Test sensitivity is important for detecting variability in pointing comprehension in canines

Péter Pongrácz[1*], Márta Gácsi[2], Dorottya Hegedüs[1], András Péter[1] & Ádám Miklósi[1,2]

[1] Department of Ethology, Institute of Biology, Eötvös Loránd University

[2] MTA-ELTE Comparative Ethology Research Group

* Corresponding author: Péter Pongrácz

Department of Ethology, Institute of Biology, Eötvös Loránd University

Budapest

Pázmány Péter s. 1/c

1114

Hungary

Tel: +3630 262 1284

Email: peter.celeste.pongracz@gmail.com

Word count: 8997

Several articles have been recently published on dogs' (*Canis familiaris*) performance in two-way object choice experiments in which subjects had to find hidden food by utilising human pointing. The interpretation of results has led to a vivid theoretical debate about the cognitive background of human gestural signal understanding in dogs, despite the fact that many important details of the testing method have not yet been standardized. We report three experiments that aim to reveal how some procedural differences influence adult companion dogs' performance in these tests. Utilising a large sample in Experiment 1 we provide evidence that neither the keeping conditions (garden/house) nor the location of the testing (outdoor/indoor) affect a dogs' performance. In Experiment 2 we compare dogs' performance using three different types of pointing gestures. Dogs' performance varied between momentary distal and momentary cross pointing but 'low' and 'high' performer dogs chose uniformly better than chance level if they responded to sustained pointing gestures with reinforcement (food reward and a clicking sound; 'clicker pointing'). In Experiment 3 we show that single features of the aforementioned 'clicker pointing' method can slightly improve dogs' success rate if they were added one by one to the momentary distal pointing method. These results provide evidence that although companion dogs show a robust performance at different testing locations regardless of their keeping conditions, the exact execution of the human gesture and additional reinforcement techniques have substantial effect on the outcomes. Consequently, researchers should standardise their methodology before engaging in debates on the comparative aspects of socio-cognitive skills because the procedures they utilise may differ in sensitivity for detecting differences.

Keywords: communication, dog; two-way object choice task; human pointing

INTRODUCTION

In the so-called two-way object choice experiments the subject has to find a hidden reward based on the directed pointing gesture of a human assistant (Anderson et al. 1995). Positive evidence of reliance on human pointing gesture was found in the case of several species (e.g. cats: Miklósi et al. 2005; goats: Kaminski et al. 2005; dolphins: Herman et al. 1999; enculturated apes: Mulcahy and Call 2009).

Dogs' ability to rely on human gestures has often been interpreted in the framework of specific behavioural adaptations to the human social environment during domestication (see for example Miklósi et al. 2004; Miklósi and Soproni 2006; Reid 2009). Some researchers have hypothesized that dogs' performance in the pointing tasks can be explained by a specific adaptation for utilizing human communicative signals (Hare et al. 2002). This possible effect of domestication has been tested by comparing the performance of dogs and wolves. The first report showed that wolves living in captivity underperformed dogs in these two-way object choice experiments (Hare et al. 2002). However, wolves' performance seems to be influenced by their rearing environment because both Miklósi et al. (2003) and Virányi et al. (2008) showed that intensively socialized young wolves display better performance in these pointing tasks than what was found in the Hare et al (2002) study. Utilizing a different population of intensively socialized four-month-old wolves, Gácsi et al. (2009b), found that these subjects were inferior to same aged dogs, but at the same time Gácsi and colleagues could not show any difference between the performances of intensively socialized adult wolves and dogs. Thus the difference between the two species may be related to their socio-cognitive development in regard to their relationship with humans. Miklósi et al. (2003) argued that intensively socialized wolves are less inclined to initiate and become engaged in gaze contact with humans and it also appears evident that wolves need more intensive exposure to social interaction with humans to be able to reach similar levels of communication skills that dogs are capable of (Gácsi et al. 2009b, Miklósi and Topál 2011).

Recently Udell et al. (2008a) offered a different hypothesis for these inter-specific differences. They argued that the dogs' superior performance can be explained by assuming that dogs living with humans gain more experience and therefore learn more about human communicative gestures which may also include exposure to a positive outcome ("reward") which follows the gestural cues. Udell et al. (2008a) supported this idea by showing that under certain conditions socialized adult wolves performed just as well as dogs. Additionally, they found that dogs from a rescue shelter did not seem to be able to utilise the Momentary Distal human pointing gesture spontaneously, though they could follow simpler forms of pointing and did learn to follow the Momentary Distal point with additional trials (Udell et al. 2010a). This finding was also interpreted by these authors as further evidence against the idea that dogs' communicative skills have been selected for in the anthropogenic environment.

Udell's work (2008a, 2010a) was followed by a debate whether the ability of dogs to follow human pointing has been driven mainly by specific selective challenges in the human environment or whether learning also plays a significant role (e.g. Udell et al. 2010b; Hare et al 2010a, Wobber et al. 2010). For example, Wobber and colleagues (2010), argued that genetic predisposition may still exist regarding this trait, because dog breeds selected for working with humans (e.g. Huskies and German shepherd dogs) show better performance in pointing tasks than breeds which were not selected for specific tasks (e.g. toy poodles or basenjis). Helton and Helton (2010) however re-analyzed the data of Wobber et al., (2010) and found that the results could be attributed to the choice of dog breeds tested. According to Helton and Helton (2010), the working breeds had a much bigger body size than the non-working breeds in the Wobber et al (2010) study, and therefore the anatomical differences between the visual apparatus of the two groups could also be a reason why the smaller (non-working) dogs underperformed in comparison to the larger (working) dogs.

In this article we examine some of the proximate factors that might contribute to this specific ability in dogs, and here we give a brief overview of a few other studies about both the ultimate and proximate causes of differences in dogs' performances in the pointing tasks.

First, different genetic factors are likely to play an influential role at different levels. Recently we have shown (Gácsi et al. 2009a) that dogs with brachycephalic skulls (e.g. Pugs, Bulldogs) perform better in pointing tasks than dogs with dolichocephalic skulls (e.g. Rough collies, Greyhounds). Furthermore, those dog breeds which have been selected for visually guided cooperation with humans (e.g. gundogs) achieve higher performance than dogs from so-called non-cooperative working breeds (e.g. terriers), and pure bred dogs seem to be more proficient with human pointing as well (Gácsi et al. 2009a). Importantly, these between-breed-group effects cannot be explained by differential experience because all the dogs tested lived as family pets and had not received any specific training. These results clearly show that there is some genetic variation behind the performance of dogs in these two-way object choice pointing tasks.

Second, environmental factors are also important. For example, deprivation of experience with humans and their behaviour may constrain performance as was shown in the case of shelter dogs (Udell et al. 2008a, 2010). The relatively slight improvement of performance during long periods of development seems to argue against extensive environmental influence (Gácsi et al. 2009c; Riedel et al. 2008). Miklósi and Topál (2011) argued that dogs living in a shelter cannot be regarded as suitable subjects for experiments that aim to test performance in a social task with humans. According to these authors, dogs need proper socialization and social environment for the full development of their socio-cognitive abilities, so sheltered dogs with an unknown and/or a troubled rearing history will most likely underperform those dogs that live in a more natural environment. More specific experience with human gestural communication (e.g. agility training) does not seem to affect the performance in this task either (Gácsi et al. 2009c), which again does not support a theory that explains this skill exclusively by environmental influence. Thus experiential social influence may be very specific and/or may play a role very early in development of dogs. Finally, Hare et al. (2010) carried out a pointing experiment on a larger sample of shelter dogs and found that these subjects performed over the chance level in this task, contrary to the earlier results of Udell et al. (2008a); however their points were repeated four times, and made from a distance of 20cm from the object as against Udell et al.'s 50cm. Hare and colleagues argued also that in the article of Udell et al, statistical analyses were performed erroneously regarding the treatment of 'no-choices' as 'faults'. Hare et al. (2010) re-analysed the data of Udell et al., treating 'no-choices' as a third category besides 'correct' and 'faulty' choices, and doing so, contrary to the conclusions of Udell et al (2008a), found no significant differences between groups (but see Udell and Wynne (2010) for continuing discussion).

Ultimately, when confirming results or conclusions of different experiments, there is often a lack of careful comparison between the effects of the procedure applied. This is very regrettable because the performance of the subjects in these inter-specific communicative experiments is very sensitive to the method used (Miklósi and Soproni 2006). In this specific case there is some evidence that the duration of the gesture, the distance between the tip of the pointing hand (and finger) and the target, and the presence or absence of an accompanying gaze (turning the head towards the target), can each have a strong influence on the performance.

Udell at al (2008a) presented two important claims about the performance of dogs and wolves. First, they argued that the performance of the subjects depends on the testing location (in their study pet dogs performed better indoors than outdoors). Second, they argued that socialized wolves' performance is comparable to that of dogs. Whilst their first finding seemed to contradict our earlier results with dogs (we have never found statistically reliable effect of testing location; unpublished data), the second observation seemed to be problematic because Udell et al. (2008a) introduced a novel form of pointing signal which could have influenced the results.

The so-called "Pointing with clicker" gesture used by Udell et al. (2008a) changed both the form of the human signal and the actual method of testing, which differed substantially from any other previously utilised version of this task. The critical differences are the following: (1) The experimenter maintains her hand in the pointing position even after the subject has started its approach toward the target, making it easier for the dog/wolf to make a choice while the signal is on, whereas in the case of the referred momentary pointing the subject is allowed to move forward only after the hand has returned to the resting position next to the body (Miklósi and Soproni 2006); (2) A correct choice is indicated by a clicking sound produced by the experimenter, whereas this has never been applied by others in this task; (3) The reward is dropped from the (previously pointing) hand of the experimenter after the correct choice has been made, while in all other protocols the dogs have to find the food in a container (there is no direct physical relationship between the emergence of the food and the hand in the testing). These differences led us to consider that the experimental trials with the “Pointing with clicker” gesture may have been easier for the subjects because this type of pointing is more pronounced. Furthermore, the correct choice by the subject is marked with an additional acoustic cue (clicker) and the subject can observe a direct physical connection between the experimenter’s hand and the food reward. This could put the whole paradigm into a different cognitive context; instead of being a communicative interaction (“the food is there”), it may be a case of associative place learning ("the subject learns to go in the direction indicated by the hand which provides food").

In the present study we report the results of three independent experiments in which we re-visit the Udell et al. (2008a) findings. In Experiment 1 we compared a large sample of companion dogs that were tested either outdoors or indoors to find out whether testing location (house or garden) affects performance.

In Experiment 2 we investigated if dogs would perform similarly with “momentary distal pointing” (e.g. Gácsi et al. 2009a) and “pointing with clicker” (Udell et al. 2008a). We hypothesized that dogs may show some variability in their performance with the momentary pointing gesture, but they would perform uniformly well with the “Pointing with clicker” gesture.

Finally in Experiment 3 we tested the possible effect of the individual components of the “Pointing with clicker” protocol as reported by Udell and colleagues (2008a) in separate experimental groups. Our goal was to discover whether application of the clicker, the sustained gesture, or the provision of food directly from the human upon correct choice, improves the performance of dogs.

GENERAL MATERIALS AND METHODS

Different dogs were used in the three experiments. The specific details of the experimental procedure are presented below, while the detailed list of participants is shown in the Appendix.

Subjects

Participation in the tests was voluntary. Subjects were recruited from public dog training schools, where they were attending basic obedience courses. Before the tests we explained to the owners what to do and how to behave during the experiment. There were no specific requirements for participating in the tests but the dogs used had to be older than one year and had to show strong motivation for food. Any dogs which were not motivated strongly by food were not tested (see later, in Pre-training phase).

The owners were requested to fill in a short questionnaire which asked for basic information about their dogs (breed, age, sex, where the dogs were kept at home (outside, inside). For Experiments 2 and 3 we also asked them how often they used a clicker during the training of the dog ('Regularly' (N=18), 'Seldom' (N=12), or 'Never' (N=16)).

Pointing Protocols

At the beginning of each trial the dog was held by its owner by the collar at the start point. The experimenter stood 2.5 m away from them. A plastic bowl (12 cm high, 15 cm wide) was placed on the floor on each side of the experimenter, at 1.5-1.6 m distance from each other, and in equal distance from the experimenter. To mask the possible effect of odour cues, both bowls were smeared inside with a piece of cold cut lunch meat shortly before the tests began. The experimenter stood 20-30 cm behind the imaginary connecting line of the two bowls. All tests were videotaped by continuous, automated recording.

Pre-training phase

This phase served a dual purpose: (a) to familiarize the dogs with the place and the experimental setup; (b) to test whether the subjects were motivated to eat food at the test location. At first we asked the owner to unleash the dog and allow it to explore the experimental site for 1.5-2 minutes. Then the owner moved to the start point, restrained the dog by its collar, and positioned the dog on the start point in front of the experimenter. The experimenter placed the two bowls on the ground. Next the experimenter put a little piece of food into one of the bowls, conspicuously enough so that the dog observed this action. After having dropped the food into the bowl, the owner let the dog free and encouraged it to eat the food. If the dog ate the food from the bowl, then experimenter put another piece of food into the other bowl, and the dog was again encouraged to eat it. This pre-training was repeated once more with both bowls. (Thus two pieces of food were placed one by one into both bowls). Commercially available cold cut lunch meat was used as reward which was previously cut to small cubes (5 mm x 5 mm).

If a dog failed to take food from the bowl and did not eat more than one piece of food during the pre-training phase, we considered it not to be food motivated and we excluded it from the experiment. Only seven dogs had to be excluded for this reason (six in Experiment 2 and one in Experiment 3).

Each specific experiment started right after the pre-training phase. The following types of pointing tests were used in this study:

Momentary distal pointing (MDP) utilised in Experiments 1, 2 and 3 (see also Soproni et al. 2002; Lakatos et al. 2009; Gácsi et al. 2009a). For a sample video of the test, go to: http://www.cmdbase.org/web/guest/play/-/videoplayer/54

Testing consisted of 20 consecutive pointing trials in Experiment 1 and 10 trials in Experiments 2 and 3. (In Experiments 2 and 3 dogs participated in more than one test, thus we lowered the number of trials from 20 to 10 in order to avoid motivational problems in the subjects.) An equal number of pointing trials were performed to the right and the left side. The order of left and right pointing was semi-random: no more than two consecutive pointing trials were performed to the same side (to avoid the development of side bias) and the experimenter did not start the session with two pointing trials to the same side (to avoid the tendency to commit perseverative errors).

At first the experimenter held both bowls in her hands in front of her body, then the experimenter put a piece of food conspicuously into one of them, then she exchanged the two bowls between her hands a few times in order to confuse the dog about the exact location of the food. After this the experimenter crouched down and with stretched arms put the two bowls simultaneously to the floor on her left and right side.

The experimenter stood up and while holding her two hands bent in front of her chest, attracted the dog's attention by calling its name. When the experimenter managed to establish eye contact with the dog, she pointed with extended ipsilateral arm and index finger in the direction of the correct location (the baited pot). The distance between the end of the pointing finger and the bowl was 1 m. The cue was displayed for approximately 1s, and then the experimenter brought her hand back in front of her chest. During the pointing gesture, the experimenter kept looking at the dog. If the dog did not leave the start position for 3s after the pointing gesture was finished, the experimenter repeated the pointing gesture one more time.

It is important to note that the owner kept the dog restrained during the pointing. The dog was released only after the experimenter's hand was again in front of her chest. If the dog approached the baited bowl first it was allowed to consume the food. After this the experimenter quickly picked up both bowls, preventing the dog from examining the other bowl. If the dog visited the empty bowl first, the experimenter did not allow it to examine the other (baited) bowl, but picked both bowls up. After the dog had made a choice and the experimenter had picked up the bowls, the owner called the dog back to the start point and the next trial started.

If the dog did not choose between the two bowls, but for example sat down in front of the experimenter, or went back to the owner, no score was given, but the trial was repeated once. If the dog did not choose again, the trial was recorded as a failure and the next trial started. In the present series of tests no dog failed to choose twice in a row and then continued to choose. However, we had some dogs that stopped choosing altogether, and these were excluded from the analysis.

Momentary cross-pointing (MCP) utilised in Experiment 2 (see also in Lakatos et al. 2009). For a sample video of the test, go to: http://www.cmdbase.org/web/guest/play/-/videoplayer/53

The setup and baiting procedure were exactly the same as used in the MDP test. The only difference was the method of pointing.

The experimenter pointed at the baited bowl as described above, but in this case she used her contralateral arm in relation to the baited bowl. Thus the pointing hand moved in front of her upper body. It should be noted that the experimenter's hand with the pointing finger protruded from her body silhouette on the side where the baited bowl was placed. Because of the configuration of this pointing gesture, the distance between the tip of the pointing finger and the bowl was somewhat further than in the MDP and pointing with clicker tests (about 1.2 m).

Pointing with clicker (PC) (see also Udell et al. 2008a) utilised in Experiment 2. For a sample video of the test, go to: http://www.cmdbase.org/group/user/edit/-/editvideo/38

In this pointing test we followed the procedure of Udell et al. (2008a) as accurately as the description of the methods in the original article made this possible (see further details in the Note). We made only one exception. Udell and colleagues inserted one control trial (in which the subject had to choose a container in the absence of any pointing signal from the experimenter) after every two test trials in their experiment. In these control trials experimenters determined in advance the "correct" choice and the subject was rewarded similarly to the test trials if it approached the 'correct' container. We decided to leave out the control trials in Experiment 2 because otherwise it would have been impossible to compare the performance of different experimental groups given the differences in the number of trials (10 vs. 15). Additionally, the increased number of trials could lead to a different rate of (mental) exhaustion resulting in differences in performance. Finally, the control trials of Udell and colleagues could have a confusing effect, because the subjects were provided with no information in a setup where the expected behavior of the human is giving a cue about whereabouts of the food. The setup was slightly different than the arrangement described above because in this case the same two bowls were turned upside down and they remained on the floor during the whole test. The position of the bowls was the same as in the other tests, as were the positions of the experimenter, the owner and dog.

Before the pointing trials the experimenter performed two pre-training trials at both bowls. The experimenter called the dog's attention and then put a piece of food conspicuously on the top of one of the bowls. When the dog approached the baited bowl and almost touched the food, the experimenter produced a clicking sound with a regular dog training clicker. After the pre-training, ten pointing trials were performed in a similar pseudo-random order as in the other tests.

The dog stood at the start point with the owner. The two bowls were not baited before the pointing. The experimenter called the dog's attention and after eye contact was established, pointed at one of the bowls with stretched arm and pointing finger. A significant note of difference from MCP and PC is that the owner had to release the dog while the pointing was still sustained. The experimenter kept on pointing motionlessly until the dog approached one of the bowls at a distance of about 0.5 m and at that time the experimenter pulled back her arm. Depending on the speed of the dogs, the average pointing gesture lasted 4 s. By using the 0.5 m distance as a threshold for terminating the pointing signal we fulfilled the criteria of the published description from Udell et al. (2008a) for this detail of the method "the experimenter returned to a neutral position before the subject reached the containers" (though subsequent personal communications from those authors show that their threshold for withdrawing the point was 2.5m). If the dog approached the signalled bowl with its snout within 10 cm (i.e. made a correct choice), the experimenter clicked the clicker and dropped a piece of food on the top of the chosen bowl and the dog was allowed to eat the food. If the dog approached the other bowl, the experimenter did not do anything and the owner had to call the dog back and the next trial started.

If the dog did not choose any of the bowls for 10s, (for example sat down in front of the experimenter, and did not move, or went back to the owner), no score was given, but the actual trial was repeated once more. Udell et al. (2008a) did not use any trial repetition. If the dog did not choose again, the trial was recorded as a failure, and the next trial was started. If a dog did not make a choice in three consecutive trials, we excluded the subject from the test. If a dog made three incorrect choices in a row, then Udell et al. (2008a) gave two pre-training trials to ensure that the dog was still motivated to obtain the food. In our experiment this procedure was not needed for any dog.

Momentary distal pointing with clicker (MDP-C) utilised in Experiment 3

The procedure used was exactly the same as described for the MDP above, however a correct choice was indicated also by a clicker. If the dog approached the indicated bowl then the experimenter provided a clicking sound at the moment when it lowered its head into the bowl.

Momentary distal pointing with food reward (MDP-F) utilised in Experiment 3.

The procedure used was exactly the same as described for the MDP above, except that the food was not hidden in any of the bowls, but it was given by the experimenter to the dog upon a correct choice. The experimenter had a piece of food hidden in her hand. She quickly dropped the food to the indicated bowl when the dog lowered its head to the bowl.

Sustained distal pointing (SDP) utilised in Experiment 3

The procedure used was exactly the same as described for the MDP above, except that the owner had to release the dog while the pointing was still displayed. Depending on the speed of the dog, the average pointing gesture lasted 4 s. The experimenter kept on pointing motionlessly, until the dog approached one of the bowls at a distance of about 0.5 m. When this happened, the experimenter pulled back her arm, independently of the correctness of the dog's choice.

Statistical Analyses

If the data deviated from the Gaussian distribution (Kolmogorov-Smirnov test) then we used nonparametric Friedman test with Dunn's post hoc test, Mann-Whitney U test and Wilcoxon signed Rank Test. If the data followed the Gaussian distribution and the error variances were equal across the groups also (Levene test for homogeneity of variance), ANOVA with Bonferroni post hoc test, or one- or two-sample t-test was employed. The proportion of successful dogs was compared among the experimental groups and within the pointing protocols with Fisher's exact tests. An individual was considered as being successful if it was correct 8 times out of 10 trials (binomial test $P<0.055$) or 15 times out of 20 (binomial $P<0.041$). Statistical analyses were performed using SPSS 16.0.

Experiment 1: Do Keeping Conditions and/or the Testing Location affect the Dogs' Performance in Pointing Tests?

Subjects and methods

Two groups of adult companion dogs were tested (outside/inside tested group: N_1=N_2=20) in a session of 20 MDP trials. The mean age for dogs kept outside was 4.13±2.97; and inside was 3.65±2.59. Both groups consisted of hunting dogs of FCI (*Fédération Cynologique Internationale*) breed groups 4, 6, 7, and 8, from a balanced variety of breeds. Half of the dogs in each group were kept in the garden and the other half lived in the garden or house. The groups were balanced for gender and age in both respects (test location and keeping condition). We used only hunting dogs because we wanted a homogenous sample represented by many breeds and from the point of view of both variety and availability, hunting dogs are the largest group of commonly encountered family dogs.

The 'Outside' group were tested in a secluded area of a dog training school which was unfamiliar for the subjects. The test area was chosen so that the actual subject would not be disturbed visually by other dogs or people. The 'Inside' group were tested in an empty experimental room (4 m x 6 m), also unfamiliar for the subjects. During the tests only the dog, the experimenter and the owner of the dog were present.

Results

The performance of both dog groups was significantly better than chance (outside group (t_{19}=7.31, P<0.001) and inside group (t_{19}=5.31, P<0.001), respectively). The success analysed at the individual level was also similar; 9 and 8 dogs out of the 20 subjects were successful in the outside and inside groups respectively (Fisher's exact test, *R*=0.92; *P*=1.00).

Taking into account the place where the dog is kept (house or garden), we analysed the results in a 2-way ANOVA (testing location x keeping condition). These results (see Figure 1) showed neither an effect of testing location ($F_{1,36}$=0.177, *P*=0.677), keeping conditions ($F_{1,36}$=0.055, *P*=0.817) nor an interaction between the two factors ($F_{1,36}$=0.966, *P*=0.334).

This suggests that companion dogs can solve the two-way object choice test independently of their keeping conditions and testing location. They are not disturbed by the relative unfamiliarity of the testing location, even if they are kept at home under different conditions.

Experiment 2: Do Dogs with Low Performance in Momentary Distal Pointing Test show better Performance in the Clicker Pointing Test?

Subjects

Fifty companion dogs from many different breeds were tested. Four dogs had to be excluded because they stopped making choices at various stages of the experiment, so 46 dogs' results were analyzed (18 males and 28 females). Dogs were at least one year old (mean age 4.2 ± 2.5 years SD). The testing locations were the same as in Experiment 1, with approximately the same number of dogs tested indoors and outdoors. (For further details see also Appendix 1.)

Procedure

Each dog participated in three tests, which were performed in the same, fixed order: (1) Momentary Distal Pointing (MDP); (2) Pointing with Clicker (PC); (3) Momentary Cross-pointing (MCP). In each test the dogs participated in 10 pointing trials. After the pre-training phase dogs participated first in the MDP test and we then continued with the PC test without delay. After a break of about 30-35 minutes the testing continued with the MCP test trials. The aim of the MDP test was to make it possible to sort the subjects into the low or high performance group depending on their success.

Experimental groups

Based on their performance in the MDP test, dogs were sorted into two groups. Dogs that were successful (at least 8 correct choices from 10), were assigned to the High Performance Group (N=23). Dogs who chose less than eight times in the MDP test, were sorted to the Low Performance Group (N=23). Accordingly, the high performance group was more successful in the MDP test than the low performance group statistically as well (Mann Whitney U-test, U=0.0; N_1=N_2=23, P<0.001). However, both groups performed above chance level (Wilcoxon test: high performance group T=276.0, N=23, P<0.001, median 8.00; low performance group T=55.0, N=23, P<0.01, median 5.00).

Results

We compared the performance in the MDP, PC and MCP test within the High and Low Performance Groups (see above). In both groups of dogs we found a significant effect of the testing condition (Friedman test: high performance group $K^2_{3,23}$=22.16, N=23, P<0.001; low performance group $K^2_{3,23}$=29.50, N=23, P<0.001, see also Figure

2). Dunn’s post hoc test showed that in both groups dogs performed significantly better in the PC test than in the MDP and MCP tests. Performances in the MDP and MCP tests did not differ significantly in either of the groups. We also compared the number of correct choices between the ‘High’ and ‘Low’ performer dogs in the MDP, PC and the MCP tests, and found difference only in the case of the MDP test (Mann-Whitney U-test: U=0.0, N_1=N_2=23, P<0.001; U=186.5, N_1=N_2=23, P=0.09; U=185.0, N_1=N_2=23, P=0.08 respectively). Note, that this result is not surprising, as the high and low performance groups were formed by sorting the dogs on the base of their performance in the MDP test.

We found that dogs in both groups performed above chance level in the PC and MCP tests (Table 1).

Next, we compared the proportions of dogs in the groups, which made at least 8 correct choices (binomial P<0.055). We did not find significant difference between the high and low performers in the case of the PC tests (Fisher’s exact test, R=1.21; P=0.11), but there were significantly more successful dogs in the MCP test from the ‘High performance’ group than from the ‘Low performance’ group (Fisher’s exact test, R=2.60; P<0.05; Figure 3). Each dog was successful from the ‘High performer’ group in the PC test and more than half of them were successful in the MCP test. ‘Low performer’ dogs also performed well in the PC test (19 out of 23 dogs were successful), but only five of them made 8 or more correct choices in the MCP test.

We also analyzed the possible effect of familiarity with the clicker on the dogs’ performance. First we performed repeated measures Friedman tests within the groups formed on the basis of familiarity with the clicker (‘often’, ‘seldom’ and ‘never’). We found a significant effect of testing condition in each group (‘often’: $K^2_{3,19}$=16.03, N=19, P<0.001; ‘seldom’: $K^2_{3,12}$=12.61, N=12, P<0.01; ‘never’:

$K^2_{3,15}$=17.10, N=15, P<0.001). With Dunn's post hoc test we found that in each group dogs performed significantly better if they received PC, but there was no difference between the MDP and the MCP conditions. Next we compared the performances in the three test conditions according to their levels of experience with the clicker. Kruskal-Wallis tests showed no significant difference in any of the cases (MDP: K^2_3=0.30, P=0.86; PC: K^2_3=0.08, P=0.96; MCP: K^2_3=1.81, P=0.40). The conclusion of these analyses was that the performance of dogs was not affected by their familiarity with the clicker which was used in the PC tests.

The results of this experiment showed that dogs that performed differently in the MDP test all invariably showed high levels of success in the PC test. To some extent the subsequent MCP tests mirrored the original difference because in this test fewer dogs chose above the chance level in the Low Performance Group than in the High Performance Group. Thus our 'Pointing with clicker' (PC) method (closely resembling to the method used by Udell et al., 2008a) was less sensitive to individual differences in dogs than the two other methods of gesturing (MDP and MCP).

Experiment 3: Do the Individual Features of PC have an Effect on the Dogs' Performance Separately?

Subjects

Seventy six companion dogs, from many different breeds, were used as subjects. One dog had to be excluded because it stopped choosing during the experiment. Thus 75 dogs' results were included in the analysis (39 males and 36 females). All dogs were at least one year old (mean age 3.8 ± 2.1 years). The testing locations were the same as in Experiment 1). (For further details see also Appendix 1).

Procedure

Four experimental groups were formed (one of them had two subgroups, see below). Each dog was assigned to one group and participated in two sessions, each consisting of ten pointing trials. In one session all dogs were tested in the MDP test, which was regarded as control. The other session was specific to each group (see below). The order of the two sessions was alternated: half of the dogs started with the MDP test, the other half started with the other test which was assigned to it.

Experimental groups

1. MDP-C Momentary distal pointing with clicker

Two subgroups were formed on the basis of the dogs' familiarity with the clicker (which was assessed by a short questionnaire filled in by the dog owner prior to the test). Subgroup 1 (*N*=15) consisted of dogs with high levels of clicker training, while the dogs in subgroup 2 (*N*=15) never received clicker training.

2. MDP-F Momentary distal pointing with food reward from the experimenter (*N*=15)

Dogs were used in this group without regard to their clicker training experience,.

3. SDP Sustained distal pointing (*N*=15)

Dogs with and without clicker training experience were used in this group.

4. Control group (*N*=15)

In this group dogs received two sessions of MDP. Dogs with and without clicker training experience were tested in this group.

Results

The data in each group followed the Gaussian distribution. Two-way mixed ANOVA was performed, where the experimental group was the between subject factor, and the type of the test (control or treatment) served as repeated (within subject) factor. While there was no difference among the experimental groups ($F_{4,70}$=1.47, P=0.22), and the interaction between the factors was also not significant ($F_{4,70}$=1.26, P=0.30), there was a significant difference between the control and treatment conditions ($F_{1,70}$=4.02, P<0.05). Dogs performed slightly better in some of the treatment sessions (MDP-C for non-clicker dogs; MDP-F; SDP) (Figure 4), while there was no difference between the performances in the first and second ten pointing trials in the MDP-C for clicker trained dogs and in the control (MDP only) group. However, when we compared the performance of the first and second test sessions in each group, we did not find any significant differences (paired t-tests: MDP-C for clicker trained dogs t_{14}=0.15; P=0.88; MDP-C for non-clicker dogs t_{14}=1.94; P=0.07; MDP-F t_{14}=1.10; P=0.29; SDP t_{14}=1.74; P=0.10; MDP (control) t_{14}=0.73; P=0.48.

We performed one-sample t-tests to compare the performance of dogs in each test session to the hypothetical expected value (5 correct choices from 10). The average performance was over the chance level in each session and each group (see Table 2 and Figure 4).

The results of this experiment showed that some of the special characteristics of the PC test (Udell et al. 2008a) can have a significant effect on dogs' performance even if added one by one to MDP. On the other hand, none of them had a substantially greater or lesser effect than the others.

DISCUSSION

The results of Experiment 1 showed that family dogs perform equally well in the human momentary distal pointing test if it is performed indoors or outdoors. Their success is also independent from their living conditions, that is, whether they are kept in the house or in the garden. This is in contrast to what was reported earlier by Udell et al. (2008a), who found that only those dogs which were tested indoors at home performed significantly above chance level. This discrepancy is difficult to explain as by our data the PC method (similar to the one that was originally used by Udell et al.) was a much easier one for any dog in our study, irrespectively of their skills to choose on the more sensitive pointing trials (MDP and MCP). The explanation of the different results between the two papers may lay probably in the larger sample size used in this study. In our experiment only nine and eight of 20 dogs were successful in the 'outside' and 'inside' groups respectively. In the case of Udell et al. (2008a) the sample size was eight in both groups, with two and three successful dogs in the separate experimental groups. The difference in the proportion of successful dogs does not differ significantly between the two experiments in either condition, however in a larger sample the proportion of successful dogs resulted in a significantly above chance group average. We should also mention that approximately half of our subjects in Experiment 2 were tested outdoors with the PC method, therefore we can conclude that this method can be performed indoors and outdoors with the same success rate.

In Experiment 2 we found that there were considerable differences in the performance of individual dogs in the momentary distal pointing (MDP) test. Most importantly, after dogs had been categorized as 'high' and 'low' performers, this difference in success disappeared in the subsequent testing session in which we used the pointing with clicker (PC) test. Furthermore, even if a dog performed at a high level in the PC tests, this experience was not readily transferred to the cross pointing (MCP) tests in which the success rate of the dogs tended to reflect their performance in the MDP test. Although the success of the 'high performer' dogs dropped in the more demanding MCP test, the proportion of successful dogs was significantly higher in this group than among the 'low performer' dogs. In Experiment 3 we found that each of the individual features (duration of the pointing gesture, method of rewarding, providing the clicking sound at correct choice) of the PC test has at least a slight effect on the dogs' performance if utilised separately. This suggests that when these elements are added together in one cueing protocol, they have the potential to improve the subjects' performance.

The present results indicate that the PC method that we employed as a close replication of the protocol used by Udell et al. (2008a) lacks the necessary sensitivity to detect differences in homogenous companion dog populations (i.e. all of our subjects were kept as pets in urban areas). This may be a decisive shortcoming if someone wants to compare wolves and dogs or reveal specific differences within certain population of dogs. It should be noted that even the MDP test utilized by Gácsi and colleagues (2009b) did not reveal differences in the performance of intensively socialized adult wolves and dogs. However, the MDP test was sensitive enough to show that young (4 months old) intensively socialised wolves are generally inferior to dogs of the same age in reading human pointing cues. In addition, the MDP method was also sensitive enough to show a differential effect of head shape and breed working history on utilizing human pointing gestures (Gácsi et al. 2009a). An even more sensitive pointing test may also reveal differences between the performances of intensively socialized adult wolves and adult dogs.

There are several possible reasons why dogs (and wolves) achieve higher performance if investigated in the PC test in comparison to the MDP test. The PC method introduces at least three additional features to the momentary distal pointing. First, the pointing arm is displayed for a longer time, which helps the subjects to attend the gesture and also guides them by local enhancement to the baited bowl. Sustained pointing was found to be more effective in the case of both cats and dogs by Miklósi et al. (2005), and our results in Experiment 3 also support these earlier findings.

Two post-cueing features could have enhanced the dogs' performance through learning. The sound of the clicker may have acted as secondary reinforcement for subjects trained with this instrument. This may have contributed to the high performance of the wolves in Udell et al. (2008a). However, the findings of Experiment 2 showed that there was no difference between the performance of clicker trained dogs and those dogs which have never been trained with this method. Clicker trained dogs had also no advantage in Experiment 3 when they were exposed to the momentary distal gesture which was combined with a clicking sound upon correct choice. Thus the high performance in the PC tests of Experiment 2 cannot be explained solely on the basis of secondary reinforcement in the case of the dogs. It is very likely that using the prior-choice and post-cueing factors together in one test enhanced the success rate of the subjects.

The food dropped by the experimenter could also have enhanced the dogs' performance because they were exposed to a direct relationship between the previously moving stimulus (the hand) and the appearance of the food. Udell et al. (2008b) argued that dogs' superior performance in this task could be the result of forming an association between the human hand and the location of the food. In the PC trials the experimenter's hand moves closer to the bowl during the rewarding and this movement makes the whole act similar to a proximal pointing cue which is considered to be a more effective gesture for inducing correct choice in subjects (e.g. Miklósi et al. 2005; Soproni et al. 2002). Although dogs observed in the PC test in Experiment 2 displayed high performance, adding the feature of dropping the food to the MDP test alone did not lead to more correct choices in Experiment 3. Most likely without the other additional cues of the PC method, dropping the food alone is not enough to raise success rates significantly.

In summary, the 'Pointing with Clicker' method tested here as a close replication of the method used by Udell et al. (2008a) contains several features that may make the task of the subjects easier. Thus one should be cautious when designing comparative tests relying on different experimental protocols because the chosen methodology will to some degree determine the outcome. In relation to the origins of dogs' superior skills in relation to relying on human communicative signals more effort should be taken to make the experiments on different subjects more comparable not just within but across laboratories. It is also likely that both genetic effects and development effects contribute to a variable degree to the communicative skills in dogs. Further, in the specific case the sensitivity of the tests could be improved by making the signals cognitively more challenging.

ACKNOWLEDGEMENTS

This paper was supported by the János Bolyai Research Scholarship from the Hungarian Academy of Sciences, and it was also funded by grants from the European Union FP7-ICT-2007 LIREC 215554, the ETOCOM project (TÁMOP-4.2.2-08/1/KMR-2008-0007) through the Hungarian National Development Agency in the framework of the Social Renewal Operative Programme supported by the EU and co-financed by the European Social Fund and the Hungarian Ministry of Education OTKA K82020. The authors are grateful to Celeste R. Pongrácz for correcting the English language of the manuscript.

[1]

[1] **NOTE**

In the case of Experiment 2, the authors did their best efforts to replicate faithfully the pointing procedure used by Udell et al. (2008a). However, possible minor discrepancies between the two methods were unfortunately impossible to avoid because there was no video material recorded of the original tests of Udell et al. (2008a). The authors are **grateful** for the kind and professional help from Monique Udell, who assessed the video footage of the testing process and the method section of this paper. **Although** personal communication between M. U. and the authors of the present article confirmed that there were **some** discrepancies between the two methodologies, **in the opinion of the authors, the** present paper can be regarded as a replication of the corresponding experiment of Udell et al (2008a).

AUTHOR'S STATEMENTS

The authors state that these experiments comply with the current laws of the Republic of Hungary regarding animal welfare.

The authors also declare that they have no conflict of interest.

REFERENCES

Anderson, J. R., Sallaberry P., & Barbier, H. (1995). Use of experimenter given cues during object- choice tasks by capuchin monkeys. *Animal Behaviour, 49*, 201–208.

Gácsi, M., McGreevy, P., Kara, E., & Miklósi, Á. (2009a) Effects of selection for cooperation and attention in dogs. *Behavioral Brain Functions, 5*, Article Number: 31.

Gácsi, M., Győri, B., Virányi, Zs., Kubinyi, E., Range, F., Belényi, B., & Miklósi, Á. (2009b). Explaining dog wolf differences in utilizing human pointing gestures: selection for synergistic shifts in the development of some social skills. *PLOS ONE 4,* e6584.

Gácsi, M., Kara, E., Belényi, B., Topál, J., & Miklósi, Á. (2009c). The effect of development and individual differences in pointing comprehension of dogs. *Animal Cognition, 12,* 471-479.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science, 298,* 1634-1636.

Hare, B., Rosati, A., Kaminski, J., Brauer, J., Call, J., & Tomasello, M. (2010). The domestication hypothesis for dogs' skills with human communication: a response to Udell et al. (2008) and Wynne et al. (2008). *Animal Behaviour, 79*, E1-E6.

Helton, W. S., Helton, N. D. (2010). Physical size matters in the domestic dog's (*Canis lupus familiaris*) ability to use human pointing cues. *Behavioural Processes, 85*, 77-79.

Herman, L. M., Abichandani, S. L., Elhajj, A. N., Herman, E. Y. K., Sanchez, J. L., & Pack, A. A. (1999). Dolphins (*Tursiops truncatus*) comprehend the referential character of the human pointing gesture. *Journal of Comparative Psychology, 113*, 347-364.

Kaminski, J., Riedel J, Call J, Tomasello M (2005) Domestic goats, *Capra hircus*, follow gaze direction and use social cues in an object choice task. Anim Behav 69:11-18.

Lakatos G, Soproni K, Dóka A, Miklósi Á (2009) A comparative approach to dogs' (*Canis familiaris*) and human infants' comprehension of various forms of pointing gestures. Anim Cogn 12:621-631.

Miklósi Á, Kubinyi E, Topál J, Gácsi M, Virányi, Zs, Csányi V (2003) A simple reason for a big difference: wolves do not look back at humans but dogs do. Curr Biol 13:763-766.

Miklósi Á, Topál J, Csányi V (2004) Comparative social cognition: what can dogs teach us? Anim Behav 67:995-1004.

Miklósi Á, Pongrácz P, Lakatos G, Topál J, Csányi V (2005) A comparative study of the use of visual communicative signals in dog-human and cat-human interactions. J Comp Psychol 119:179-186.

Miklósi Á, Soproni K (2006) A comparative analysis of animals' understanding of the human pointing gesture. Anim Cogn 9:81-93.

Miklósi Á, Topál J (2011) On the hunt for the gene of perspective taking: pitfalls in methodology. Learn Behav DOI 10.3758/s1340-011-0038-2.

Mulcahy, N., Call, J. (2009) The performance of bonobos (*Pan paniscus*), chimpanzees (*Pan troglodytes*) and orangutans (*Pongo pygmaeus*) in two versions of an object choice task. J Comp Psychol 123:304-309.

Reid PJ (2009) Adapting to the human world: Dogs' responsiveness to our social cues. Behav Proc 80:325-333.

Riedel J, Schumann K, Kaminski J, Call J, Tomasello M (2008) The early ontogeny of human-dog communication. Anim Behav 75:1003-1014.

Soproni K, Miklósi Á, Topál J, Csányi V (2002) Dogs' (*Canis familiaris*) responsiveness to human pointing gestures. J Comp Psychol 116:27-34.

Udell MAR, Dorey NR, Wynne CDL (2008a) Wolves outperform dogs in following human social cues. Anim Behav 76:1767-1773.

Udell MAR, Giglio RF, Wynne CDL (2008b) Domestic dogs (*Canis familiaris*) use human gestures but not nonhuman tokens to find hidden food. J Comp Psychol 122:84–93.

Udell, MAR., Dorey, N. R. & Wynne, C. D. L. 2010a. The performance of stray dogs (*Canis familiaris*) living in a shelter on human-guided object-choice tasks. Animal Behaviour, 79, 717-725.

Udell MAR, Dorey NR, Wynne CDL (2010b) What did domestication do to dogs? A new account of dogs' sensitivity to human actions. Biol Rev 85:327-345.

Virányi Zs, Gácsi M, Kubinyi E, Topál J, Belényi B, Ujfalussy D, Miklósi Á (2008) Comprehension of human pointing gestures in young human-reared wolves (*Canis lupus*) and dogs (*Canis familiaris*). Anim Cogn 11:373-387.

Wobber V, Hare B, Koler-Matznick J, Wrangham R, Tomasello M (2010) Breed differences in domestic dogs' (*Canis familiaris*) comprehension of human communicative signals. Interact Stud 10:SI 206-224.

TABLES

Table 1. Experiment 2: Results of one sample Wilcoxon tests on group level performance of dogs in two different tests (Pointing with Clicker and Momentary Cross Pointing). Chance level was 5 correct out of 10 pointing trials in each case.

Group	Session	T	P
‘High performer’	Pointing with Clicker	$T=276$,	$P<0.001$
$N=23$	Momentary Cross Pointing	$T=203$,	$P<0.001$
‘Low performer’	Pointing with Clicker	$T=253$,	$P<0.001$
$N=23$	Momentary Cross Pointing	$T=129$,	$P<0.01$

Table 2. Experiment 3: Results of the one-sample t-tests. The mean number of correct choices was compared to the expected value (5 from 10) in each experimental group. The order of the control and test sessions was balanced in the groups.

Group	Session	*Df, t*	*P*
MDP-C clicker trained	(MDP)	14, 4.43	<0.001
	Test (MDP-C)	14, 3.06	<0.01
MDP-C no clicker training	(MDP)	14, 2.82	<0.05
	Test (MDP-C)	14, 5.91	<0.001
MDP-F	(MDP)	14, 4.83	<0.001
	Test (MDP-F)	14, 5.87	<0.001
SDP	(MDP)	14, 2.80	<0.05
	Test (SDP)	14, 5.21	<0.001
Control (MDP)	(MDP)	14, 4.79	<0.001
	Test (MDP)	14, 3.76	<0.01

FIGURE CAPTIONS

Fig. 1 The effect of keeping and testing locations on the performance of dogs in Experiment 1. All dogs were tested away from home, inside of a building or outside on an open, grassy area. Keeping conditions (‘house’ or ‘garden’) and testing locations (‘in’ or ‘out’) are marked below the bars. The horizontal line shows the level of random choices. All groups performed over the chance level (one sample t-test, $p<0.05$), and their results did not differ from each other (2-way ANOVA)

Fig. 2 The effect of three different pointing tests on the performance of dogs in Experiment 2. Dogs were sorted into the ‘High or Low performance’ group based on their results in the MDP (Momentary distal pointing) test. Asterisks over the box plots mark significant differences between the test results (Friedman repeated test with Dunn’s post hoc test). The results of ‘High’ and ‘Low performance’ dogs were analyzed separately. ‘High performance’ dogs had at least 8 correct choices out of 10 in the MDP test. ‘Low performance’ dogs had less than 8 correct choices in the MDP test. ***: $P<0.001$, **: $P<0.01$

Fig 3 The proportion of dogs that chose correctly in at least 8 trials and which had less than 8 correct choices in Experiment 2. The ratios of these dogs in the PC (Pointing with Clicker) and MCP (Momentary Cross Pointing) tests were compared with pair-wise Fisher’s exact tests. *: $P<0.05$

Fig 4 The performance of the five experimental groups in two 10-trial sessions of pointing in Experiment 3. The control session was always MDP (Momentary Cross

Pointing) test and the testing sessions were different among the groups. The order of the two sessions was balanced in the experimental groups. ‘Clicker’ refers to dogs that are familiar with clicker training, ‘No clicker’ indicates dogs that have never participated in clicker training. Mixed two-way ANOVA for repeated measures found an overall significant effect of ‘treatment’, where the dogs performed better in the ‘test’ groups than in the control trials. There were no significant differences between the control and test sessions within groups (separate ANOVAs for repeated measures). Each group exceeded significantly the level of random choices (one-sample *t* tests). The horizontal line shows the level of random choices. MDP-C = Momentary distal pointing with clicker; MDP-F = Momentary distal pointing with food reward from the experimenter; SDP = Sustained distal pointing.

APPENDIX

Participants of Experiment 2 and 3. Participation in a specific experiment is marked with ‘x’ after the dogs’ name.

Abbreviations: M=male; F=female; H=kept in the house; G=kept in the garden; H/G=kept in the house and in the garden; experience with the clicker: ‘yes’=regularly; ‘no’=never; ‘no+’=seldom

Experiments: 2=Experiment 2; 3a=Momentary Distal Pointing with Clicker for clicker trained dogs; 3b=Momentary Distal Pointing with Clicker for dogs with no clicker experience; 3c=Momentary Distal Pointing with food dropping; 3d=Sustained Distal Pointing; 3e=Momentary Distal Pointing (control)