TITLE: Lab cognition going wild: Implementing a new portable touchscreen system in vervet monkeys

Rachel A. Harrison[a, b*†], Tecla Mohr [a, b, c, †] & Erica van de Waal [a , b, c]

[a] Department of Ecology & Evolution, University of Lausanne, Lausanne, Switzerland

[b] The Sense Innovation and Research Center, Lausanne and Sion, Switzerland

[c] Inkawu Vervet Project, Mawana Game Reserve, South Africa

[†] These authors contributed equally to this paper

* Corresponding author: rachelh363@gmail.com / rachel.harrison@unil.ch (R.A.H)

Abstract

1. Touchscreen technology has provided researchers with opportunities to conduct well-controlled cognitive tests with captive animals, allowing researchers to isolate individuals, select participants based on specific traits, and control aspects of the environment.

2. In this study, we aimed to investigate the potential utility of touchscreen technology for the study of cognition in wild vervet monkeys. We assessed the viability of touchscreen testing by comparing rates of participation between wild and sanctuary-housed vervets. Additionally, we compared performance on a simple associative learning task in order to verify that wild participants are able to engage meaningfully with a touchscreen task presented in their natural environment.

3. We presented eight groups of vervet monkeys (four wild and four sanctuary groups, totalling 240 individuals) with a portable touchscreen device. The touchscreen displayed tasks in which food rewards could be gained by touching a stimulus displayed on the screen. We assessed individuals' likelihood of interacting with the touchscreen, their frequency of participation, and their performance on a simple associative learning task.

4. We found that sanctuary-housed monkeys were more likely to interact with the touchscreen. Participation in wild vervet monkeys was influenced by sex and age. However, monkeys in the two contexts (sanctuary versus wild) did not differ in their performance on a simple associative learning task.

5. This study demonstrates that touchscreen technology can be successfully deployed in a population of wild primates. This gives us a starting point to test animal cognition under natural conditions that include varying group composition, environmental challenges, and ongoing activities such as foraging, which are challenging to recreate in captivity. While rates of participation were lower than those found in captivity, reasonable sample sizes can be achieved, and wild primates can successfully learn touchscreen tasks in a manner comparable to their captive counterparts.

Keywords:

Touchscreen field experiment - Captivity effect - Free time hypothesis - Participation

Cognitive task - Vervet monkey

Introduction

In recent years, while many experiments have been done to quantify both within and between-species differences in cognition, most have been conducted in captive environments such as zoos (see Hopper, 2017 for review) or laboratories, with relatively few conducted in the wild (Martin et al., 2022; Pritchard et al., 2016). In captivity, researchers can control environmental factors, animals can be isolated from the rest of the group, and can be sampled based on individual characteristics. Research conducted in captivity raises two key questions; how representative is the cognition of captive animals in comparison to their wild counterparts, and if there are differences, from where do they stem? Identifying the ways in which captive animals' cognitive abilities may differ from those of wild animals could provide an insight into the extrinsic and intrinsic factors influencing performance on cognitive tests. To answer these questions, it is necessary to find ways to study animal cognition in the wild in a manner that is comparable to captive tests of cognition.

Studying cognition in the wild

There has been an increase in cognitive experiments conducted in the wild since Matsuzawa's pioneering 'outdoor laboratory' at the chimpanzee fieldsite Bossou (Matsuzawa, 1994), along with improvements in the methods and technology used (Szabo et al., 2022). Cognitive experiments have been conducted with wild subjects across a wide range of taxa, from spatial cognition in wild rufus hummingbirds (*Selasphorus rufus*, Healy & Hurly, 2013), innovative problem solving in wild meerkats (*Suricata suricatta*: Thornton & Samson, 2012), to reversal learning in wild vervet monkeys (*Chlorocebus pygerythrus*: Kumpan et al., 2020). These studies have the benefit not only of complementing existing research into cognition in captive animals, but additionally testing animals within their natural context, with the ecological pressures that it entails (Harrison & van de Waal, 2022). A challenge of conducting cognitive research in wild populations is finding methods which can be deployed in the field and produce comparable data to captive studies. One possibility, explored in the current study, is the use of touchscreen technology.

Touchscreens in cognitive testing

Touchscreens have been used in captivity to examine cognitive processes including memory, decision making, associative and reversal learning, and collaboration (Egelkamp & Ross, 2019). There are already many examples of cognitive testing using touchscreens in captive primates (Martin et al., 2022). For example, touchscreens have been used to test working memory in captive chimpanzees (Inoue & Matsuzawa, 2007; Kawai & Matsuzawa, 2000), risk-taking in chimpanzees, gorillas (*Gorilla gorilla gorilla*) and Japanese macaques (*Macaca fuscata*) (Leinwand et al., 2020) and in-group recognition in capuchin monkeys (*Sapajus apella*) (Pokorny & de Waal, 2009). While touchscreens could be an innovative device to test wild animals, they bring technical challenges, namely the potential lack of electricity and Wi-Fi connection outside of captive environments (Schmitt, 2019). On the other hand, studying animals under unnatural conditions, such as small group sizes, isolation, and close contact with humans, affects their behaviour (Forss et al., 2022; Seferta et al., 2001; Woolverton et al., 1989), suggesting that researchers should explore the potential of using innovative testing methodologies to compare wild and captive performance in cognitive tasks. In order to validate new methods for testing cognition comparatively in the wild and captivity, it is important to understand likely participation levels as well as comparing cognitive performance, as cognitive performance can only be assessed in individuals which participate in testing (van Horik et al., 2017).

Participation in experiments: the 'free time' and necessity hypotheses

There are alternative hypotheses to explain different rates of participation in experimental tasks between individuals. The 'free time' or 'excess energy' hypothesis suggests that individuals will engage in exploration when they are under less pressure to find key resources or evade predators (Kummer & Goodall, 1985). The 'necessity hypothesis' in contrast, predicts that animals will become more motivated to explore and innovate when resources are lacking, forcing them to find alternative solutions to access food sources (Grund et al., 2019). The day-to-day behaviour of wild animals is strongly influenced by needs introduced by their social and physical environment which vary seasonally and may not be present to the same extent in captive animals (Cauchoix et al., 2017). These needs may restrict the 'free time' available to participate in cognitive testing. Furthermore, in many primate species, individuals need to travel to find daily resources, giving researchers less time in a fixed location to present an experimental apparatus. In captivity, animals often have free access to food and water, reducing the need for foraging time, which may give them more opportunity to interact with the experimental paradigm.

Some studies have indeed shown higher rates of exploration of novel objects and higher persistence in captive animals in comparison to wild ones (Benson-Amram et al., 2013; Forss et al., 2015), suggesting that captive individuals should interact with and explore experimental paradigms more than wild individuals. Rates of participation in wild populations may also be low, with Morand-Ferron et al. (2015) finding that in a population of wild great tits (*Parus major*) only 8% visited and interacted with an artificial feeder. Alongside differences between wild and captive populations, experiments have also shown high inter-individual variability in interaction with experimental tasks, with factors such as age and level of distraction influencing an individual's likelihood of interacting with a task (Martina et al., 2021).

Individual differences in participation: Age and sex

Research across multiple species has shown that juveniles are more likely to explore novel objects than adults (Biondi et al., 2010; Morand-Ferron et al., 2015) and are more persistent when solving novel problems (Benson-Amram et al., 2013), spending more time in proximity to novel tasks (Kendal et al., 2005). Results such as these suggest that juveniles may be less neophobic and more motivated to explore and discover their environments than adults, and therefore show higher levels of participation in novel cognitive tasks.

Other studies have focused on differences in exploration and innovation between the sexes. Many species are sexually dimorphic in body mass, with males being larger and stronger than females, which results in task monopolization by males (Bean, 1999; van Horik et al., 2017). In a scrounging test in vervet monkeys, males, who are larger and stronger than females, obtained more food by participating more and displacing others from artificial food patches (Li et al., 2021). In group-testing paradigms, in which individuals must compete to access a task rather than being offered it individually, the larger sex may therefore be more likely to participate.

Understanding and quantifying differences in levels of participation in novel tasks is critical for two reasons: firstly, this can aid in study design, allowing researchers to predict which age classes are more likely to participate in a task, and potentially design paradigms to encourage participation in less well-represented groups. Secondly, this gives an insight into intra-species differences in factors such as neophobia, risk-taking, and motivation.

The captivity effect and performance in cognitive tests

Beyond participation in cognitive tests, some studies have found differences in performance between captive and wild subjects (see McCune et al., 2019, for review). Whilst these studies are limited in number, the majority have found enhanced performance in captive subjects in comparison with wild subjects (e.g. Benson-Amram et al., 2013), and within primates, it has been suggested that increased exposure to humans enhances problem-solving capacities

(Damerius et al., 2017; Forss et al., 2020). However, some studies have found the opposite, with wild subjects performing equally well as (Cauchoix et al., 2017) or outperforming (McCune et al., 2019) wild-caught subjects temporarily held in captivity, though it is unclear how the stress of time spent in captivity might influence the performance of wild-caught subjects.

The tests deployed in comparisons of wild and captive cognition have so far been varied. Associative learning appears to be a good target for those interested in the evolution of cognition, as it is highly conserved but also demonstrates large inter-specific and intra-specific differences in performance, and is likely to have fitness consequences (Morand-Ferron, 2017; Morand-Ferron et al., 2016; Raine & Chittka, 2008). Currently, there is little evidence regarding the causal links between cognitive ability and fitness in wild populations; measuring performance in associative learning tasks and then relating this to functional mechanisms would be one route to elucidating these causal relationships (Cole et al., 2012; Morand-Ferron et al., 2015). Equally, examination of whether the associative learning abilities of wild animals are similar to those of animals tested in captivity may give some insight into the developmental processes and evolutionary pressures that influence this ability, including the role of social context and sociality. Testing both captive populations and wild populations with varying group sizes may provide some insight into the impact of social context upon problem solving. If links can also be made between cognition and fitness in wild populations, this could provide support for the social intelligence hypothesis that posits large group sizes drive cognitive evolution (Dunbar, 1998), as demonstrated in a study of the effect of group size on cognition in wild pinyon jays by Ashton et al. (2018).

Aims and hypotheses

The current study had two major aims; firstly, to disentangle the effects of environmental factors and individual traits on participation rates of wild and captive vervet monkeys in cognitive testing, and secondly, to provide the first results of a simple associative learning task comparing the performance of captive and wild vervet monkeys. To do this, we adopted a multi-step training procedure using portable touchscreen technology (Schmitt, 2018). We conducted an experiment with the aim of training both wild and captive vervet monkeys living in social groups in South Africa (wild: four groups at the Inkawu Vervet Project; captive: four groups at the Wild Animals Trauma Centre & Haven) to interact with the screen (using a visual stimulus, a blue square, that had to be touched to gain a food reward). In this paper, we analyse rates of participation across multiple tasks presented on the touchscreen and present the first results stemming from one of these tasks; a simple associative learning task. Our hypotheses are outlined below and summarised in Table 1.

First, we investigated the likelihood of participation in two environmental contexts (wild and captivity). Our hypothesis was in line with the ‘free time’ hypothesis (Kummer & Goodall, 1985). The lack of danger and the free time available in captivity should provide captive vervet monkeys with more opportunities to interact with the touchscreen compared to wild vervet monkeys. The captive groups tested also had, in general, smaller group sizes than the wild groups, and so reduced competition for the task is also likely to result in increased likelihood of participation.

Secondly, we investigated the impact of two individual traits: age and sex, on the likelihood of wild monkeys participating in the experiment. In vervet monkeys, juveniles are more explorative and less neophobic than adults (Forss et al., 2021). In line with this evidence, we expected higher curiosity and greater participation from juvenile vervet monkeys compared to adults. Following findings in previous studies described above (Bean, 1999; Li et al., 2021; van Horik et al., 2017), adult male vervet monkeys, being larger than adult females, could be expected to have higher rates of participation. However, co-dominance has been found in these study groups at IVP (Hemelrijk et al., 2020), and adult females as core group members have been trained as models in multiple field experiments (Borgeaud & Bshary, 2015; Botting et al., 2018; Gareta García et al., 2021; van de Waal et al., 2015). Thus, we also have evidence leading us to expect a high monopolization of the apparatus by adult females and for this reason we do not have a clear hypothesis concerning higher likelihood of participation in one sex or the other.

In the wild population, we also explored the effect of age and sex upon the *rate* of participation. Age was expected to influence the number of attempts made at the task; we expected juveniles to make more attempts than adults. However, an individual’s ability to monopolise the task is also expected to influence the number of attempts made in a session, and adults may be better able to outcompete others for access than juveniles.

Finally, we tested the cognitive abilities of both wild and captive monkeys on a simple associative test (the speed of learning to touch a blue square presented on the screen to attain a reward). Following evidence of a captivity effect in problem solving in primates (Forss et al., 2020), we expected that captive monkeys would require fewer trials to reach criterion on the associative task.

Table 1. Summary of comparisons and predictions

Outcome	**Comparison**	**Prediction**	**Sample size**
Likelihood of participation	Wild vs. captive	Higher in captive	$N_{Wild} = 178$ $N_{Captive} = 62$
	Sex and age classes within wild population	Higher in juveniles	N = 178 (wild only): $N_{AdultFemale} = 53$* $N_{AdultMale} = 40$* $N_{JuvenileFemale} = 43$* $N_{JuvenileMale} = 55$*
Rate of participation	Sex and age classes within wild population	Higher in juveniles	N = 86 (wild only): $N_{AdultFemale} = 23$[†] $N_{AdultMale} = 19$[†] $N_{JuvenileFemale} = 20$[†] $N_{JuvenileMale} = 26$[†]
Task performance – simple association learning	Wild vs. captive	Faster in captive	$N_{Wild} = 15$ $N_{Captive} = 8$

* Note: Thirteen individuals were tested as both juveniles and adults over the course of the study.

[†] Note: Three individuals participated as both juveniles and adults over the course of the study.

Materials and Methods

Study site and species

Data were collected from May 2019 to January 2022 on four groups of wild vervet monkeys (Ankhase, Baie Dankie, Kubu, and Noha) at the Inkawu Vervet Project (IVP) in Mawana Game reserve, South Africa, and from August 2018 to January 2022 on four groups of captive vervet monkeys (Boeta, Cowen, Liffie, and Poena) at the Wild Animals Trauma Centre & Haven (W.A.T.C.H), South Africa. Group size varied from three to 65 individuals (for detailed group composition see Table S1). Females were defined as adults when they reached five years old or when they first gave birth (whichever occurred first); we defined males as adults when they reached five years old or upon dispersal from their natal group (whichever occurred first). Individuals were defined as juveniles from the age of four months until they reached adulthood. Babies (individuals aged less than four months during testing) were excluded from the dataset as they were not independent from their mothers. For total sample composition, see Table S2.

Ethical statement

Our study adhered to the ASAB/ABS Guidelines for the use of animals in research (ASAB/ABS, 2020) and was approved by the relevant local authority, Ezemvelo KZN Wildlife, South Africa.

Subjects

Since 2010, six neighbouring groups of wild vervet monkeys have been habituated to humans and their artefacts at the IVP. Researchers individually identify individuals by face and body characteristics. The size of the four studied wild groups at the IVP during the study period varied between 15 and 65 individuals (see Table S3). Vervet monkeys are semi-terrestrial, which, coupled with their opportunistic nature, allows researchers to observe them easily and to test them with field experiments (Mertz et al., 2019). Monkeys in the four groups which participated in the current study have previously participated in behavioural experiments involving artificial foraging tasks (e.g. Bono et al., 2018; Canteloup et al., 2020; van de Waal et al., 2015).

At the sanctuary, groups are composed of individuals with different backgrounds (including orphans, monkeys rescued from roadsides or street-markets, or injured individuals) and live in four large outdoor enclosures in social groups of three to 21 individuals (see Table S3). The majority of individuals arrived at the sanctuary as infants and were initially cared for by humans before being integrated into mixed-age social groups at three months old. Water and food are continuously available during the day. One group (Liffie) was released after the first year of experiments (November 2019), while the other three groups (Boeta, Cowen, Poena) were studied until February 2022. The last three groups participated in novel object experiments (Forss et al., 2021). While other behavioural research has previously been conducted at the sanctuary (van de Waal et al., 2013; van de Waal & Whiten, 2012) none of the groups included in the current study had participated in other studies beyond that of Forss et al. (2021).

Material and Procedure

We used a portable touchscreen (**Z**oo-based **A**nimal-**C**omputer-**I**nteraction System, ZACI; Figure 1S) built to conduct research on apes in zoos (Schmitt, 2018). Whilst originally intended for studying animals in captivity, we adapted this portable touchscreen for field usage. To allow distance between the researcher and the animals, a convertible laptop was connected by a hotspot to an operating tablet from which we could control the program. The experiment was written in Matlab using Psychophysics Toolbox extensions (Brainard, 1997). The code records the identity of the individual participating (manual input), the type of stimuli used in the task (manual input), the number of trials attempted, the area in which each stimuli was presented on the screen, and whether the response was correct or not (coded as 1, 0). To reward participants for each correct touch, the code sends an input to an electronic control unit (ECU) composed of a rechargeable battery and attached to a food dispenser (Model ENV-203-190IR, by Med Associates Inc St. Albans) ejecting soaked corn kernels. Researchers carried the touchscreen into the field for the experiment, attached it to a tree, and removed it immediately following each day of testing (Figure 1). The experimenter(s) maintained a distance of at least five metres from the task. At the sanctuary, we attached the portable touchscreen to the enclosure using hooks. The experiment started immediately after attaching the device. We recorded each experiment using one camera (JVC EverioR Quad Proof GZ-R415BE) fixed on a tripod. All sanctuary data was collected by TM, with wild data collected by TM and six field assistants who followed the same protocol as TM. TM and all field assistants involved in data collection were required to pass an identification test prior to the start of the study in which they had to repeatedly and accurately identify every monkey in the group they were working with.

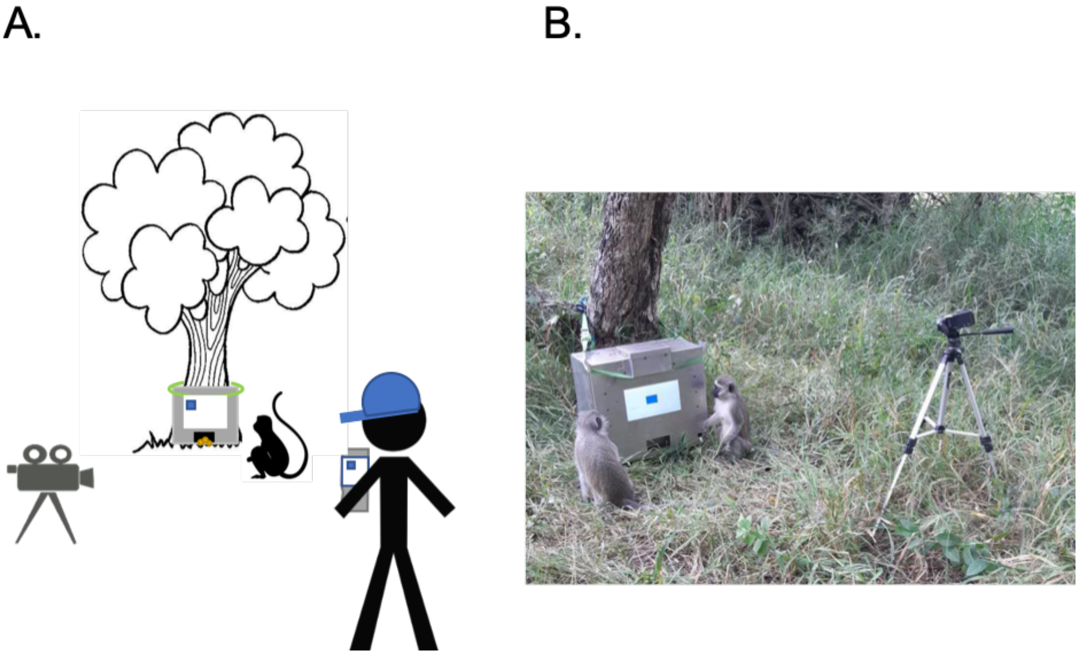

Figure 1: A) A schematic representation of an experimental setup in the field. B) Monkeys interacting with the touchscreen at their sleeping site.

We trained monkeys using three training tasks based upon the presentation of a blue square on a white background (Figure 2). In Task 1 (Habituation phase) we presented a blue square in the centre of a screen, with a white background. Monkeys received a reward if they touched either the blue square or the white background. We considered this first habituation phase completed when subjects sat in front of the screen and interacted with the touchscreen. Once the monkeys completed Task 1, we used the same image (central blue square on a white background) to test Task 2. In Task 2, monkeys were rewarded only when they touched the blue square. In Task 3, the blue square changed position in each trial, and again monkeys were only rewarded for touching the blue square. From Task 2 onwards, monkeys had to reach the learning criterion of seven correct touches in three consecutive sets, eight correct touches in two consecutive sets, or nine correct touches in one set (learning criterion from Paula et al., 2019; Salwiczek et al., 2012). When individuals made incorrect touches in Tasks 2 and 3 the blue square remained in the same location on the screen. The training was followed by two classical associative learning tasks (classical associative learning, CAL; reversal classical associative learning, RCAL) based upon the presentation of two stimuli of different shape, colour, and patterns on the screen (see Supplement for further information – CAL and RCAL testing made up 16% of recorded attempts at the task in the current dataset; 424 / 2547 attempts). These tests were included when measuring participation, but more detailed analyses of performance are not within the scope of this paper. For each correct touch, subjects received three to four corn kernels as a reward. Individuals were allowed to participate with up to 30 touches divided into three sets of 10 touches in the first three tasks, while they were allowed to participate only up to ten touches in the CAL and RCAL tasks, after which a black screen was displayed to prevent further interaction. All trials (correct and incorrect) were automatically recorded. There was an inter-trial interval of two seconds regardless of whether the previous choice was correct (there was no punishment in the form of increased inter-trial interval for an incorrect choice).

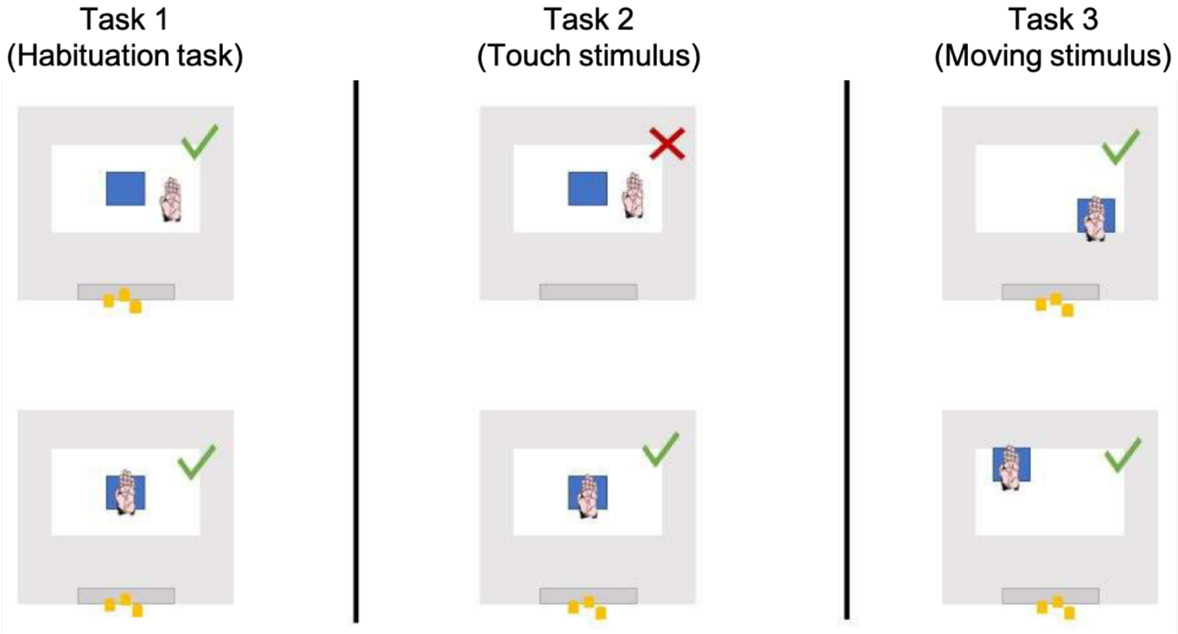

Figure 2. Three training tasks: from left to right: In Task 1 (Habituation task) monkeys had to touch anywhere on the screen (white background and blue square); In Task 2 (Touch stimulus), monkeys had to touch the blue square in the middle of the screen. In Task 3 (Moving stimulus), monkeys had to touch the blue square as in Task 2, but for each correct touch the square changed position on the screen. The hand shown in the pictures represents the monkeys' touches (potentially correct or incorrect, as shown by ticks and crosses) and the yellow dots represent the corn reward. Figure: Lucas Zermatten

We presented the touchscreen approximately twice a week both at the sanctuary and in the wild. For wild groups, the experiment was stopped when all members of the group moved away or stopped interacting. Experimental sessions were occasionally stopped due to technological problems, interruptions by sanctuary staff, storms, or other possible factors that could distract the monkeys' attention. Excluding these sessions, the wild monkeys were exposed to the device for a minimum session duration of 6.82 minutes, and a maximum duration of 187.43 minutes (mean session duration = 61.67 minutes). At the sanctuary the monkeys were exposed to the device for a minimum test session duration of 5.9 minutes and a maximum duration of 193.43 minutes (mean session duration = 31.95 minutes). See Supplemental Table S4 for a summary of the total number of test sessions conducted per group, and Supplemental Table S5 for the total presentation time per group. Due to the COVID-19 pandemic, there were occasionally longer breaks in testing, meaning that the interval between test sessions ranged from zero to 310 days in the wild and one to 426 days in the sanctuary (the length of time between test sessions did not significantly impact individuals' likelihood of participation, see Supplemental Information: Additional analyses: Effect of breaks in testing for details).

Pilot testing in the sanctuary groups

Three of the four sanctuary-housed groups (Liffie, Poena, and Boeta) participated in pilot testing from August 2018 – November 2018. During these pilot sessions (N = 155 test sessions), monkeys were exposed primarily to Task 2, with three individuals participating in Task 3. Following this pilot, changes were made to the size of the touchscreen and to the positioning of the infrared device to achieve greater accuracy in recording correct and incorrect touches. Data from this pilot testing are included in the current dataset when comparing participation rates between wild and captive groups in which participation is a binary measure, thus allowing us to compare likelihood of participation from individuals' first exposure to the task onwards. For this comparison, accurate counts of correct and incorrect touches are not required, and so the data collected in the pilot sessions is comparable with data collected during the main test period. Individuals which received Task 3 during these pilot sessions were excluded from our analysis of performance on this task, due to potential differences in their training experience in comparison with individuals who received this task for the first time during the main test period.

Statistical analyses

All analyses were performed using R (R Core Team, 2020) and RStudio (RStudio Team, 2020). Generalised linear mixed models (GLMMs) were used to analyse individuals' likelihood of participation in the task and the number of trials individuals required to reach criterion in Task 3. Binary participation in the touchscreen task was assessed using binomial GLMMs with logit link function (function 'glmer' in the R package lme4; Bates et al., 2014) and the optimizer "bobyqa". The number of attempts individuals made per session was analysed using a Cox proportional hazards model (function “coxme” in the R package coxme; Therneau, 2015), to account for right-censored data. The number of attempts each individual made before reaching criterion in Task 3 was assessed using a negative binomial GLMM with a log link (function ‘glmmTMB’ in the R package glmmTMB; Magnusson et al., 2017). A negative binomial distribution was used to account for significant overdispersion in the data.

In two analyses (likelihood of participation and trials required to pass Task 3), comparison was made between wild and sanctuary populations. The wild and sanctuary populations were not comparable in terms of age and sex classes, and so further analyses exploring the impact of these factors upon task participation were limited to the wild population only.

A binomial GLMM (Analysis 1) was used to compare the likelihood of captive and wild individuals participating in the tasks (with participation in each session coded as 1 or 0) with Context (Wild versus Sanctuary), and Session Duration (z-transformed) as predictor variables. This analysis included all touchscreen tasks (Tasks 1 – 3, CAL and RCAL) and included all individuals present in the groups at the time of testing ($N_{Total} = 240$; $N_{Wild} = 178$; $N_{Sanctuary} = 62$) with random effects for both Individual and Group (unnested, as some individuals moved between groups over the course of the experiment). Sessions which were terminated early due to external factors were not included in this analysis, as individuals were potentially prevented from participating during these sessions and so they may not reflect how many individuals would have participated had the session continued uninterrupted. Eighty-seven sessions were excluded for this reason, leaving 675 experimental sessions in the analysis. See Table S4 for a breakdown by group of the number of uninterrupted test sessions included in this analysis.

A binomial GLMM (Analysis 2) was used to investigate the impact of individual factors such as age and sex upon participation in the task in the wild population only. Participation in each session was coded as 1 or 0, with Age Class (Juvenile versus Adult), Sex, Group (N = 4), and Session Duration (z-transformed) as predictor variables, with an interaction between Age Class and Sex. Individual identity was included as a random effect. This analysis included all touchscreen tasks (Tasks 1 – 3, CAL and RCAL) and included all individuals present in the groups at the time of testing (N = 178). As in GLMM 1, sessions which were terminated early due to external factors were not included.

A mixed-effects survival model (Analysis 3) was used to investigate the impact of individual factors upon the number of attempts made by individuals per session within wild groups, with Sex, Age Class, and Group as predictor variables, and a random effect of Individual. This analysis included data from Tasks 2 and 3 only, as the experimental procedure of allowing only 30 attempts per session was applied most consistently for these tasks. Only individuals that participated in Tasks 2 and 3 were included in this analysis (N = 86). Prior to running the analysis, any individuals who had participated beyond 30 attempts had their number of attempts truncated at 30. We fitted a Cox proportional hazards model as these models are appropriate for right-censored data, in which an experimental cut-off point prevents further data collection. In our case, as many individuals were prevented from making more than 30 attempts per session, we do not have data showing the upper limit of how many attempts an individual would have made without this limit. This model therefore analysed the likelihood of an individual ceasing to participate prior to making 30 attempts. The Cox model produces hazard ratios (HR), in which an HR > 1 indicates a positive relationship between a variable and event probability, and an HR < 1 indicates a negative relationship between a variable and event probability. In our case, an HR above one indicates that an individual was more likely to stop participating sooner (i.e. to make fewer attempts), while an HR below one indicates an individual was more likely to approach 30 attempts without ceasing to participate.

A negative binomial GLMM (Analysis 4) was used to analyse the number of trials taken to pass Task 3 in the 23 individuals ($N_{\text{Wild}} = 15$; $N_{\text{Sanctuary}} = 8$) who achieved this, with Context as a predictor variable, and a random effect of Group. Some individuals in the Sanctuary (N = 3) had been trained with Task 3 during an initial pilot testing period (2018), during which the size of the blue square presented on the screen was larger and the infrared technology used to count the number of correct and incorrect touches was still being calibrated. These individuals were excluded from the analysis due to the difference in their training experience and concerns regarding the accuracy of measurement of the number of trials they had completed.

For full details of all diagnostic checks performed on the above models, see Supplemental Information “Model Assessment”.

Results

Analysis 1: Likelihood of participating in the task, Sanctuary vs. Wild

The full model (Analysis 1) was a significantly better fit to the data than a null model containing only the random effects structure ($\chi^2 = 187.54$, $p < 0.0001$). Context had a significant effect upon individuals' likelihood of participating in the touchscreen task (Figure 3, Table 2). A main effect of Context indicates that Sanctuary individuals were significantly more likely to participate than Wild individuals ($\beta = 1.96$, $p = 0.002$).

Table 2. Results of a GLMM predicting individual participation in the task in wild and captive groups.

Effect	**Estimate**	**Wald 95% confidence interval**	***p*-value**
Intercept	-4.05	-4.82; -3.28	
Context (Sanctuary)	1.96	0.74; 3.18	0.002
Session Duration (z-transformed)	0.48	0.41; 0.55	< 0.0001

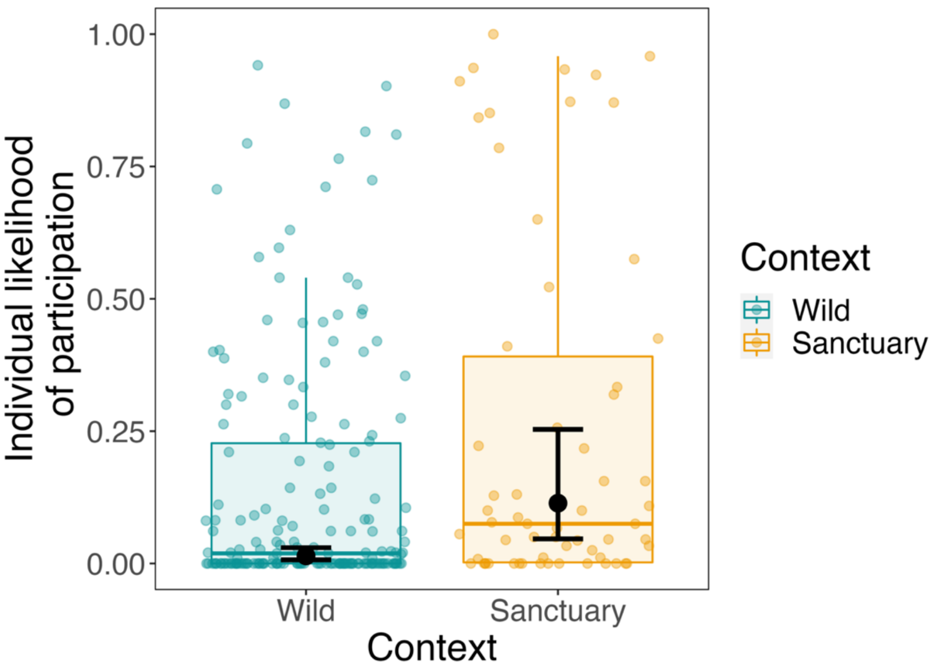

Figure 3. Model predictions (estimated marginal means) of the effect of Context on likelihood of participation in the task. The solid black point shows the prediction for each Context, with black error bars showing the 95% confidence interval. The prediction is at the mean session duration. Boxplots show the distribution of the observed likelihood of participation. The solid horizontal line, coloured by Context, shows the median proportion of sessions in which individuals participated. Upper and lower limits of the box show the first and third quartiles, and whiskers extend to the highest and lowest values at 1.5 times the interquartile range. Coloured points show the observed proportion of sessions in which each individual participated, with each point representing one individual.

Analysis 2: Likelihood of participation in the task, individual factors in the wild population

The full model was a significantly better fit to the data than a null model containing only the random effect structure ($\chi^2 = 229.1$, $p < 0.0001$). The likelihood of participation varied by both age and sex (see Table 3, Figure 4). Adult females were more likely to participate than juvenile females ($\beta = 0.77$, $p = 0.014$), while the contrary effect was seen in males, who were less likely to participate as adults than as juveniles (Age Class * Sex interaction: $\beta = -3.42$, $p < 0.0001$). Juvenile males were more likely to participate than juvenile females ($\beta = 1.58$, $p = 0.016$).

Table 3. Results of a GLMM predicting the likelihood of participation by individuals per session in the task in the wild population.

Effect	**Estimate**	**Wald 95% confidence interval**	***p*-value**
Intercept	-3.31	-4.68; -1.95	
Age Class (Adult)	0.77	0.16; 1.38	0.014
Sex (Male)	1.58	0.29; 2.86	0.016
Group (Baie Dankie)	-2.37	-3.84; -0.90	0.002
Group (Kubu)	-2.27	-3.78; -0.76	0.003
Group (Noha)	-1.14	-2.42; 0.14	0.080
Session duration (z-transformed)	0.4	0.40; 0.55	< 0.0001
Age Class * Sex	-3.42	-4.62; -2.21	< 0.0001

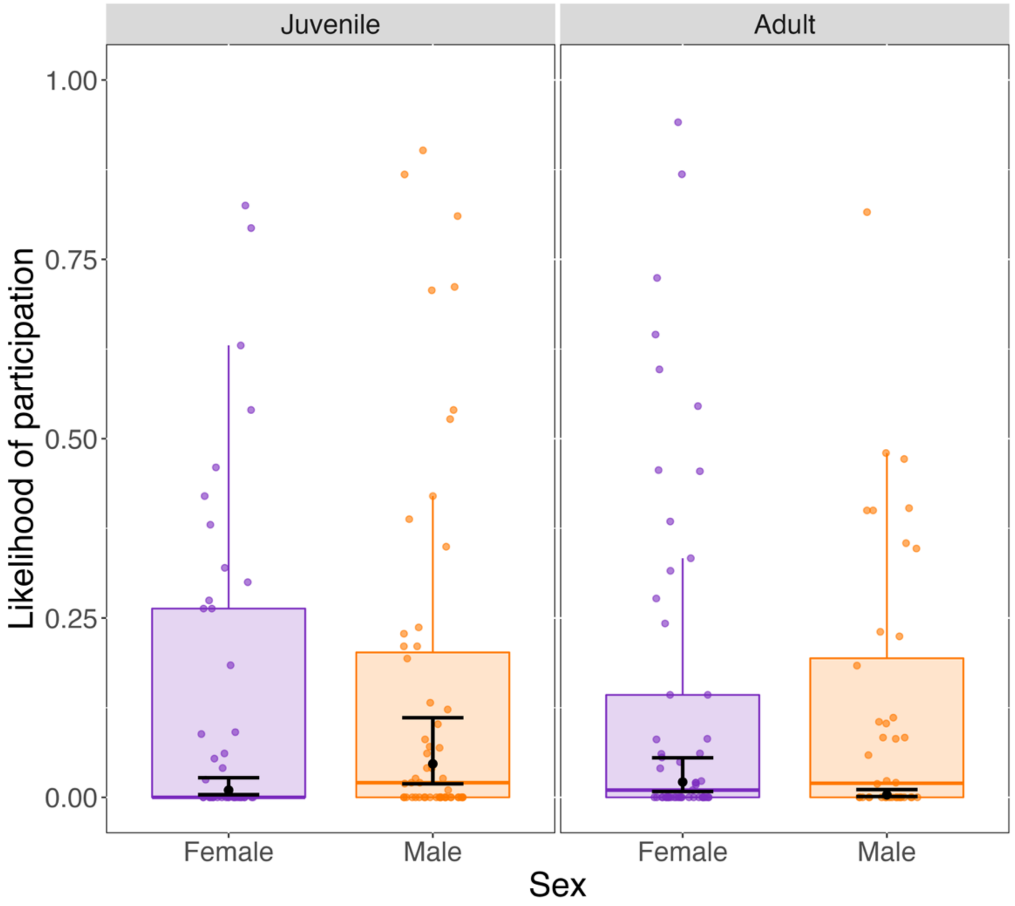

Figure 4. Model predictions (estimated marginal means) of the effect of Age Class and Sex on likelihood of participation in the task in the wild population, shown by age sex class. The predicted value for each Sex and Age Class category is shown by the solid black point, with error bars showing the 95% confidence interval around this prediction. The prediction is at the mean Session Duration. Boxplots show the distribution of the proportion of sessions participated in by each age-sex class, with solid horizontal lines showing the median proportion, upper and lower limits of the box showing the first and third quartiles, and whiskers extending to the highest and lowest values at 1.5 times the interquartile range. Points show the proportion of sessions participated in by each individual, coloured by Sex.

Group also impacted individuals' likelihood of participation (overall significance calculated using the 'Anova' function in the package 'car': $\chi^2 = 14.23$, $p = 0.003$). A post-hoc Tukey test revealed that individuals in Baie Dankie ($\beta = -2.37$, $p = 0.008$) and Kubu ($\beta = -2.27$, $p = 0.016$) were significantly less likely to participate than those in Ankhase. No other significant between-group differences were found.

Analysis 3: Number of attempts made in the task, individual factors in the wild population

The full model was a significantly better fit to the data than a null model containing only the random effect of individual ($\chi^2 = 14.83$, $p = 0.02$). No effect of Sex or Age Class was found, but there was a main effect of Group ($\chi^2 = 8.73$, $p = 0.033$). A post-hoc Tukey test, however, indicated that there was only a marginal, non-significant, difference between Noha and Baie Dankie (HR = 0.53, $p = 0.060$), indicating that individuals in Noha tended to make more attempts at the task than individuals in Baie Dankie.

While not significant, both Sex and Age Class influenced the number of attempts made per session by individuals. Male juveniles were more likely to reach 30 attempts in a session than female juveniles (HR = 0.59, $p = 0.061$, see Figure 5). There was also a non-significant interaction between Age and Sex, such that male adults were more likely to stop participating before reaching 30 attempts than were male juveniles (HR = 2.00, $p = 0.086$). Caution should be taken in interpreting these non-significant effects, but it is possible that with a larger sample size significant sex and age differences in individuals' rate of participation would be found.

Table 4. Results of a Cox mixed effects model predicting the number of attempts made by individuals per session in the task in the wild population.

Effect	**Hazard Ratio**	**Standard error**	***p*-value**
Sex (Male)	0.59	0.28	0.061
Age Class (Adult)	0.80	0.28	0.430
Group (Baie Dankie)	1.89	0.29	0.029
Group (Kubu)	1.72	0.37	0.140
Group (Noha)	1.00	0.29	0.990
Interaction: Age Class x Sex	2.00	0.41	0.086

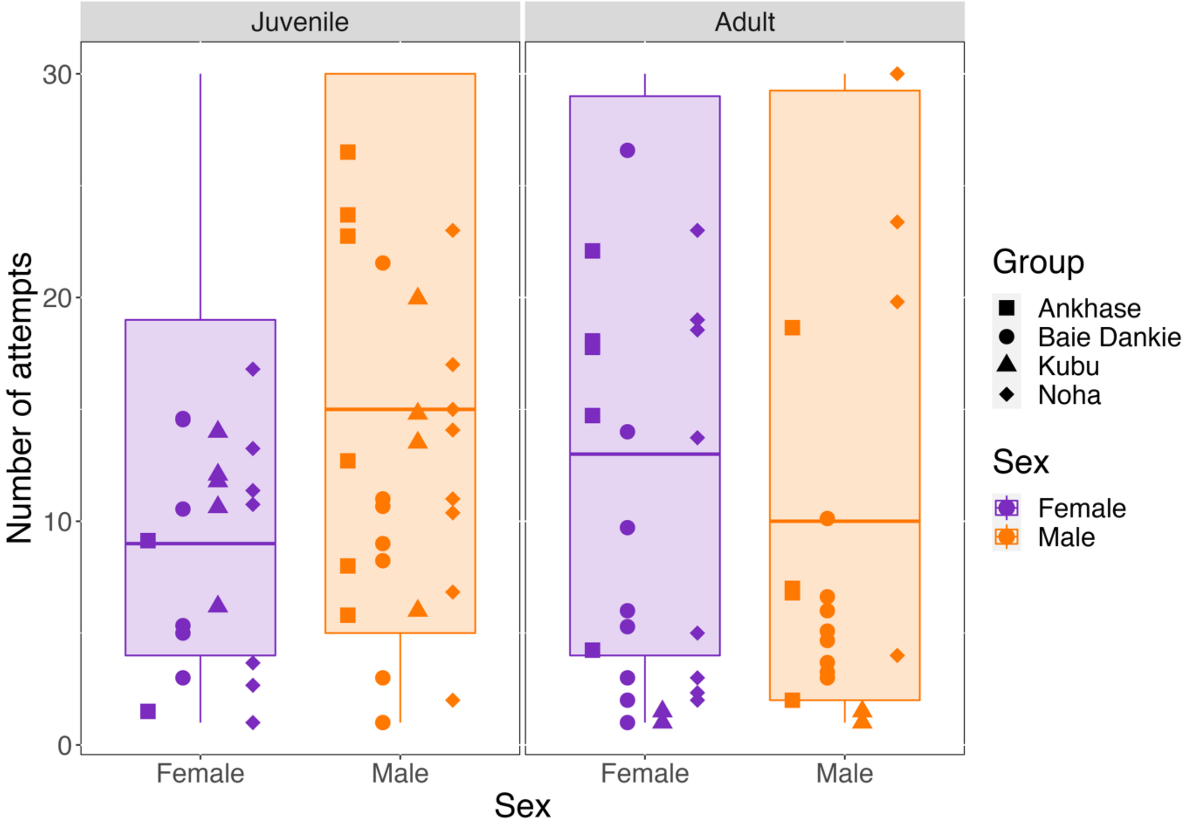

Figure 5. The number of attempts made by individuals in the wild population. Boxplots, colored by sex, show the distribution of the number of attempts made by each age-sex class, with solid lines showing the median number, upper and lower limits of the box showing the first and third quartiles, and whiskers extending to the highest and lowest values at 1.5 times the interquartile range. Solid points show the mean number of attempts made per session by each individual, coloured by Sex with shape indicating Group membership.

Analysis 4: Trials required to pass Task 3

The full model (Analysis 4) analysing the number of attempts individuals required to pass Task 3 was not a better fit to the data than a null model containing only the random effect structure ($\chi^2 = 0.46$, $p = 0.50$; see Table S6 for full model output). Context had no impact upon the number of attempts an individual made before reaching criterion (Figure 6; see Table S6 for full model output).

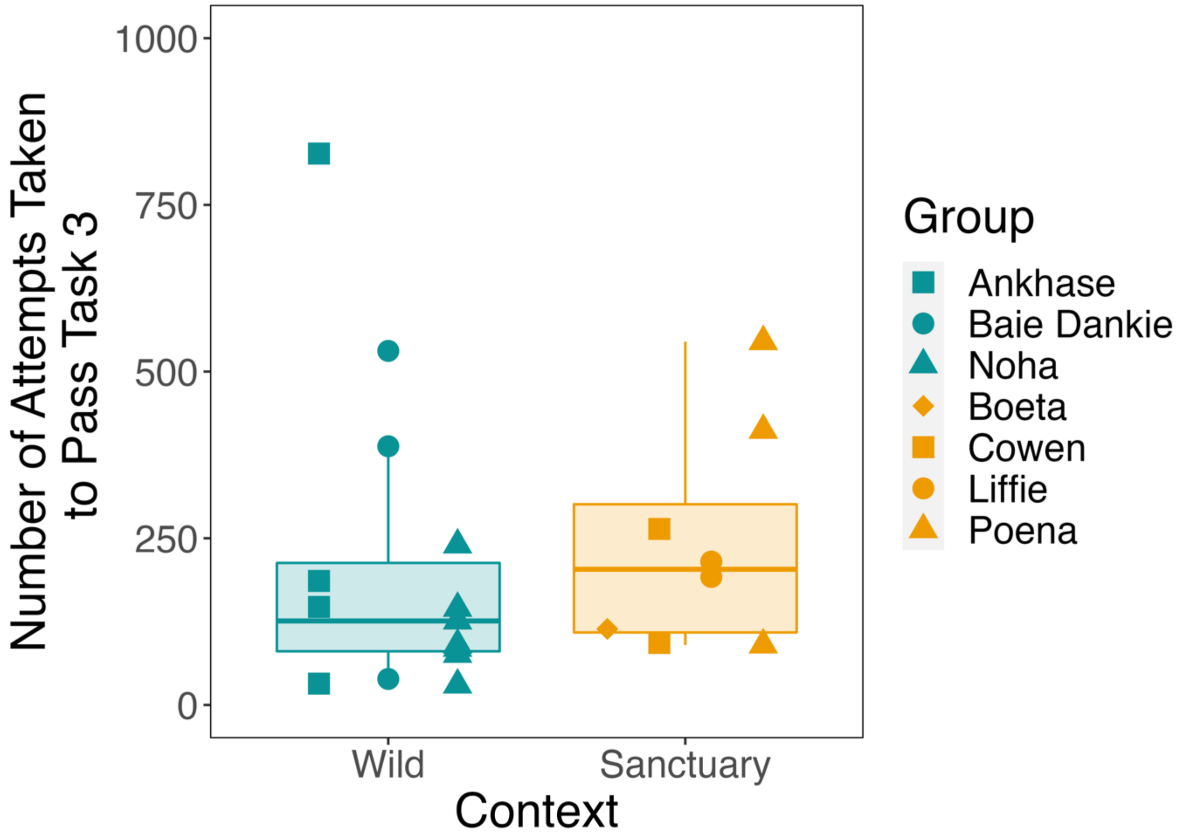

Figure 6. The number of attempts made by individuals prior to reaching criterion on Task 3. Points show the number of attempts made by each individual, coloured by Context, with shape indicating Group membership. Boxplots show the median in a solid line, first and third quartiles at the upper and lower box edges, and whiskers show the largest and smallest values at 1.5 times the interquartile range of the distribution of observations for Wild and Sanctuary contexts.

Discussion

The influence of context on participation

In this study, we compared the participation rates of wild and sanctuary-housed vervet monkeys when offered a touchscreen device displaying various cognitive tasks. According to the free time hypothesis (Kummer & Goodall, 1985), we expected different rates of participation between the captive and wild vervet populations. Our results supported our hypothesis; participation level was influenced by environment. Captive individuals were significantly more likely to interact with the touchscreen than their wild conspecifics. This result supports the free time hypothesis, which would predict higher rates of participation in captive vervet monkeys which have more time and energy to spend interacting with experiments (Kummer & Goodall, 1985). Conversely, wild individuals may have had less free time to interact with the touchscreen due to the distraction of necessary activities such as foraging or being vigilant towards predators. Test sessions in wild groups ended when the group moved away from the test site; in line with the free time hypothesis, pressure to undertake daily activities such as foraging likely reduced the amount of time groups spent in proximity to the task. It should be noted that group sizes were quite disparate between the wild and captive groups tested in this study (see Table S3), with three of the wild groups consistently containing more individuals than any of the captive groups. It is therefore possible that group size influenced individuals’ access to the task, with individuals in captive groups having greater opportunities to interact due smaller group sizes resulting in reduced competition.

Despite significant differences in their likelihood of participation, individuals which reached criterion in Task 3 (learning to reliably touch a blue square, displayed in different locations on the screen, in order to get a reward) in both the wild and sanctuary-housed contexts did so with no difference in the number of attempts required. This result allows us to conclude that any differences in methodology between the two different environments did not affect the amount of time required for individuals to learn the association between touching the blue square and receiving a food reward, and also indicates that individuals in both contexts interacted meaningfully with the task. This result is in line with the findings of Cauchoix et al. (2017), who found no differences in reversal-learning performance between great tits tested in the wild and captivity and appears to be counter to the argument that captive primates display a ‘captivity effect’ in their cognitive abilities (Forss et al., 2020). However, this task was primarily a training task designed to facilitate later testing with more complex problems and is therefore relatively simple. It is possible that any captivity effect, or other differences in cognitive performance between wild and captive populations, may become apparent only when using more challenging tests of cognition.

Group differences in the wild population

Within the wild population, group membership significantly influenced individuals' likelihood of participating in the task. Individuals in Baie Dankie (the largest group, N = 57 - 65) and Kubu (the smallest group, N = 15 - 18) were less likely to participate in testing than those in Ankhase (a medium-sized group, N = 23 - 26). Individuals in Kubu also made fewer attempts than those in Ankhase, and individuals in both Baie Dankie and Kubu made fewer attempts than those in Noha (a medium-sized group, N = 32 - 40). It is possible that membership of a larger group suppressed participation in the case of monkeys in Baie Dankie, due to higher rates of competition. This may have allowed high ranking individuals to monopolise the touchscreen; the effect of rank on participation should be explored in future studies.

Individuals in the smallest group (Kubu) were also less likely to participate. This group was habituated in 2013, more recently than the other wild groups tested in the study, and additionally has previously been found to have a lower habituation index than the other groups tested (Forss et al., 2021). It is therefore possible that in this group, while inter-individual competition for the task was reduced due to the small group size, reduced habituation to humans suppressed participation. Levels of participation may also have been influenced by variation in individual levels of habituation, along with factors such as personality (Webster & Rutz, 2020). Different groups may also contain individuals with differing dominance styles and resource holding potential, potentially rendering monopolisation a greater issue in some groups than others (as has been suggested as an explanation for differing levels of social tolerance in neighbouring chimpanzee groups, Cronin et al., 2014).

While we believe testing multiple groups of wild individuals represents a crucial step forward in our understanding of the impact of social factors upon cognition and behaviour, our interpretation of any differences remains limited somewhat by sample size when it comes to pinpointing the cause of between-group differences, as the four groups tested here are likely to differ in a number of parameters, group size and habituation being only two. Future studies could begin to tease this apart both by more detailed examination of interactions at and around the task (which is outside the scope of the current paper), for example examining the number of individuals who approached the task but were outcompeted, and by incorporating parallel tests of social tolerance and habituation to quantify existing group differences.

Phenotypic traits influenced participation in the wild population

Although many studies have shown that juveniles are less neophobic than adults (Bergman & Kitchen, 2009; Biondi et al., 2010; Miller et al., 2015; Visalberghi et al., 2003), and therefore we expected juveniles to participate and interact with the touchscreen more frequently than adults, our findings regarding this were impacted by sex. Male juveniles were indeed more likely to participate than adult males, but female juveniles were less likely to participate than adult females. Vervet monkeys have a complex social system based on a linear hierarchy organised by matrilines. In our context, it is very likely that juveniles had to wait for the end of monopolization by the high-ranking adult females before interacting with the experiment. Nonetheless, male juveniles were significantly more likely to participate than adult males. When they reach sexual maturity, male vervets disperse from their natal group. When juvenile males approach this milestone, they become less central and more independent (Young et al., 2019). It is therefore possible that the age class effect seen in wild juvenile males in our study is driven by highly explorative older juveniles and high competition levels between juvenile males and other group members. Our sample size did not allow detailed exploration of age effects and competition, but future research should explore these possibilities.

Regarding sex differences, we found that adult females made more attempts at the touchscreen than males. One possible explanation for this can derive from physiological needs. Females, especially during the gestation and lactating period, need higher energy intake and there were females in our sample who were pregnant and had offspring during the period of the experiment. We also found that adult females participated more than males in the experiments. In our study population, it has been shown that females are often dominant over males and many groups have a female as the most dominant individual in the group (Hemelrijk et al., 2020). Therefore, our results could be explained by the fact that dominant adult females might have the greatest opportunity to monopolize and interact with the touchscreen.

Within this study, we found that the likelihood of participation increased with the number of rewards a monkey had received the last time they participated (see Supplemental

Information: Effect of previous success on participation in the wild population). Additionally, regardless of the length of time between experiments, monkeys in both the wild and captivity were motivated to participate again (see Supplemental Information: Effect of breaks in testing). Both these findings could influence the design of future studies.

While our study demonstrates that it is possible to conduct cognitive tests using touchscreen technology in the wild, some sampling biases are likely to apply (Webster & Rutz, 2020) which may impact the generalisability of results, particularly regarding self-selection according to personality and habituation. Our findings indicate that some age and sex classes (adult females and juvenile males) were more likely to participate than others, suggesting that future studies may need to engage in some selection of individuals to achieve balanced samples in this regard – we suggest a potential method to achieve this below.

Future directions

To increase the number of individuals who can interact with touchscreens presented in the wild, it would be useful in the future to integrate more advanced technological methods such as automatic facial recognition (Schofield et al., 2019) or individual radio frequency identification (RFID: Bridge et al., 2019; Fagot & Paleressompoulle, 2009). We note that fitting RFID tags would require sedation of the study subjects and therefore may not be appropriate in many cases (Soulsbury et al., 2020), but could be an option in populations in which individuals are already sedated to fit GPS collars as a standard protocol, rather than sedating animals for the sole purpose of fitting RFID tags. These methods would allow the touchscreen to automatically present the appropriate task according to the individual interacting, and would allow the process to be automatized, with the touchscreen remaining in place in the field for prolonged periods. This would not prevent monopolisation, and so might result in larger amounts of data but not necessarily larger sample sizes. Automatization would also potentially facilitate the presentation of multiple touchscreens in the field at the same time, which likely would reduce the impact of monopolisation upon participation. To further address monopolisation, individuals could also be trained to recognize a visual pattern on the screen indicating that it is their turn to interact with the task, with the screen blocked for all individuals except the target participant. This procedure could also be used to generate more balanced samples in terms of age and sex class (a technique successfully used with artificial foraging boxes by Borgeaud and Bshary, 2015). We found that across the four wild groups, the total number of individuals interacting per session was relatively similar regardless of group size (see Supplemental Information: Number of individuals participating per session in wild groups), and so methods to increase presentation time by automating aspects of the process may help to increase sample sizes.

Further research could also explore in more detail how social factors influence participation in experimental tasks. Our findings point towards potential roles for group size and resulting competition, and inter-individual tolerance or monopolisation by dominant individuals. These represent potential constraints that testing in social contexts may place upon cognitive testing of wild animals. However, it is also likely that social factors facilitate participation and interaction with novel objects such as the touchscreen, through processes such as social facilitation (Dindo et al., 2009; Forss et al., 2017; Miller et al., 2014). There is also the potential for social learning to influence interaction with the task, via local or stimulus enhancement, or more high-fidelity forms of social transmission. As adult females were key participants in our task, and are the preferred models for social learning in wild vervet monkeys (van de Waal et al., 2010; Bono et al., 2018), the presence of adult females interacting with the task may well facilitate participation by other age sex classes in this species.

We believe that touchscreen testing with wild populations has the potential to provide a step forward in studying non-human cognition. However, the potential ethical implications of this type of research must be carefully considered (Gruber, 2022). Like other types of field experiment, providing touchscreens for testing animals in their natural habitat removes the need to bring animals into captivity for testing, opening up the possibility of testing animals for which capturing would be unethical. This benefit aside, it is unlikely that touchscreen testing will be appropriate in all situations. The current set-up requires the relatively close presence of an experimenter, and food rewards are provided. The possibility of individuals forming an association between humans and food rewards may be dangerous in some species (for example, the great apes), and provisioning can carry risks (Fedigan, 2010). Exposure to human artefacts in the form of screens may also be considered unacceptable for some species (though we note that video demonstrations have previously been shown to wild primates e.g. Gunhold et al., 2014). The population of monkeys tested in the current paper do not live in close proximity to humans, and therefore the risk of habituation to the touchscreen spilling over into increased attempts to interact with screens in other contexts is limited, but this should be considered if testing urban populations, for example. Additionally, the potential for exposure to the touchscreen to alter the natural behavioural repertoire of the subjects should be considered. Vervet monkeys are not endangered, and only a subset of groups resident at the IVP were included in touchscreen testing. However, for endangered species in which the natural behavioural repertoire must be carefully conserved, exposure to the touchscreen and other artificial experimental tasks may not be appropriate.

Conclusion

This study explored rates of participation at a novel touchscreen task in both wild and sanctuary-housed vervet monkeys. While participation rates were higher in captive monkeys, in line with the 'free time' hypothesis, performance as measured by the number of trials taken to reach criterion in a task was the same across both contexts. Age and sex differences in participation of wild monkeys were observed, which may reflect different levels of neophobia and motivation between juveniles and adults and males and females, as well as the ability to compete for access to and monopolise the touchscreen. The study demonstrates that touchscreen technology can be deployed successfully in a wild primate population, and that while levels of participation may differ, results comparable to the performance of captive populations can be achieved. This opens the door for comparative studies examining the ways in which cognitive abilities may differ between wild and captive primate populations and offers the opportunity to validate results from captive studies in a wild population using identical methodology.

Acknowledgements

This study was supported by the Swiss National Science Foundation (PP03P3_170624 and PP00P3_198913) and the Branco Weiss Fellowship—Society in Science granted to Erica van de Waal. We are grateful to the van der Walt family for permission to conduct the study at Mawana game reserve and Sandy Cronk, owner of the rehabilitation centre for permission to conduct the study at W.A.T.C.H. We thank Ezemvelo KZN Wildlife who gave us permission to conduct our research on vervet monkeys in South Africa. We are also grateful to M. Henshall for managing the Inkawu Vervet Project and his help in the data collection. We thank all the team from the IVP for their support and help in data collection, with a special thanks to T. Henshall, M. German, and S. Glosenger-Thrasher, who provided great help in carrying out the experiments. We are grateful to V. Schmitt for her help with the touchscreen procedure and for her participation in the design of the study. We thank students from the EPFL (École Polytechnique Féderale de Lausanne) for writing the test code. Vervet illustrations in the graphical abstract are courtesy of Chen Chi-Hsin. Finally, we are grateful to the reviewers and editors for their constructive comments.

Conflict of interest

The authors have no conflict of interest to declare.

Author contributions

Tecla Mohr participated in the design of the study, collected field data, compiled the dataset and drafted the manuscript. Erica van de Waal coordinates the research at IVP in South Africa, designed the study, financed the study and corrected the manuscript. Rachel Harrison conducted the data analysis, drafted and corrected the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability statement

All relevant data and analysis code are available from the OSF digital repository https://doi.org/10.17605/OSF.IO/P2XGQ (Harrison et al., 2022).

References

ASAB/ABS. (2020). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour*, *159*, I–XI. https://doi.org/10.1016/j.anbehav.2019.11.002

Ashton, B. J., Thornton, A., & Ridley, A. R. (2018). An intraspecific appraisal of the social intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170288. https://doi.org/10.1098/rstb.2017.0288

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting Linear Mixed-Effects Models using lme4* (arXiv:1406.5823). arXiv. https://doi.org/10.48550/arXiv.1406.5823

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., & Grothendieck, G. (2011). Package 'lme4'. *Linear Mixed-Effects Models Using S4 Classes. R Package Version*, *1*(6).

Bean, A. (1999). Ecology of sex differences in great ape foraging. In P. C. Lee (Ed.), *Comparative Primate Socioecology* (1st ed., pp. 339–362). Cambridge University Press. https://doi.org/10.1017/CBO9780511542466.017

Benson-Amram, S., Weldele, M. L., & Holekamp, K. E. (2013). A comparison of innovative problem-solving abilities between wild and captive spotted hyaenas, Crocuta crocuta. *Animal Behaviour*, *85*(2), 349–356. https://doi.org/10.1016/j.anbehav.2012.11.003

Bergman, T. J., & Kitchen, D. M. (2009). Comparing responses to novel objects in wild baboons (Papio ursinus) and geladas (Theropithecus gelada). *Animal Cognition*, *12*(1), 63–73. https://doi.org/10.1007/s10071-008-0171-2

Biondi, L. M., Bó, M. S., & Vassallo, A. I. (2010). Inter-individual and age differences in exploration, neophobia and problem-solving ability in a Neotropical raptor (Milvago chimango). *Animal Cognition*, *13*(5), 701–710. https://doi.org/10.1007/s10071-010-0319-8

Bono, A. E. J., Whiten, A., van Schaik, C., Krützen, M., Eichenberger, F., Schnider, A., & van de Waal, E. (2018). Payoff- and Sex-Biased Social Learning Interact in a Wild Primate Population. *Current Biology*, *28*(17), 2800-2805.e4. https://doi.org/10.1016/j.cub.2018.06.015

Borgeaud, C., & Bshary, R. (2015). Wild Vervet Monkeys Trade Tolerance and Specific Coalitionary Support for Grooming in Experimentally Induced Conflicts. *Current Biology*, *25*(22), 3011–3016. https://doi.org/10.1016/j.cub.2015.10.016

Botting, J., Whiten, A., Grampp, M., & van de Waal, E. (2018). Field experiments with wild primates reveal no consistent dominance-based bias in social learning. *Animal Behaviour*, *136*, 1–12. https://doi.org/10.1016/j.anbehav.2017.11.025

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Bridge, E. S., Wilhelm, J., Pandit, M. M., Moreno, A., Curry, C. M., Pearson, T. D., Proppe, D. S., Holwerda, C., Eadie, J. M., Stair, T. F., Olson, A. C., Lyon, B. E., Branch, C. L., Pitera, A. M., Kozlovsky, D., Sonnenberg, B. R., Pravosudov, V. V., & Ruyle, J. E. (2019). An Arduino-Based RFID Platform for Animal Research. *Frontiers in Ecology and Evolution*, *7*. https://doi.org/10.3389/fevo.2019.00257

Canteloup, C., Hoppitt, W., & van de Waal, E. (2020). Wild primates copy higher-ranked individuals in a social transmission experiment. *Nature Communications*, *11*(1), 459. https://doi.org/10.1038/s41467-019-14209-8

Cole, E. F., Morand-Ferron, J., Hinks, A. E., & Quinn, J. L. (2012). Cognitive Ability Influences Reproductive Life History Variation in the Wild. *Current Biology*, *22*(19), 1808–1812. https://doi.org/10.1016/j.cub.2012.07.051

Cronin, K. A., van Leeuwen, E. J. C., Vreeman, V., & Haun, D. B. M. (2014). Population-level variability in the social climates of four chimpanzee societies. *Evolution and*

Human Behavior, *35*(5), 389–396. https://doi.org/10.1016/j.evolhumbehav.2014.05.004

Damerius, L. A., Forss, S. I. F., Kosonen, Z. K., Willems, E. P., Burkart, J. M., Call, J., Galdikas, B. M., Liebal, K., Haun, D., & van Schaik, C. P. (2017). Orientation toward humans predicts cognitive performance in orang-utans. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/srep40052

Dindo, M., Whiten, A., & de Waal, F. B. M. (2009). Social facilitation of exploratory foraging behavior in capuchin monkeys (Cebus apella). *American Journal of Primatology*, *71*(5), 419–426. https://doi.org/10.1002/ajp.20669

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, *6*(5), 178–190. https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Egelkamp, C. L., & Ross, S. R. (2019). A review of zoo-based cognitive research using touchscreen interfaces. *Zoo Biology*, *38*(2), 220–235. https://doi.org/10.1002/zoo.21458

Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, *41*(2), 396–404. https://doi.org/10.3758/BRM.41.2.396

Fedigan, L. M. (2010). Ethical issues faced by field primatologists: Asking the relevant questions. *American Journal of Primatology*, *72*(9), 754–771. https://doi.org/10.1002/ajp.20814

Forss, S. I. F., Koski, S. E., & van Schaik, C. P. (2017). Explaining the Paradox of Neophobic Explorers: The Social Information Hypothesis. *International Journal of Primatology*, *38*(5), 799–822. https://doi.org/10.1007/s10764-017-9984-7

Forss, S. I. F., Motes-Rodrigo, A., Dongre, P., Mohr, T., & van de Waal, E. (2022). Captivity and habituation to humans raise curiosity in vervet monkeys. *Animal Cognition*, *25*(3), 671–682. https://doi.org/10.1007/s10071-021-01589-y

Forss, S. I. F., Motes-Rodrigo, A., Hrubesch, C., & Tennie, C. (2020). Chimpanzees'(Pan troglodytes) problem-solving skills are influenced by housing facility and captive care duration. *PeerJ*, *8*, e10263. https://doi.org/10.7717/peerj.10263

Forss, S. I. F., Schuppli, C., Haiden, D., Zweifel, N., & van Schaik, C. P. (2015). Contrasting responses to novelty by wild and captive orangutans: Novelty Response in Orangutans. *American Journal of Primatology*, *77*(10), 1109–1121. https://doi.org/10.1002/ajp.22445

Gareta García, M., Lemieux, D., & Bshary, R. (2021). Factors affecting tolerance persistence after grooming interactions in wild female vervet monkeys, Chlorocebus pygerythrus. *Animal Behaviour*, *177*, 135–145. https://doi.org/10.1016/j.anbehav.2021.04.025

Gruber, T. (2022). An ethical assessment of the use of old and new methods to study sociality in wild animals. *Methods in Ecology and Evolution*, *n/a*(n/a). https://doi.org/10.1111/2041-210X.13988

Grund, C., Neumann, C., Zuberbühler, K., & Gruber, T. (2019). Necessity creates opportunities for chimpanzee tool use. *Behavioral Ecology*, *30*(4), 1136–1144. https://doi.org/10.1093/beheco/arz062

Gunhold, T., Whiten, A., & Bugnyar, T. (2014). Video demonstrations seed alternative problem-solving techniques in wild common marmosets. *Biology Letters*, *10*(9), 20140439. https://doi.org/10.1098/rsbl.2014.0439

Harrison, R. A., & van de Waal, E. (2022). The unique potential of field research to understand primate social learning and cognition. *Current Opinion in Behavioral Sciences*, *45*, 101132. https://doi.org/10.1016/j.cobeha.2022.101132

Harrison, R. A., Mohr, T. & van de Waal, E. (2022). Data from: Lab cognition going wild: implementing a new portable touchscreen system in vervet monkeys. Open Science Framework digital repository, https://doi.org/10.17605/OSF.IO/P2XGQ

Hemelrijk, C. K., Wubs, M., Gort, G., Botting, J., & van de Waal, E. (2020). Dynamics of Intersexual Dominance and Adult Sex- Ratio in Wild Vervet Monkeys. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00839

Hopper, L. M. (2017). Cognitive research in zoos. *Current Opinion in Behavioral Sciences*, *16*, 100–110. https://doi.org/10.1016/j.cobeha.2017.04.006

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17*(23), R1004–R1005. https://doi.org/10.1016/j.cub.2007.10.027

Kawai, N., & Matsuzawa, T. (2000). Numerical memory span in a chimpanzee. *Nature*, *403*(6765), 39–40. https://doi.org/10.1038/47405

Kendal, R., Coe, R., & Laland, K. (2005). Age differences in neophilia, exploration, and innovation in family groups of callitrichid monkeys. *American Journal of Primatology 66*(2), 167–188. https://doi.org/10.1002/ajp.20136

Kumpan, L. T., Smeltzer, E. A., & Teichroeb, J. A. (2020). Animal cognition in the field: Performance of wild vervet monkeys (Chlorocebus pygerythrus) on a reversal learning task. *Animal Cognition*, *23*(3), 523–534. https://doi.org/10.1007/s10071-020-01356-5

Leinwand, J. G., Huskisson, S. M., Egelkamp, C. L., & Hopper, L. M. (2020). Within-and between-species variation in the responses of three primate species to a touchscreen gambling task. *Learning and Motivation*, *71*, 101635. https://doi.org/10.1016/j.lmot.2020.101635

Li, M. F., Arseneau-Robar, T. J. M., Smeltzer, E. A., & Teichroeb, J. A. (2021). Be early or be tolerated: Vervet monkey, Chlorocebus pygerythrus, foraging strategies in a dispersed resource. *Animal Behaviour*, *176*, 1–15. https://doi.org/10.1016/j.anbehav.2021.03.010

Martin, C. F., Muramatsu, A., & Matsuzawa, T. (2022). Apex and ApeTouch: Development of a Portable Touchscreen System and Software for Primates at Zoos. *Animals*, *12*(13), 1660. https://doi.org/10.3390/ani12131660

Martina, C., Cowlishaw, G., & Carter, A. J. (2021). Individual differences in task participation in wild chacma baboons. *Animal Behaviour*, *172*, 73–91. https://doi.org/10.1016/j.anbehav.2020.11.020

Matsuzawa, T. (1994). Field experiments on use of stone tools by chimpanzees in the wild. *Chimpanzee Cultures*, 351–370.

McCune, K. B., Jablonski, P., Lee, S., & Ha, R. R. (2019). Captive jays exhibit reduced problem-solving performance compared to wild conspecifics. *Royal Society Open Science*, *6*(1), 181311. https://doi.org/10.1098/rsos.181311

Mertz, J., Surreault, A., van de Waal, E., & Botting, J. (2019). Primates are living links to our past: The contribution of comparative studies with wild vervet monkeys to the field of social cognition. *Cortex*, *118*, 65–81. https://doi.org/10.1016/j.cortex.2019.03.007

Miller, R., Bugnyar, T., Pölzl, K., & Schwab, C. (2015). Differences in exploration behaviour in common ravens and carrion crows during development and across social context. *Behavioral Ecology and Sociobiology*, *69*(7), 1209–1220. https://doi.org/10.1007/s00265-015-1935-8

Miller, R., Schiestl, M., Whiten, A., Schwab, C., & Bugnyar, T. (2014). Tolerance and Social Facilitation in the Foraging Behaviour of Free-Ranging Crows (Corvus corone corone; C. c. Cornix). *Ethology*, *120*(12), 1248–1255. https://doi.org/10.1111/eth.12298

Morand-Ferron, J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Current Opinion in Behavioral Sciences*, *16*, 73–79. https://doi.org/10.1016/j.cobeha.2017.03.008

Morand-Ferron, J., Cole, E. F., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: A review of practical and conceptual challenges. *Biological Reviews*, *91*(2), 367–389. https://doi.org/10.1111/brv.12174

Morand-Ferron, J., Hamblin, S., Cole, E. F., Aplin, L. M., & Quinn, J. L. (2015). Taking the operant paradigm into the field: Associative learning in wild great tits. *PloS One*, *10*(8), e0133821. https://doi.org/10.1371/journal.pone.0133821

Paula, J. R., Baptista, M., Carvalho, F., Repolho, T., Bshary, R., & Rosa, R. (2019). The past, present and future of cleaner fish cognitive performance as a function of CO2 levels. *Biology Letters*, *15*(12), 20190618. https://doi.org/10.1098/rsbl.2019.0618

Pokorny, J. J., & de Waal, F. B. (2009). Monkeys recognize the faces of group mates in photographs. *Proceedings of the National Academy of Sciences*, *106*(51), 21539–21543. https://doi.org/10.1073/pnas.0912174106

R Core Team. (2020). *R: A Language and Environment for Statistical Computing.*

Raine, N. E., & Chittka, L. (2008). The correlation of learning speed and natural foraging success in bumble-bees. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1636), 803–808. https://doi.org/10.1098/rspb.2007.1652

RStudio Team. (2020). *RStudio: Integrated Development Environment for R.* www.rstudio.com

Salwiczek, L. H., Prétôt, L., Demarta, L., Proctor, D., Essler, J., Pinto, A. I., Wismer, S., Stoinski, T., Brosnan, S. F., & Bshary, R. (2012). Adult Cleaner Wrasse Outperform Capuchin Monkeys, Chimpanzees and Orang-utans in a Complex Foraging Task

Derived from Cleaner – Client Reef Fish Cooperation. *PLOS ONE*, *7*(11), e49068. https://doi.org/10.1371/journal.pone.0049068

Schmitt, V. (2019). Implementing portable touchscreen-setups to enhance cognitive research and enrich zoo-housed animals. *Journal of Zoo and Aquarium Research*, *7*(2), 50–58. https://doi.org/10.19227/jzar.v7i2.314

Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., & Carvalho, S. (2019). Chimpanzee face recognition from videos in the wild using deep learning. *Science Advances*, *5*(9), eaaw0736. https://doi.org/10.1126/sciadv.aaw0736

Seferta, A., Guay, P., Marzinotto, E., & Lefebvre, L. (2001). Learning differences between feral pigeons and zenaida doves: The role of neophobia and human proximity. *Ethology*, *107*(4), 281–293. https://doi.org/10.1046/j.1439-0310.2001.00658.x

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal Behaviour*, *61*(2), 277–286. https://doi.org/10.1006/anbe.2000.1606

Soulsbury, C. D., Gray, H. E., Smith, L. M., Braithwaite, V., Cotter, S. C., Elwood, R. W., Wilkinson, A., & Collins, L. M. (2020). The welfare and ethics of research involving wild animals: A primer. *Methods in Ecology and Evolution*, *11*(10), 1164–1181. https://doi.org/10.1111/2041-210X.13435

Therneau, T. M. (2015). Package ‘coxme’. *R Package Version*, *2*(5).

Thornton, A., & Samson, J. (2012). Innovative problem solving in wild meerkats. *Animal Behaviour*, *83*(6), 1459–1468. https://doi.org/10.1016/j.anbehav.2012.03.018

van de Waal, E., Claidière, N., & Whiten, A. (2013). Social learning and spread of alternative means of opening an artificial fruit in four groups of vervet monkeys. *Animal Behaviour*, *85*(1), 71–76. https://doi.org/10.1016/j.anbehav.2012.10.008

van de Waal, E., Claidière, N., & Whiten, A. (2015). Wild vervet monkeys copy alternative methods for opening an artificial fruit. *Animal Cognition*, *18*(3), 617–627. https://doi.org/10.1007/s10071-014-0830-4

van de Waal, E., Renevey, N., Favre, C. M., & Bshary, R. (2010). Selective attention to philopatric models causes directed social learning in wild vervet monkeys. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1691), 2105–2111. https://doi.org/10.1098/rspb.2009.2260

van de Waal, E., & Whiten, A. (2012). Spontaneous Emergence, Imitation and Spread of Alternative Foraging Techniques among Groups of Vervet Monkeys. *PLOS ONE*, *7*(10), e47008. https://doi.org/10.1371/journal.pone.0047008

van Horik, J. O., Langley, E. J., Whiteside, M. A., & Madden, J. R. (2017). Differential participation in cognitive tests is driven by personality, sex, body condition and experience. *Behavioural Processes*, *134*, 22–30. https://doi.org/10.1016/j.beproc.2016.07.001

Visalberghi, E., Janson, C. H., & Agostini, I. (2003). Response Toward Novel Foods and Novel Objects in Wild Cebus apella. *International Journal of Primatology*, *24*(3), 653–675. https://doi.org/10.1023/A:1023700800113

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, *582*(7812), 337–340. https://doi.org/10.1038/d41586-020-01751-5

Woolverton, W., Ator, N., Beardsley, P., & Carroll, M. E. (1989). Effects of environmental conditions on the psychological well-being of primates: A review of the literature. *Life Sciences*, *44*(14), 901–917. https://doi.org/10.1016/0024-3205(89)90489-X

Young, C., McFarland, R., Ganswindt, A., Young, M. M. I., Barrett, L., & Henzi, S. P. (2019). Male residency and dispersal triggers in a seasonal breeder with influential females. *Animal Behaviour*, *154*, 29–37. https://doi.org/10.1016/j.anbehav.2019.06.010