Domestic dogs' (*Canis familiaris*) understanding of projected video images of a human demonstrator in an object choice task

András Péter, Ádám Miklósi & Péter Pongrácz

András Péter (corresponding author), Department of Ethology, Eötvös Loránd University, Budapest, Pázmány Péter sétány 1/C, H-1117, Hungary. E-mail: bioandras@gmail.com

Ádám Miklósi, Department of Ethology, Eötvös Loránd University, Budapest, Pázmány Péter sétány 1/C, H-1117, Hungary and MTA-ELTE Comparative Ethological Research Group. E-mail: amiklosi62@gmail.com

Péter Pongrácz, Department of Ethology, Eötvös Loránd University, Budapest, Pázmány Péter sétány 1/C, H-1117, Hungary. E-mail: peter.celeste.pongracz@gmail.com

Running title: Domestic dogs' understanding of projected video images

Word count: **6002**

Abstract

Presenting animals with artificial visual stimuli is a key element of many recent behavioral experiments largely because images are easier to control and manipulate than live demonstrations. Determining how animals process images is crucial for being able to correctly interpret subjects' reactions towards these stimuli. In this study we aimed to use the framework proposed by Fagot and colleagues (2010) to classify how dogs perceive life-sized projected videos. First we tested whether dogs can use pre-recorded and hence non-interactive, video footage of a human to locate a hidden reward in a three-way choice task. Secondly, we investigated whether dogs solve this task by means of referential understanding. To achieve this we separated the location of the video projection from the location where dogs had to search for the hidden reward. Our results confirmed that dogs can reliably use pre-recorded videos of a human as a source of information when the demonstration and the hiding locations are in the same room. However they did not find the hidden object above the chance level when the hiding locations were in a separate room. Still, further analysis found a positive connection between the attention paid to the projection and the success rate of dogs. This finding suggests that the factor limiting dogs' performance was their attention, and that with further training they might be able to master tasks involving referential understanding.

Introduction

The use of photographs, slides or video films is widespread in behavioral experiments with non-human animals of various species (D'Eath 1998; Bovet & Vauclair 2000). The use of artificial visual stimuli (images) in such experiments has two obvious benefits: a) it enables presentation of an invariable stimulus, thus allowing stricter control of the experimental conditions; b) it enables manipulation of the stimulus in ways, which would be difficult or impossible to achieve with real objects or actors. However, it raises the question whether the animal is able to recognize the content of the picture.

Fagot et al. (2010) distinguishes between three modes of picture processing: independence, confusion and equivalence. Independence defines those cases when the animal makes no connection between the picture and its content, but processes the picture as a combination of features or patterns independently of what the picture might represent. Confusion defines those conditions in which the animal confuses the image and its referent, thus reacting the same way to the picture as to the real object. Equivalence defines instances where the animal understands that the picture is a representation of the depicted object. This latter level corresponds to referential understanding, which is the ability to perceive an object (e.g.: picture, video, replica, scale model) as standing for another entity in the world (DeLoache 1991; Gliga & Csibra 2009).

There are a number of different ways how understanding images can be tested in animals. Some studies investigated the subjects' spontaneous responses towards artificial stimuli (e.g. social behavior shown towards the picture (Fox 1971); preference shown for different pictures (Fujita 1993) or videos (Rosenthal et al. 1996)). However, the interpretation of these types of experiments is not clear, as the observed behavior could also be triggered by some key perceptual elements that the images shared with the real objects. Other studies are based on acquired responses where the animal is trained to discriminate between stimuli, for example by first training the animal either to discriminate between the real objects or the visual representations and then testing the transfer to the other modality (Bovet & Vauclair 1998). But as with the other methods, in this case the animals might also largely rely on a set of common perceptual features. Also, these tests require extensive training, which makes it difficult to draw inferences about the animal's spontaneous capacities. Additionally both types of methods described here are not suitable to differentiate between confusion and equivalence mode.

A method that is specifically designed to test for the presence of equivalence mode (referential understanding of images) is based on the method of DeLoache (1987), who tested children's referential abilities using a scale model. In such a tests the subject is presented with a picture or video of a room on which the position of a hidden reward is shown (Poss & Rochat 2003). If the subject can find the reward in the real room based on this demonstration, and without extensive training, then it can be assumed that it was able to connect the content of the picture with its real world referent, thus it is capable of referential understanding. Until now, only humans (children of 2-3 years of age (DeLoache & Burns 1994; Troseth & DeLoache 1998)) and chimpanzees (Menzel et al. 1978), could successfully solve such tests.

Dogs live among humans and are constantly exposed to referential artifacts that inhabit the human world (e.g.: pictures, television, mirrors) and anecdotal evidence from dog owners also suggests that dogs react appropriately to these artifacts. Additionally, artificial visual stimuli (pictures: Range et al. 2008; Faragó et al. 2010; Racca et al. 2010, as well as videos: Pongrácz et al. 2003; Harr et al. 2009; Téglás et al. 2012) were used in numerous recent experiments that were conducted with dogs. However, how dogs process these stimuli has only been sporadically investigated.

In an experiment Kaminski et al. (2009) demonstrated that dogs are able to use iconic signs (life sized replicas, miniature replicas or photographs) to correctly retrieve the corresponding object from a pool of objects. The authors of this paper argue that mastering this ability without previous training proved that dogs understood the referential nature of iconic signs. Still it is unclear whether dogs would be able to pass the test of locating an object using a picture or a video of the room where that object was hidden.

One earlier study utilizing life sized projected videos found that dogs can reliably follow the pointing gesture of the projected human (Pongrácz et al. 2003) to choose from two containers. However in this experiment a live video feed was used which enabled feedback between the dog and the human on the video, and also the question how dogs understand the projected video was not investigated. Due to their everyday exposure to referential artifacts and their apparent cognitive ability to understand the referentiality of pictures (Kaminski et al. 2009), it is conceivable that dogs comprehended the referential aspect of the video demonstration. Therefore the principal aim of the present study was to investigate whether dogs understand the referential nature of projected videos and additionally to find out whether dogs can utilize information from a pre-recorded footage to locate an object in the real world.

To answer these questions, we designed a visible displacement task (Triana & Pasnak 1981) similar to ones used for testing the referential abilities of children and chimpanzees (Poss & Rochat 2003). In our test we used pre-recorded videos as stimuli for the dogs. In the videos a human hid an object behind one of three different hiding locations. After the video demonstration, the subject could choose from the corresponding real hiding locations.

Dogs were tested in two conditions labeled either one-room or two-room condition. In the one-room condition the video demonstration and the real containers were placed in the same room. In the two-room condition the video demonstration and the real containers were placed in separate rooms. Solving the task in the one-room condition would mean that dogs can use the information on the video footage to find the object, but because of the lack of spatial separation, this would not necessarily mean they are capable of referential understanding. If dogs solved the task in the two-room condition then one could argue that, similarly to 2-3 year old children and chimpanzees, they also rely on referential understanding. Dogs participated first in the one-room and then in the two-room condition. The rationale for this fixed order design was to start with the simpler one-room condition and to introduce subjects to the nature of the search task with video demonstration.

A separate set of dogs was additionally tested in a control condition. The control condition was intended to control for the delay between the demonstration and the start of the search in the two-room condition, which occurred due to the dogs moving between the two rooms. The control set-up was similar to the one-room condition, but a pre-set time delay was introduced between the end of the video demonstration and the start of the search.

Methods

Subjects

Pet dogs (N = 36) and their owners were recruited on a voluntary basis. The dogs had to be highly motivated to retrieve a ball. Dogs were older than 1 year and represented various pure or mixed breeds (Gagnon and Doré (1992) showed that domestic dogs from various breeds showed equal performance in a visible displacement task).

Half of the subjects (N = 18, 9 females and 9 males, mean age = 3.2 years, range 1-5 years) participated first in the one-room condition and subsequently in the two-room condition. The other half of the subjects (N = 18, 9 females and 9 males, mean age = 2.8 years, range 1-6 years) participated in the control condition only.

Setup

All tests were performed indoors, in the experimental rooms at the Department of Ethology, Eötvös Loránd University, in Budapest. The two rooms used in this study had the same dimensions (3m x 6m). In all conditions three hiding locations were used. Each one was composed of a blue plastic panel (30cm×30cm) and a plastic flower pot (diameter 12cm) which was fixed behind the panel. Each of the three panels had a different geometric shape: triangle, square or pentagon. The hiding locations were arranged along a line at a distance of 1m from each other and approximately 3m from the starting position (SP) of the dog. The position of the individual geometric shapes was randomized for each subject in each condition. We used a small ball as the target object.

In room 1, the projector screen was placed opposite to the door: 2m wide, 1.8m high (Figure 1). Behind the screen were two loudspeakers. The projector was fixed near the ceiling on the other end of the room. In room 1 four cameras were recording the experiment. One of the cameras was an infrared camera, which recorded the dogs' orientation during the video demonstration. An array of infrared LEDs were directed towards the dog to increase efficiency of the **infra-red** camera. In room 2 one camera directed towards the hiding locations recorded the dogs' choices (Figure 2).

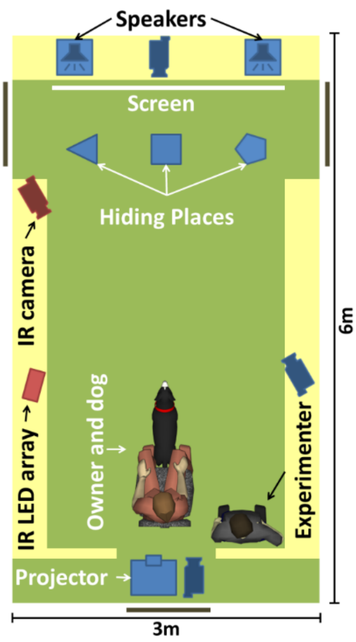

Figure 1 Arrangement of experimental room 1 in the one-room condition.

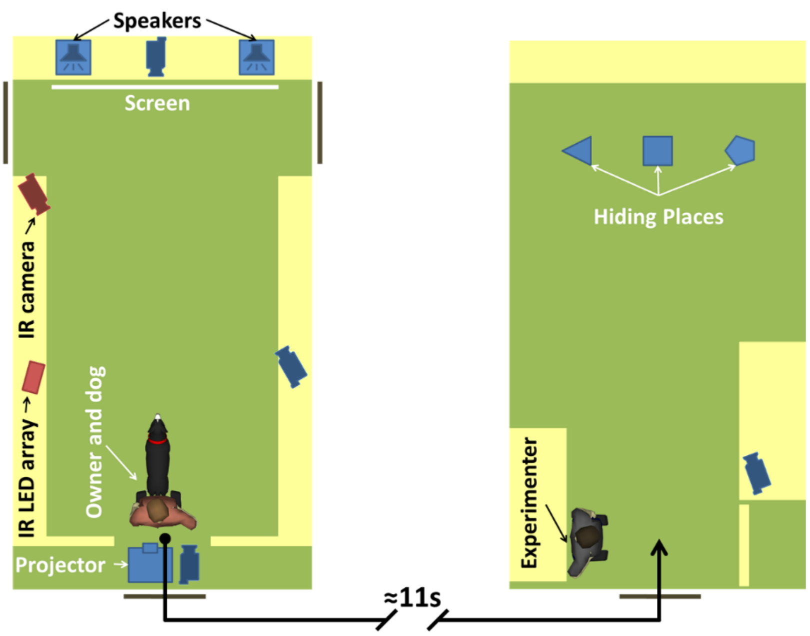

Figure 2 Arrangements of the experimental rooms in the two-room condition (room 1 on left side, room 2 on right side). The black arrow on the bottom shows the path of the owner and the dog between the two rooms and the time it took. The distance between the doors of the two rooms was 9m.

Procedure

The experiment consisted of three conditions: one-room, two-room and control. In the one-room and control conditions, both the video demonstration and the hiding locations were in room 1, in the two-room condition the video demonstration was in room 1 and the hiding locations were in room 2. Dogs participating in the one-room condition were subsequently tested in the two-room condition (after at least one week of delay). Dogs participating in the control condition were tested in that condition only. All conditions consisted of 3 warm-up and 9 test trials, each dog participated in a given condition only once. A video showing examples of warm-up and test trials of each condition can be found in the video supplement.

1) Warm-up phase

The aim of this phase was to familiarize the dogs with the hide-and-search task. However **as** our goal was to test the subjects' spontaneous performance in the oncoming test, we kept the number of warm-up trials as low as possible, to minimize the chance of any kind of learning occurring during these trials. For the same reason there was also no criterion set to pass this phase. The procedure of the warm-up phase was identical in all three conditions. In the one-room and control conditions it took place in room 1, and in the two-room condition in room 2. Each warm-up trial started with the dog, the owner (O) and the experimenter (E) being at the SP. The E showed the target object to the dog, went straight to one of the hiding locations, stopped behind it facing the SP, called the dogs' attention, raised and waved the object, put it into the pot behind the hiding location and finally returned to the SP next to the O. After this the O released the dog with one command "You can go!" to search for the target object. The dog was allowed to search until the object was found. Lastly the O called the dog back, praised it and took the object from the dog. During the 3 trials the target was placed behind each hiding location once in a random order.

2) Test phase

a) One-room condition: After the warm-up trials the O covered the dog's eyes by hand and the E turned off the lights in the room. The windows of the room were covered, therefore the room was semi-dark. The E took the object to the actual hiding location and put the object into the pot behind it, then returned to a location behind the O where he hid behind a curtain. After returning, the E started the video projection by pressing a button on the wall, and the O uncovered the eyes of the dog. The pre-recorded video was projected onto the canvas behind the hiding locations (Figure 3).

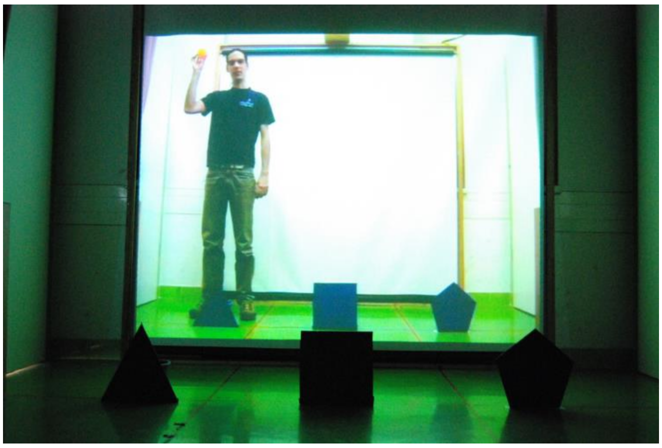

Figure 3 The video demonstration from the perspective of the dog in the one-room condition.

The video demonstration consisted of three phases: 1) attention getting phase (3s): the E stood still for 1s behind one of the hiding locations holding the tennis ball in his hand, then he greeted the dog saying “Hello!”, and he waved the ball saying “Look, look, look!”; 2) hiding phase (3s): the E crouched down and placed the ball to the actual hiding location, then he stood up; 3) conclusion phase (1s): the E displayed his empty hands while standing still. The E on the video placed the object to the same location where it was placed in reality. The arrangement of the hiding location shapes on the video was identical to the arrangement in the room.

After the video was over, the canvas turned black and the E turned the lights on in the room. The O released the dog and with one command allowed it search for the object. The dog could search for the object until it was found. Then the O called the dog back and praised the dog. During the nine trials the object was placed behind each hiding location three times in a semi-random order, so that it was never at the same location in two consecutive trials.

b) Two-room condition: The three hiding locations were in room 2, but the video projection took place in room 1. After the warm-up trials in room 2, the O and the dog went to room 1, and the E stayed in room 2 and placed the target object to the actual hiding location. During this the O and the dog entered room 1 and positioned themselves in front of the door. The O ensured that the dog was facing the canvas and started the projection with a wired remote located next to the door. The video was identical to the ones used in the one-room condition, except that it showed the E in room 2 placing the target object to the actual hiding location.

After the video ended, the screen turned black and the O led the dog back to room 2. Upon entering room 2 the O released the dog, and with the one command let it search for the object. The dog was allowed to search until it found the object. Throughout the nine trials the location of the target object was randomized in the same way as in the one-room condition.

c) Control condition: The procedure was mostly identical to the one-room condition, therefore we only highlight the differences here. After the warm-up trials, the O covered the eyes of the dog and the E placed the target object to one of the hiding locations. Next the E left the room through the open door behind the O and went into an adjacent room. Upon leaving the room, the E pushed a button on the wall which started the video projection with a 5s delay. When the video started the O uncovered the eyes of the dog. After the video ended, the O and the dog left the room and took a short walk in the hallway outside of the room for the amount of time it would have taken to walk to room 2. During this time the E remotely turned on the lights of room 1. The O and the dog returned to room 1 and the O released the dog with one command to search for the target object. After the dog found the object and returned it to O, the E entered room 1 through the door behind the O, and the next trial started. Throughout the nine trials the location of the target object was randomized in the same manner used in the other two conditions.

Data collection and analysis

All trials were video recorded and the recordings coded with Solomon Coder beta (© 2012 by András Péter). Dogs' first location choices were coded in each trial: location choice was defined as the first pot the dog looked into behind a hiding location.

We also coded whether dogs oriented towards the projector canvas during the hiding phase of the video demonstration. Orienting towards the screen was defined by the head of the dog having an angular deviation less than 45° from perpendicular to the screen. We considered the hiding phase the main section of the demonstration because in this phase the object disappeared from sight. The trials where dogs were orienting towards the screen during the entire hiding phase were labeled "complete attention" trials. Trials where dogs broke eye contact with the screen were labeled "incomplete attention".

Data were analyzed with IBM® SPSS® Statistics 20. In each of the three conditions 4 out of the 18 videos were coded by an independent coder who was naïve regarding the aim of the study. In case of location choices there was a 100% agreement between the two coders, whereas in the case of complete / incomplete attention trials the interrater reliability was found to be: Kappa=0.82, $p<0.001$, 95% CI: 0.68-0.95. According to one-sample Kolmogorov–Smirnov tests, our data did not follow the normal distribution, therefore we used nonparametric tests. For each condition we tested whether there is a difference in the number of correct trials and in the number of trials with complete attention between female and male dogs (Mann-Whitney U tests). We found no difference between the sexes in any of the conditions, therefore we pooled the data for further analysis.

When comparing the number of correct trials, or the number of trials with complete attention, among the three experimental conditions, we always carried out three pairwise comparisons. The one-room and two-room conditions were compared with a related-samples Wilcoxon signed rank test, and the control condition was compared with the one-room and two-room conditions with two independent-samples Mann-Whitney U tests. We chose to analyze our data this way to account for the repeated nature of measurements in the one-room and two-room condition, and the fact that the dogs included in the control condition did not participate in either of the other two conditions. When testing for learning effects we compared the number of correct trials in the first three and the last three trials for each condition with related-samples Wilcoxon signed rank tests.

Wilcoxon and Mann-Whitney U tests require data having homogeneous variances across groups. The Brown-Forsythe test (done with R 2.15.3 and the lawstat package; Gastwirth et al. 2013) did not detect evidence for heterogeneity either in the case of the number of correct trials ($F=2.66$, $p=0.08$) and the number of trials with complete attention ($F=1.63$, $p=0.21$) when compared between conditions. Also the number of correct trials in the first three and last three trials met this criteria in all conditions (one-room: $F=3.78$, $p=0.06$; two-room: $F=1.45$, $p=0.24$; control: $F=0.51$, $p=0.48$). In both the Wilcoxon and Mann-Whitney U tests SPSS handled ties in the dataset by assigning an average rank to them, and by using normal approximation.

When analyzing the effect of attention on performance, we used a generalized linear mixed model with a binary logistic link, correct/incorrect choices as the target variable, complete/incomplete attention paid to the hiding phase as the fixed effect and dog ID as a random factor and number of trial (1 to 9) set as the repeated variable. We allowed the degree of freedom to vary between tests because the differing number of trials with complete/incomplete attention resulted in an unbalanced data set. Also to compensate for potential deviations from the model's assumptions, we used robust covariance estimates.

To account for the increased chance of type-one errors due to multiple comparisons, we adjusted the p values in each test battery using the method by Hochberg (1988) as was described by Wright (1992). The adjusted p values are marked as p_{Hoch}.

Results

First we analyzed the dogs' performance in the three experimental conditions. We compared the number of correct trials to the level expected by chance (3 in a three choice task with nine trials) to determine whether the dogs could reliably solve the tasks (Figure 4). According to the one-sample Wilcoxon signed rank test, in the one-room condition the dogs had significantly more correct trials ($N=18$, $Z=3.43$, $p_{Hoch}<0.01$) than expected by chance. However in the two-room condition the number of correct trials ($N=18$, $Z=0.31$, $p_{Hoch}=0.76$) did not differ from chance. Finally, in the control condition the number of correct trials ($N=18$, $Z=2.57$, $p_{Hoch}<0.05$) was significantly higher than that expected by chance.

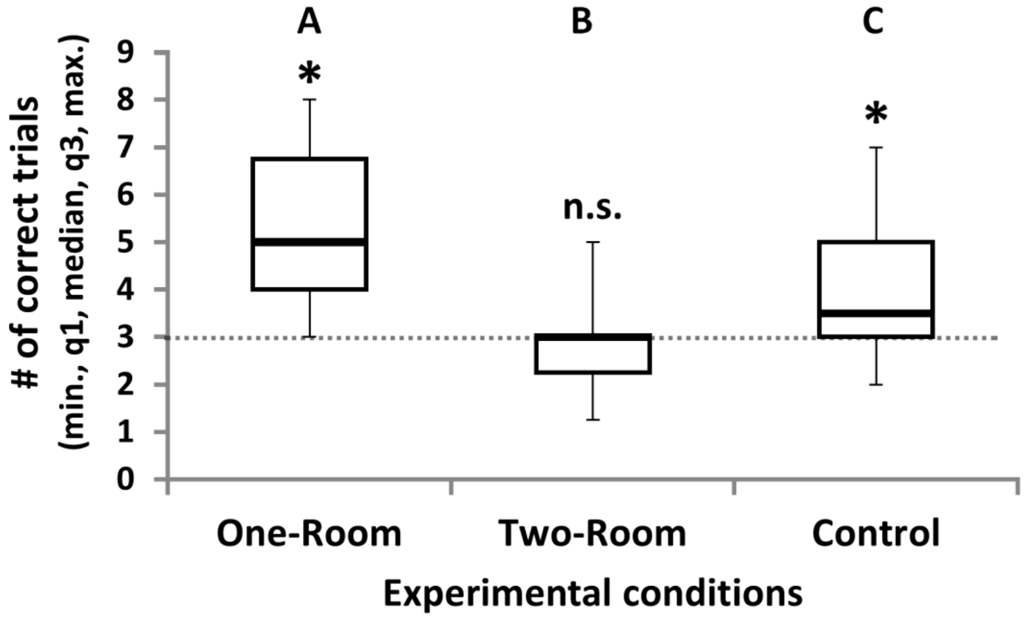

Figure 4 Number of correct trials compared to the chance level (dashed line) in the three conditions: one-sample Wilcoxon signed rank test (*: $p_{Hoch}<0.05$; n.s.: $p_{Hoch}\geq0.05$). Number of correct trials compared between the three conditions: related-samples Wilcoxon signed rank test (conditions differing significantly are labeled with different letters).

We also performed three pair-wise comparisons of the number of correct trials among the three conditions. When comparing the one-room and two-room conditions with a related-samples Wilcoxon signed rank test ($N=18$, $Z=3.24$, $p_{Hoch}<0.01$), we found that the dogs performed significantly worse in the two-room condition than in the one-room condition. After comparing the control condition with an independent-samples Mann-Whitney U test to the one-room ($N=36$, $Z=2.22$, $p_{Hoch}<0.05$) and the two-room conditions (N=36, Z=2.27, p_{Hoch}<0.05), we found that in the control condition the dogs had significantly more correct choices than in the two-room but significantly less than in the one-room condition.

Because of the pre-recorded nature of the video presentation, the dogs could have paid different amounts of attention to the demonstration in the three conditions, which could have caused the observed performance difference between the three conditions. Therefore we compared the number of trials with complete attention paid to the hiding phase of the demonstration between the three conditions. According to the related-samples Wilcoxon signed rank test (N=18, Z=1.78, p_{Hoch}=0.15), there was no significant difference in the number of trials with complete attention between the one-room and two-room conditions. A comparison of the control condition, with an independent-samples Mann-Whitney U test, to the one-room (N=36, Z=2.18, p_{Hoch}=0.09) and the two-room conditions (N=36, Z=1.17, p_{Hoch}=0.25), detected no significant differences between these groups either.

We also tested whether attention (orienting towards the screen) had an effect on performance in the three conditions (Figure 5) with a generalized linear mixed model. In the one-room condition the test found no differences in the number of correct trials between those with complete and incomplete attention (N=18, $F_{1,160}$=0.26, p_{Hoch}=0.61). However in the two-room condition the dogs found the object significantly more often after paying complete attention to the hiding phase of the demonstration than when they broke eye contact with the screen during this phase (N=18, $F_{1,119}$=6.50, p_{Hoch}<0.05). In the two-room condition, in the trials where they had paid complete attention to the hiding phase, the dogs had performed better (median: 33%) than when they had not watched the complete hiding phase (median 17%). In the control condition, similar to the one-room condition, we found no difference between the trials with complete and incomplete attention (N=18, $F_{1,80}$=0.26, p_{Hoch}=0.68).

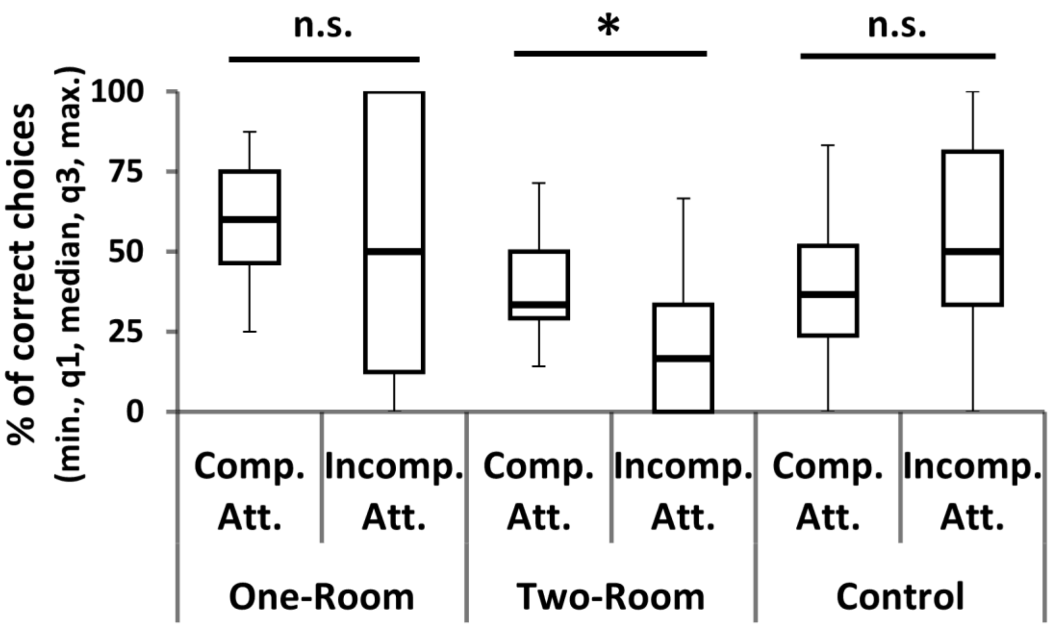

Figure 5 Ratio of correct choices in trials where the dogs oriented the entire time towards the screen during the video demonstration's hiding phase (Comp. Att.) and in trials where they oriented less (Incomp. Att.) in the three experimental conditions: generalized linear mixed model with binary logit link (*: $p_{Hoch}<0.05$; n.s.: $p_{Hoch}\geq 0.05$).

Finally we analyzed whether the dogs' performance increased during the trials in the three test conditions. We compared the number of correct trials in the first three and the last three trials for each condition. The related-samples Wilcoxon signed rank tests indicated no significant differences for the one-room (N=18, Z=0.37, p_{Hoch}=0.71), two-room (N=18, Z=0.52, p_{Hoch}=0.61) and control (N=18, Z=2.00, p_{Hoch}=0.14) conditions.

Discussion

The dogs in our study were able to reliably find the target object in the one-room and control conditions without any pre-training, except for three warm-up trials. In the two-room condition the dogs' performance was significantly lower than in the two other conditions, although they did not orient significantly less towards the screen during the hiding phase of the demonstration. The low performance in the two-room condition cannot be attributed to the delay between the end of the video demonstration and the start of the search either, because the control condition had a similar delay, but the dogs’ performance was still significantly higher than in the two-room condition. This result is in accordance with previous findings showing that dogs can reliably find a target object in a multi-well choice task with 10s or 30s of delay (Fiset et al. 2003). Being able to find the object in the control condition means that the dogs could memorize the physical position of the hiding location and retrieve it during search.

The result that dogs could not reliably find the object in the two-room condition indicates that they process the videos in confusion mode according to Fagot et al.'s (2010) classification. This outcome is in line with the observation of Fox (1971), who found that dogs react the same way to life sized painted dogs as to a real conspecific.

However, we also found that in the two-room condition, paying attention to the video demonstration’s hiding phase makes it more likely that the dogs find the hidden object in the same trial. This suggests that they memorized the relative position of the object’s disappearance on the video and transferred this information to the relative position of the hiding locations in room 2. Nevertheless, even in those trials where dogs paid attention to the demonstration, the median of successfully retrieving the ball was only 33%. This indicates that the effect of the information transfer is fairly small, and might only be enough to compensate for factors that would otherwise decrease the observed performance (e.g.: choosing the location where the object was in the previous trial).

We found no association between attention and performance in the one-room and control conditions. This does not mean that attention affects performance differently in these conditions than in the two-room condition, since a larger sample size could have yielded an association in these conditions, too. On the other hand dogs might have found the hidden object without the need to pay attention to the critical section of the demonstration, because in these conditions the hiding locations on the video were in close proximity to the real ones. Consequently, in this case dogs could have found the correct location by relying on simple local enhancement cues.

These results do not support the notion unambiguously that dogs process videos only in confusion mode. Earlier, the results of Kaminski et al. (2009) suggested that dogs can understand the referential nature of pictures. However, in that study some of the dogs underwent considerable training before the test, which suggests that with additional training dogs might have shown a clear sign of referential understanding in our study too. Although currently we were interested in dogs' spontaneous reaction to videos, it could be a topic of future studies to find out whether training can improve dogs' performance in a referential understanding task.

In our present study we found that dogs could use information from pre-recorded videos, if the location of the video demonstration and the location referred by the video were in close proximity. Dogs were able to extract, memorize and retrieve location information from the video demonstration.

On the other hand we did not find evidence that dogs would process life sized videos in equivalence mode, which means that to date only humans (Troseth & DeLoache 1998) and chimpanzees (Menzel et al. 1978) were shown to referentially understand videos. However the majority of experiments utilizing artificial visual stimuli is built on the assumption that dogs process the stimuli in confusion mode (eg.: Faragó et al. 2010; Téglás et al. 2012). Therefore our findings open up many possibilities for further studies. Using this paradigm, experiments could be conducted where the presented stimulus is uniform across trials. Alternatively, stimuli could be presented which would be impossible in a real life setup. For example experiments of physical cognition using the violation of expectation paradigm (e.g.: Pattison et al. 2010) can benefit from such a method, as in such studies actions often have to be presented which do not occur in reality.

Also experiments on social cognition could profit from such a methodology, because the non-interactive nature of the stimulus presentation could eliminate many sources of the Clever Hans effect (Pfungst 1911). Although it has been shown that this effect is not as powerful as assumed earlier (Schmidjell et al. 2012; Pongrácz et al. 2013), it is also known that the precise timing of ostensive cues can have a dramatic effect on dogs performance (Range et al. 2009) and using video demonstrations provides the means to have the timing of cues under perfect control. Finally life sized videos could be used as stimuli for a number of other species with visual perception suited to process projected images (for a detailed review see: Fleishman et al. 1998). Also the method used in this article is a straightforward way to test how animals process projected videos and to validate the use of such stimuli in further experiments.

Acknowledgements

This paper was supported by the János Bolyai Research Scholarship from the Hungarian Academy of Sciences. The study was funded by a grant from the Hungarian Ministry of Education OTKA K82020. The support of the Hungarian Academy of Science is acknowledged (MTA-ELTE 01-031), Á.M. is also a member of the Comparative Ethology Research Group.

References

Bovet, D. & Vauclair, J. 1998: Functional Categorization of Objects and of Their Pictures in Baboons (Papio anubis). Learn. Motiv. **29**, 309–322.

Bovet, D. & Vauclair, J. 2000: Picture recognition in animals and humans. Behav. Brain Res. **109**, 143–165.

D'Eath, R. B. 1998: Can video images imitate real stimuli in animal behaviour experiments? Biol. Rev. Camb. Philos. Soc. **73**, 267–292.

DeLoache, J. S. 1987: Rapid change in the symbolic functioning of very young children. Science **238**, 1556–7.

DeLoache, J. S. 1991: Symbolic Functioning in Very Young Children: Understanding of Pictures and Models. Child. Dev. **62**, 736–752.

DeLoache, J. S. & Burns, N. M. 1994: Early understanding of the representational function of pictures. Cognition **52**, 83–110.

Fagot, J., Thompson, R. K. R. & Parron, C. 2010: How to read a picture: lessons from nonhuman primates. Proc. Natl. Acad. Sci. U.S.A. **107**, 519–20.

Faragó, T., Pongrácz, P., Miklósi, Á., Huber, L., Virányi, Z. & Range, F. 2010: Dogs' expectation about signalers' body size by virtue of their growls. PLOS ONE **5**, e15175.

Fiset, S., Beaulieu, C. & Landry, F. 2003: Duration of dogs' (Canis familiaris) working memory in search for disappearing objects. Anim. Cogn. **6**, 1–10.

Fleishman, L. J., McClintock, W. J., D'Eath, R. B., Brainard, D. H. & Endler, J. A. 1998: Colour perception and the use of video playback experiments in animal behaviour. Anim. Behav. **56**, 1035–1040.

Fox, M. W. 1971: Socio-infantile and Socio-sexual signals in Canids: a Comparative and Ontogenetic Study. Zeitschrift für Tierpsychologie **28**, 185–210.

Fujita, K. 1993: Role of Some Physical Characteristics in Species Recognition by Pigtail Monkeys **34**, 133–140.

Gagnon, S. & Doré, F. Y. 1992: Search behavior in various breeds of adult dogs (Canis familiaris): Object permanence and olfactory cues. J. Comp. Psychol. **106**, 58–68.

Gastwirth, J. L., Gel, Y. R., Hui, W. L. W., Lyubchich, V., Miao, W. & Noguchi, K. 2013: lawstat: An R package for biostatistics, public policy, and law.

Gliga, T. & Csibra, G. 2009: One-year-old infants appreciate the referential nature of deictic gestures and words. Psychological Science **20**, 347–53.

Harr, A. L., Gilbert, V. R. & Phillips, K. A. 2009: Do dogs (Canis familiaris) show contagious yawning? Anim. Cogn. **12**, 833–7.

Hochberg, Y. 1988: A sharper Bonferroni procedure for multiple tests of significance. Biometrika **75**, 800–802.

Kaminski, J., Tempelmann, S., Call, J. & Tomasello, M. 2009: Domestic dogs comprehend human communication with iconic signs. Dev. Sci. **12**, 831–7.

Menzel, E. W., Premack, D. & Woodruff, G. 1978: Map Reading by Chimpanzees. Folia Primatologica **29**, 241–249.

Pattison, K. F., Miller, H. C., Rayburn-Reeves, R. & Zentall, T. R. 2010: The case of the disappearing bone: Dogs' understanding of the physical properties of objects. Behav. Processes **85**, 278–82.

Pfungst, O. 1911: Clever Hans (The horse of Mr. von Osten): A contribution to experimental animal and human psychology. Henry Holt and Company, New York.

Pongrácz, P., Miklósi, Á., Bálint, A. & Hegedüs, D. 2013: Owners fail to influence the choices of dogs in a two-choice, visual pointing task. Behaviour **150**, 427–43.

Pongrácz, P., Miklósi, Á., Dóka, A. & Csányi, V. 2003: Successful Application of Video-Projected Human Images for Signalling to Dogs. Ethology **109**, 809–821.

Poss, S. R. & Rochat, P. 2003: Referential understanding of videos in chimpanzees (Pan troglodytes), orangutans (Pongo pygmaeus), and children (Homo sapiens). J. Comp. Psychol. **117**, 420–8.

Racca, A., Amadei, E., Ligout, S., Guo, K., Meints, K. & Mills, D. S. 2010: Discrimination of human and dog faces and inversion responses in domestic dogs (Canis familiaris). Anim. Cogn. **13**, 525–33.

Range, F., Aust, U., Steurer, M. M. & Huber, L. 2008: Visual categorization of natural stimuli by domestic dogs. Anim. Cogn. **11**, 339–47.

Range, F., Heucke, S. L., Gruber, C., Konz, A., Huber, L. & Virányi, Z. 2009: The effect of ostensive cues on dogs’ performance in a manipulative social learning task. Appl. Anim. Behav. Sci. **120**, 170–178.

Rosenthal, G. G., Evans, C. S. & Miller, W. L. 1996: Female preference for dynamic traits in the green swordtail, Xiphophorus helleri. Anim. Behav. **51**, 811–820.

Schmidjell, T., Range, F., Huber, L. & Virányi, Z. 2012: Do owners have a clever hans effect on dogs? Results of a pointing study. Frontiers in Psychology **3**, 558.

Téglás, E., Gergely, A., Kupán, K., Miklósi, Á. & Topál, J. 2012: Dogs’ gaze following is tuned to human communicative signals. Curr. Biol. **22**, 209–12.

Triana, E. & Pasnak, R. 1981: Object permanence in cats and dogs. Anim. Learn. Behav. **9**, 135–139.

Troseth, G. L. & DeLoache, J. S. 1998: The Medium Can Obscure the Message: Young Children’s Understanding of Video. Child. Dev. **69**, 950–965.

Wright, S. P. 1992: Adjusted P-values for simultaneous inference. Biometrics **48**, 1005–1013.