



Exploring the use of projected videos to test action matching from different perspectives in dogs

Claudia Fugazza^{1,2} · Fumi Higaki¹

Received: 15 December 2023 / Accepted: 18 April 2024
© The Author(s) 2024

Abstract

Dogs trained with the Do as I Do method can imitate human actions upon request, but their ability to match actions observed from different perspectives remains unknown. The use of 2D video stimuli may enable researchers to systematically manipulate the perspective from which demonstrations are observed, thereby widening the range of methods available to study cognitive skills related to imitation. In this study, we explore the possibility of using 2D stimuli to test action matching in dogs, including when demonstrations are seen from different perspectives. We examined two dogs' imitative performance using videos projected on a screen; while, the owner interacted with the dog remotely through an online meeting software. The dogs were first trained to match human actions seen on a screen frontally, and then were tested when the projected demonstrations were seen frontally, from the side, and from above. Results revealed that both dogs matched the demonstrated actions from frontal and, notably, also from side perspectives, at least to some extent, consistent with familiarity of their daily interactions with humans. However, action matching from an above perspective presented challenges, indicating the potential influence of observational experience and highlighting the importance of perspective manipulation when investigating imitation abilities. These findings show that it is possible to use 2D videos to test imitation in dogs, thereby expanding the potential methodologies to study imitation and other related cognitive skills.

Keywords Dog · Imitation · Perspective taking · Projected videos · 2D stimuli

Introduction

Dogs trained with the Do as I Do method can reproduce human actions upon request (Topál et al. 2006; Huber et al. 2009). Imitative skills can be used to test several cognitive abilities that would be otherwise elusive in non-human species, such as mapping others' body parts into one's own, representation of others' goals, and different forms of memory (e.g., Fugazza et al. 2016, 2019). Video projected stimuli are widely used in cognitive studies, not only with human subjects but also with non-human animals (e.g., Myowa-Yamakoshi et al. 2012), and their use is increasing also in cognitive research with dogs (e.g., Pongrácz et al. 2018). The main advantage of the use of 2D stimuli is that they

can be experimentally manipulated to expose the subjects to conditions that would otherwise be difficult to present or systematically modify.

Dogs were shown to be able to use pre-recorded videos as a source of information (Péter et al. 2013) but familiarity with the depicted scene affected the way they process it (Pongrácz et al. 2018). Infants are known to also imitate televised models (e.g., Seehagen & Herbert 2011; Sommer et al. 2021) although there is a decrease in their imitative performance compared to live models (e.g., Barr and Hayne 1999; Barr 2010). This decrease has been suggested to arise from the perceptual difficulty of transferring 2D information to a three-dimensional real-world setting (Barr 2010). Chimpanzees were shown to successfully learn how to solve a task by observing video demonstrations, even though, also in the case of this species, there seems to be a reduced action matching accuracy from video demonstrations, compared to real life models (Hopper et al. 2012). Some evidence of social learning from video demonstration has been provided for other species as well (e.g., monkeys: Price and Caldwell

✉ Claudia Fugazza
claudia.happydog@gmail.com

¹ Department of Ethology, Eötvös Loránd University, Budapest, Hungary

² ELTE NAP Comparative Ethology Research Group, Budapest, Hungary

2007; birds: Hämäläinen et al. 2019; Guillette and Healy 2019).

To imitate, the observer needs to encode the model's actions with reference to one's role or perspective (Carpenter et al. 2005). Therefore, it has been proposed that there is a relationship between imitation and perspective-taking (e.g., Santiesteban et al. 2012). For example, it has been suggested that the difficulties of individuals with autism in imitation tasks are due to a reduced perspective-taking capacity (Meyer and Hobson 2004; Yu et al. 2011).

Whether non-human animals are able to take the perspective of another individual is a topic that has received a lot of attention in comparative cognition research. The topic is typically investigated by applying the Guesser–Knower task (Povinelli et al. 1990), in which successful subjects choose to follow the pointing of a human who witnessed a food hiding event, over that of a human who did not. Povinelli et al. (1990) showed that a few individual chimpanzees could succeed in this task after extensive experience, but subsequent studies mostly failed to prove that they consistently preferred to follow the cues of the “Knower” over those of the “Guesser” (e.g., Povinelli et al. 1994), also in other species (e.g., pigs: Held et al. 2001). In some cases, subjects only reached some success after training (e.g., capuchin monkeys: Kuroshima et al. 2002). Since the test requires subjects to respond on the basis of cues provided by human informants, it has been suggested that these failures may be due to the tested species lacking extensive interaction with humans (Maginnity and Grace 2014). Dogs may be at some advantage in this sense, thanks to their evolution and development in the human environment. This species, indeed, showed some success in the Guesser–Knower task, suggesting that geometrical gaze following and perspective-taking may be within dogs' cognitive abilities (Catala et al. 2017). However, this topic remains controversial (Huber and Lonardo 2023), and more solid methods would be needed to gain a clearer understanding of the cognitive processes involved in dogs' success in solving tasks relying on perspective-taking (Abdai and Miklósi 2023).

For both, perspective-taking and imitation, the viewpoint from which the demonstration is observed (visual perspective) may be crucial as it may affect the successful mapping of the model's actions into one's own. Using video projected demonstrations may constitute a promising method to test how visual perspective affects the representation of others' actions in different species, because it provides the possibility to systematically manipulate the perspective from which the demonstration is observed. Hopper et al. (2012) showed a model seen from different perspectives in the video demonstration observed by chimpanzees. However, the demonstrations were all presented in a single video, shown to the chimpanzees before they had a chance to try to solve the demonstrated task. Thus, it is not possible to disentangle the

effect of demonstrations seen from different perspectives. In the present exploratory study, we aimed at experimenting the possibility of using 2D videos in the Do as I Do paradigm, also including conditions in which the observers' perspective was experimentally manipulated. We tested two dogs, previously trained with the Do as I Do method to match human actions on command “Do it!” based on Fugazza and Miklósi (2014). Before the study began, the dogs were also trained to imitate 6 actions projected frontally on a screen, using an online meeting software. We aimed at exploring the capacity of dogs to match human actions seen on a 2D screen when the owner was frontal to the camera, and also when the demonstrated actions were seen from different perspectives: from the side and from above.

Based on dogs' skills in imitating human actions (e.g., Topál et al. 2006) and on the capacity of dogs to use projected images to solve other tasks (e.g., Péter et al. 2013), we expected the training to be successful—i.e., we expected that dogs imitated the trained actions when the owner was frontal to the camera, at least to some extent. By interacting with humans on a daily basis, dogs gain extensive experience of observing humans' actions from lateral perspectives too. Thus, we expected dogs to succeed also in the condition where they observed the demonstrations from the side, despite not being specifically trained for this. In contrast, dogs may not typically have opportunities to observe humans from above. Thus, we expected dogs to be able to generalize the imitation of the demonstrated actions to when the actions were shown from above only to a lesser extent.

Methods

Subjects

We tested $N=2$ dogs, Tara, a 2-year-old male Golden retriever, and Franc, a 6-year-old female Labrador, both previously trained by their owners to imitate human actions with the Do as I Do method, as described in detail in Fugazza and Miklósi (2014).

Setup

During the tests, the owner and the dog were in two separate rooms (owner's room and dog's room), on two different floors of the same building. The dog was with a familiar helper. The owner was connected to an online software (Zoom), which live-streamed her on a PC in the dog's room. This PC was connected to a projector that projected the images on a screen (2214 mm wide, 1245 mm tall). Two loudspeakers, also connected to the PC, were placed behind the screen, in the center of it. A mat, on which the dogs were previously trained to sit and pay attention to the screen, was

placed in front of the screen, at 1.5 m from it for Franc and at 2 m for Tara (this positioning took into account the dogs' height difference and thus prevented the dogs' head shadows from being cast on the projector screen). The computer in the dog's room was placed on the floor, in front of the screen, in the center of it, so that its camera would record and livestream the dog's behavior to the owner's device in the owner's room during the tests. The screen of the PC in the dog's room was obscured with a cardboard sheet, leaving only the camera out, so that only the projected images, but not the computer screen, were visible for the dog. A remotely controlled food dispenser (Treat & Train™: Remote Reward Dog Trainer produced by PetSafe®) was placed on the floor, in front of the screen, next to the computer (Fig. 1). The Experimenter sat approximately 3 m behind the dog and communicated with the owner, instructing her through the Zoom call. The Experimenter did not interact with the dog.

Training to imitate video projected human actions

Before the actual tests, the dogs received additional training to imitate 6 actions, seen on a projector screen when the actions were demonstrated by their owners who were connected to a livestream software and were standing frontally in front of the camera (see Supplementary file S1

for a detailed description of the training and Table S2 for detailed report of the number and content of each training session for each dog). The setup for the training was the same as the one described above. The success of the training was assessed during the tests, in the Frontal camera condition (see below).

Testing

The dogs participated in three testing conditions in which they were tested on their ability to imitate actions projected on a screen, and a Baseline condition, to assess their capacity to imitate the actions used in the tests when no projector was used—i.e., when the actions were demonstrated by their owner who was standing in front of them.

The conditions were administered in the order described below. The order of the six familiar actions demonstrated in each trial within condition was randomized within trials 1–6 and trials 7–12.

Frontal camera condition: the owner was facing the camera of her device, so that she was visible from the front in the projected video, similar to during the training (Fig. 2A). This condition also served to assess the success of the training.

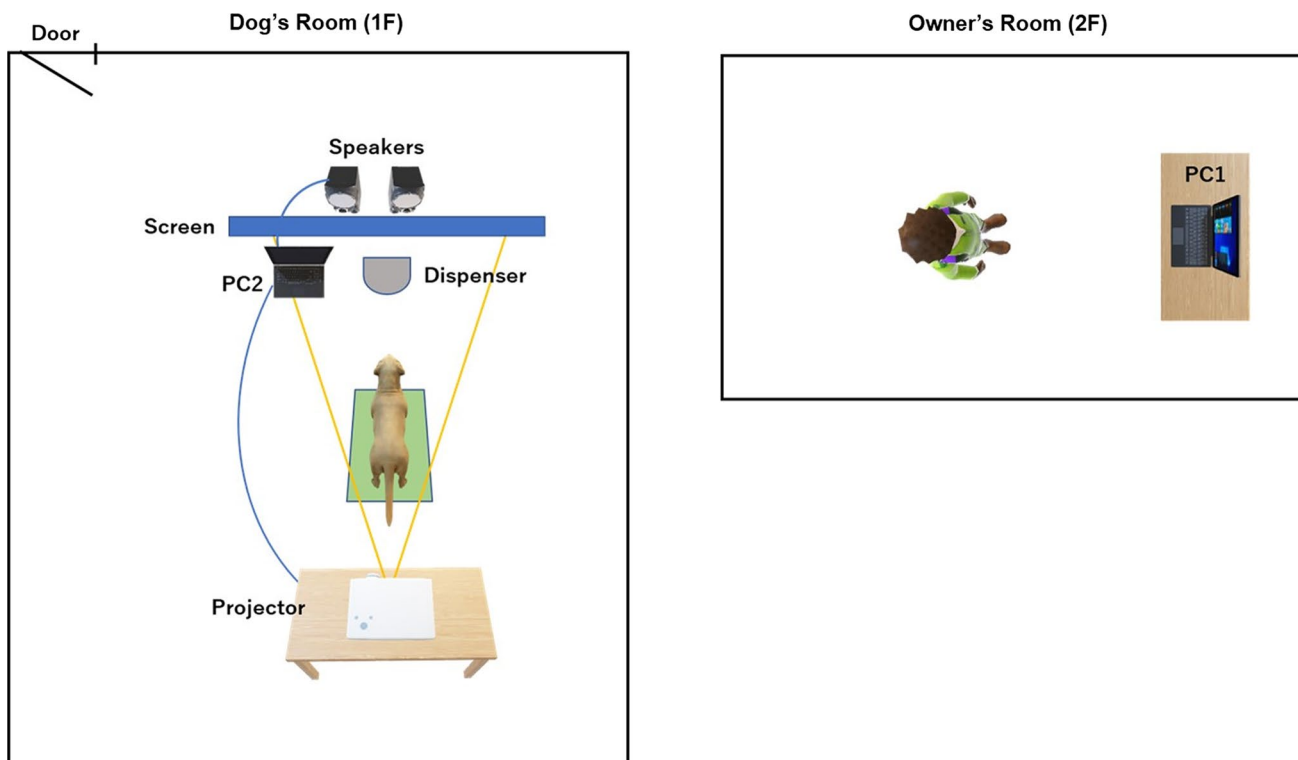


Fig. 1 Setup for the experiment



Fig. 2 Examples of one of the demonstrations, cross one arm, seen on the screen from different perspectives (**A** Frontal; **B** Lateral; **C** From above) and **D** one of the subjects, Tara, looking at the projected demonstration as seen from the screen of the owner's device

Side camera condition: the owner was turned 90 degrees to the side from the camera of her device, so that she was visible from the side in the projected video (Fig. 2B).

Above camera condition: the owner's device was attached to the ceiling, exactly above her, so that she was visible from above in the projected video (Fig. 2C).

Baseline: only the dog's room was used in this condition, as the owner demonstrated the actions live, in front of the dog. During this test, the dog sat on the mat as in the other conditions, and the owner stood in front of him/her, with the screen on her back. The projector and loudspeaker were turned off.

Detailed procedure

At the beginning of each trial, the owner attracted the dog's attention (from the screen and loudspeaker in the experimental conditions with projector, and live in the Baseline) and asked the dog to sit in his starting position on the mat, using signals known by the dog. In the conditions with projector, if necessary, a helper led the dog to his starting sitting position on the mat, pointed to the screen to ensure that the dog paid attention to its owner projected on it, then walked away, approximately 6 m to the side.

The owner demonstrated the randomly chosen action (based on a predetermined randomized schedule), and gave the "Do it!" command. If the dog turned his head away from the screen during the demonstration, or went away from the starting position, the Experimenter instructed the owner to ask the dog to sit again in his starting position and pay attention and the demonstration was repeated. After the "Do it!" command was given, the dog typically performed an action and was rewarded with food from the remotely controlled food dispenser activated by the Experimenter, if the action matched the demonstrated one. If the dog performed a different action, the same demonstration was repeated one more time, but the repeated trials were not included in the statistical analysis. After another mistake or after the dog was rewarded with food for performing a matching action, the owner asked him/her to sit again in the starting position (eventually with the assistance of the helper), and the next trial started.

In total, 6 familiar (i.e., trained) actions and 3 novel actions were demonstrated twice in a random order in each condition, so that in each condition, 12 trials with six familiar actions and two trials with the novel action were carried out. To avoid the novel actions becoming familiar with repeated trials, in each condition, only one of the novel actions was included and different ones were

demonstrated in the other conditions. The following actions were demonstrated: spin, down, paw cross, walk back, jump back, nose touch, press, pick up, drop, and house (see Table 1 for detailed description of the demonstrated actions and dog actions considered as matching those).

Statistical analysis

To explore how the performance of dogs in the conditions with projected demonstrations lagged from their imitative performance with live demonstrations, we compared each dog’s performance in the conditions with projector, to their performance in the Baseline, using Fisher Exact test.

In addition, in the case of familiar actions being demonstrated, we conducted separate Binomial tests for each condition and each dog to determine if the dogs' responses matched the demonstrations above chance level. The repeated trials that were carried out in case of the dogs’ mistake were not included in the analysis. We set chance level at 0.17 because there were 6 different actions on which the dogs were trained and that were demonstrated in the tests. We note that this is a conservative analysis because the dogs could potentially perform any other behavior, not only the trained/demonstrated actions (so that the applied 0.17 chance level is an upper limit).

In the case of novel actions, in the different perspective conditions, since only one novel action was demonstrated in each condition, we report whether the dogs performed an action that matched the demonstrated one.

Results

Comparison of performance in conditions with projector to baseline condition

The proportion of successful trials of the dogs in each condition with projector was significantly lower than in the Baseline, with the exception of Tara’s performance in the Side camera condition (Fisher Exact test, Frontal camera condition, Tara: $P=0.037$; Franc: $P=0.005$; Side camera condition, Tara $P=0.217$; Franc $P=0.014$; Above camera condition, Tara $P=0.001$; Franc $P=0.005$).

Frontal camera condition

Tara matched the familiar actions in 7 out of 12 trials. Thus, the dog matched the projected human demonstrations above chance (Binomial probability $P=0.001$).

Tara did not match the novel action demonstrated of touching a foam cylinder with a nose in the first trial. Instead, he stood up and went to the food dispenser. When the demonstration was repeated, Tara matched the demonstrated action.

Franc matched familiar actions in 5 out of 12 trials in which familiar actions were demonstrated. Thus, the dog matched the projected human demonstrations above chance (Binomial probability $P=0.03$).

Franc did not match the novel action demonstrated of picking up a toy neither in the first trial, nor in the second when this action was demonstrated again. Instead, she lied down in both trials.

Table 1 The table reports the actions included in the tests, the description of their demonstration, the description of the dog's behaviour considered as matching and reports whether the actions were novel for the dogs

Tara		Franc		Action	Description of human demonstration	Description of dog's action considered as matching
Trained	Novel	Trained	Novel			
✓		✓		Spin	The handler lays its body horizontally on the floor, with her chest downward	The dog lays its body horizontally on the floor, with his chest on the floor
✓		✓		Down	The handler turns its body around the vertical axis	The dog turns its body around the vertical axis
✓		✓		Paw cross	The handler raises one arm, placing it in front of the other, across its chest	The dog raises one paw, across its chest
			✓	Walk Back	The handler walks backward	The dog walks backward
✓			✓	Jump Back	The handler jumps backward	The dog jumps backward
	✓	✓		Nose Touch	The handler touches a cylinder made of foam placed on the table with its nose	The dog touches a cylinder made of foam placed on the floor with its nose
✓		✓		Press	The handler presses a button on the floor to emit a buzzer sound	The dog presses a button on the floor to emit a buzzer sound
✓			✓	Pick up	The handler picks up a toy	The dog picks up a toy
	✓		✓	Drop	The handler knocks the bottle down	The dog knocks the bottle down
	✓		✓	House	The handler goes into the crate	The dog goes into the crate

Side camera condition

Tara matched familiar actions in 9 out of 12 trials in which familiar actions were demonstrated. Thus, the dog matched the projected human demonstrations above chance (Binomial probability $P < 0.001$).

Tara also matched the novel action demonstrated of going inside a crate in the first trial.

Franc matched familiar actions in 6 out of 12 trials in which familiar actions were demonstrated. Thus, the dog matched the projected human demonstrations above chance (Binomial probability $P = 0.007$).

Franc did not match the novel action demonstrated of going inside the crate neither in the first trial, nor in the second when this action was demonstrated again. Instead, she lied down in both trials.

Above camera condition

Tara matched familiar actions in 4 out of 12 trials in which familiar actions were demonstrated. Thus, the dog did not match the projected human demonstrations above chance (Binomial probability $P = 0.09$). However, Tara matched the novel action demonstrated of dropping over a plastic bottle in the first trial.

Franc matched familiar actions in 5 out of 12 trials in which familiar actions were demonstrated. Thus, the dog matched the projected human demonstrations above chance (Binomial probability $P = 0.03$).

Franc did not match the novel action demonstrated of dropping over a plastic bottle neither in the first trial, nor in the second when this action was demonstrated again. Instead, she lied down in the first trial and bowed in the second.

Baseline

Both dogs matched all the 6 demonstrated familiar actions in all 12 trials. Both dogs also matched the 3 novel actions that were demonstrated in the other conditions.

Results data file with the results for each action is provided as Supplementary material S2.

Discussion

In this exploratory study, we investigated the possibility to use 2D videos, with manipulated camera angles, to test dogs' action matching abilities when they observe some demonstrations from different perspectives. The results showed that dogs previously trained with the Do as I Do method were capable, to some extent, of matching human actions projected on a screen, also from different

perspectives. This provides a novel potential method to test imitative skills and related cognitive abilities, while systematically manipulating several factors, such as, in this case, the perspective from which the demonstration is observed.

The significant drop observed in the dogs' performance in most conditions with projected demonstrations, compared to the imitation of live demonstrations in the Baseline, suggests that imitation from 2D projected videos is more challenging for dogs than imitation of a live model. It is possible that this is due to the more extensive experience of the dogs with live Do as I Do settings than with the 2D one, and that performance with projected demonstrations may improve with more extensive training. Thus, future studies may consider training the dogs until they perform similarly to live conditions, before testing variations of the 2D presentations. Dogs' decreased success in 2D conditions is also in line with the results of previous studies in human children (e.g., Barr 2010) and chimpanzees (Hopper et al. 2012) and may suggest a perceptual challenges in the task of transferring 2D information to a three-dimensional real-world setting.

The results of the Frontal camera condition showed that both Tara and Franc matched familiar actions displayed on the screen above chance, indicating that the training was successful, at least to some extent, and dogs can match human actions when the owner's frontal view is projected in 2D, to some degree. The dogs' capacity to use 2D stimuli as samples against which to match their own behavior expands the findings of previous studies on dogs' imitative skills (e.g., Topál et al. 2006) and is in line with dogs' capacity to rely on 2D stimuli in an object search and object choice tasks (Pongrácz et al. 2018; Péter et al. 2013).

In the Side camera condition, both dogs continued to exhibit similarly successful imitation of familiar actions, supporting the hypothesis that dogs generalize their imitative abilities to lateral perspectives, even in the case of 2D stimuli. Dogs regularly interact with humans from lateral angles, which likely contributes to their ability to imitate actions demonstrated from the side without specific training.

However, the Above camera condition presented a different pattern. Tara showed a reduced success rate in matching familiar actions from an above perspective; whereas, Franc exhibited a comparable performance to the Side camera condition. These results suggest that dogs might face challenges when attempting to imitate actions observed from an overhead view, which is an uncommon perspective for them to encounter during their daily interactions with humans. The difficulties of imitating demonstrated actions seen from above may also include a limitation in recognizing the demonstrator's body parts as seen from this perspective and, therefore, a difficulty in mapping those in the dog's own body parts.

The dogs' overall better performance in the Frontal and Side conditions is in line with previous studies showing that familiarity with the observed scene is fundamental for dogs to successfully process the stimuli (Pongrácz et al. 2018). Despite this, Tara's successful imitation of a novel action in the Above camera condition may suggest some flexibility in his action matching abilities. It is possible that the process of recognizing the demonstrated actions, in the trials when those were object-related, was facilitated by seeing objects in the projected videos, which were identical to those present for the dogs to act upon. The objects may have acted as recall cues for the action demonstrated on them. However, although comparisons between different types of actions are not possible due to limited number of trials, it does not seem that object-related actions were easier to imitate for the dogs. For example, the action of pressing a button was never imitated in any of the conditions with projected videos (but it was imitated in the Baseline with live demonstrations).

In this regard, it must be noted that there appears to be variability in the imitation success of the different human demonstrations, but conclusions on the potential effect of the demonstrated actions cannot be drawn with such a limited number of trials and subjects.

The results of the Baseline condition, where the owner demonstrated the actions live, show that both dogs successfully matched familiar actions in all trials, reaffirming dogs' well-established imitative skills in direct interactions with humans. Both dogs were also able to imitate the novel actions in the Baseline condition, supporting their capacity for flexible learning and generalization.

Some limitations should be acknowledged. The exploratory nature of this study, with its small sample size, the use of relatively few demonstrated actions and the possible variability of dogs' responses to different actions, may affect the generalizability of the results. It is thus important that future studies with larger sample sizes confirm and strengthen the conclusions drawn from this research.

Importantly, however, by demonstrating the possibility to use 2D projected demonstrations seen from different perspectives with dogs, these findings provide the foundation for a method to further investigate imitation, visual perspective and perspective-taking abilities in canines. This method also has the potential of being used with other species, thereby expanding the methods available to study cognitive skills related to imitation and perspective-taking.

Conclusions for future biology

In conclusion, our study provides evidence that dogs possess cognitive abilities to match human actions based on 2D visual stimuli, at least to some extent, and suggest that their performance varies with the familiarity of the

demonstration perspective. These results highlight the significance of using 2D stimuli and manipulated perspectives in investigating dogs' imitative abilities, and shed light on the role of familiarity of the perspective from which the demonstration is observed. Importantly, the results of this study pave the way for a novel method to be implemented in investigations involving projected images observed from different perspectives, to explore dogs'—and potentially other species'—cognitive capacities in the domains of imitation and perspective taking, among others.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42977-024-00222-6>.

Acknowledgements We are grateful to Nobuyo Konoha for her valuable contribution in training her dog Franc and helping during the tests. We thank Yukiko Kawakami, Chie Kikuchi, Ai Nakamura, and Yoshie Tanaka for their help during data collection.

Authors contribution C.F. devised the experiments, F.H. trained the dogs, C.F. and F.H. collected data, C.F. wrote the main manuscript and all authors contributed to and reviewed it.

Funding Open access funding provided by Eötvös Loránd University. CF was supported by the National Brain Research Program NAP 3.0 of the Hungarian Academy of Sciences (NAP2022-I-3/2022) and by the Hungarian Ethology Foundation (METAL).

Availability of data and material Data files are available upon request.

Declarations

Conflict of interest The authors have no competing interests to declare.

Ethical approval The authors affirm that the experiments reported in this paper are in accordance with the current Hungarian laws regarding animal protection. Ethical permission for conducting this study was obtained from The Institutional Committee of Eötvös Loránd University (N. PE/EA/692-5/2019).

Consent to participate The dog owners gave their consent to participate in the study with their dogs.

Consent for publication All authors and dog owners gave their consent for publication.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdai J, Miklósi Á (2023) After 150 years of watching: is there a need for synthetic ethology? *Anim Cogn* 26:261–274. <https://doi.org/10.1007/s10071-022-01719-0>
- Barr R (2010) Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Dev Rev* 30:128–154. <https://doi.org/10.1016/j.dr.2010.03.001>
- Barr R, Hayne H (1999) Developmental changes in imitation from television during Infancy. *Child Dev* 70:1067–1081. <https://doi.org/10.1111/1467-8624.00079>
- Carpenter M, Tomasello M, Striano T (2005) Role reversal imitation and language in typically developing infants and children with autism. *Infancy* 8(3):253–278. https://doi.org/10.1207/s15327078i0803_4
- Catala A, Mang B, Wallis L, Huber L (2017) Dogs demonstrate perspective taking based on geometrical gaze following in a Guesser–Knower task. *Anim Cogn* 20:581–589. <https://doi.org/10.1007/s10071-017-1082-x>
- Fugazza C, Miklósi Á (2014) Deferred imitation and declarative memory in domestic dogs. *Anim Cogn* 17:237–247. <https://doi.org/10.1007/s10071-013-0656-5>
- Fugazza C, Pogány Á, Miklósi Á (2016) Recall of others' actions after incidental encoding reveals episodic-like memory in dogs. *Curr Biol* 26:3209–3213. <https://doi.org/10.1016/j.cub.2016.09.057>
- Fugazza C, Petro E, Miklósi Á, Pogány Á (2019) Social learning of goal-directed actions in dogs (*Canis familiaris*): imitation or emulation? *J Comp Psychol* 133:244–251. <https://doi.org/10.1037/com0000149>
- Guillette LM, Healy SD (2019) Social learning in nest-building birds watching live-streaming video demonstrators. *Integr Zool* 14:204–213. <https://doi.org/10.1111/1749-4877.12316>
- Hämäläinen L, Rowland HM, Mappes J, Thorogood R (2019) The effect of social information from live demonstrators compared to video playback on blue tit foraging decisions. *PeerJ* 7:e7998. <https://doi.org/10.7717/peerj.7998>
- Held S, Mendl M, Devereux C, Byrne R (2001) Behaviour of domestic pigs in a visual perspective taking task. *Behaviour* 138:1337–1354. <https://doi.org/10.1163/156853901317367627>
- Hopper LM, Lambeth SP, Schapiro SJ (2012) An evaluation of the efficacy of video displays for use with chimpanzees (*Pan troglodytes*). *Am J Primatol* 74:442–449. <https://doi.org/10.1002/ajp.22001>
- Huber L, Lonardo L (2023) Canine perspective-taking. *Anim Cogn* 26:275–298. <https://doi.org/10.1007/s10071-022-01736-z>
- Huber L, Range F, Voelkl B et al (2009) The evolution of imitation: what do the capacities of non-human animals tell us about the mechanisms of imitation? *Phil Trans R Soc B* 364:2299–2309. <https://doi.org/10.1098/rstb.2009.0060>
- Kuroshima H, Fujita K, Fuyuki A, Masuda T (2002) Understanding of the relationship between seeing and knowing by tufted capuchin monkeys (*Cebus apella*). *AnimCogn* 5:41–48. <https://doi.org/10.1007/s10071-001-0123-6>
- Maginnity ME, Grace RC (2014) Visual perspective taking by dogs (*Canis familiaris*) in a Guesser–Knower task: evidence for a canine theory of mind? *Anim Cogn* 17:1375–1392. <https://doi.org/10.1007/s10071-014-0773-9>
- Meyer JA, Hobson RP (2004) Orientation in relation to self and other: the case of autism. *Interact Stud* 5(2):221–244. <https://doi.org/10.1075/is.5.2.04mey>
- Myowa-Yamakoshi M, Scola C, Hirata S (2012) Humans and chimpanzees attend differently to goal-directed actions. *Nat Commun* 3:693. <https://doi.org/10.1038/ncomms1695>
- Péter A, Miklósi Á, Pongrácz P (2013) Domestic dogs' (*Canis familiaris*) understanding of projected video images of a human demonstrator in an object-choice task. *Ethology* 119:898–906. <https://doi.org/10.1111/eth.12131>
- Pongrácz P, Péter A, Miklósi Á (2018) Familiarity with images affects how dogs (*Canis familiaris*) process life-size video projections of humans. *Quarterly J Exp Psychol* 71:1457–1468. <https://doi.org/10.1080/17470218.2017.1333623>
- Povinelli DJ, Nelson KE, Boysen ST (1990) Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *J Comput Psychol* 104:203–210. <https://doi.org/10.1037/0735-7036.104.3.203>
- Povinelli DJ, Rulf AB, Bierschwale DT (1994) Absence of knowledge attribution and self-recognition in young chimpanzees (*Pan troglodytes*). *J Comput Psychol* 108:74–80. <https://doi.org/10.1037/0735-7036.108.1.74>
- Price E, Caldwell CA (2007) Artificially generated cultural variation between two groups of captive monkeys, *Colobus guereza* kikuyuensis. *Behav Proc* 74:13–20. <https://doi.org/10.1016/j.beproc.2006.09.003>
- Santisteban I, White S, Cook J, Gilbert SJ, Heyes C, Bird G (2012) Training social cognition: from imitation to theory of mind. *Cognition* 122(2):228–235. <https://doi.org/10.1016/j.cognition.2011.11.004>
- Seehagen S, Herbert JS (2011) Infant imitation from televised peer and adult models. *Infancy* 16:113–136. <https://doi.org/10.1111/j.1532-7078.2010.00045.x>
- Sommer K, Redshaw J, Slaughter V et al (2021) The early ontogeny of infants' imitation of on screen humans and robots. *Infant Behav Dev* 64:101614. <https://doi.org/10.1016/j.infbeh.2021.101614>
- Topál J, Byrne RW, Miklósi Á, Csányi V (2006) Reproducing human actions and action sequences: “Do as I Do!” in a dog. *Anim Cogn* 9:355–367. <https://doi.org/10.1007/s10071-006-0051-6>
- Yu Y, Su Y, Chan R (2011) The relationship between visual perspective taking and imitation impairments in children with autism. *Compr Book Autism Spectrum Disorders*. <https://doi.org/10.5772/19717>