ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/372156837

Dogs assess human competence from observation alone and use it to predict future behaviour

Article *in* Learning and Motivation · July 2023

DOI: 10.1016/j.lmot.2023.101911

CITATION
1

READS
30

4 authors:

Rebecca Hassall
University of Auckland
5 PUBLICATIONS **33** CITATIONS
SEE PROFILE

Patrick Daniel Neilands
University of Auckland
11 PUBLICATIONS **65** CITATIONS
SEE PROFILE

Amalia Bastos
Johns Hopkins University
18 PUBLICATIONS **97** CITATIONS
SEE PROFILE

Alex H Taylor
University of Auckland
117 PUBLICATIONS **2,521** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Cross-Species Comparison of Contrafreeloading in Parrots View project

Project The signature testing approach to investigating animal cognition. Focusing on canine cognition. View project

All content following this page was uploaded by Patrick Daniel Neilands on 19 July 2023.

The user has requested enhancement of the downloaded file.

Dogs assess human competence from observation alone and use it to predict future behaviour

Rebecca Susan Hassall[1*†], Patrick Neilands[1†] (http://orcid.org/0000-0002-8880-8674), Amalia Pinkusfeld Medeiros Bastos[1] (https://orcid.org/0000-0003-3037-687X), Alex Harwood Taylor[1,2,3,4] (https://orcid.org/0000-0003-3492-7667)

---

[1] School of Psychology, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
[2] ICREA, Pg. Lluís Companys 23,Barcelona, Spain
[3] Institut de Neurociències, Universitat Autònoma de Barcelona, 08193, Bellaterra, Barcelona, Spain
[4] School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch 8140,New Zealand
* Correspondence: patrickneilands@gmail.com

[†] Both authors contributed equally to this work.

Abstract

One key aspect of social cognition is the ability to assess the competence of other agents and then use this information to predict future behaviour. Current research shows that humans assess competence in a highly sophisticated manner. Our species is capable of taking third-party, relative observations of agent's competencies and then using them to make future predictions about the different behaviours of a single agent. To date, few studies have examined if non-human animals show these same features of competence assessment. Here, we find suggestive evidence that domestic dogs can use third-party observations of the relative competence of humans (at throwing or kicking a ball) to predict how far an experimenter will propel a ball in the future. This was despite the actual behaviour produced at test (a fake throw or kick) being identical, and dogs needing to assign different, action-specific competencies to the same agent for these two behaviours. Dogs therefore appear to display four behavioural signatures of human competence judgement, which suggests that the ability to make sophisticated competence judgements may not be unique to humans.

Keywords: domestic dog, social cognition, competence judgement, competence prediction

One central aspect of social intelligence is the ability to predict the behaviour of other agents (FeldmanHall & Shenhav, 2019; Frith & Frith, 2006; Sciutti et al., 2015). There is potentially high adaptive value in assessing both if another agent intends to be a friend or foe, and whether this agent has the competence to carry out these intentions. In humans, these two dimensions of social evaluation, which have been labelled as 'warmth' and 'competence', account for 82% of the variance in the everyday inferences humans make about the social behaviours of others (Fiske et al., 2007; Wocjciszke et al., 1998). While there has been a sustained research effort focused on the development and evolution of the social evaluations of prosociality (the warmth dimension) across infants (Hamlin, 2015; Hamlin et al., 2007, 2010, 2011; Hamlin & Wynn, 2011; Buon et al., 2014; Surian & Franchin, 2017; Paquette-Smith & Johnson, 2015; Hamlin, 2013a, 2013b) and animals (Abril-de-Abreu et al., 2015; Anderson, Kuroshima, et al., 2013; Anderson, Takimoto, et al., 2013; Bshary & Grutter, 2006; Chijiiwa et al., 2015; Herrmann et al., 2013; Jim et al., 2020; Krupenye & Hare, 2018; Russell et al., 2008; Trösch et al., 2020), there has been less focus on the competence dimension, despite its evolutionary importance.

Being able to judge competence seems likely to be adaptive in many ecological situations, from tracking which conspecifics are particularly vigilant or skilled at extracting food, to knowing which escape or hunting strategies predators and prey in the environment excel at. It would therefore be useful to not only be able to track the overall competence of other agents (e.g., generally competent or incompetent), but to track action-specific competencies by assigning different competencies for different behaviours an agent performs (e.g., competent at vigilance, incompetent at foraging).

Current research shows that human children assess competence in a highly sophisticated manner. Infants (1.5-2.5 years old) can discriminate between more and less competent agents, preferring to play with a competent puppet that they have observed causing a toy to play music by pushing a button once, compared to an incompetent puppet that takes 6-8 attempts to push the button (Jara-Ettinger et al., 2015). Children can use these competence inferences to make different predictions and judgements about the agent’s subsequent behaviour, even going so far as to interpret identical behaviours differently, depending on the relative competencies they have previously observed (Jara-Ettinger et al., 2015; Pasquini et al., 2007) For example, 3-4 year old children prefer to use an unfamiliar object label suggested by an agent who has previously correctly labelled familiar objects 75% of the time, rather than a label suggested by an agent who has labelled familiar objects correctly only 25% of the time (Pasquini et al., 2007). Finally, 2.5-5 year old children show the ability to assign action-specific competencies to the same agent, such as when choosing a tall agent over a short agent when an object needs to be moved from a high shelf but choosing the short agent over the tall when the agent needs to move through a small door (Paulus & Moore, 2011).

These results demonstrate that even at a young age, there are a number of key behavioural components that characterise the competence judgements of humans. The following components can be considered as behavioural signatures of the cognitive mechanisms underpinning competence attribution in humans (Taylor, 2014). First, these judgements can be made from third-party information alone, i.e., from the observation of other agents, rather than from directly interacting with another agent (Signature 1). Second, judgements can be made from the observation of relative task performance, in terms of how good an agent is at a particular task, rather than absolute performance, in terms of whether an agent is simply capable of a behaviour or not (Signature 2). Third, these judgements can be used to interpret identical behaviours by agents differently, based on the past level of competence each agent has shown for the behaviour in question (Signature 3). Finally, children can assign action-specific competencies to the same agent for different behaviours, such as when choosing a tall agent over a small one for a reaching task, but a small agent over a tall one for a task where the agent needs to enter a small door (Signature 4).

To date, there has been limited work on the competence attribution of nonhuman animals, despite its evolutionary significance and its everyday use in humans. Melis et al. (2006) showed that chimpanzees had a preference for a more competent string puller over a less competent one in a cooperative string-pulling task. However, this preference arose only after direct interaction with each string puller, rather than via third party observation, and appears to have been based on using a ‘win-stay, lose-shift’ strategy, where the chimpanzees switched partners if they experienced failure on a trial. It is therefore unclear if chimpanzees were truly assigning some level of competence to each string puller or simply associating each agent with receiving a reward or not. Outside of chimpanzees, a similar study showed that coral trout were more likely to recruit effective rather than ineffective moray eel collaborators (Vail et al., 2014). However, due to the analogous experimental design, the same difficulties in interpreting the result remains.

Dogs are a key model species for understanding the evolution of social intelligence (Hare et al., 2002; MacLean et al., 2017; Miklósi et al., 2003, 2004) due to their domestication by humans over the last 30,000 years (Hare et al., 2002; Miklosi et al., 2004; Miklósi et al., 2003). Despite being more distantly related to humans than other primates, their long association with humans may have led to them evolving similar socio-cognitive abilities to our own (MacLean et al., 2017; Miklosi et al., 2004). In particular, two elements of dogs' evolution history make them an ideal species for exploring the evolution of competence attribution. Firstly, both dogs and their ancestors took part in group hunting. Secondly, dogs' evolutionary success is rooted in the ability to form social partnerships with humans. In both circumstances, the ability to assess the competence of potential partner or prey could be highly advantageous. Furthermore, as well as having important theoretical implications, understanding whether dogs can assess competence may have important practical implications in areas such as the training of disability support dogs.

However, there have been few attempts to examine if dogs assess competence and the results have been inconclusive. In one study, where dogs were able to access an apparatus to retrieve a reward, one experimenter rebaited an apparatus (the 'filler') and another experimenter would unblock the food-dispensing mechanisms (the 'helper'). When given the opportunity to 'look back' (a form of requesting assistance) towards either the 'filler' or a 'helper', dogs did not differentially look back to the correct person depending on whether the apparatus was unbaited or blocked (Horn et al., 2012). However, dogs did spend more time closer to the appropriate person, which, coupled with doubts about whether looking back is a problem solving strategy in dogs (Lazzaroni et al., 2020), suggests further testing is required. Furthermore, while suggestive that dogs are sensitive to whether humans differ in how helpful their behaviour would be to the dog, this study did not assess competence *per se*. A more recent study directly studied competence-assessment but had similarly inconclusive findings. Dogs did not look back more at a skilful demonstrator than an unskilful one when presented with an unsolvable task, but did show a non-significant trend to look longer at a skilful demonstrator when no longer required to choose between two human demonstrators (Piotti et al., 2017).

Here, we designed a paradigm to test if dogs show the four signatures of human competence attribution outlined above, while using a more naturalistic situation than past studies. To do this, we examined if dogs could predict how far one of two experimenters would throw or kick a ball, based on their past observations of the two experimenters throwing or kicking the ball to another human. If humans and dogs use similar cognitive mechanisms for assessing competence, dog should also show these signatures during a competence-attribution task.

In Experiment 1, we gave dogs experience with two unfamiliar agents who they observed throwing or kicking a ball back and forth between themselves. Each agent was demonstrated to be competent at one task, but incompetent at the other, i.e., the competent thrower was an incompetent kicker, and vice versa. Dogs then observed a fake kick or throw by each of these agents. We took the distances the dogs pre-emptively ran for each fake throw/kick as a measure of their prediction of the individuals' competence, i.e., how far they expected the ball to move given the agents' skill level. Given that the experimenters never directly threw or kicked the ball for the dogs, making such a prediction requires the dogs to be capable of i) reacting differently to an identical behaviour at test (the fake throw or kick) based on the third-party interaction they had previously observed between two humans throwing or kicking the ball relatively different distances, and ii) assigning action-specific competencies to each agent for each behaviour, as good throwers were bad kickers and vice versa.

However, there remains a simpler explanation than competence attribution for the dogs running further for the competent experimenter. Namely, that the dogs merely learnt to associate the competent experimenter with the ball moving further. As such, the competent thrower or kicker holding the ball in the test phase may have been sufficient to cause the dogs to run further without requiring any competence attribution. We controlled for this association hypothesis by running a second experiment where, in the observation trials, the experimenters stood beside a ball-throwing machine that either threw the ball a long or short distance. In the test trials, the experimenters then gave a fake throw as in Experiment 1. If the distance that the dogs ran was merely due to the association between the ball moving further and the experimenter in proximity to the ball at the time, the dogs should have run further in the long-distance condition compared to the short-distance condition.

The purpose of this study is to investigate whether dogs are capable of assigning action-specific competencies to agents. If they are, we would predict that dogs would run further for the competent experimenter in Experiment 1 but would show no difference in how they ran in the two conditions in Experiment 2.

Methods

Subjects

For Experiment 1, we tested 22 dogs (14 female, 8 male) between the ages of two and ten years old (Table S1). A further eight dogs were excluded (4 due to a loss of motivation to chase the ball during the motivation routine, 1 due to a refusal to return to the starting position and 3 due to experimenter error). For Experiment 2, we tested 22 new dogs (7 female, 15 male) between the age of two and ten years old (Table S2). A further four dogs were excluded, 2 due to a loss of motivation during the motivation routine, 2 due to experimenter error). The dogs were unfamiliar with the experimenters before taking part in this study. All dogs were pet dogs, whose owners had registered them to take part in sessions at the Clever Canine Lab. Dogs were selected for this study based on their willingness to chase a ball that was both kicked and thrown after being handled and brought to the testing spot. This was measured by having Experimenter 1 bring the dog to the testing spot where another experimenter not involved in the study propelling the ball four times (two throws and two kicks). This selection ensured all dogs ran for the ball in the experimental condition (all dogs ran at least 0.25m, aside from one trial by Participant 3 in the incompetent kicker condition). Our work was carried out under the approval of the University of Auckland Animal Ethics Committee (reference no. 001826).

Experiment 1

Dogs took part in a total of four conditions: Competent Thrower, Incompetent Thrower, Competent Kicker, Incompetent Kicker. Each condition consisted of three phases (Fig 1).

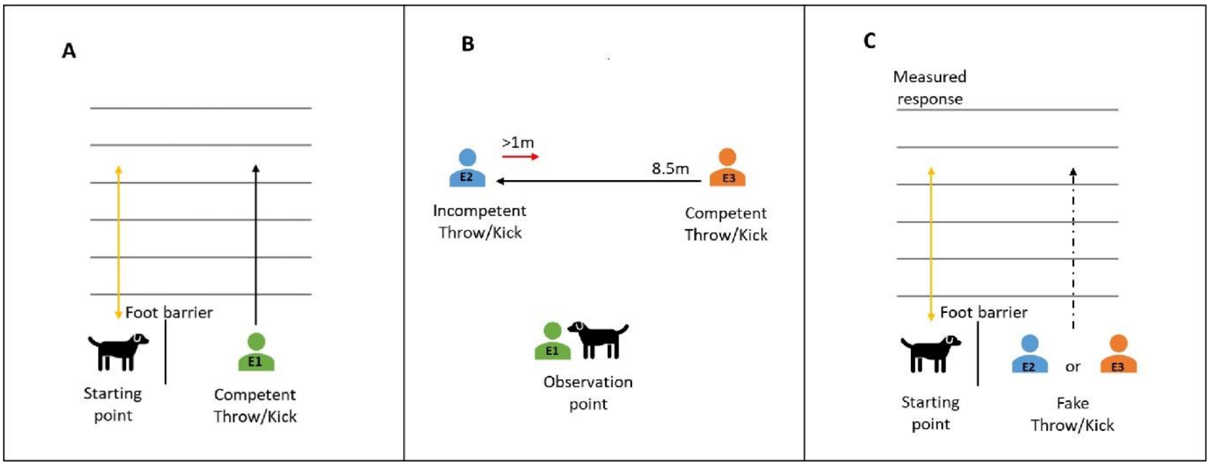

***Fig 1: Summary of Experiment 1 trial phases. A) Pre-test phase**: Dogs were led to the starting point by a handler. E1 then stood beside the dog and threw or kicked the ball normally for the dog (depending on condition). In kicking trials, the foot barrier was present during the test phase. **B) Observation phase**: E1 settled the dog at the observation point, equidistant from experimenters E2 and E3. The dog would watch one experimenter throw or kick the ball competently (propelling it 8.5m across the room)*

and the other experimenter throw/kick the ball incompetently (propelling it <1m across the room). The experimenters would throw the ball between each other twice. ***C) Test phase****: E1 would then return the dog to the original starting point, and depending on condition, either the competent or incompetent thrower or kicker would give a fake throw or kick. At the same time, the fake action was performed, E1 release the dog and the distance it ran in anticipation of the ball was measured.*

Pre-test phase: Dogs were first habituated to the experimental set up. To do this, dogs were led to a predefined starting point by a handler, at which point an experimenter (E1) approached and stood next to the dog. E1 called the dog's name twice then threw or kicked the ball (matching the activity of the condition) at the same time as the handler released the dog. This motivation routine was performed twice before each condition, in order to ensure the dog had an expectation that the ball would be thrown or kicked at this spot, and so keeping it motivated to chase the ball on release across the duration of the experiment.

Observation Phase: After pre-test, dogs were positioned at an observation point, equidistant between two experimenters, E2 and E3, who stood facing each other at a distance of 8.5 meters apart (Fig 2). Experimenters stood on the same side of the dog across all four conditions, with the sides counter-balanced across dogs. Dogs observed these experimenters either throwing (Competent and Incompetent Thrower conditions) or kicking (Competent and Incompetent Kicker conditions) a ball between themselves. During the observation phase, the competent experimenter would always propel the ball to the other experimenter, ensuring that the ball travelled the 8.5m distance between the experimenters. In contrast, the incompetent experimenter would fail to successfully propel the ball to the other experimenter. While there was some variation in how far the incompetent experimenter propelled the ball, it always moved less than 1m towards the other experimenter. Whichever experimenter was the competent thrower in the throwing conditions was the incompetent kicker in the kicking conditions (and vice-a-versa). Dogs only observed E2 and E3 and were never released to chase their throws/kicks. Experimenter identity (E2 / E3) was counterbalanced within activity and across dogs.

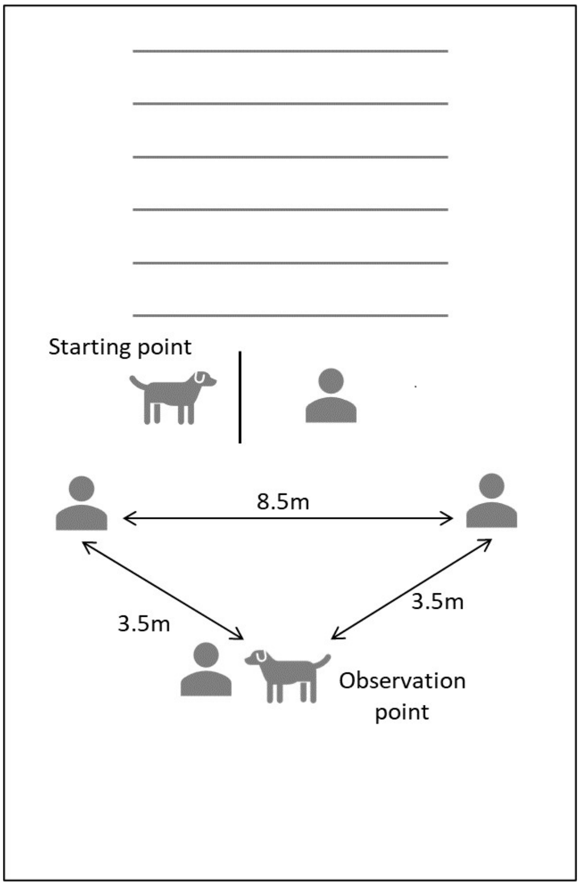

Fig 2: Set up of experimental room. *All phases of the experiment took place in the same room. The observation phase occurred at the bottom of the room and the pre-test and test phases took place approximately 3m away from the area used for the observations.*

Each demonstration consisted of either E2 or E3 calling the dog’s name twice to get their attention, then either throwing or kicking the ball towards the other experimenter. The other experimenter then retrieved the ball and called the dog’s name twice before throwing or kicking the ball back to the first experimenter. This procedure was repeated so that, in a single observation phase, both experimenters would throw/kick the ball twice. Therefore, in a single observation phase, the dog would see the ball thrown or kicked four times in total.

During the demonstration phase, dogs were assumed to be attending to the experimenter if they looked towards the experimenter after two calls. For consistency, the experimenters began their demonstration after two calls. Theoretically, this could have led to some dogs not attending to the experimenter the start of the demonstration. However, in practice, all dogs looked towards the experimenter after the two calls.

Test: After the observation phase was complete, dogs were led back to the starting point and held by the handler. Note the starting point was in a different location of the room, approximately 3 away from the area used for the observation phase. Either E2 or E3 then took the position next to the subject, called the dog’s name twice, and then performed either a fake kick or throw. The dog was released immediately after the fake action.

The fake throw involved the experimenter moving their arm as if to throw the ball but then not releasing the ball, and instead putting it out of sight behind their back. The fake kick involved a small barrier (approx. 40cm x 40cm) being placed on the floor between the experimenter and the dog. The dog then watched the experimenter appear to place the ball on the ground behind the barrier before attempting to kick it. In reality, while the experimenter did place the ball on the ground, they did not make contact with the ball when attempting to kick it.

For each condition, we measured how far the dogs expected the ball to go by measuring how far they ran before stopping. After the dog had stopped running, the demonstrator passed the ball to the handler (E1), who then threw or kicked it for the dog (depending on condition) so dogs did not stop anticipatory running in later trials. Crucially, E2 and E3 never directly threw or kicked the ball for the dog.

Dogs received one trial of each of these four conditions. Dogs took part in only one trial of each condition to minimise the possibility of the dogs learning that the ball was never released during the test phase. This avoided the risk of dogs learning whether the experimenter was willing or unwilling to throw the ball, which could confound the distance that the dogs ran due to their anticipation of how far the experimenter could throw the ball.

The order in which the dog received the throwing and kicking conditions was counterbalanced. However, dogs always received either two throwing condition trials, or two kicking condition trials first, so that the type of actions tested were not mixed, so as to avoid excessive memory constraints. In each trial, the first experimenter to throw or kick the ball in the observation phase was the experimenter who performed the fake throw/kick. This ensured that the last demonstration the dog witnessed was not from the experimenter about to perform the fake throw, reducing the chance that last movement of the ball could explain the dogs' sensitivity to the agents' competence.

Analyses

In Experiment 1, we first constructed Bayesian repeated-measures ANOVA models using three fixed factors: activity, condition, and trial order. The resulting models contained each combination of the factors and the interactions between them (see Table S1 for full details on all 14 models). All models also contained Participant as a random effect to account for the within-subject nature of the experimental design. Each of the models were compared with a simple, null model which just contained Participant as a random factor and we selected the model which best explained the data (i.e., had the highest Bayes Factor (BF)). Each model was constructed with objective priors of prior width r=1 for fixed effects and r=0.5 for random effects. Having established that the Activity+Trial model was the best fitting model to the data, we used Bayesian paired t-tests to investigate the effects of the individual factors (Activity and Condition) on the distance that dogs ran in anticipation of chasing the ball. For both t-tests, objective priors were used, with a positive half-Cauchy distribution (r=0.707) centred on an effect size of zero.

The fake throws and kicks during the test phase were conducted identically between conditions. However, it is possible that subtle differences in fake throwing and kicking behaviours could have affected how far the dogs ran. In order to rule out this possibility, a naïve coder, blind to condition, attempted to categorise clips of the test phase as being part of competent or incompetent trials. If the experimenters' actions were cueing the dog, there should be a non-independent relationship between condition and the competence evaluation. We ran a Bayesian contingency test, using an independent multinomial sampling plan and a prior a=1, to assess the dependence of condition and competence evaluation.

All analyses were carried out using the "Bayes Factor" package in R using the lmBF and ttestBF functions. Distance run was coded from video using lines drawn on the floor at 0.25m intervals, with distance run based on the furthest line crossed by the lead front paw of a dog. The distance run was first coded by an experimenter blind to condition and then re-coded by a volunteer who was blind to hypotheses and condition. For inter-observer reliability we ran an intraclass correlation coefficient (two-way mixed effects, consistency), which indicated there was high agreement between coders (ICC = 0.992; 95%, CI = [0.984,0.996]).

Experiment 1 Results

The best fitting model for the data from Experiment 1 was the Activity + Condition model (repeated measures ANOVA: BF=12712; see Table S3 for all model BFs), indicating that both activity and condition influenced how far dogs ran in anticipation of chasing the ball (Figure 1). This model includes activity and condition as main effects (with no interaction effect) and suggests both factors influenced how far dogs ran in anticipation of chasing the ball in an additive manner. Dogs ran substantially further for throwing trials compared to kicking trials (paired t-test: BF=2,369, Cohen's D=0.76) and also ran substantially further (paired t-test: BF=9.08, Cohen's D=0.29) for the competent experimenter than the incompetent experimenter (competent thrower: mean ± SE: 4.52 ± 0.51m; competent kicker: mean ± SE: 2.69 ± 0.33m; incompetent thrower: mean ± SE: 3.60 ± 0.40; incompetent kicker: mean ± SE: 2.30 ± 0.30m). While the distance that dogs ran in anticipation of chasing the ball was clearly affected primarily by the type of action performed by the experimenter, our results show that, independent of activity, whether the experimenter was competent or incompetent at the action had an additional, additive effect on how far the dogs ran in the experiment.

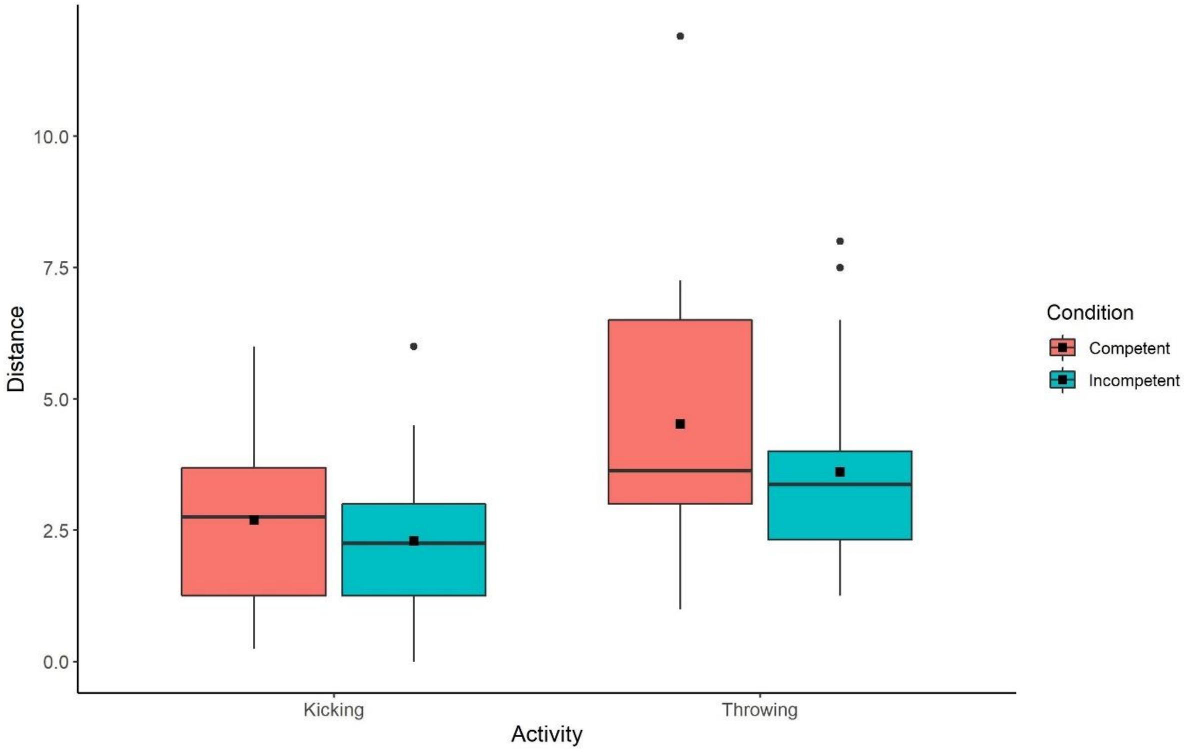

Figure 3: Average distance (meters) that dogs pre-emptively ran in Experiment 1*, following observations of competent and incompetent ball throwers and kickers. In each boxplot, the horizontal bar is the median, the box ranges from the 25th to 75th percentile, and the whiskers show the range of data points, and black squares indicate means.*

There was substantial evidence against the alternative hypothesis of non-independence between the actual condition of the trial and the naïve coder's evaluation (BF=0.27). Therefore, the difference in how far dogs ran in the competent and incompetent trials do not appear to be explainable by subtle difference in the experimenters' throwing and kicking actions.

Experiment 1 Discussion

Dogs ran further when the competent experimenter threw or kicked the ball compared to when the incompetent experimenter threw or kicked the ball. These results are consistent with the hypothesis that dogs are capable of attributing competence to human agents and that they can recognize a human can be competent at one task and incompetent at another.

However, there is a simpler alternative explanation for these results.. Rather than the dogs attributing competence to the experimenter, they may simply associate ball moving further with the sight of the competent thrower or kicker preparing to throw or kick the ball. When the competent experimenter prepares to give the false action, this association may be sufficient to cause the dogs to run further without the dogs having to anticipate that the ball will move further due to the competence of the experimenter.

To rule out this alternative hypothesis, we carried out a second experiment where the experimenters stood beside a ball-throwing machine during an observation phase when it was throwing the ball either a long distance or short distance. In the subsequent test stage, the experimenter gave a fake throw for the dogs in the same way as the throwing trials in Experiment 1.

If dogs' pre-emptive running was driven by them assigning varying levels of competence to these agents, we predicted anticipatory runs would not differ significantly across conditions. In contrast, if pre-emptive running was driven by lower-level associative explanations, then we predicted that the dogs would run further for the agent who launched the ball further.

Experiment 2

Dogs were first given experience with the ball launcher to ensure they were comfortable observing the machine launch a ball. We used the PetSafe Automatic Ball Launcher, a roughly spherical device (approximately 310cm circumference) with a hopper at the top, where tennis balls could be loaded. An experimenter placed a ball in the hopper and then, after a 2 second delay, the device played a sound before an automatic spring loading mechanism propelled the ball across the room, at a 45-degree angle. As in Experiment 1, each experiment trial has three phases (Fig 4). The pre-test phase was the same as Experiment 1, in that dogs were led to the same predefined starting point by a handler, at which point E1 approached and stood next to the dog. E1 called the dog's name twice, then threw the ball at the same time as the handler released the dog. This routine was performed twice before each condition, in order to habituate the dog to the release/ball chasing part of the experiment, and to keep them motivated to chase the ball on release across the duration of the experiment.

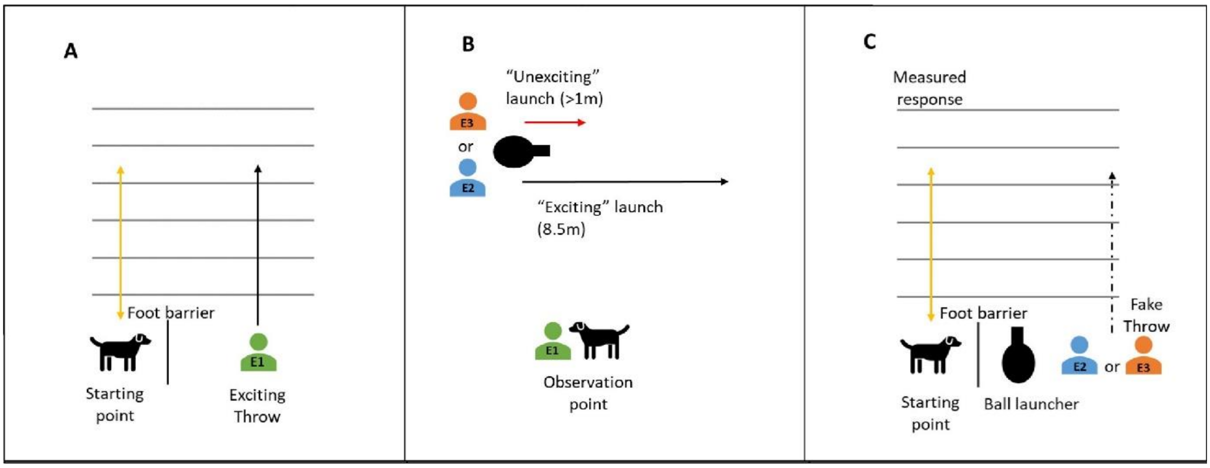

***Fig 4: Summary of Experiment 2 trial phases.* *A) Pre-test phase**: Dogs were led to the starting point by a handler. E1 then stood beside the dog and threw the ball normally for the dog.* ***B) Observation phase**: E1 settled the dog at the observation point. The dog would watch the experimenter place the ball in the ballthrowing machine and the machine either launch the ball a long distance (Long-distnace condition) or short distance (Short-distnace condition). Dogs watched the ball being launched twice in each condition.* ***C) Test phase**: E1 would then return the dog to the original starting point, and depending on condition, with the machine present and experimenter on the other side of the barrier. After a recording of the machine played, the experimenter would give a fake throw. At the same time,*

the fake action was performed, E1 release the dog and the distance it ran in anticipation of the ball was measured.

In the observation phase the dogs witnessed, from the same observation point as in Experiment 1, a demonstrator (E2 or E3) using the ball launcher to propel the ball across the room. Each demonstration consisted of the experimenter holding up a ball, then calling the dog's name to get their attention. Once the dog was attending, the experimenter placed the ball in the launcher. The ball launcher played a tone, then as the ball launched the experimenter called the dog's name twice as the signal for release during the test phase (as in Experiment 1). This procedure was then repeated once more. Dogs were never permitted to chase after the ball during the observation phase. Each condition included one of the two experimenters (E2/E3) who performed either two long-distance launches, or two short-distance launches depending upon the condition. Dogs therefore observed two launches of the same type by the same person before each test. This was designed to reduce memory constraints and amplify any potential effect of association and arousal state which these observations might be thought to confer. In the long-distance condition, the ball launcher launched the ball 8.5m, matching the competent thrower and kicker distances from Experiment 1. In the short-distance condition the ball launched less than 1 meter, matching the distance thrown and kicked in the incompetent conditions of Experiment 1. The order in which the dog received each condition was counterbalanced between subjects.

At test, the demonstrator took the ball launcher and positioned it next to themselves at the point where the motivating throws had previously occurred. As in Experiment 1, the dog was led to the starting point, and held in place by the handler. The dog then observed an identical sequence of events to what they had previously observed: the experimenter called the dog's name once, holding up the ball to get their attention. A speaker hidden behind the ball launcher then played the usual ball launching noise followed by the experimenter calling the dog’s name twice in quick succession (as they had during the observed launches) and performing a fake throw. The distance the dog ran before stopping was again taken as measure of how far they predicted the ball would go. If dogs had simply associated competent agents with the ball moving further during the demonstration phase, we predicted that dogs would run further in the long-distance condition at test, because all the cues associated with the ball being thrown far were present (the presence of the experimenter and launcher, the noise of the launcher and the experimenters calling the dog’s name). Experimenter identity (E2 / E3) was counterbalanced within activity and across dogs. Similarly to Experiment 1, dogs received one trial in each of the two conditions.

Analyses

In Experiment 2, we constructed Bayesian repeated-measures ANOVA models using two fixed factors: condition and trial order. Similarly to Experiment 1, the four resulting models (Condition-Only; Trial-Only; Condition+Trial; Condition*Trial) also included Participant as a random effect to account for the within-subject nature of the experimental design. All four models were compared with a null model, with only Participant as a random effect. For all analyses, Bayes factors >3 indicate substantial support for the alternative hypothesis, whilst Bayes Factors <0.333 indicate substantial support for the null hypothesis.

Experiment 2 Results

Unlike our results in Experiment 1, none of the four models were a better fit to the data for Experiment 2 than the null model. Crucially, all three models containing condition were substantially worse fits to the data (Condition: BF=0.300; Trial: BF=0.377; Condition+Trial: BF=0.12; Condition*Trial: BF=0.097), showing that condition had no effect on how far dogs ran in anticipation of chasing the ball (Figure 2). Dogs ran similar distances (Cohen’s D=0.06) in both the long-distance condition (mean ± SE: 3.70 ± 0.50m) and the short-distance condition (mean ± SE: 3.57 ± 0.46; see fig 2), providing strong evidence against the possibility that dogs' running patterns can be explained in terms of lower-level arousal cues.

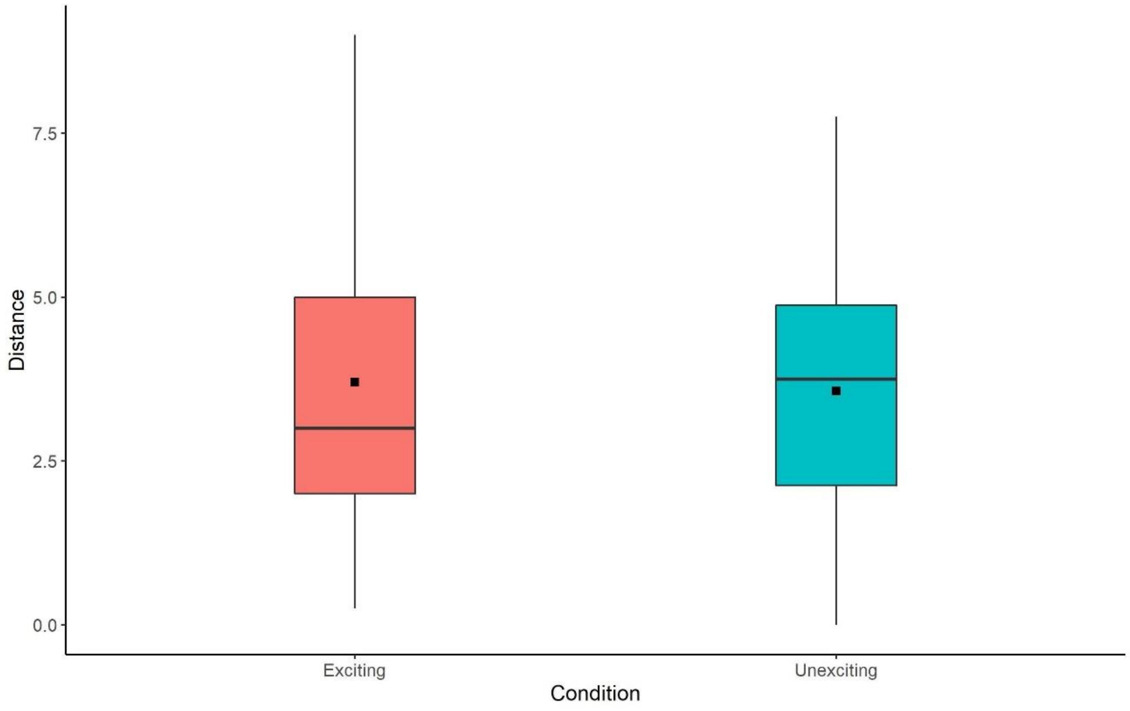

Figure 5. Average distance (meters) that dogs pre-emptively ran in Experiment 2*, following observations of a ball being launched either over a long distance or a short distance. In each boxplot, the horizontal bar is the median, the box ranges from the 25th to 75th percentile, and the whiskers show the range of data points Black squares indicate means.*

Discussion

Our results show that dogs pre-emptively ran further after either a fake kick or throw if they had previously observed the agent responsible throwing or kicking a ball competently, compared to a second agent who threw or kicked the ball incompetently. This was despite each agent: (i) performing the same actions at test (a fake kick or fake throw), (ii) demonstrating competence at a different location from where the test action was performed; and (iii) demonstrating only differences in relative performance (i.e., the distance they were able to move the ball) during the observation phase. To react in this way, dogs had to assign competence to the experimenters based on third-party observations of the individuals interacting with the ball alone rather than based on direct interactions with the experimenters. That is, dogs did not directly interact with the two experimenters or the ball during the observation stage. Therefore, there was no opportunity for the dogs' anticipatory running behaviour to be directly shaped or reinforced. Furthermore, the dogs had to attribute action-specific competencies to the same agent, as good throwers were bad kickers and vice versa. Thus, dogs needed to assign and maintain two behavior-specific values of competence to each agent, rather than assigning only one value based on their first observation of the agent (e.g. competent vs incompetent), or averaging across experiences (e.g. observing a good kick in one condition and then a bad throw in the next and then predicting an average movement of the ball when observing a fake throw).

It is notable that there was individual variation in how far the dogs ran in response to the fake throws, and the difference between how far individual dogs ran across conditions (See Supp Fig 1 and 2). Therefore, it is possible that dogs vary in how well they can attribute behaviour-specific competencies. Future research, using a larger sample of dogs and with more trials per dog, could explore individual differences in competence attribution across dogs.

As well as finding that dogs ran further for the competent experimenter, our results showed that dogs ran further in throwing trials than in kicking trials. This may be due to dog owners being more likely to throw rather than kick the ball. Similarly, the set-up with the barrier in kicking trials may have affected the dogs' behaviour. Dogs do show object permanence (Zentall & Pattison, 2016) and so should be aware that the ball is still present behind the barrier. However, the barrier and unfamiliarity of the set-up may still increase cognitive load and reduce the distance that the dogs run. Critically, however, including the competence-activity interaction term substantially reduced model fit and so this unfamiliarity appears to have not affected the dogs' assessment of competence.

We also found clear evidence against the hypothesis that dogs simply associated one experimenter with the ball moving further in a particular set of trials. This association hypothesis predicts that dogs should run further for the experimenter in closest proximity to the ball when it is moved a longer distance. However, in Experiment 2 we found no difference in running distance in how far dogs ran in response to fake throws from two experimenters, after observing them move a ball through the use of a ball launcher. Dogs observed one experimenter using the ball launcher to launch the ball a very short distance, while the other was observed using the launcher to send the ball far across the room. At test, although all the auditory and visual cues associated with these observations were present, dogs did not run different distances when the experimenter performed a fake throw. This suggests that rather than simply associating the experimenters with how far the ball moves, the dogs appear to be sensitive to the actions of the competent and incompetent agents. Additionally, differences in the fake actions of the experimenters in the test phase do not appear to explain differences in how far the dog ran as the actions were indistinguishable to a blind coder.

These results demonstrate that, for at least throwing and kicking behaviours, dogs exhibit four behavioural signatures of human competence attribution. As in humans, dogs took third-party observations (Signature 1) of relative competence (Signature 2) and then used them to predict future behaviour differentially, even though the actual behaviour produced at test (a fake throw or kick) was identical (Signature 3) and required assigning action-specific competencies to the same agent (Signature 4).

However, in humans, competence attribution goes beyond simply learning that an agent is more capable at a particular task than another agent. Instead, humans show some understanding of which properties of the agent make it more competent at the task. For example, work with infants showed that they understood that an agent being tall made them competent at lifting objects from a high shelf but incompetent at fitting though a small door (Paulus & Moore, 2011). While both experiment 1 and 2 suggest that the dogs in our current study showed an understanding that the experimenters differed in how far they could propel the ball using different actions, it is not clear whether dogs understand these task-specific competencies as being underpinned by more general skills and abilities of the experimenters.

Rather, instead of understanding that an experimenter is good or bad at throwing or kicking, the dogs may simply be recognizing that the ball moves further when one experimenter throws the ball, or the other experimenter kicks the ball. Whilst being able to make that distinction would meet the criteria for minimal task-specific competence attribution, it does not require any understanding of the properties underpinning competence at a particular task. Instead, despite there being no differential reinforcement of dogs' anticipatory running during this study, such minimal competence attribution could also be explained by conditional differential reinforcement of the dogs' running behaviour in the past.

As such, while our current results show that dogs show the behavioural signatures required for minimal task-specific competence attribution, future work is required to determine whether dogs are capable of more sophisticated competence attribution and to investigate the role of conditional reinforcement in dogs' competence attribution. The competence attribution hypothesis predicts that dogs should be able to generalize an agent's competence at one task to another task if similar skills and traits are required to succeed in both tasks. For example, as arm strength is important for both throwing and opening jars, we might predict that a good thrower might be more likely to be good at opening jars or lifting heavy objects as well. Future research should aim to carry out transfer tasks to see if dogs transfer competence across tasks within the same skill domain but not tasks which lie in different skill domains.

These findings also open up a number of other research questions. First, it is not yet clear whether the competence attribution demonstrated by the dogs in our study might extend to other human and animal behaviours. Given the potential fitness benefits of assessing the behaviours of conspecifics when group hunting, or more generally of assessing key behaviours of predators and prey, it seems possible that competence attribution would have been an evolutionary advantage for dogs' ancestors. Whilst modern dogs experience a very different lifestyle, assessing competence may still improve fitness in ecologically relevant situations as well. In particular, given their close relationship with humans, it could be advantageous for dogs to be capable of more fine-grained assessments of human competence. Exploring this potential capacity may have important applied applications in areas such as training dogs to work with people that are disabled. Understanding more about how best to demonstrate human competence in different behaviours to dogs, perhaps through demonstrations of relative competence as in our study, offers a potential line of future enquiry.

There is increasing evidence that, even as infants, humans' intuitions about agent's goals and competencies align with Naïve Utility Calculus; a model which assumes agents act to optimise their utility by maximizing benefits and minimizing costs (Jara-Ettinger et al., 2016). A sensitivity for the costs, rewards, and goals of an agent's actions can lead to inferences about that agent's competence at a specific task. For example, if an observer knows that an agent places high value on a reward but does not attempt to obtain it when given the chance, they may infer that the agent is not competent enough to obtain the reward without incurring a substantial cost (Jara-Ettinger et al., 2016, 2020; Leonard et al., 2019). It is possible that dogs may use Naïve Utility Calculus in a similar way to make inferences about agents' competencies. This could be further explored by increasing the complexity of the cognitive attributions dogs need to do, such as by manipulating task difficulty and agent speed (Leonard et al., 2019). We hope that future research focused on addressing these questions in dogs and other species can provide insight into this under-studied area of social evaluation. In particular, a better understanding of whether dogs and other species are capable of competence contribution and the cognitive models behind such attribution could provide a better understanding of the cognitive and neurological underpinnings of competence attribution in humans.

Competing Interests

The authors declare no competing interests.

Funding

AHT was supported by a Rutherford Discovery Fellowship and a Prime Minister's McDiarmid Emerging Scientist Prize from the Royal Society of New Zealand.

Author's Contributions

All authors designed the experiments; RSH, PN, and APMB collected the data; PN carried out statistical analyses; all authors were involved in writing the manuscript.

Data Statement

All data used in our analysis is attached to this paper as supplementary materials. The R code used to analysis the data is also attached to this paper as supplementary materials. Demonstration of the method can be found as a video in the supplementary materials.

Acknowledgements

We thank the owners and their dogs for giving their time to contribute to our research. We also thank Ivy Ren, Olivia Kingsley-Smith, Frederique Derks, and Byung Chan Lim, and the Clever

Canine Lab volunteers for assisting with data collection, and Patrick Wood for blind-coding the video data.

References

Abril-de-Abreu, R., Cruz, J., & Oliveira, R. F. (2015). Social Eavesdropping in Zebrafish: Tuning of Attention to Social Interactions. *Scientific Reports*, *5*(1), 12678. https://doi.org/10.1038/srep12678

Anderson, J. R., Kuroshima, H., Takimoto, A., & Fujita, K. (2013). Third-party social evaluation of humans by monkeys. *Nature Communications*, *4*(1), 1561. https://doi.org/10.1038/ncomms2495

Anderson, J. R., Takimoto, A., Kuroshima, H., & Fujita, K. (2013). Capuchin monkeys judge third-party reciprocity. *Cognition*, *127*(1), 140–146. https://doi.org/10.1016/j.cognition.2012.12.007

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, *441*(7096), 975–978. https://doi.org/10.1038/nature04755

Buon, M., Jacob, P., Margules, S., Brunet, I., Dutat, M., Cabrol, D., & Dupoux, E. (2014). Friend or Foe? Early Social Evaluation of Human Interactions. *PLOS ONE*, *9*(2), e88612. https://doi.org/10.1371/journal.pone.0088612

Chijiiwa, H., Kuroshima, H., Hori, Y., Anderson, J. R., & Fujita, K. (2015). Dogs avoid people who behave negatively to their owner: Third-party affective evaluation. *Animal Behaviour*, *106*, 123–127. https://doi.org/10.1016/j.anbehav.2015.05.018

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*(5), 426–435. https://doi.org/10.1038/s41562-019-0590-x

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Research*, *1079*(1), 36–46. https://doi.org/10.1016/j.brainres.2005.12.126

Hamlin, J. K. (2013a). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*(3), 451–474. https://doi.org/10.1016/j.cognition.2013.04.004

Hamlin, J. K. (2013b). Moral Judgment and Action in Preverbal Infants and Toddlers: Evidence for an Innate Moral Core. *Current Directions in Psychological Science*, *22*(3), 186–193.

Hamlin, J. K. (2015). The case for social evaluation in preverbal infants: Gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/article/10.3389/fpsyg.2014.01563

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. https://doi.org/10.1016/j.cogdev.2010.09.001

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. https://doi.org/10.1038/nature06288

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). 3-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*(6), 923–929. https://doi.org/10.1111/j.1467-7687.2010.00951.x

Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences*, *108*(50), 19931–19936. https://doi.org/10.1073/pnas.1110306108

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The Domestication of Social Cognition in Dogs. *Science, 298*(5998), 1634–1636. https://doi.org/10.1126/science.1072702

Herrmann, E., Keupp, S., Hare, B., Vaish, A., & Tomasello, M. (2013). Direct and indirect reputation formation in nonhuman great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus) and human children (Homo sapiens). *Journal of Comparative Psychology*, *127*(1), 63–75. https://doi.org/10.1037/a0028929

Horn, L., Virányi, Z., Miklósi, Á., Huber, L., & Range, F. (2012). Domestic dogs (Canis familiaris) flexibly adjust their human-directed behavior to the actions of their human partners in a problem situation. *Animal Cognition*, *15*(1), 57–71. https://doi.org/10.1007/s10071-011-0432-3

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604. https://doi.org/10.1016/j.tics.2016.05.011

Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The Naïve Utility Calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology*, *123*, 101334. https://doi.org/10.1016/j.cogpsych.2020.101334

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not So Innocent: Toddlers' Inferences About Costs and Culpability. *Psychological Science*, *26*(5), 633–640. https://doi.org/10.1177/0956797615572806

Jim, H.-L., Marshall-Pescini, S., & Range, F. (2020). Do dogs eavesdrop on human interactions in a helping situation? *PLOS ONE*, *15*(8), e0237373. https://doi.org/10.1371/journal.pone.0237373

Krupenye, C., & Hare, B. (2018). Bonobos Prefer Individuals that Hinder Others over Those that Help. *Current Biology*, *28*(2), 280-286.e5. https://doi.org/10.1016/j.cub.2017.11.061

Lazzaroni, M., Marshall-Pescini, S., Manzenreiter, H., Gosch, S., Přibilová, L., Darc, L., McGetrick, J., & Range, F. (2020). Why do dogs look back at the human in an impossible task? Looking back behaviour may be over-interpreted. *Animal Cognition*, *23*(3), 427–441. https://doi.org/10.1007/s10071-020-01345-8

Leonard, J. A., Bennet-Pierre, G., & Gweon, H. (2019). Who is better? Preschoolers infer relative competence based on efficiency of process and quality of outcome. *CogSci*, 639–645.

MacLean, E. L., Herrmann, E., Suchindran, S., & Hare, B. (2017). Individual differences in cooperative communicative skills are more similar between dogs and humans than chimpanzees. *Animal Behaviour*, *126,* 41–51. https://doi.org/10.1016/j.anbehav.2017.01.005

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees Recruit the Best Collaborators. *Science*, *311*(5765), 1297–1300.

Miklósi, Á., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z., & Csányi, V. (2003). A Simple Reason for a Big Difference: Wolves Do Not Look Back at Humans, but Dogs Do. *Current Biology*, *13*(9), 763–766. https://doi.org/10.1016/S0960-9822(03)00263-X

Miklósi, Á., Topál, J., & Csányi, V. (2004). Comparative social cognition: What can dogs teach us? *Animal Behaviour*, *67*(6), 995–1004. https://doi.org/10.1016/j.anbehav.2003.10.008

Paquette-Smith, M., & Johnson, E. K. (2015). I Don’t Like the Tone of Your Voice: Infants Use Vocal Affect to Socially Evaluate Others—Paquette-Smith—2016—Infancy—Wiley Online Library. *Infancy*, *21*(1), 104–121.

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Paulus, M., & Moore, C. (2011). Whom to ask for help? Children’s developing understanding of other people’s action capabilities. *Experimental Brain Research, 211*(3), 593. https://doi.org/10.1007/s00221-011-2676-1

Piotti, P., Spooner, R. M., Jim, H.-L., & Kaminski, J. (2017). Who to ask for help? Do dogs form an opinion on humans based on skilfulness? *Applied Animal Behaviour Science*, *195*, 93–102. https://doi.org/10.1016/j.applanim.2017.05.024

Russell, Y. I., Call, J., & Dunbar, R. I. M. (2008). Image scoring in great apes. *Behavioural Processes*, *78*(1), 108–111. https://doi.org/10.1016/j.beproc.2007.10.009

Sciutti, A., Ansuini, C., Becchio, C., & Sandini, G. (2015). Investigating the ability to read others’ intentions using humanoid robots. *Frontiers in Psychology, 6*. https://www.frontiersin.org/article/10.3389/fpsyg.2015.01362

Surian, L., & Franchin, L. (2017). Infants reason about deserving agents: A test with distributive actions. *Cognitive Development, 44*, 49–56. https://doi.org/10.1016/j.cogdev.2017.08.009

Taylor, A. H. (2014). Corvid cognition. *WIREs Cognitive Science*, *5*(3), 361–372. https://doi.org/10.1002/wcs.1286

Trösch, M., Pellon, S., Cuzol, F., Parias, C., Nowak, R., Calandreau, L., & Lansade, L. (2020). Horses feel emotions when they watch positive and negative horse–human interactions in a video and transpose what they saw to real life. *Animal Cognition*, *23*(4), 643–653. https://doi.org/10.1007/s10071-020-01369-0

Vail, A. L., Manica, A., & Bshary, R. (2014). Fish choose appropriately when and with whom to collaborate. *Current Biology*, *24*(17), R791–R793. https://doi.org/10.1016/j.cub.2014.07.033

Wocjciszke, B., Bazinska, R., & Jaworski, M. (1998). On the Dominance of Moral Categories in Impression Formation. *Personality and Social Psychology Bulletin*, *24*(12), 1251–1263.

Zentall, T. R., & Pattison, K. F. (2016). Now You See It, Now You Don’t: Object Permanence in Dogs. *Current Directions in Psychological Science*, *25*(5), 357–362.

643

View publication stats