scientific reports

OPEN

Human recognition of emotional valence and arousal of zoo animals

Laura Hiisivuori[1]✉, Helena Telkänranta[2,3], Emma Vitikainen[4] & Sonja E. Koski[4,5]

Correctly identifying emotions of other species is central to the welfare of animals in our care; yet, the factors underlying variation in our ability to recognise animal emotions remain unclear. Here, we investigated the human ability to recognise emotional valence and arousal in zoo-living Barbary macaques, tigers, and markhors, in short video clips from which contextual information and other clues were removed. Visitors at Korkeasaari Zoo were recruited to rate valence and arousal, and their ratings were analysed for correct identification of positive vs. negative valence and low vs. high arousal. Overall, arousal was more accurately recognised than valence, low arousal more accurately than high arousal, and negative valence more accurately than positive valence. Moreover, recognition accuracy varied among species being rated. Valence was recognised most accurately in macaques, arousal most accurately in markhors, and tigers' emotional states were recognised the least accurately, both for valence and arousal. The results suggest that while we can recognise non-domesticated species' emotional states, accuracy varies depending on the species and the emotional state in question, which highlights the importance of training in assessing animal welfare. Overall, considering animal welfare, it is crucial that we improve education for the identification of animal emotions, both positive and negative ones.

Keywords Human-animal-interaction, Animal, Emotion recognition, Valence, Arousal, Animal welfare

Rapid identification and comprehension of the emotional states of conspecifics are crucial for humans as well as other social species[1]. Additionally, cross-species recognition can be beneficial, for example if it provides crucial information regarding danger or resources[2,3]. In modern humans, the ability to recognise non-human species' emotions is particularly central to the welfare of the animals in our care[4]. However, what affects a person's ability to recognise animal emotional states is still poorly understood. Several factors may influence cross-species emotion recognition. They may be related to external effects such as culture and experience[5], personal factors including gender, age, personality, or empathy[3,6–8] and to the assessed animal, its behaviour and surroundings, which directly affect emotional states and may also make the same experienced emotion either easier or more difficult to recognise[9,10]. Understanding such effects and what makes us either understand or misunderstand animal emotions is useful for education and can be used to inform animal welfare practices where animals are under our care. Zoos for example are increasingly focused on assessing and maintaining animal welfare[11,12] and studies have been conducted on visitors' perceptions of the welfare of animals living in zoos[13].

The study of animal emotions and their expression in faces and body language traces back to the seminal book by Charles Darwin[14]and since then, large body of research has investigated emotions in humans and non-human animals. Whereas carving a common description for emotional states across taxa may be meaningless or difficult at best (e.g[15].), a fruitful approach in the study of emotional states in non-human animals has been to use objectively identifiable facets of emotional processing. One such widely used *prescriptive* definition of emotional states in humans as well as in nonhuman animals is the approach where emotional states are described in two dimensions: valence, which ranges from positive to negative, and arousal, ranging from low to high[16–18]. Discrete emotions, including fear, sadness, and happiness, can then be understood as situated within this two-dimensional space, even though their mere spatial positioning does not comprehensively capture their subjective characteristics[16].

Investigating emotional states with respect to their valence and arousal is a useful approach, as it allows for analytical comparison of factors that affect our ability to recognise aspects of emotional experience in others. For example, when looking at still images of emotional states, human negative emotions are recognised

[1]Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, P.O. Box 65, Helsinki 00014, Finland. [2]Arador Innovations Ltd, Helsinki, Finland. [3]Research Centre for Animal Welfare, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland. [4]The Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland. [5]Department of Biology, University of Turku, Turku, Finland. ✉email: laura.hiisivuori@helsinki.fi

faster and more accurately than positive ones, implicating that there may be a greater cross-cultural similarity in expressions of negative states as compared to positive ones[5,19–23]. However, it is unclear whether the same negativity bias applies to human recognition of emotions in other species. For example, in some studies, humans have recognised dogs' positive emotional states, such as joy or kindness better than negative states such as fear or anger[24–26].

Vocal indicators and human recognition of emotional arousal, in turn, have been relatively well studied[27]. In a broad comparative study, participants recognised high arousal relatively accurately across all included vertebrates, from amphibians to mammals, although accuracy varied across species[28]. Similarly, when comparing the recognition of emotional states in domestic animals and their wild relatives, respondents more accurately recognised arousal compared to valence, and domestic animals were more accurate than wild ones[3]. This may be due to a common, inherited vocal emotional system present in at least mammals, which is more pronounced in arousal compared to valence, and expressions of which seem to vary more between species[28–30]. Indeed, studies of emotion recognition have often presented auditory signals[2,3,29,31].

Experience with a species may improve the recognition of that species' emotions: for example, in pigs, more experienced individuals and ethologists achieved better recognition than students[36]. In some studies, having a dog or cat appeared to enhance the recognition of their emotions[8,20,32,33], while in other studies it did not[24,34] or even impaired the correct assessment of negative emotions[35]. Finally, our ability to recognise animal emotional states may also vary with individual traits such as gender and age, but effects vary between studies and emotional states being assessed. Young women, in particular, have shown a better ability to accurately identify animal emotional states than men or older women[37,38]. As older age groups are less accurate at perceiving other people's emotions[38], they may also be less accurate at recognising animals' emotions[3].

The majority of studies on emotion recognition have been conducted on domesticated animals, particularly dogs[25,30,34,36,40–43] with only a few assessing wild or captive, non-domesticated species. Some studies have assessed the recognition of non-human primate emotional states[5,44–47], as well as wild counterparts of farm animals, such as wild boars or Przewalski horses[3]. In general, most studies on human ability to recognise animal emotions have focused on animal sounds, with species ranging from amphibians[28] to tree shrews[48].

In this study, we assessed zoo visitors' ability to identify animal emotional states in a zoo setting. We subjected the respondents to video footage of three captive-living, non-domesticated animal species: the Barbary macaque (*Macaca sylvanus*), the Siberian tiger (*Panthera tigris altaica*), and the Turkmenian markhor (*Capra falconeri heptneri*). The species were chosen to represent different mammalian groups. The Barbary macaque, a primate, was chosen as an evolutionarily close relative to humans. The tiger was selected for its similarity to a species (the domestic cat) with which many participants would have personal experience. To our knowledge, only one study has assessed recognition of tiger emotional states; keepers of zoo-housed big cats, including a tiger, did not change their evaluations of the animals' emotional states between situations where behaviour specialists noticed significant differences in emotional states[49]. Finally, markhor was selected as a more distantly related species, related species of which the majority of zoo visitors were assumed not to have direct experience. We found two prior studies on human ability to recognise goats' emotions: in one study, participants recognised domestic goat's valence and arousal on their calls just a little bit above chance level[3], and in another, a very recent one, significantly above chance level[51]. Additionally, the emotions of dairy goats have been evaluated with the Qualitative Behaviour Assessment (QBA) with professional observers. QBA method based on expert observers to interpret details of behaviour, posture, and context into descriptions of an animal's emotions, such as 'relaxed' or 'frustrated'[50]. QBA has been developed and tested for dairy goats[52,53] and using it has been evaluated under different housing conditions[54].

We investigated respondents' ability to recognise emotional states using visual signals from mute video clips. Visual signals from motion pictures might offer the possibility of a more accurate assessment of emotional states compared to still images[47]. Additionally, videos have been successfully used in a few previous studies on human intraspecific emotion perception[26,55,56] and many QBA analyses with pet and farmed animals[57–59], and zoo-housed animals like primates[60] and tigers[61].

We formulated the following predictions regarding respondents' ability to recognise animal emotional states: (1) Participants would be more accurate in recognising arousal than valence in animal emotional expressions. (2) Participants would demonstrate greater accuracy in recognising negative emotional states compared to positive, and high arousal emotional states compared to low. (3) Participants with prior experience with pet cats would be more accurate in recognising tiger emotional states, given potential familiarity with feline behaviour. (4) Women would exhibit higher overall accuracy in recognition of emotional states than men.

Results

Altogether, we obtained 5306 valence and arousal rating scores from 1082 respondents. There were 2160 scores from 432 respondents in 2021, collected over 21 days (1–3 sessions per day), and 3250 scores from 650 respondents in 2022 (1–5 sessions per day), collected over 48 days (Table 1). 63% of the respondents self-identified as women, 35% as men, and 2% (N = 19) as non-binary or chose not to specify their gender; this last group was excluded from the analyses. The most respondents were in the age group 31–45 years (46%). Compared to the overall demographic of the Helsinki Zoo visitor profile, the study had a bias toward women (in our data women 65%, in the Korkeasaari visitor profile in the year 2021 52%). Overall, our respondents fall into the category of so-called WEIRD (Western, Industrialised, Rich, Democratic) populations[62].

Valence and arousal

Overall, valence (positive vs. negative) was recognised correctly at 55% and arousal (low vs. high) at 67% accuracy; thus, arousal was recognised more accurately than valence (Wilcoxon signed-rank test: V = 827 K, p < .0001).

Year	2021	%	2022	%
Age groups				
18–30	145	34	152	23
31–45	200	**46**	297	**46**
46–60	57	13	118	18
61 and over	22	5	72	11
No answer	8	2	11	2
Gender				
Woman	271	**63**	412	**63**
Man	152	35	226	35
Other and not-want-to-say	8	2	11	2
No answer	1	0	1	0
Language				
Finnish	432	**100**	601	**92**
English	-	-	49	8
Having a pet				
Yes	226	**52**	536	**82**
No	198	46	102	16
No answer	8	2	12	2
Having a cat				
Yes	83	19	-	-
No	341	**79**	-	-
No answer	8	2	-	-
All respondents				
	432	100	650	100

Table 1. Demographic data of the respondents. Note that missing answers are here counted per question, but for the analyses, including these variables as predictors, all answers that had any background variables as missing were excluded.

Accuracy of Valence recognition

We found that respondents recognised negative valence more accurately (Fig. 1). Additionally, recognition accuracy differed between species. Respondents were most accurate when evaluating valence of Barbary macaques (predicted probability = 0.82, 95% CI [0.78, 0.86]), followed by markhors (predicted probability = 0.68, 95% CI [0.63, 0.73]) and least accurate with tigers (predicted probability = 0.23, 95% CI [0.20, 0.27], Fig. 2). Respondents who completed forms in Finnish performed better than those who completed them in English, but the difference in the number of forms was quite significant (Finnish 738, English 40) and the effect relatively small (63% vs. 54%).

We did not find any influence of gender, age or pet ownership on valence recognition accuracy (Table 2). When assessed only for ownership of a cat vs. no cat, respondents with a pet cat were no more accurate in recognising tigers' valence, compared to persons who did not have a cat ($X^{21} = 0.02$, $p = .876$). Finally, video compilation had a significant effect on the accuracy of valence recognition (Table 2): Valence was recognised correctly most often in compilation 2 (predicted probability = 0.74, CI [0.68, 0.79]), then compilation 3 (predicted probability = 0.62, CI [0.57, 0.67]) and then compilation 1 (predicted probability = 0.39, CI [0.35, 0.44]). Average accuracy of recognition varied between the video clip compilations, which was to be expected, and confirms that the ease at which emotional states can be correctly identified varies across species and emotional states. Although we were unable to balance the clips so that the emotional states of the three species would be equally easy or difficult to evaluate, we don't expect any systematic bias in our interpretation of the results, as the participants were assigned one of the three compilations at random with respect to the potential predictors of recognition accuracy.

Accuracy of arousal recognition

Respondents recognised low arousal more accurately (predicted probability = 0.95, 95% CI [0.94, 0.96]) than high arousal (predicted probability = 0.56, 95% CI [0.51, 0.61], Fig. 1). Similar to valence recognition, species influenced the accuracy of arousal recognition, with arousal in markhors recognised with the highest accuracy (predicted probability = 0.91, 95% CI [0.89, 0.93]), followed by Barbary macaques (predicted probability = 0.81, 95% CI [0.77, 0.84]), and lowest – still higher than chance – in tigers (predicted probability = 0.74, 95% CI [0.69, 0.78], Table 2; Fig. 2).

In contrast to valence recognition, women were better at arousal recognition, but similar to valence, neither age nor pet ownership had an effect (Table 2). The language forms were completed similarly showing no effect in arousal recognition. Respondents with a pet cat were not better in recognising arousal in tigers, as compared to persons who did not have a cat ($X^{21} = 0.000$, $p = .988$). Arousal was recognised correctly most often in compilation 3 (predicted probability = 0.92, CI [0.89–0.94]), then compilation 1 (predicted probability = 0.78, CI

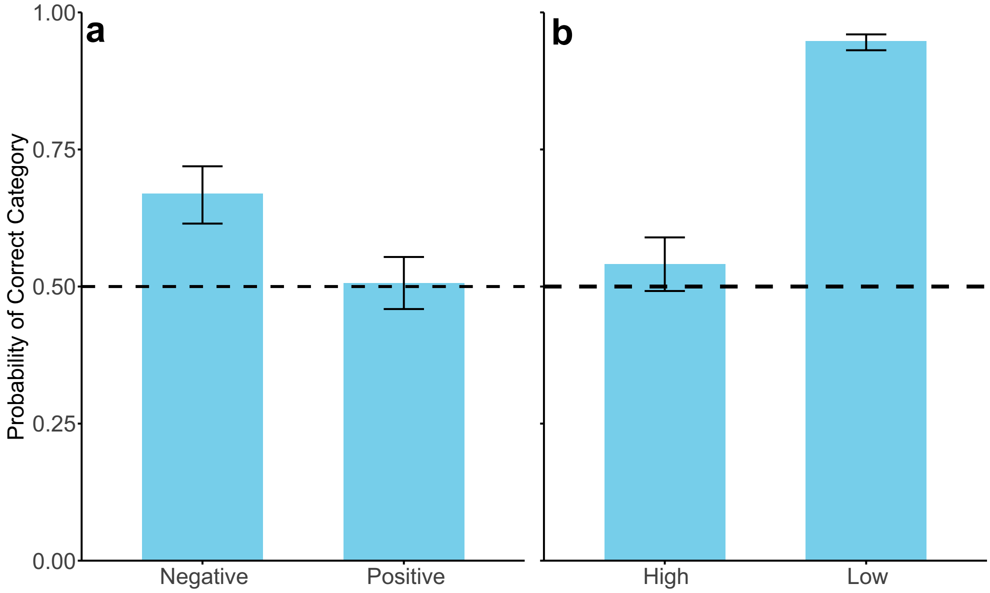

Fig. 1. Bars show the estimated probability of correct valence **(a)** and arousal **(b)** category recognition for the emotion category. Error bars represent 95% asymptotic confidence intervals. The dashed horizontal line indicates chance-level performance (0.50). Probabilities have been calculated based on estimated marginal means, and then the log-odds have been converted to odds.

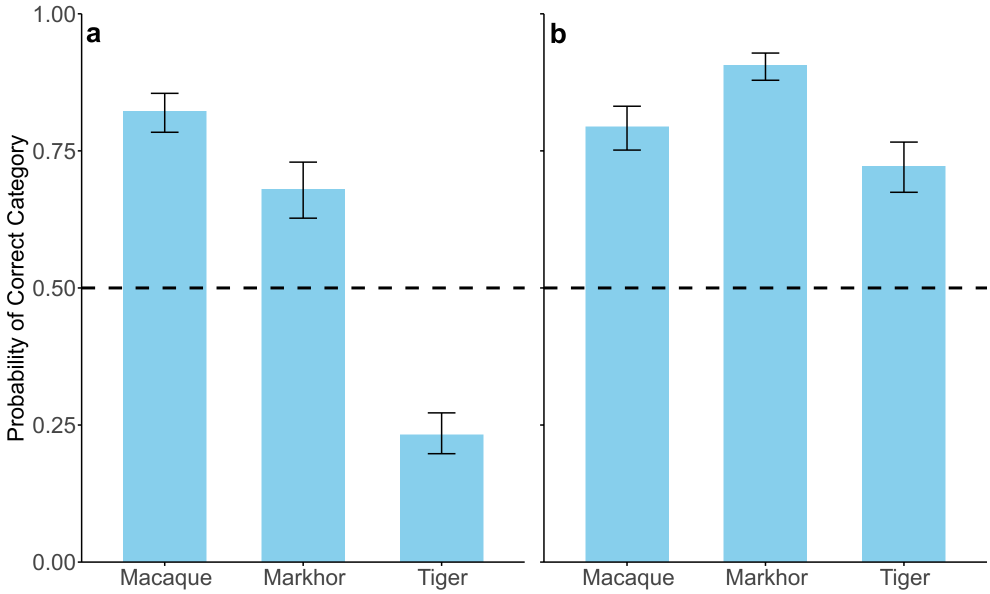

Fig. 2. Bars show the estimated probability of correct valence **(a)** and arousal **(b)** category recognition for each species. Error bars represent 95% asymptotic confidence intervals. The dashed horizontal line indicates chance-level performance (0.50). Probabilities have been calculated based on estimated marginal means, and then the log-odds have been converted to odds. Macaque = Barbary macaque.

[0.74–0.82]) and then compilation 2 (predicted probability = 0.76, CI [0.71–0.80]) and respondents accuracy in arousal recognition was higher in 2021 (predicted probability = 0.86, 95% CI [0.82, 0.89]) than in 2022 (predicted probability = 0.81, 95% CI [0.77, 0.84], Table 2).

Discussion

In this study, we assessed the ability of zoo visitors to accurately recognise the emotional states of three species of zoo animals, as seen in short video clips. We asked visitors to rate video clips for valence and arousal without much contextual information. An independent expert evaluation of the emotional state was based on known

Variable	Estimate	± SE	X^2	*p*
Accuracy of valence recognition				
Intercept	−32.61	162.11		
Gender: man (woman)	− 0.09	0.08	1.19	0.28
Age group (18–30)			11.61	0.01
Age group 31–45	0.22	0.09		
Age group 46–60	0.24	0.11		
Age group 61+	− 0.12	0.14		
Language: Finnish (English)	0.40	0.18	5.07	0.02
Owing a pet: yes (no)	0.14	0.09	2.56	0.11
Valence: positive (negative)	− 0.68	0.10	43.45	< 0.0001
Species (barbary macaque)			709.16	< 0.0001
Species: tiger	−2.72	0.12		
Compilation (1)			173.19	< 0.0001
Compilation: 2	1.47	0.14		
Compilation: 3	0.92	0.09		
Year: 2022 (2021)	0.02	0.08	0.04	0.84
Model fit: McFadden's R^2 = 0.16				
Accuracy of arousal recognition				
Intercept	718.35	180.29		
Gender: man (woman)	− 0.24	0.09	7.58	0.01
Age group (18–30)			5.59	0.13
Age group 31–45	− 0.10	0.10		
Age group 46–60	− 0.08	0.13		
Age group 61+	− 0.37	0.15		
Language: Finnish (English)	− 0.26	0.19	1.85	0.17
Owing a pet: yes (no)	0.07	0.09	0.49	0.48
Arousal: low (high)	2.74	0.12	796.89	< 0.0001
Species (barbary macaque)			124.63	< 0.0001
Species: markhor	0.93	0.14		
Species: tiger	− 0.40	0.10		
Compilation (1)			154.03	< 0.0001
Compilation: 2	− 0.13	0.12		
Compilation: 3	1.10	0.13		
Year: 2022 (2021)	− 0.36	0.09	15.99	< 0.0001
Model fit: McFadden's R^2 = 0.23				

Table 2. Effects of predictors on recognition accuracy (0/1) of emotional Valence (positive vs. negative) and arousal (low vs. high), as observed in video clips. Model estimates of a generalised linear model (GLM) with a binomial family are given with standard error, and p-values are based on a chi-square likelihood ratio test. For factors with multiple levels, estimates are provided for each level, and the intercept gives the value for the level in brackets. There was a total of 3889 responses from $N = 778$ respondents.

emotional indicators for positive and negative valence and low and high arousal for each species, as well as the context in which the video clips were filmed. We then analysed the visitor's ability to recognise accurately emotional states, based on whether the respondent had correctly identified low versus high and positive versus negative valence.

People recognised valence only slightly over chance level (correct valence in 56%) and arousal in 68% of the cases. Confirming our first prediction, we found that overall, respondents recognised arousal more accurately than valence. These findings align with previous results on other mammalian species[3,28], despite our study being limited to the visual modality. Expressions of arousal may be evolutionarily more conserved across species[2] and be therefore more universally interpretable and recognisable, whereas expressions of valence may have evolved more species-specifically and recently[3,27].

Second, we examined whether the valence or arousal depicted in the video clip influenced the accuracy of recognition. Negatively valenced emotions appeared easier to recognise than positive ones. Our findings align with previous findings and support the idea that negative emotions, such as anger and fear, are the most salient across species, because the ability to recognise them tends to be more crucial for survival than recognising positive valence states in other animals[28]. There is also a longer tradition in studying recognition of negative than positive emotions[12,63,64]. A consistent finding in human studies is the so called negativity bias, i.e. the tendency to recognise and react to negative states in others more strongly than to positive ones (e.g[65,66]). Our findings further contribute to our understanding of the negativity bias in emotion recognition by demonstrating that humans are better at recognising negative states also in other animals.

We also found that calm, low-aroused emotional states were almost always recognised correctly (overall 92% of time), but high-aroused ones just above chance level (56% of time). The more accurate recognition of low arousal states compared to high arousal states was somewhat surprising: high arousal is often associated with vigorous movement, which could be easier to identify[34]. In some of the video clips included in the present study, low arousal states may have however been easier to recognise because of the readily identifiable behaviours present in the video, such as a Barbary macaque being groomed. Lack of movement in video clips with low arousal states may have also made it easier for the participants to focus their attention to the subtler details in facial expressions and body language. The hypothesis that participants may associate lack of movement with low arousal states is further supported in that also the high arousal states in non-moving animals, such as a standing markhor waiting for food or a tiger that is frightened of another individual, were difficult to recognise and tended to be scored as low arousal in our dataset.

Out of the three species included here tigers' valence appeared the most difficult to recognise. In fact, the proportion of respondents correctly identifying the valence was just 23%, significantly lower than at chance level, which implies that the valence was not only difficult to recognise, but even actively misunderstood. Like the more precise identification of negative emotional states, this observation also highlights the importance of training in assessing animal welfare. Emotional states in large cats may be particularly difficult to interpret: an earlier study with zoo-housed tigers, reported that keepers' evaluations of their emotional states did not change between situations where behaviour specialists noticed significant differences[49].

To further investigate the factors influencing recognition of emotional states, we also tested whether the owners of domestic cats, another felid, were more accurate in their recognition than non-cat owners in recognising emotional states in tigers. However, we found no such effect and cat owners were no better at identifying valence or arousal of tigers, suggesting that cat ownership doesn't improve people's ability to read emotions in big cats. It has also been noted that overall, emotional states of the domestic cat appear to be rather poorly recognised[3467]. Importantly, we do not know how well cat owners in our data understand their pets in the first place, and whether they would also misinterpret their pets' emotional states.

The video featuring a tiger lying on its back and baring its teeth, which depicts an aroused and negative state in an aggressive interaction with a conspecific, encapsulates much about our ability to interpret animal emotional states. Almost 90% of respondents misinterpreted the situation as being positive. The end of the video, where another tiger arrives and the focal individual continues baring its teeth, was removed from the assessed video clip. Not seeing the end of the interaction possibly leads to a misjudged interpretation of behaviour as being playful, and we may initially interpret animal emotions from our own perspectives, assuming for example that an animal lying on its back wants to play. Additionally, perhaps our cultural perception of the tiger as a powerful and brave animal might make it less likely to be associated with emotions like fear.

A technical issue that limits our ability to interpret species differences is that we were unable to remove other individuals from some of the videos of the social species – Barbary macaques and markhors – whereas tigers were depicted alone in all the videos, which removes the social context that could have facilitated recognition of emotional states.

The respondents recognised most accurately the valence of Barbary macaques, which may be explained by the phylogenetic closeness between humans and other primates. The shared evolutionary origin of emotional expressions and possibly similarity of facial morphology may facilitate recognition of emotions in macaques[5,44,47,68]. Prior research has found that laypeople have difficulties in recognising aggressive, distressed, or friendly facial expressions in Barbary macaques[44] and indeed, some expressions may only appear similar, despite reflecting different emotional states. An example (not present in our video clips) is the human "bared-teeth" smile, which is similar to the chimpanzee fear grin, and they may share a common origin[71], although their functions and underlying emotions differ[69,71]. Another example of evolutionary distance not being the sole explanation for differences in recognition between species is that in our study, the arousal state of markhors was recognised more accurately than that of Barbary macaques.

We also examined whether pet ownership correlated with ability to recognise emotional states. Pet ownership has been shown to correlate with thinking of animals as sentient beings with mental states and the capacity of feeling emotions[20,72]. However, in our study owning a pet did not improve emotion recognition accuracy in either valence or arousal. This supports some previous studies where dog ownership[5] or cat ownership[34] did not influence the emotion recognition accuracy of pet owners and highlights that simple exposure to animals or interaction with them does not necessarily increase interspecific understanding.

In line with our prediction, men were slightly less accurate in their recognition of arousal than women, but there was no difference in valence recognition. Some studies have shown women to score animal emotional states more accurately than men[73,747576], while others have not found a difference[3,20]. Studies, therefore, do not support an overall gender difference in animal emotion recognition accuracy. In line with previous studies, the effect of gender on recognition accuracy was also small in our study: women correctly recognised the category for arousal 85% vs. men 82%.

We found no significant differences between age groups in recognition of neither valence nor arousal, contrary to some previous studies in which younger people (aged 18–30) recognised emotional states more accurately than older age groups[3,39]. Overall, findings on age effects in emotion recognition remain inconsistent, pointing to a more complex interplay between age and context and other factors[787779].

It is well established from human studies that language plays a crucial role in enabling humans to perceive others' emotions[8180]. Despite this, the influence of different languages on the recognition of emotions seems to be somewhat limited, with tone of voice and its other characteristics being more important[82,83]. We found only one study that included the respondents' language in animal emotion recognition, where Filippi et al[28]. found no association between language group and recognition accuracy. In our study, Finnish speakers recognised valence more accurately than non-Finnish respondents (55% vs. 46%). Perhaps the difference is explained by the shared language and cultural background of the research assistants conducting the surveys and the respondents, for example, making it easier for the respondents to understand the finer nuances of the instructions. Additionally, the Finnish speakers were completing the questionnaires in their native language, in contrast to the English language group, the majority of which were probably using a non-native language. As the sample for English speakers was also very small, a larger and more representative sample would be required to make any conclusions about effect of language in this context.

Zoo visitors' perceptions of animal emotional states and their welfare are not necessarily based on knowledge, and can be influenced, for example, by visitors' own emotional states, with visitors' own positive emotions being associated with higher welfare ratings[85]. Assumptions about the well-being of individuals living in zoos have also found to be linked to behavioural activity, with more active individuals seen as being in a more positive emotional state[13,84] For example, a previous study found that despite knowing that tigers are mostly inactive both in the wild as well as in zoos, visitors linked well-being to behavioral activity, with more active individuals seen as being in a more positive emotional state. However, stereotypical behaviour, such as pacing in tigers, has also been found to be noticeable by some visitors, and lead to lowered perceptions of welfare[86].

Education provided by zoos is known to influence the staff's perceptions of animal welfare and enhance their ability to assess it[87]. Our research provides a starting point for zoos, particularly for the educational work of Korkeasaari Zoo, where this study was conducted, and the training of animal caretakers, to educate both staff and public about common misconceptions regarding animal emotional states. According to our observations, some animal caretakers also have partially incorrect perceptions of the expressions of emotional states in these animals, which our findings might help correct.

When we humans improve our ability to observe and identify animal signals correctly, we gain accurate information about how the animal truly feels. This enables us to direct our efforts to better animal welfare in all circumstances. Furthermore, in lieu of the current post-humanistic movement to acknowledge our position in an inter-species network of interactions, we should improve our ability to empathise with other species and move away from anthropocentric interpretations. As a result, our understanding of animals evolves, as do our attitudes and behaviours toward them.

Limitations

Our study has several limitations. To minimise disturbance to the animals, all filming was done from public viewing platforms capturing what was happening in the enclosures, which limited the comparability of situations across species, as well as the context and visibility of conspecifics in the video clips. The year of data collection also had an impact on recognition of arousal, which was more often correctly identified in 2021 as compared to 2022. However, recognition probability of valence did not differ between the years and whereas the result regarding arousal was statistically significant, the effect is minimal (predicted probability 2021: 86% and 2022: 81%,) and likely to reflect random variation between the samples, with significance driven by the large sample size rather than meaningful differences between populations.

Methods

Ethics considerations

All video footage was obtained by following the animals' daily activities and filmed from the audience's vantage point to minimise any potential disturbance. We did not stimulate or disturb the animals in any way. Our research follows all guidelines set by the Finnish National Board on Research Integrity (TENK; https://tenk.fi/en*)* and the institutional review board. Upon evaluation, the institutional review board, Research Ethics Committee in the Humanities and Social and Behavioural Sciences at University of Helsinki, stated that a formal license was not necessary. This was based on the following principles, which we followed: (1) Participation was entirely voluntary, followed principles of informed consent, and respondents could remove their consent at any point without consequences. (2) No sensitive data were collected nor data that would allow identification of individuals. (3) The research did not involve intervening in the physical integrity of research respondents in any way. (4) All respondents were adults over 18 years old. No minors took part in this research. (5) Research did not expose respondents to exceptionally strong stimuli. (6) The research did not involve a risk of causing mental harm that would exceed the limits of normal daily life to the research respondents or their family members or others closest to them. All videos shown to the respondents were of situations and animals that the respondents could also see as part of their normal visit to the zoo, were they not been taking part in this research. (7) Conducting the research did not involve any threat to the safety of respondents or their family members or others closest to them.

Study location and respondents' recruitment

We conducted the study at the Korkeasaari Zoo, Helsinki, Finland. The respondents were recruited by direct requests in situ as well as by advertising the study on the Zoo's social media, website, and posters around the zoo. The study was conducted on two consecutive summers, 2021 and 2022, from June to September. The study took place in an auditorium at the zoo, closed from bystanders (door closed and curtains drawn) for the duration of the session. Each session lasted 15–25 min, with 2–10 participants. The study began with a brief explanation of its aims, procedure, informed consent and instructions (including information that an arrow in some videos indicates the animal to be evaluated). After that, the respondents were asked to remain silent, so they do not to disturb others and to fill in the questionnaire. They were then shown five videos to evaluate. The session ended by collecting the questionnaire and evaluation sheets. The video procedure is explained in detail below.

	Clip Compilation	1st	2nd	3rd	4th	5th
Species	*1*	*tiger*	*markhor*	*macaque*	*markhor*	*markhor*
Valence		neg	neut	pos	pos	neg
Arousal		high	low	ow	high	high
Species	*2*	*macaque*	*markhor*	*tiger*	*macaque*	*tiger*
Valence		neg	neg	neut	pos	pos
Arousal		high	average	low	average	high
Species	*3*	*macaque*	*macaque*	*markhor*	*tiger*	*tiger*
Valence		neg	neut	pos	pos	neg
Arousal		high	low	low	low	high

Table 3. Final selection of video clips with reference categories. Macaque = Barbary macaque; valence: negative, neutral, positive and arousal: low, middle, high.

In the end, we gave the respondents the possibility to remain seated to listen to the study team's explanation of the events and emotional states in the videos with some background information about the stimulus animals. The aim was to keep every session as same as possible. If any respondents had given the English version of the form, everything was also explained in English. The introduction to the videos was identical in content, regardless of the language used in the sessions, and participants were given the same background information and instructions, whether in English or Finnish.

Questionnaire

The questionnaires included questions about the respondent's age group (18–30, 31–45, 46–60, 61–75, 76 and over), gender (woman, man, other, prefer not to say), and animal experience (currently having a pet/not), in the year 2021 there were separate question of owning a cat. It also contained additional questions relevant to another part of the study, not included here. The year 2022 questionnaire was also available in English.

Animal emotion videos

We collected video clips depicting various emotional states of three species, all living at the Korkeasaari Zoo: the Barbary macaque, the tiger, and the markhor. We obtained the video footage by ourselves and filmed all material from the same viewing areas that the visitors had access to during zoo opening hours. The animals were not disturbed in any way for the purpose of obtaining the footage.

The raw video footage was then screened by HT and SEK to select clips depicting emotional states ranging from low to high arousal state and negative to positive valence state. The selection criteria included: the target individual being clearly visible, no food visible and if possible, no other individuals visible, few contextual clues available, and the animal's face visible to the camera at least most of the time. From the first pool of selected film material, we edited clips that were 3–10 s in duration. The length was restricted by our choice of not providing the viewers with too much contextual information, such as a carcass hanging in a tree for the tigers or a physical conflict breaking out among the Barbary macaques. We aimed at having equally many clips of each species in each emotional state.

Then, to derive numeric values for each clip, HT and SEK, both experienced scientists in fields of animal behaviour and emotions, rated the emotional states of the target animals on each video clip on a 7-step Likert scale for both valence and arousal separately, based on the behaviour and facial expressions visible in each clip. Then they carried out a consensus rating for each clip. Finally, for the statistical analyses, the rates were recoded into emotion categories, where valence was scored as 'positive' and arousal as high for scores 5–7, valence as neutral and arousal as average for reference score 4, and valence as 'negative' and arousal as 'low' for reference scores 1–3.

Finally, there were 15 video clips, 5 of each species. They were divided into three compilations, each containing clips from all three species across different levels of arousal and valence (Table 3). Each clip was included in the compilation three times: in real-life speed, as 0.5x slow motion, and as a still image lasting 10 s of the target animal. The total duration of the compilations was 3.1, 3.6 and 3.14 min.

Emotional state recognition by the respondents

We showed the respondents short video clips depicting the target animal exhibiting an emotion in a real-life situation. The respondents were shown one of the three video clip compilations. The video was paused after each clip to allow respondents time to rate it. The respondents rated each clip for arousal and valence, respectively, on a 7-point Likert scale. The explanations were given as follows: 1 = negative, 7 = positive for valence, and 1 = low, 7 = high for arousal. The video compilations shown to study respondents are available at FigShare https://doi.org/10.6084/m9.figshare.c.7807931.v1.

Analyses

Altogether, we obtained 5410 ratings for valence and arousal, respectively, from 1082 respondents. We excluded from analyses the respondents who had not rated one or more of the clips or had left questions concerning gender, age, or having a pet unanswered (N = 85). Regarding gender, respondents who categorised themselves as 'other' or chose the 'not want to answer' option were excluded from the analysis due to the small number (respectively 9 and 10). Eleven respondents who reported their age as over 75 years were combined into the 61–75 age group.

First, we recalculated the accuracy of the valence and arousal ratings given by the respondents, by assessing whether the answer matched the emotion category ratings done by the animal behaviour experts as explained above ('positive'/'high' = reference rates 5–7, vs 'negative'/'low' = reference rates 1–3). Comparing these yielded the accuracy of each video clip from each respondent. The rates for neutral valence and average arousal, indicated by value 4, were omitted from analyses. This resulted in a final dataset of 3889 responses.

Next, we assessed the factors predicting the accuracy of the respondents' recognition by fitting generalised linear models (GLM), with binomial error distribution and a logit link function. We used separate models to evaluate the effects on valence and arousal recognition. The target variable was the accuracy of the evaluation (correct/not correct) and the predictors were the reference category of the emotional state being evaluated (negative/positive valence, or low/high arousal), species (Barbary macaque, tiger, markhor), gender (woman/man), age group (18–30, 31–45, 46–60, 61 and over) and pet ownership (yes or no). The year of data collection (2021/2022) and video clip compilation were included as additional predictors. We also modelled a subset of the 2021 data for recognition accuracy of tiger emotional states only, adding cat ownership into the models as a predictor (yes, no) and the response variable was the accuracy of valence or arousal recognition as above.

Initially, we attempted to use Generalised linear mixed models (GLMM) with the respondent identity as a random variable to account for pseudoreplication in our sample, as each respondent provided ten scores to the data (five for valence and five for arousal). However, these models failed to converge. The high number of total responses (> 5000) versus the number of answers provided by a single respondent (5 per model) is the likely reason for this, as total variation in our data explained by respondent identity is minuscule, in comparison to that explained by the variables included in the models. When running the models with respondent ID number as a random factor despite the warnings, variation explained by the random factor was estimated as 0.0000, both for valence and arousal.

All analyses were conducted in R statistical software (v4.2.2[88]) using generalised linear models (GLMs) implemented in the lme4 package (v1.1–26[89]). Model fit was evaluated using McFadden's pseudo-R^2, which describes the proportional improvement in model fit compared to the null model. For factors with multiple levels, estimated marginal means were used to test pairwise differences between levels using the emmeans package (v1.10.4.900003[90]). The emmeans package was also used to calculate log-odds and convert them into predicted probabilities.

Data availability

The dataset used in the current study is available in FigShare at https://doi.org/10.6084/m9.figshare.c.7807931.v1.

Received: 26 July 2024; Accepted: 10 November 2025
Published online: 24 December 2025

References

1. Pritsch, C., Telkemeyer, S., Mühlenbeck, C. & Liebal, K. Perception of facial expressions reveals selective affect-biased attention in humans and orangutans. *Sci. Rep.* **7**, 7782 (2017).
2. Filippi, P., Gogoleva, S. S., Volodina, E. V., Volodin, I. A. & de Boer, B. Humans identify negative (but not positive) arousal in silver Fox vocalizations: implications for the adaptive value of interspecific eavesdropping. *Curr. Zool.* **63**, 445–456 (2017).
3. Greenall, J. S., Cornu, L., Maigrot, A. L., de la Torre, M. P. & Briefer, E. F. Age, empathy, familiarity, domestication and call features enhance human perception of animal emotion expressions. *R Soc. Open. Sci.* **9**, 221138 (2022).
4. Dawkins, M. S. A user's guide to animal welfare science. *Trends Ecol. Evol.* **21**, 77–82 (2006).
5. Amici, F., Waterman, J., Kellermann, C. M., Karimullah, K. & Bräuer, J. The ability to recognize dog emotions depends on the cultural milieu in which we grow up. *Sci. Rep.* **9**, 16414 (2019).
6. Apostol, L., Rebega, O. L. & Miclea, M. Psychological and Socio-demographic predictors of attitudes toward animals. *Procedia - Soc. Behav. Sci.* **78**, 521–525 (2013).
7. Furnham, A., McManus, C. & Scott, D. Personality, empathy and attitudes to animal welfare. *Anthrozoös* **16**, 135–146 (2003).
8. Kujala, M. V., Somppi, S., Jokela, M., Vainio, O. & Parkkonen, L. Human Empathy, personality and experience affect the emotion ratings of dog and human facial expressions. *PLOS ONE*. **12**, e0170730 (2017).
9. Alba, A. C., Ferrie, G. M., Schook, M. W., Leahy, M. & Cronin, K. A. Gender and Age, but not visual Context, impact wildlife attitudes and perceptions of animals in zoos. *J. Zool. Bot. Gard*. **4**, 118–133 (2023).
10. Bell, C., Rogers, S., Taylor, J. & Busby, D. Improving the recognition of equine affective States. *Animals* **9**, 1124 (2019).
11. Farmer, H., Holmes, L., Spiezio, C. & Binding, S. EAZA animal welfare special Issue - Editorial. *J. Zoo Aquar Res.* **12**, 195–195 (2024).
12. Jones, N., Sherwen, S. L., Robbins, R., McLelland, D. J. & Whittaker, A. L. Welfare assessment tools in zoos: from theory to practice. *Vet. Sci.* **9**, 170 (2022).
13. Melfi, V. A., McCormick, W. & Gibbs, A. A preliminary assessment of how zoo visitors evaluate animal welfare according to enclosure style and the expression of behavior. *Anthrozoös* **17**, 98–108 (2004).
14. Darwin, C. The Expression of the Emotions in Man and Animals. John Murray, London (1872).
15. Barrett, L. F. Are emotions natural kinds? *Perspect. Psychol. Sci.* **1**, 28–58 (2006).
16. Mendl, M., Burman, O. H. P. & Paul, E. S. An integrative and functional framework for the study of animal emotion and mood. . *Proc. R. Soc. B Biol. Sci.* **277**, 2895–2904 (2010).
17. Paul, E. S. & Mendl, M. T. Animal emotion: descriptive and prescriptive definitions and their implications for a comparative perspective. *Appl. Anim. Behav. Sci.* **205**, 202–209 (2018).
18. Russell, J. A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**, 1161–1178 (1980).
19. Dijksterhuis, A. & Aarts, H. On wildebeests and humans: the Preferential detection of negative stimuli. *Psychol. Sci.* **14**, 14–18 (2003).
20. Hawkins, R. D., Hatin, B. D. M. & Revesz, E. O. Accuracy of Canine vs. Human Emotion Identification: Impact of Dog Ownership and Belief in Animal Mind. *Hum.-Anim. Interact. Bull.* **2021**. (2021).

21. Kensinger, E. & Kensinger Remembering the Details: Effects of Emotion - Elizabeth A. (2009). https://journals.sagepub.com/doi/abs/10.1177/1754073908100432 (2009).
22. Mortillaro, M. & Dukes, D. Jumping for joy: the importance of the body and of dynamics in the expression and recognition of positive emotions. *Front Psychol.* **9** 763 (2018).
23. Törnqvist, H., Höller, H., Vsetecka, K., Hoehl, S. & Kujala, M. V. Matters of development and experience: evaluation of dog and human emotional expressions by children and adults. *PLOS ONE*. **18**, e0288137 (2023).
24. Bloom, T. & Friedman, H. Classifying dogs' (Canis familiaris) facial expressions from photographs. *Behav. Processes*. **96**, 1–10 (2013).
25. Donnier, S., Kovács, G., Oña, L. S., Bräuer, J. & Amici, F. Experience has a limited effect on humans' ability to predict the outcome of social interactions in children, dogs and macaques. *Sci. Rep.* **10**, 21240 (2020).
26. Lakestani, N. N., Donaldson, M. L. & Waran, N. Interpretation of dog behavior by children and young adults. *Anthrozoös* **27**, 65–80 (2014).
27. Briefer, E. F. Coding for 'Dynamic' information: vocal expression of emotional arousal and Valence in Non-human animals. In *Coding Strategies in Vertebrate Acoustic Communication* (eds Aubin, T. & Mathevon, N.) 137–162 (Springer International Publishing, 2020). https://doi.org/10.1007/978-3-030-39200-0_6.
28. Filippi, P. et al. Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: evidence for acoustic universals. *Proc. Biol. Sci.* **284**, 20170990 (2017).
29. Briefer, E. F. et al. Segregation of information about emotional arousal and Valence in horse whinnies. *Sci. Rep.* **5**, 9989 (2015).
30. Briefer, E. F. et al. Classification of pig calls produced from birth to slaughter according to their emotional Valence and context of production. *Sci. Rep.* **12**, 3409 (2022).
31. McGrath, N., Phillips, C. J. C., Burman, O. H. P., Dwyer, C. M. & Henning, J. Humans can identify reward-related call types of chickens. *R Soc. Open. Sci.* **11**, 231284 (2024).
32. McComb, K., Taylor, A. M., Wilson, C. & Charlton, B. D. The cry embedded within the purr. *Curr. Biol.* **19**, R507–R508 (2009).
33. Somppi, S. et al. Dog–Owner Relationship, owner interpretations and dog personality are connected with the emotional reactivity of dogs. *Animals* **12**, 1338 (2022).
34. Dawson, L., Cheal, J., Niel, L. & Mason, G. Humans can identify cats' affective States from subtle facial expressions. *Anim. Welf.* **28**, 519–531 (2019).
35. Demirbas, Y. S. et al. Adults' ability to interpret canine body Language during a Dog–Child interaction. *Anthrozoös* **29**, 581–596 (2016).
36. Tallet, C., Špinka, M., Maruščáková, I. & Šimeček, P. Human perception of vocalizations of domestic piglets and modulation by experience with domestic pigs (Sus scrofa). *J. Comp. Psychol.* **124**, 81–91 (2010).
37. Hampson, E., van Anders, S. M. & Mullin, L. I. A female advantage in the recognition of emotional facial expressions: test of an evolutionary hypothesis. *Evol. Hum. Behav.* **27**, 401–416 (2006).
38. Sullivan, S., Campbell, A., Hutton, S. B. & Ruffman, T. What's good for the Goose is not good for the gander: age and gender differences in scanning emotion faces. *J. Gerontol. Ser. B*. **72**, 441–447 (2017).
39. West, J. T. et al. Age effects on emotion recognition in facial displays: from 20 to 89 years of age. *Exp. Aging Res.* **38**, 146–168 (2012).
40. Battini, M., Agostini, A. & Mattiello, S. Understanding Cows' Emotions on Farm: Are Eye White and Ear Posture Reliable Indicators?. *Animals* **9**, 477 (2019).
41. Bremhorst, A., Sutter, N. A., Würbel, H., Mills, D. S. & Riemer, S. Differences in facial expressions during positive anticipation and frustration in dogs awaiting a reward. *Sci. Rep.* **9**, 19312 (2019).
42. Hall, C., Randle, H., Pearson, G. & Preshaw, L. Waran, N. Assessing equine emotional state. *Appl. Anim. Behav. Sci.* **205**, 183–193 (2018).
43. Hazel, S. J., O'Dwyer, L. & Ryan, T. Chickens are a lot smarter than I originally thought: changes in student attitudes to chickens following a chicken training class. *Animals* **5**, 821–837 (2015).
44. Clark, L. *Understanding Human Perception of Primate Facial Expressions and its Impacts on Human-nonhuman Primate Interactions* (University of Lincoln, 2020).
45. Kelly, T. et al. Adult human perception of distress in the cries of bonobo, chimpanzee, and human infants. *Biol. J. Linn. Soc.* **120**, 919–930 (2017).
46. Linnankoski, I., Laakso, M., Aulanko, R. & Leinonen, L. Recognition of emotions in macaque vocalizations by children and adults. *Lang. Commun.* **14**, 183–192 (1994).
47. Maréchal, L., Levy, X., Meints, K. & Majolo, B. Experience-based human perception of facial expressions in barbary macaques (Macaca sylvanus). *PeerJ* **5**, e3413 (2017).
48. Scheumann, M., Hasting, A. S., Kotz, S. A. & Zimmermann, E. The voice of emotion across species: how do human listeners recognize animals' affective states? *PLoS ONE*. **9**, e91192 (2014).
49. Phillips, C. J. C., Tribe, A., Lisle, A., Galloway, T. K. & Hansen, K. Keepers' rating of emotions in captive big cats, and their use in determining responses to different types of enrichment. *J. Vet. Behav.* **20**, 22–30 (2017).
50. Wemelsfelder, F., Millard, F., De Rosa, G. & Napolitano, F. *Qualitative Behaviour assessment. In Welfare Quality* (GBR, 2009).
51. Celozzi, S. et al. Human perception of goats' vocal expression of emotions. *Appl. Anim. Behav. Sci.* **292**, 106833 (2025).
52. Battini, M. et al. The use of qualitative behaviour assessment for the On-Farm welfare assessment of dairy goats. *Animals* **8**, 123 (2018).
53. Muri, K. & Stubsjøen, S. M. Inter-observer reliability of qualitative behavioural assessments (QBA) of housed sheep in Norway using fixed lists of descriptors. *Anim. Welf.* **26**, 427–435 (2017).
54. Grosso, L. et al. On-farm qualitative behaviour assessment of dairy goats in different housing conditions. *Appl. Anim. Behav. Sci.* **180**, 51–57 (2016).
55. Tami, G. & Gallagher, A. Description of the behaviour of domestic dog (Canis familiaris) by experienced and inexperienced people. *Appl. Anim. Behav. Sci.* **120**, 159–169 (2009).
56. Wan, M., Bolger, N. & Champagne, F. A. Human perception of fear in dogs varies according to experience with dogs. *PLOS ONE*. **7**, e51775 (2012).
57. Bokkers, E. M., de Vries, M., Antonissen, I. & de Boer, I. Inter- and intra-observer reliability of experienced and inexperienced observers for the qualitative behaviour assessment in dairy cattle. *Anim. Welf.* **21**, 307–318 (2012).
58. Munsterhjelm, C., Heinonen, M. & Valros, A. Application of the welfare Quality® animal welfare assessment system in Finnish pig production, part II: associations between animal-based and environmental measures of welfare. *Anim. Welf.* **24**, 161–172 (2015).
59. Travnik, I. C., Machado, D. S. & Sant'Anna, A. C. Do you see the same Cat that I see? Inter- and intra-observer reliability for qualitative behaviour assessment as temperament indicator in domestic Cats. *Anim. Welf.* **31**, 319–327 (2022).
60. Laméris, D. W. et al. Qualitative behavioural assessment of Bonobo emotional expressivity across observer groups and zoo housing environments. *Anim. Welf.* **33**, e28 (2024).
61. Erzinçlioğlu, T. S. & Rutherford, K. M. D. Using qualitative behavior assessment to investigate the effect of tourist presence on the welfare in captive tigers (Panthera tigris) in three tourism facilities in Thailand. *Zoo Biol.* **43**(1), 42-54 (2024).
62. Baek, C., Doleck, T. & Learning Analytics A comparison of Western, Educated, Industrialized, Rich, and Democratic (WEIRD) and Non-WEIRD research. *Knowl. Manag E-Learn*. **16**, 217–236 (2024).
63. Boissy, A. et al. Emotions and cognition: A new approach to animal welfare. *Anim. Welf.* **16**, 37–43 (2007).

64. Nelson, X. J. et al. Joyful by nature: approaches to investigate the evolution and function of joy in non-human animals. *Biol. Rev.* **98**, 1548–1563 (2023).
65. Ito, T. A., Larsen, J. T., Smith, N. K. & Cacioppo, J. T. Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J. Personal. Soc. Psychol.* **75**, 887–900 (1998).
66. Rozin, P. & Royzman, E. B. Negativity Bias, negativity Dominance, and contagion. *Pers. Soc. Psychol. Rev.* **5**, 296–320 (2001).
67. Holden, E. et al. Evaluation of facial expression in acute pain in cats. *J. Small Anim. Pract.* **55**, 615–621 (2014).
68. Sullivan, S. K., Kim, A., Vinicius Castilho, L. & Harris, L. T. Comparing emotion inferences from dogs (Canis familiaris), Panins (Pan troglodytes/Pan paniscus), and humans (Homo sapiens) facial displays. *Sci. Rep.* **12**, 13171 (2022).
69. Preuschoft, S. Laughter and smile in barbary macaques (Macaca sylvanus). *Ethology* **91**, 220–236 (1992).
70. Graziano, M. S. A. The origin of smiling, laughing, and crying: the defensive mimic theory. *Evol. Hum. Sci.* **4**, e10 (2022).
71. Leopold, D. A. & Rhodes, G. A comparative view of face perception. *J. Comp. Psychol.* **124**, 233–251 (2010).
72. Kupsala, S., Vinnari, M., Jokinen, P. & Räsänen, P. Public perceptions of mental capacities of nonhuman animals: Finnish population survey. *Soc. Anim.* **24**, 445–466 (2016).
73. Faragó, T., Takács, N., Miklósi, Á. & Pongrácz, P. Dog growls express various contextual and affective content for human listeners. *R Soc. Open. Sci.* **4**, 170134 (2017).
74. Merkies, K., Crouchman, E. & Belliveau, H. Human ability to determine affective States in domestic horse whinnies. *Anthrozoös* **35**, 483–494 (2022).
75. Mestre, M. V., Samper, P., Frías, M. D. & Tur, A. M. Are women more empathetic than men? A longitudinal study in adolescence. *Span. J. Psychol.* **12**, 76–83 (2009).
76. Paul, E. S. Empathy with Animals and with Humans: Are They Linked?. *Anthrozoös* **13**, 194–202 (2000).
77. Norring, M., Wikman, I., Hokkanen, A. H., Kujala, M. V. & Hänninen, L. Empathic veterinarians score cattle pain higher. *Vet. J.* **200**, 186–190 (2014).
78. Taylor, N. & Signal, T. D. Empathy and attitudes to animals. *Anthrozoös* **18**, 18–27 (2005).
79. Mill, A., Allik, J., Realo, A. & Valk, R. Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion* **9**, 619–630 (2009).
80. Scherer, K. R. & Scherer, U. Assessing the ability to recognize facial and vocal expressions of emotion: construction and validation of the emotion recognition index. *J. Nonverbal Behav.* **35**, 305–326 (2011).
81. Lindquist, K. A. & Gendron, M. What's in a word? Language constructs emotion perception. *Emot. Rev.* **5**, 66–71 (2013).
82. Pell, M. D., Paulmann, S., Dara, C., Alasseri, A. & Kotz, S. A. Factors in the recognition of vocally expressed emotions: A comparison of four languages. *J. Phon.* **37**, 417–435 (2009).
83. Sauter, D. A., Eisner, F., Calder, A. J. & Scott, S. K. Perceptual cues in nonverbal vocal expressions of emotion. *Q. J. Exp. Psychol.* **63**, 2251–2272 (2010).
84. Salas, M., Laméris, D. W., Depoortere, A., Plessers, L. & Verspeek, J. Zoo visitor attitudes are more influenced by animal behaviour than environmental enrichment appearance. *Animals* **11**, 1971 (2021).
85. Vera-Concha, M. I. et al. How happy do these animals look? Exploring factors influencing children's perceptions of animal welfare at the zoo. *Anim. Open. Access. J. MDPI.* **15**, 1595 (2025).
86. Miller, L. J. Visitor reaction to pacing behavior: influence on the perception of animal care and interest in supporting zoological institutions. *Zoo Biol.* **31**, 242–248 (2012).
87. Ogle, B. & DeSmet, A. The perception of felid welfare by zookeepers in North America and the implications for zoo managers. *Zoo Biol.* **42**, 651–660 (2023).
88. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL (2022). https://www.R-project.org/
89. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear Mixed-Effects models using lme4. *J Stat. Softw* **67**(1), 1-48. (2015).
90. Lenth, R. V., Least-Squares & Means The R package Lsmeans. *J. Stat. Softw.* **69**, 1–33 (2016).

Acknowledgements

This research was funded by the Kone Foundation to PI's Koski, Telkänranta & Vitikainen, grant no 202010130 and the Finnish Society of Sciences and Letters to Laura Hiisivuori. Arador Innovations Ltd provided additional support. We thank the personnel of Korkeasaari Zoo for collaboration throughout the project, for providing access to the visitors and for all the logistic and marketing help which made this study possible. We like to especially acknowledge Research Coordinator Kirsi Pynnönen-Oudman, Director of Animal Care and Conservation Nina Trontti and the animal keepers Jonne Stenroth, Outi Hannikainen, Merja Wahlroos and Lotta Kivalo, Mari Lehmonen and Annika Sorjonen from Zoo Communications. We also thank the M.Sc. students Pinja Nivalainen, Anssi Vainio and Katri Alavalkama as well as summer assistants Pilvi Eronen, Nea Hamara, Tea Harilainen, Kaisla Kareoja, Essi Kumpumäki, Hilla Laakkonen and Eemeli Lounela who made the data collection at this scale possible. We also thank Jussi Eronen for his thoughtful comments on the manuscript. With great sadness and gratitude, we dedicate this work to the memory of our beloved colleague and co-author Sonja Koski, who passed away this spring. Her intellectual leadership was vital to this study, and her wisdom and commitment continue to inspire us.

Author contributions
SEK, HT, KP, and EV designed and coordinated the study. LH collected the data, SEK and HT procured and evaluated the video material, and LH and EV carried out the statistical analyses. LH, EV, and SEK wrote the paper; all authors have approved the final submitted version, and EV takes full responsibility over the scientific contributions of SEK to this work.

Funding

This research was funded by Kone Foundation, Helsinki, Finland (grant number 202010130). Open access was funded by Helsinki University Library.

Declarations

Competing interests
The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.H.