Univ Calif Publ Zool* 82:1–64
- Struhsaker TT (1967b) Social structure among vervet monkeys (*Cercopithecus aethiops*). *Behaviour* 29:83–121
- Tapp PD, Siwak CT, Estrada J, Head E, Muggenburg BA, Cotman CW, Milgram NW (2003) Size and reversal learning in the beagle dog as a measure of executive function and inhibitory control in aging. *Learn Mem* 10:64–73
- Teichroeb J (2015) Vervet monkeys use paths consistent with context-specific spatial movement heuristics. *Ecol Evol* 5:4706–4716
- Thornton A, Samson J (2012) Innovative problem solving in wild meerkats. *Anim Behav* 83:1459–1468
- Tomasello M, Call J (1997) *Primate cognition*. Oxford University Press, New York
- Turner TR, Anapol F, Jolly CJ (1997) Growth, development, and sexual dimorphism in vervet monkeys (*Cercopithecus aethiops*) at four sites in Kenya. *Am J Phys Anthropol* 103:19–35
- van de Waal E, Bshary R (2010) Contact with human facilities appears to enhance technical skills in wild vervet monkeys (*Chlorocebus aethiops*). *Folia Primatol* 81:282–291
- van Horik JO, Langley EJ, Whiteside MA, Madden JR (2017) Differential participation in cognitive tests is driven by personality, sex, body condition and experience. *Behav Proc* 134:22–30
- Vauclair J (1990) Processus cognitifs élaborés: Etude des représentations mentales chez le babouin. In: Roder JJ, Anderson JR (eds) *Primates*. Kluwer Academic Publishers-Plenum, Paris, *Recherches actuelles*, pp 170–180
- Voytko ML (1999) Impairments in acquisition and reversals of two-choice discriminations by aged rhesus monkeys. *Neurobiol Aging* 20:617–627
- Washburn DA, Rumbaugh DM (1991) Rhesus monkey (*Macaca mulatta*) complex learning skills reassessed. *Int J Primatol* 12:377–388
- Washburn DA, Rumbaugh DM (1992) Testing primates with joystick-based automated apparatus: lessons from the Language Research Center's computerized test system. *Behav Res Methods Instrum Comput* 24:157–164
- Washburn DA, Hopkins WD, Rumbaugh DM (1989) Video-task assessment of learning and memory in macaques (*Macaca mulatta*): effects of stimulus movement on performance. *J Exp Psychol Anim Behav* 15:393–400
- Webster SJ, Lefebvre L (2001) Problem solving and neophobia in a columbiform-passiform assemblage in Barbados. *Anim Behav* 62:23–32

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.