Animal Cognition
https://doi.org/10.1007/s10071-022-01668-8

ORIGINAL PAPER

Human interventions in a behavioural experiment for Asian Elephants (*Elephas maximus*)

Pui Ching Chu[1] · Kaja Wierucka[1,2] · Derek Murphy[1,3] · Hannah Bethany Tilley[1] · Hannah Sue Mumby[1]

Received: 25 October 2021 / Revised: 21 July 2022 / Accepted: 1 August 2022

Abstract
Experiments are widely used to investigate the behaviour and cognition of animals. While the automation of experiments to avoid potential experimenter bias is sometimes possible, not all experiments can be conducted without human presence. This is particularly true for large animals in captivity, which are often managed by professional handlers. For the safety of the animals and experimenters, a handler must be present during behavioural studies with certain species. It is not always clear to what extent cues provided by handlers affect the animals, and therefore the experimental results. In this study, we investigate handler interventions during the training process for a behavioural experiment with Asian elephants (*Elephas maximus*) in Nepal. We show that elephant handlers (mahouts) intervened to guide elephants in performing the learning task using vocal and behavioural cues, despite experimenters requesting minimal intervention. We found that although the frequency of mahout interventions did not decrease as the training progressed, the nature of their interventions changed. We also found more non-verbal than verbal cues across the training. Our results suggest that guidance from handlers may be common in behavioural studies, and continued consideration should be put into experimental design to reduce or account for cues that animals may receive from humans. This study also emphasises the need to take into account the presence of humans in interpreting the results of animal behavioural experiments, which not only presents challenges to behavioural research, but also represents opportunities for further study.

Keywords Choice task · Experimenter bias · Animal behaviour · Captivity · Animal handler · Mahout

Introduction

Behavioural experiments are commonly used to investigate animal cognition and behaviour (Cuthill 1991). This widespread use has instigated discussions of the responsibilities of humans towards animals used in research and the impacts on animals when using these methods, given they involve animals that cannot consent, or have studies explained to them in the way that experiments involving human participants must (Barnard 2007). The presence of human researchers or animal handlers in such studies, while often necessary, may impose limitations or biases on the study design and data collection, and may affect the generalisability of the findings or even the ability to address the research question (Marsh and Hanlon 2004). Therefore, a primary focus has been to reduce the possibility of bias that humans could introduce to a study (Burghardt et al. 2012). For example, human beliefs, assumptions, and other aspects of their identity or role may introduce confirmation bias into data collected through observations of animal behaviour (Marsh and Hanlon 2007; van Wilgenburg and Elgar 2013). The use of double blind trials to reduce these impacts has been proposed and sometimes carried out (Tuyttens et al. 2014; Holman et al. 2015). Equally, if not more concerning is the possibility that humans directly influence the behaviour of the animals that are the focus of behavioural studies, for example, through behavioural cues expressed consciously or otherwise (Samhita and Gross 2013). This was most famously demonstrated in the impressive feats of apparently human-like cognitive abilities performed by the horse

✉ Hannah Sue Mumby
hsmumby@hku.hk

1 Area of Ecology and Biodiversity, School of Biological Sciences, University of Hong Kong, Pok Fu Lam, Hong Kong

2 Present Address: Department of Anthropology, University of Zurich, Zurich, Switzerland

3 Present Address: Cognitive Ethology Laboratory, German Primate Centre, Göttingen, Germany

Published online: 20 August 2022

Clever Hans in the early twentieth century, who appeared able to solve arithmetic problems, tell the time and date, and spell names of artists when presented with their work. In fact, he was rather (and perhaps similarly impressively) responding to subtle behavioural cues from his handler and other humans (Samhita and Gross 2013). These issues represent important points of discussion, because the central goal of most animal behaviour experiments is to objectively observe, measure, and analyse animal behaviour, with limited impact of human observers, handlers, or experimenters.

Researchers have attempted to reduce the impact of humans in experiments in various ways, such as by automating behavioural experiments, for example, through the use of operant chambers, or Skinner Boxes (Mueller-Paul et al. 2014). Automated experiments have been commonly conducted on model species, including rodents, primates, and pigeons (Mueller-Paul et al. 2014), and other species in captivity such as the American black bear (*Ursus americanus*), bottlenose dolphin (*Tursiops truncatus*), and red-footed tortoise (*Chelonoidis carbonaria*) (Egelkamp and Ross 2019). However, not all experiments can be automated, and often, experiments with animals in captivity require keepers or handlers to be present or even involved to ensure the safety of the animals and experimenters. For instance, in experiments investigating relative quantity judgement in South American sea lions (*Otaria flavescens*), bottlenose dolphins (*Tursiops truncatus*), and beluga whales (*Delphinapterus leucas*), the presence of trainers was built into the experimental design to maintain the animal before the start of the experiment, present them with food rewards, and indicate to the experimenter when the animals were tired or not willing to participate in the experiment (Abramson et al. 2011, 2013). In another relative quantity judgement experiment, where dogs (*Canis familiaris*) judged the quantity using olfaction, the owner had the role of releasing their dog at the start of the experiment (Horowitz et al. 2013), while in an experiment on spatial memory in Asian small-clawed otters (*Aonyx cinereus*), a keeper was responsible for ushering the animals to different test areas (Perdue et al. 2013).

This is particularly important when working with large mammals, such as Asian elephants (*Elephas maximus*), where experimenters have relied on working closely with animal handlers to successfully conduct behavioural experiments (e.g., Plotnik et al. 2011, 2014, 2019; Dale and Plotnik 2017). However, handlers and experimenters present during experiments may (intentionally or inadvertently) give cues to the animals (Miklösi et al. 1998; Watters and Krebs 2019). Therefore, it is important for researchers to consider and account for handler presence, as they may influence the behaviour of captive animals during experiments. Furthermore, if data on this are collected systematically, they can also provide insights into how humans interact with animals. In circumstances where humans cannot be removed from a study, these points represent two options that are not mutually exclusive for researchers: attempting to measure human–animal interactions in experiments as we would measure other environmental factors, and using the inevitable interactions as an area of investigation in itself.

Approximately one-third of Asian elephants live in captive conditions (Sukumar 2006), where they can range in their natural habitat and are managed by specialised handlers called mahouts (Locke 2011). Traditionally, the occupation is passed down through the male family line (Hart and Sundar 2000), where mahouts learn handling skills through observation and apprenticeship from a young age (Crawley et al. 2019). They often work with the same elephant over many years, sometimes even decades (Locke 2011). Mahouts are responsible for feeding, cleaning, training, and driving their elephants (Hart 1994), in addition to participating in religious ceremonies, tourism activities, and transportation of materials (Mumby 2019). The elephants have contact with humans without barriers, and mahouts are physically close to the elephants they work with during training and other daily tasks (Highfill et al. 2016). This system, common throughout south and southeast Asia, requires mahouts to communicate regularly with elephants, where communication may be verbal or non-verbal, for example involving visual signals or physical touch (Lainé 2016). Training elephants in any new task or introducing a new elephant to the group involves extensive interactions with the mahout (Hart and Sundar 2000); thus, such interactions are fundamental to the relationship between elephant and mahout.

Conducting behavioural experiments on captive Asian elephants often requires the presence of their mahouts for the safety of both the animals and experimenters (e.g., Plotnik et al. 2011, 2014, 2019; Dale and Plotnik 2017), which provides an excellent opportunity to investigate how and when handlers provide cues to their animals during behavioural experiments. As the risks of working with an animal weighing more than two tonnes are high (Mumby et al. 2015), the presence of their handler is essential. Mahouts can manage the elephants' behaviour when needed, as the elephants are trained to respond to cues and commands used by mahouts, such as lying down, getting up, and walking forward and backwards (Hart and Sundar 2000). Although measures have been taken to minimise the effect of mahouts on elephant behaviour during experiments, such as restricting the frequency or timing of vocal commands that can be given during the experiment (Plotnik et al. 2011; Dale and Plotnik 2017), we cannot eliminate the possibility that behavioural cues from mahouts may influence the animals' behaviour. Therefore, as captive Asian elephants have a complex relationship with their mahouts, who must be present during experiments, the species makes an excellent study system for investigating handler influence in behavioural studies.

In this study, we investigated the training phase of an olfactory learning task (choice test) involving captive Asian elephants to determine the type and frequency of cues given by mahouts during the task, and the response of elephants to these cues. This setting highlighted a particular challenge, where the mahouts' presence was essential for successful completion of experimental trials and animal safety, yet the presence of handlers can generate bias regarding the behaviour of the animal. This tension between the necessity of handler presence and incompatibility with experimental design is central to our analysis. We sought to investigate (1) which elephant behaviours elicited verbal or behavioural cues from mahouts and caused them to intervene with the training; (2) how elephants responded to these mahout interventions; and (3) whether the nature of mahout intervention was consistent throughout the training. Our findings provide insight into handler–animal interactions and contribute to improving the experimental design of future research studies.

Methods

Study site

The learning task was conducted at Tiger Tops Tharu Lodge (27°34′10.5′N 84°06′06.9′E) from January to March 2020. Situated outside of Chitwan National Park in Nepal, the lodge owns and houses ten female Asian elephants for wildlife tourism. Eight of the elephants were able to participate in this study (Table 1). The elephants spend the majority of time in corrals where they were kept in pairs ($n=6$) or alone ($n=2$) but within visual, auditory, and olfactory contact with at least one other elephant (Mumby 2019). Each elephant had a first mahout, who interacts with it throughout the day and is responsible for training and preparation of the elephant to participate in tourism activities (Mumby 2019). Their daily duties include feeding and bathing their elephant, cleaning corrals, 'driving' their elephant using verbal and physical commands with their feet during forest trekking and grass cutting, and checking on the elephants when they are asleep (Mumby 2019). Each elephant also has a second mahout, who oversees caretaking responsibilities when the first mahout is on leave (Mumby 2019). The close relationships each elephant has with two human caretakers offer opportunities to investigate how elephants respond to each of them. Throughout these activities, all mahouts interact closely with the elephants and give them verbal and physical commands, for example to direct them to move, stop, or lie down. The specific nature of these commands varies between mahout-elephant dyads, but in all cases, they serve to mediate the interaction of the elephants with the environment (Lainé 2016), and are the primary method of communication from humans to elephants in this system. The elephants might participate in tourism activities (which did not include riding by tourists), for example walks through the forest or bathing in the river (Mumby 2019).

Table 1 Approximate age of each elephant and time spent with the first and second mahout (if they participated in the study) in years

Elephant	Age	Years with first mahout	Years with second mahout
1	40	12	NA
2	45	3	15
3	50	21	9
4	55	9	1
5	60	12	3
6	55	8	15
7	45	3	5
8	35	28	NA

Video recordings

The video recordings used in this study were collected during the training phase of an experiment, in which eight captive female elephants at the lodge were being trained to take part in olfactory food-based choice tests. The initial training had four training stages. For each stage, the elephants repeated sets of six trials until they passed the criterion of the stage (Table 2).

In each trial, an elephant investigated the contents of two buckets and was required to indicate their preference by removing the bucket lid with their trunk (Fig. 1). They received the food reward inside of the bucket they indicated. Each trial began with a curtain drawn to conceal the experimental setup from the elephant and mahout. A food item was placed into each bucket and the lids were placed securely, following which the curtains were opened. The elephant was then allowed to move forward and investigate the contents of both buckets through small holes drilled into the locked bucket lids. In the first training stage, if the elephant remained inactive for 10 s and had not approached either bucket, then the experimenter (standing behind the facing camera) came forward and tapped both buckets with their hand at the same time. In the same training stage, if the elephant approached one bucket and made no attempts to touch the other, the experimenter would tap on the untouched bucket after 10 s. After this had been completed, the mahout was instructed to give the command 'back', and the curtains were redrawn. Finally, the bucket lids were unfastened before reopening the curtains, and the elephants were able to make an 'active' choice by removing the lid of one bucket using their trunk and accessing the food reward inside. This design means that

Table 2 Summary of each training stage with the mean (min–max) and number of sets (*n* set) each elephant completed

Training stage	Food choice	Criterion for early completion	Mean sets (min–max)
1	Papaya or papaya	Complete 4 sets with 30 s between bucket touches in the last 2 consecutive sets	5 (4–8)
2	Papaya or watermelon	Complete 4 sets with 5/6 of the same food in the last 2 consecutive sets	5 (3–7)
3	Watermelon or straw	Complete 2 of the sets with 5/6 of the same food to indicate preference	4.4 (3–5)
4	Empty bucket or banana	Complete 2 of the sets where 7/8 of the baited (banana) bucket to indicate preference	2 (2–2)

Elephants repeated sets until they completed the early criteria, or reached the fixed maximum of sets for each training stage (stage 1 = maximum of 8, stage 2 = maximum of 7, stage 3 = maximum of 5, stage 4 stopped after maximum of 2 sets because of COVID outbreak), or were unable to continue due to welfare or safety reasons (e.g., mahout absence)

Fig. 1 Photo showing experimental setup from the front. The two buckets lodged into the table are highlighted in white

the elephants had to hold the content of the buckets in their working memory. The training process was filmed from the front and the side with two JVC GZ-R495 Everio camcorders.

Mahouts were present during the training to ensure the safety of the elephants and experimenters. However, the task was designed with the expectation of minimal mahout intervention, where mahouts were instructed only to bring the elephant to the training, encourage the elephant into the training corral, and give the command 'back' after elephants investigated both buckets when the lids were closed or accessed the food reward inside when the lids were open. Otherwise, they were asked not to intervene. They were verbally informed of how the experiment would be conducted in an initial briefing session where a camp manager translated the instructions, but they were not aware of the specific aims of the experiment. They and the elephants had also not been involved in behavioural studies before. The second mahout took part in the training when the first mahout was on leave, which occurred for six of the elephants in the study.

Data collection

We constructed an ethogram of elephant behaviours (Table 3) and used Behavioural Observation Research Interactive Software (BORIS; Friard and Gamba 2016) to manually annotate the video recordings. An observation period started when the curtains were completely opened and ended when the elephant investigated the first bucket. As all roles of the mahout built into the experimental design were outside of our observation period, all mahout behaviours during our observation period were regarded as interventions. To investigate the elephant behaviours which elicited mahout interventions and the behavioural responses of elephants to interventions, we noted all instances of mahout interventions (Table 4). We defined mahout interventions as single mahout behaviour or consecutive mahout behaviours in quick succession directed at the elephant within the observation timeframe, i.e., a single intervention can include one single mahout behaviour or multiple consecutive mahout behaviours. We also noted the elephant behaviour based on the ethogram

Table 3 Elephant ethogram

Behaviour Group	Behaviour	Description
Locomotion	Backward	Walk backward
	Forward	Walk forward
	Turn	Turn to face another direction
	Stand	Stand in place, facing bucket
Bucket	Pull	Wrap trunk around the bucket and pull upwards, possibly dislodging the bucket from the experimental setup
Sounds	Vocalisation	Produce a loud vocalisation
	Exhale	Produce an audible exhale
Others	Face away	Stand in place, facing away from bucket
	Scratch	Rub trunk or body on object in the experimental area (e.g., bucket, table, and tree)

Definitions of all elephant behaviours observed directly prior to and following mahout intervention

Table 4 Mahout ethogram

Modality	Behaviour	Description
Non-verbal	Pull ear	Hold or pull elephant ear with one hand
	Pull tail	Pull elephant tail with one hand
	Hand on trunk	Place hand on elephant trunk, with or without force
	Hand on body	Place hand on elephant torso or legs, with or without force
	Touch trunk	Tap elephant trunk with hand
	Touch body	Tap elephant torso or legs with hand, with or without force
	Object to head	Tap elephant head with unsharpened stick or mat mahouts sit on, with or without force
	Object to body	Tap elephant torso or legs with object, with or without force
	No contact	Approach elephant with confidence stance while facing elephant, causing elephant to move in the direction the mahout walking towards
Verbal	Move	Commands for come and move left or right in Nepalese, 'agath' and 'ae mar'
	Back	Commands for move back in Nepalese, 'peechu' and 'peeche hat'
	Eat	Commands for eat in Nepalese, 'kha' and 'dhar'
	Stop	Commands for stay or leave it in Nepalese, 'ra', 'chow' and 'chee'
	Command	Other vocalisations directed at elephant

Definition of all mahout behaviours directed at the elephant or mahout interventions. Commands in Nepalese reproduced from verbal commands used by mahouts

in Table 3 observed directly prior to and following each mahout intervention.

Statistical analysis

All statistical analyses were conducted in R version 4.0.3 (R Core Team 2020) using RStudio version 1.4.1103 (RStudio Team 2020).

Elephant behaviours prior to/following mahout interventions

To investigate which elephant behaviours occurred immediately prior to interventions from mahouts ('prior behaviour'), we used a generalised linear mixed effect model (GLMM). We first tallied instances of each prior behaviour for each elephant in each observation period ('prior tally'). The tally for a behaviour in the ethogram that did not occur in a given observation period was zero. We excluded behaviours that were observed fewer than 5 times in the data (exhale: $n=3$, trumpeting: $n=2$, pull: $n=1$). As we found the data to be zero-inflated (check_zeroinflation function, performance package; Lüdecke et al. 2020) and overdispersed (overdisp_func; Bolker 2017) when fit with a Poisson GLMM, we fitted a constant zero-inflation GLMM with a negative binomial distribution (glmmTMB package; Brooks et al. 2017) to model the prior tally as a response variable, with prior behaviour as the fixed effect (factor with 9 levels). We included the observation period nested within elephant identity as a random effect. Then, to compare the likelihood of each behaviour occurring prior to mahout intervention, we conducted Tukey's post hoc comparisons (emmeans package; Lenth 2022).

To investigate how elephants responded to mahout interventions, we repeated the above analysis using the tally of elephant behaviours recorded immediately after mahout interventions as the response variable. We excluded behaviours that were observed fewer than 5 times in the data (exhale: $n = 4$, scratch: $n = 4$, trumpeting: $n = 3$, and pull: $n = 1$). To explore the response of elephants to verbal and non-verbal mahout behaviours, respectively, we added an interaction term between behaviour and modality (factor with 2 levels: verbal and non-verbal). As an intervention may consist of multiple behaviours, which can include both verbal and non-verbal behaviours, we classify the intervention as verbal when the mahout behaviour immediately prior to the elephant behaviour was verbal and non-verbal when it was not.

Mahout interventions across trainings and sets

To investigate whether the number of mahout interventions changed as the training progressed, we modelled the frequency of mahout interventions across training and set number. We excluded training stage 4 from this analysis as only seven of the eight elephants participated in this training stage and each elephant only completed two training sets, because we had to halt the training at the start of the COVID-19 pandemic. We first tallied the number of mahout interventions in each training set within each training stage, where training sets with no interventions had a tally of zero. As the data were zero-inflated (check_zeroinflation function, performance package; Lüdecke et al. 2020), we fitted a zero-inflated negative binomial GLMM with the mahout intervention tally as a response variable (glmmTMB package; Brooks et al. 2017) after testing for overdispersion (overdisp_func; Bolker 2017). The model had training stage (factor with 3 levels) and set number (numeric variable from 1 to 8) as an interaction term for the fixed effects, elephant identity ($n = 8$) as a random effect, and the logarithmic duration of each set as an offset term to standardise the number of interventions per unit of time in each set. To compare the occurrence of mahout interventions across trainings and sets, we conducted Tukey's post hoc comparisons (emmeans package; Lenth 2022).

To determine whether the nature of the interventions changed as the training progressed, we repeated this analysis with the tally of individual mahout behaviours (as opposed to interventions) as the response variable. We also explored how the use of verbal and non-verbal mahout behaviours changed as training progressed. We did so by fitting the same zero-inflated negative binomial GLMM with number of mahout behaviours as the response variable, but we replaced the fixed effect with a three-way interaction term between training stage, set, and modality.

We compared each model with its equivalent intercept-only model using a likelihood ratio test (LRT; Anova function, car package; Fox and Weisberg 2019) to determine the statistical significance of each fixed effect term.

Results

A total of 131 sets from the 4 training stages, which amounted to 19.45 h of footage, were analysed. The average set length was 10.33 min (range 5.63–19.83 min). We found 997 instances of mahout interventions, comprising 2439 instances of individual mahout behaviours in total across all training stages, with 522 verbal behaviours and 1917 non-verbal behaviours (Table 5).

Elephant behaviours prior to mahout intervention

Model results suggested that the number of interventions differed across elephant behaviours ($\chi^2 = 478.43$, $p < 0.001$). Tukey's post hoc comparisons revealed that 'Stand' was the most frequent behaviour prior to interventions, which was 4.82 times as likely to occur as the next most frequent behaviour, 'Forward' ($p < 0.001$) (Fig. 2, Supplementary Information Table S1).

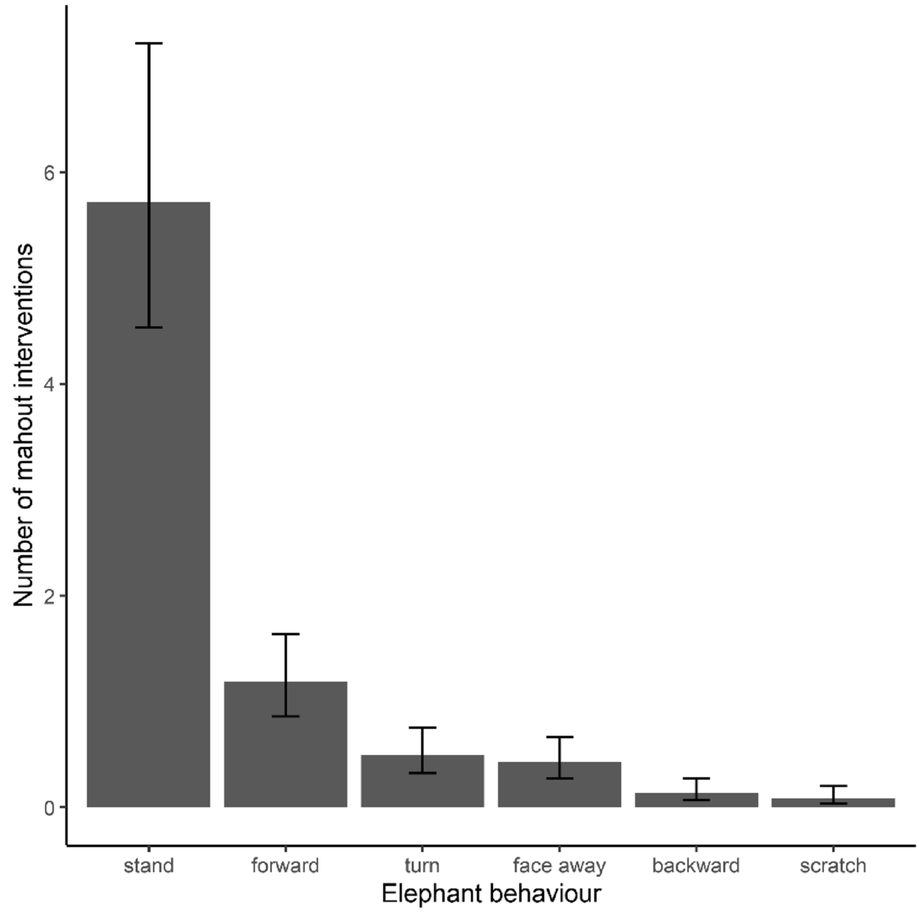
Elephant behaviours following mahout intervention

The model results suggested a significant interaction between elephant behaviour and intervention modality (LRT: $\chi^2 = 22.13$, $p < 0.001$), meaning that mahout use of

Table 5 Occurrence of each mahout behaviour and its percentage out of all occurrences across 4 training stages

Modality	Behaviour	Occurrence	Percentage (%)
Non-verbal	Hand on body	586	24
	Touch body	442	18
	Pull ear	403	17
	Object to body	262	11
	Hand on trunk	146	6
	Touch trunk	36	1
	Pull tail	20	1
	Object to head	18	1
	No contact	4	0.2
Verbal	Eat	345	14
	Command	91	4
	Back	60	2
	Stop	18	0.7
	Move	8	0.3

Fig. 2 Model predicted number (with 95% confidence intervals) of mahout interventions occurring immediately after each elephant behaviour (for a single mahout-elephant dyad)



verbal and non-verbal interventions differed across elephant behaviours. Tukey's post hoc comparisons revealed that 'Forward' was the most frequent behaviour immediately after both verbal and non-verbal interventions, and 'Stand' was the next most frequent behaviour. After verbal interventions, 'Forward' was 2.1 times as likely as 'Stand', while after non-verbal interventions, it was 2.074 times as likely as 'Stand' (Fig. 3, Supplementary Information Table S2). Post hoc analysis also suggested that non-verbal interventions were used more frequently before 'Forward' (4.6 times more likely, $p < 0.001$) and 'Stand' (4.7 times more likely, $p < 0.001$).

Mahout interventions and behaviours across training and sets

Our first model showed no effect of training stage or set ($p > 0.05$) on the number of mahout interventions. Our second model, with number of individual mahout behaviours instead of number of mahout interventions (which we defined as a single mahout behaviour, or multiple mahout behaviours occurring in quick succession) as a response variable. In this second model, we also included modality as an interaction term with training stage and set, and found a significant interaction between modality and training stage (LRT: $\chi^2 = 10.617$, $p = 0.005$). Post hoc analyses showed that non-verbal behaviours occurred more frequently than verbal behaviours in training stages 1 (β = 1.856, $p < 0.001$, p values adjusted for multiple comparisons; Fig. 4) and 2 ($\beta = 0.998$, $p = 0.002$), but not in training stage 3 ($p = 0.303$). Post hoc contrasts also showed that the mahouts used fewer verbal behaviours in training stage 1 compared to training stage 3 ($\beta = -1.052$, $p < 0.001$), but there was no difference in the occurrence of verbal behaviours between training stages 1 and 2 ($p > 0.05$), and between training stages 2 and 3 ($p = 0.611$). Non-verbal behaviours did not differ across training stages (1–2: $p = 0.941$, 1–3: $p = 0.951$, 2–3: $p = 1$). Controlling for training stage and modality, we found no significant effect of set on the number of mahout behaviours ($p > 0.1$).

Discussion

In this study, we investigated mahout intervention during the training phase of a behavioural experiment with Asian elephants. We examined the impact of interventions on elephant behaviours occurring immediately prior to and following

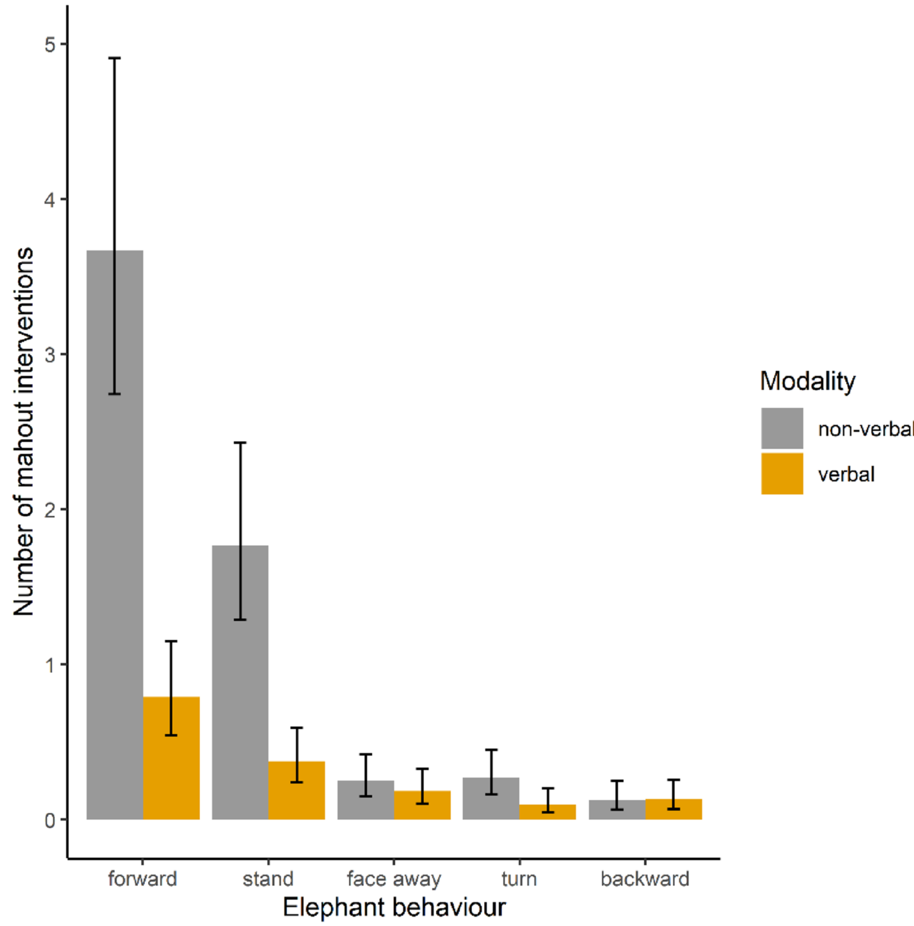


Fig. 3 Model predicted number (with 95% confidence intervals) of mahout interventions occurring immediately before each elephant behaviour (for a single mahout-elephant dyad)

the interventions. We found that mahouts frequently intervened verbally and non-verbally despite being instructed not to do so. Mahouts most commonly provided cues to the elephants by giving verbal commands (e.g., 'Back', 'Stop', and 'Move') and non-verbal commands through physical contact (e.g., touching elephant body with an object or their hands), and used more non-verbal cues than verbal ones. The mahouts most frequently intervened when an elephant was standing still or, much less frequently, when the elephant was moving forward. Similarly, the elephants responded to mahout cues most frequently by moving forward or, much less frequently, standing still.

As the training required elephants to stand still until the curtains were completely opened before approaching the buckets, it is unsurprising that elephants standing still elicited the majority of interventions from mahouts, and elephants in turn responded to most of these interventions by moving forward. However, mahouts also intervened frequently when elephants moved forward after the curtains had been opened, which was the desired behaviour for the training. This seems to be because the elephants had already started moving forward before the curtains were fully opened and the mahouts intervened to stop such behaviour. The most common mahout behaviours were "touch body" and the command "eat". The former is likely to be a command for the elephants to stay in place and not go forward before the curtains open, and the latter is likely to encourage the elephants to go to the buckets and obtain the food reward. It is also interesting to note that non-verbal behaviours were used more frequently than verbal behaviours before elephants moved forward or stood still, and across training stages and sets in general. This may be because although animals are more likely to follow non-verbal cues, as previous research has shown in working dogs (D'Aniello et al. 2016; Scandurra et al. 2017). The mahouts were specifically told not to intervene during the training and therefore might have preferred non-verbal interventions, which they might have viewed as less disruptive than verbal interventions.

Elephant behaviour can be affected by many factors and potentially influenced by cues that we did not measure. Therefore, we should be cautious in interpreting the elephants' responses to mahout interventions. Plotnik et al. (2013) found that Asian elephants did not follow pointing cues from mahouts to locate hidden food, although mahouts often reported successfully using pointing to direct elephants to pick up objects for tourists. This suggests that rather than

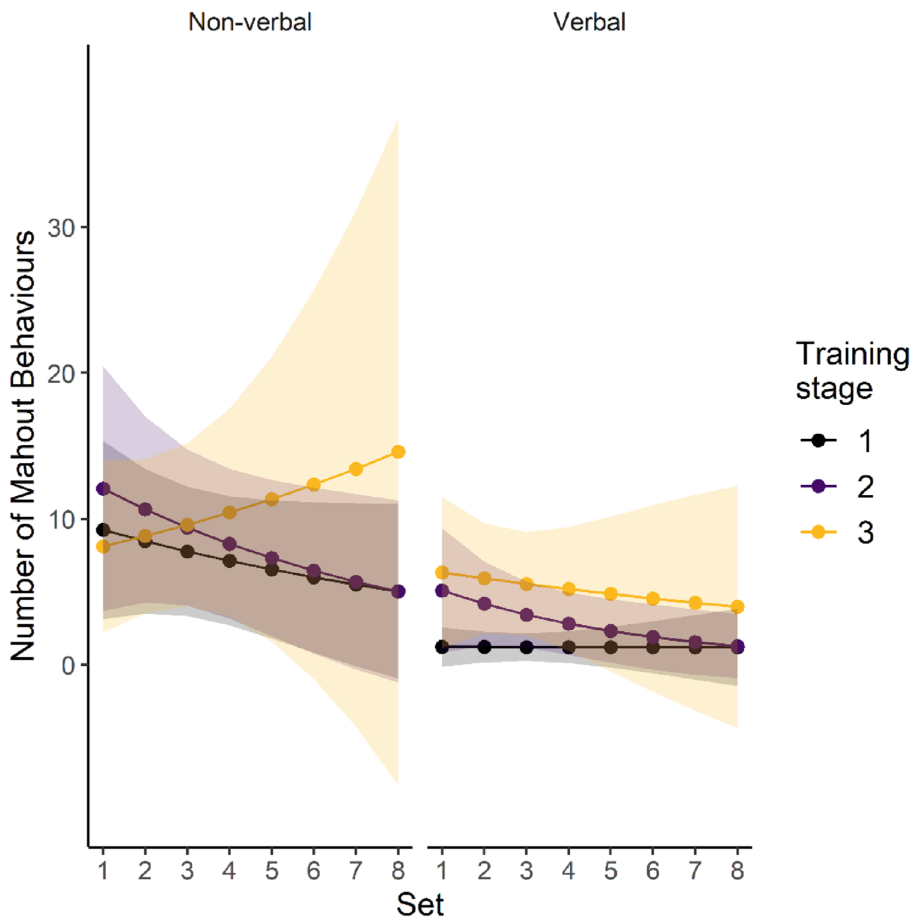


Fig. 4 Model predicted change of verbal and non-verbal mahout behaviours across sets of six trials in the three training stages (with 95% confidence intervals)

responding to the pointing behaviour itself, the elephants may have been responding to other cues, such as the orientation of the mahout's body, verbal cues, and other contextual cues. Apart from responding to commands differently than expected, the elephants may also have been responding to cues that we did not record in our training, for example the emotional states of their mahouts. In the same study, Plotnik et al. (2013) found that the elephants' performance in the experiment was negatively affected when mahouts showed visible frustration to subject noncompliance, which suggests that elephants may be sensitive to the emotional state of their mahouts. This finding is consistent with previous research in zoo stockmanship, which demonstrates that animals respond more positively and readily to handlers with a positive attitude (Ward and Melfi 2015).

We expected mahouts to intervene less as the elephants became more experienced with the experimental design. It was surprising that the frequency of mahout behaviours (individually categorised actions) and interventions (which could be constituted of single behaviours or several consecutive behaviours in quick succession) did not decrease as the training progressed. This may be because mahouts were accustomed to their role as a trainer, which requires them to initially introduce elephants to daily tasks and later continue to direct them in their activities. Therefore, even when the elephants became more experienced with the training task, mahouts still intervened with the same frequency. However, when we analysed the number of verbal and non-verbal behaviours in each training stage, we found that the nature of mahout behaviours was different across the training stages in terms of intervention modality, and the frequency of verbal mahout behaviours was different across training stages. Non-verbal behaviours occurred more frequently than verbal behaviours in training stages 1 and 2, but not in training stage 3; and the average number of verbal mahout behaviours in training stage 3 was higher than in training stage 1. This may have been due to the increased difficulty and reduced average pay off in training stage 3, meaning that although the mahouts were told not to intervene and may have been less likely to use verbal cues in previous stages, they resorted to them in this stage. In training stage 1, both buckets contained the same food reward as the purpose of this stage was simply to familiarise the elephants with the experimental setup. While in training stages 2 and 3, the elephants had to make a choice between two different food rewards, both fruit in stage 2 and fruit or grass in stage 3. We suggest that the elephants may have behaved with less certainty in these stages, prompting mahouts to guide the behaviour of their elephants. Furthermore, with previous research suggesting that animals perform better when given larger rewards (Ferrucci et al. 2019), consideration should be given to reward type and quantity during experimental design to maximise the motivating effects of the reward during behavioural experiments. For example, in the case of our study, fewer interventions may have been made by the mahouts at the start of the experiments if the elephants were highly motivated by the food rewards. We note again that despite the focus on reducing potential for handler impact on experiments, we do not intend to frame all mahout interventions here as negative. In fact, their presence is necessary for the elephants to be able to participate in any experiments and training elephants to interact with the experimental apparatus with no interventions from the researcher or mahout is impossible. Further research may be conducted to examine whether and how handler interventions may affect the performance of animals in the actual experimental stage after moving on from the training stage.

This study highlights some of the difficulties faced by experimenters seeking to reduce the impact of human presence in animal cognitive and behavioural studies, particularly when animal handlers are necessary for the safety of both the animals and the humans involved. We have shown here that handler behaviour is associated with animal behaviour in a choice test and therefore may affect the results of the behavioural experiment. We suggest that experimenters should record instances of handler intervention and take account of them when analysing results of animal behavioural experiments where handlers are required, as it may be difficult to for handlers to completely avoid providing cues to their animals when this is an integral part of their relationship. Extra measures such as better communication with the handlers may be needed to minimise interventions and other possible handler biases. Furthermore, studies could be designed to minimise intervention points. In our case, interventions took place around times elephants were meant to approach or move back from the experimental equipment. One way to reduce this could be to take and withdraw the apparatus from them, for example with the help of a sliding table as in a choice test by Plotnik et al. (2019), rather than having the subjects' movement built into the study design. With reduced possible intervention points in experimental design, further research may also be conducted to investigate the effect of human presence without any interventions on behavioural studies. It is also important to note that animals may be responding to less salient cues from the environment and their handlers, and their performance during experiments may be influenced by their affective states or other individual motivations such as the rewards given. Therefore, overgeneralization and simplification should be avoided when interpreting behavioural responses from animals. Although it may be challenging in outdoor settings like elephant camps, which are not designed for experiments, careful consideration should be taken during experimental design to eliminate the effect of other environmental factors when possible, for example, by reducing the potential people and objects in the surrounding the animals may respond to.

More broadly, our results are relevant to any behavioural studies conducted on animals in captivity with human presence. As captive settings such as zoos allow researchers to control for many variables, many behavioural experiments have been conducted with animals in captivity, largely on mammals, particularly on primates (Anderson et al. 2008; McEwen et al. 2022), but also on other taxa such as reptiles (Egelkamp and Ross 2019). Nevertheless, studying captive animals that are trained to respond to human cues requires careful measures to avoid experimenter bias, the effects of which are expected to vary with species and keeping style. As cognitive abilities in perceiving cues from humans vary greatly across different species and captive animals are trained for different purposes (Schusterman et al. 2013; Wynne 2016; Melfi et al. 2020), differences in the type and effect of handler intervention can be expected.

Our study may also be of interest in the light of human–animal interactions in behavioural experiments. We have highlighted the challenges associated with the presence of human caretakers in animal behaviour experiments. We note that substantial efforts are made to reduce the potential impact of these caretakers to ensure that animals are influenced by handlers as little as possible during experiments, for example using the sliding table design in Plotnik's study to reduce the potential for handlers to encourage the elephants to move towards or away from equipment (Plotnik et al. 2019). However, an alternative way of approaching the challenge of having handlers present is to use it as an opportunity to investigate other research questions that do not require the mahout and animal to be separated, such as how handlers and animals interact when presented with a novel task, or how animals learn from their trainers. For example, Lit and colleagues (2011) were able to investigate whether information they gave to handlers of dogs trained for drug and/or explosive detection affected how and whether the human and dog teams alerted for drugs or explosives when cues for neither were present. They found that handler beliefs that a scent associated with drugs or explosives was present potentiated handler identification of detection dog alerts (Lit et al. 2011). We note that in our system, with captive social animals that have a high degree of interaction with handlers, the presence of handlers during tasks such as those used in behavioural experiments is a much more frequent scenario in their lives than being separated from both their handler and other animals. Our study shows that there is further potential for investigating human–animal interactions through behavioural studies (Davis and Balfour 1992), including analysing what, when, and how the interactions take place and how the animals respond. Further research in quantifying human–animal interactions may provide more insight into these questions, in addition to focusing on reducing human influences in behavioural studies with the aim of reducing experimenter bias (Tuyttens et al. 2014).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10071-022-01668-8.

Acknowledgements We thank the University of Hong Kong for funding the student and the Applied Behavioural Ecology and Conservation Lab, including a PhD studentship to HBT. HSM thanks the Branco Weiss Society- in Science Fellowship administered by the ETH Zürich, which allowed her to pay salaries to KW and DM. HSM thanks the Association for Animal Behaviour Christopher Barnard Award for funding computing equipment. The authors would like to thank Sagarika Phalke (PhD student Mumby Laboratory) for aiding HBT with the data collection and methodological setup for this project and undertaking the behavioural experiments. We would like to thank Tiger Tops for allowing us to conduct research at their site, particularly all the staff who supported this work.

Author contributions HSM, KW, DM, and HBT designed the study. HBT conducted the behavioural experiments and video recorded all training sessions. PCC collected the data from the videos and conducted the analyses under the supervision of HSM, KW, and DM. PCC wrote the first draft of the manuscript. HSM acquired funding for the research. All authors contributed revisions and all authors approved the final manuscript.

Funding Funding for this research was provided by grants from the University of Hong Kong and the Branco Weiss—Society in Science Fellowship to HSM.

Availability of data and materials All data will be made available upon publication via the Dryad data depository.

Code availability Associated code will be included as a supplementary file in the publication.

Declarations

Conflicts of interest/competing interests The authors declare no competing interests.

Ethics approval We received ethical permission to conduct this work from the University of Hong Kong Committee on the Use of Live Animals in Teaching and Research (CULATR) under Project No. 5226-19.

Consent to participate Mahouts also consented to participate in our research and participated in recordings. This was approved through the University of Hong Kong Human Research Ethics Committee (application EA2002001).

Consent for publication In line with our research agreement with Tiger Tops Ltd and the consent of mahouts, we have consent for publication.

References

Abramson JZ, Hernández-Lloreda V, Call J, Colmenares F (2011) Relative quantity judgments in South American sea lions (*Otaria flavescens*). Anim Cogn 14:695–706. https://doi.org/10.1007/s10071-011-0404-7

Abramson JZ, Hernández-Lloreda V, Call J, Colmenares F (2013) Relative quantity judgments in the beluga whale (*Delphinapterus leucas*) and the bottlenose dolphin (*Tursiops truncatus*). Behav Process 96:11–19. https://doi.org/10.1016/j.beproc.2013.02.006

Anderson US, Kelling AS, Maple TL (2008) Twenty-five years of Zoo Biology: a publication analysis. Zoo Biol 27:444–457. https://doi.org/10.1002/zoo.20177

Barnard C (2007) Ethical regulation and animal science: why animal behaviour is special. Anim Behav 74:5–13. https://doi.org/10.1016/j.anbehav.2007.04.002

Bolker BM (2017) Glmm Faq. https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html. Accessed 5 Apr 2021

Brooks ME, Kristensen K, van Benthem KJ et al (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 9:378–400. https://doi.org/10.32614/rj-2017-066

Burghardt GM, Bartmess-Levasseur JN, Browning SA et al (2012) Perspectives—minimizing observer bias in behavioral studies: a review and recommendations. Ethology 118:511–517. https://doi.org/10.1111/j.1439-0310.2012.02040.x

Crawley JAH, Lahdenperä M, Seltmann MW et al (2019) Investigating changes within the handling system of the largest semi-captive population of Asian elephants. PLoS ONE 14:e0209701. https://doi.org/10.1371/journal.pone.0209701

Cuthill I (1991) Field experiments in animal behaviour: methods and ethics. Anim Behav 42:1007–1014. https://doi.org/10.1016/S0003-3472(05)80153-8

D'Aniello B, Scandurra A, Alterisio A et al (2016) The importance of gestural communication: a study of human–dog communication using incongruent information. Anim Cogn 19:1231–1235. https://doi.org/10.1007/s10071-016-1010-5

Dale R, Plotnik JM (2017) Elephants know when their bodies are obstacles to success in a novel transfer task. Sci Rep 7:46309. https://doi.org/10.1038/srep46309

Davis H, Balfour D (1992) The inevitable bond: Examining scientist–animal interactions. Cambridge University Press

Egelkamp CL, Ross SR (2019) A review of zoo-based cognitive research using touchscreen interfaces. Zoo Biol 38:220–235. https://doi.org/10.1002/zoo.21458

Ferrucci L, Nougaret S, Brunamonti E, Genovesio A (2019) Effects of reward size and context on learning in macaque monkeys. Behav Brain Res 372:111983. https://doi.org/10.1016/j.bbr.2019.111983

Fox J, Weisberg S (2019) An R companion to applied regression, third. Sage, Thousand Oaks

Friard O, Gamba M (2016) BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol Evol 7:1325–1330. https://doi.org/10.1111/2041-210X.12584

Hart LA (1994) The Asian elephants-driver partnership: the drivers' perspective. Appl Anim Behav Sci 40:297–312. https://doi.org/10.1016/0168-1591(94)90070-1

Hart L, Sundar (2000) Family traditions for mahouts of Asian elephants. Anthrozoos 13:34–42. https://doi.org/10.2752/089279300787000055

Highfill LE, Spencer JM, Fad O, Arnold AM (2016) Performance on a means-end task by Asian elephants (*Elephas maximus*) in a positive reinforcement-based protected-contact setting. Int J Comp Psychol. https://doi.org/10.46867/ijcp.2016.29.00.12

Holman L, Head ML, Lanfear R, Jennions MD (2015) Evidence of experimental bias in the life sciences: Why we need blind data recording. PLoS Biol 13:e1002190. https://doi.org/10.1371/journal.pbio.1002190

Horowitz A, Hecht J, Dedrick A (2013) Smelling more or less: investigating the olfactory experience of the domestic dog. Learn Motiv 44:207–217. https://doi.org/10.1016/j.lmot.2013.02.002

Lainé N (2016) Conduct and collaboration in human–elephant working communities of Northeast India. In: Conflict, negotiation, and coexistence. Oxford University Press, pp 180–204

Lenth RV (2022) emmeans: estimated marginal means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans

Lit L, Schweitzer JB, Oberbauer AM (2011) Handler beliefs affect scent detection dog outcomes. Anim Cogn 14:387–394. https://doi.org/10.1007/s10071-010-0373-2

Locke P (2011) The ethnography of captive elephant management in Nepal: a synopsis. Gajah 34:32–40

Lüdecke D, Makowski D, Waggoner P, Patil I (2020) performance: assessment of regression models performance. CRAN. https://doi.org/10.5281/zenodo.3952174

Marsh DM, Hanlon TJ (2004) Observer gender and observation bias in animal behaviour research: Experimental tests with red-backed salamanders. Anim Behav 68:1425–1433. https://doi.org/10.1016/j.anbehav.2004.02.017

Marsh DM, Hanlon TJ (2007) Seeing what we want to see: confirmation bias in animal behavior research. Ethology 113:1089–1098. https://doi.org/10.1111/j.1439-0310.2007.01406.x

McEwen ES, Warren E, Tenpas S et al (2022) Primate cognition in zoos: reviewing the impact of zoo-based research over 15 years. Am J Primatol. https://doi.org/10.1002/ajp.23369

Melfi V, Dorey NR, Ward SJ (2020) Zoo animal learning and training. Wiley

Miklösi Á, Polgárdi R, Topál J, Csányi V (1998) Use of experimenter-given cues in dogs. Anim Cogn 1:113–121. https://doi.org/10.1007/s100710050016

Mueller-Paul J, Wilkinson A, Aust U et al (2014) Touchscreen performance and knowledge transfer in the red-footed tortoise (*Chelonoidis carbonaria*). Behav Process 106:187–192. https://doi.org/10.1016/j.beproc.2014.06.003

Mumby HS (2019) Mahout perspectives on Asian elephants and their living conditions. Animals 9:879. https://doi.org/10.3390/ani9110879

Mumby HS, Chapman SN, Crawley JAH et al (2015) Distinguishing between determinate and indeterminate growth in a long-lived mammal. BMC Evol Biol 15:214. https://doi.org/10.1186/s12862-015-0487-x

Perdue BM, Snyder RJ, Maple TL (2013) Cognitive research in Asian small-clawed otters. Int J Comp Psychol 26:105–113. https://doi.org/10.46867/ijcp.2013.26.01.01

Plotnik JM, Lair R, Suphachoksahakun W, De Waal FBM (2011) Elephants know when they need a helping trunk in a cooperative task. Proc Natl Acad Sci U S A 108:5116–5121. https://doi.org/10.1073/pnas.1101765108

Plotnik JM, Pokorny JJ, Keratimanochaya T et al (2013) Visual cues given by humans are not sufficient for asian elephants (*Elephas maximus*) to find hidden food. PLoS ONE 8:e61174. https://doi.org/10.1371/journal.pone.0061174

Plotnik JM, Shaw RC, Brubaker DL et al (2014) Thinking with their trunks: elephants use smell but not sound to locate food and exclude nonrewarding alternatives. Anim Behav 88:91–98. https://doi.org/10.1016/j.anbehav.2013.11.011

Plotnik JM, Brubaker DL, Dale R et al (2019) Elephants have a nose for quantity. Proc Natl Acad Sci 116:12566–12571. https://doi.org/10.1073/pnas.1818284116

R Core Team (2020) R: a language and environment for statistical computing. R foundation for statistical computing, Vienna

RStudio Team (2020) RStudio: Integrated Development for R. http://www.rstudio.com/. Accessed Oct 2020

Samhita L, Gross HJ (2013) The "Clever Hans Phenomenon" revisited. Commun Integr Biol 6:e27122. https://doi.org/10.4161/cib.27122

Scandurra A, Alterisio A, Marinelli L et al (2017) Effectiveness of verbal and gestural signals and familiarity with signal-senders on the performance of working dogs. Appl Anim Behav Sci 191:78–83. https://doi.org/10.1016/j.applanim.2017.02.003

Schusterman R, Thomas J, Wood F (2013) Dolphin cognition and behavior. Psychology Press

Sukumar R (2006) A brief review of the status, distribution and biology of wild Asian elephants *Elephas maximus*. Int Zoo Yearb 40:1–8. https://doi.org/10.1111/j.1748-1090.2006.00001.x

Tuyttens FAM, de Graaf S, Heerkens JLT et al (2014) Observer bias in animal behaviour research: can we believe what we score, if we score what we believe? Anim Behav 90:273–280. https://doi.org/10.1016/j.anbehav.2014.02.007

van Wilgenburg E, Elgar MA (2013) Confirmation bias in studies of nestmate recognition: a cautionary note for research into the behaviour of animals. PLoS ONE. https://doi.org/10.1371/journal.pone.0053548

Ward SJ, Melfi V (2015) Keeper-animal interactions: differences between the behaviour of zoo animals affect stockmanship. PLoS ONE. https://doi.org/10.1371/journal.pone.0140237

Watters JV, Krebs BL (2019) Assessing and enhancing the welfare of animals with equivocal and reliable cues. Animals 9:680. https://doi.org/10.3390/ani9090680

Wynne CDL (2016) What is special about dog cognition? Curr Dir Psychol Sci 25:345–350. https://doi.org/10.1177/0963721416657540