

2013 *Applied Animal Behaviour Science*, 149: 45-54.

Association between subjective rating and behaviour coding and the role of experience in making video assessments on the personality of the domestic dog (*Canis familiaris*)

Erika Mirkó¹, Antal Dóka¹, Ádám Miklósi^{1,2}

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

²MTA-ELTE Comparative Ethology Research Group

Corresponding author: Erika Mirkó, E-mail address: erika.mirko@gmail.com

Abstract

In this paper our first aim was to investigate association between behaviour coding based on a test battery called the FIDO Personality test for dogs (*Canis familiaris*) and a subjective rating given by dog owners. For the latter we asked dog owners to fill in the Dog Personality Questionnaire (DPQ) which established four personality traits for dogs: *Stranger-directed Sociability, Activity, Aggressiveness and Trainability*. The FIDO Personality test consists of a series of subtests in which we utilised an ethogram and measured different aspects of dog behaviour. Prior to this study we established a set of possible associations between the behavioural and the questionnaire variables.

With regard to *Trainability* no correlation could be detected between any of the questionnaire variables and the hypothetically corresponding test battery variables. In the case of *Stranger-directed Sociability*, we revealed correlation between owners' reports and the dogs' behaviour during the test situation. Those dogs which were reported 'initiative' and 'not mistrustful with unfamiliar humans' approached the experimenter with shorter latency ($r = -0.25, p < 0.01$; $r = 0.24, p < 0.01$) and played more with her than those who were scored lower on these questionnaire items ($r = -0.27, p < 0.01$). For the factor *Activity*, no correlation was

found between the owners' report and the behaviour of dogs displayed in the Spontaneous activity phase of the test battery. Association was revealed between the owners' report regarding the dogs' preference for ball games and playing and their behaviour in the corresponding subtest. Those dogs which were reported to like ball games and like playing with the owner played with the ball as well as with a tug more in the test situation than those which were scored lower on the questionnaire variable 'likes fetching balls' ($r = 0.47, p < 0.001$; $r = 0.25, p < 0.01$). As regards the factor *Aggressiveness*, the questionnaire variable 'shows a tendency to bark' correlated with the dogs' aggressive reactions (growl, bite, attack) ($r = 0.25, p < 0.01$) in the Bone take-away subtest and with the 'latency of getting the DNA sample' ($r = 0.30, p < 0.001$) in the Getting DNA sample subtest. The results show that despite of the investigated associations there is a need for more effort to refine both questionnaire-based and behaviour-based assessment of dog personality.

The second aim of our experiments was to reveal whether owner ratings on the above personality factors converged with the ratings of independent groups of observers based on the video recordings of the FIDO Personality test. First, in the case of each factor, we compared the degrees of ratings given by the owners and by the three rater groups with different levels of hands-on experience with dogs. We found no difference in the case of *Stranger-directed Sociability*, *Activity* and *Trainability* ($F(3,128)=1.17, p=0.17$; $F(3,128)=0.99, p=0.4$; $F(3,128)=1.6, p=0.18$ respectively) but there was a significant variability in judging *Aggressiveness* ($F(3,128)=2.86, p=0.04$). In the case of *Stranger-directed Sociability*, *Activity* and *Aggressiveness* the owner's report correlated with the assessments of all the three rater groups and the assessments of the rater groups correlated with each other as well. As regards *Trainability*, the owners' report did not correlate with the assessments of any of the rating groups, but the assessments of all the rating groups correlated with each other.

Our investigations provide a contribution to the field of research on comparing behaviour coding based on test batteries and subjective rating based on questionnaire studies as methods for establishing the personality factors of individual animals.

Key words: dog; personality; test battery; coding; subjective rating

1. Introduction

The last few decades have witnessed the refreshing re-emergence of the study of dog behaviour. There are two main methods to record information about the behaviour of individual animals. One is behaviour coding, which is based on units of behaviour (ethograms: Martin and Bateson, 2007; Lehner, 1996), the other one is subjective rating of behavioural tendencies on ordinal scales (e.g. Wemelsfelder et al., 2001; Gosling et al., 2003; Ley et al., 2008; Kubinyi et al., 2009).

In the case of behaviour coding, the occurrence of discrete, well-defined behaviour units without reference to their function are recorded on a ratio scale in a behavioural test battery to capture in a quantitative way (overall duration or frequency) what an animal does on that particular occasion. Researchers, for example, may count how many times the target animal lifts its leg etc. This method is thus considered objective, unaffected by the observers' perceptual and interpretational biases (Gosling, 2001).

Subjective ratings focus on the aggregated character of animal behaviour and observers pay little attention to small, specific behaviour units. There are two types of ratings, behaviour rating and adjective rating, the latter is also referred to as trait rating (Uher et al., 2008). The items of behaviour rating describe actions and the observers make frequency assessments (e.g. „gets frightened by noises” can be rated from ‘rarely’ to ‘often’). The items of adjective rating comprise adjectives only, which summarize a range of behaviours (e.g. an aspect of an animal’s behavioural history is condensed on a scale ranging from ‘unaggressive’ to ‘aggressive’). Since both behaviour and adjective rating rely on the subjective judgement of the observers, they are sometimes considered less reliable than behaviour coding (e.g. Stevenson-Hinde and Zunz, 1978; Gosling, 2001).

However, examination of inter-observer agreement, internal consistency and test-retest reliability have proven that the reliabilities of subjective ratings can be strong (for a review see Weinstein et al., 2008) and some even argued subjective ratings are more reliable than behaviour coding (Vazire et al., 2007). Though both methods provide grounds for considering that they are valid, few studies have studied the possible correlations between subjective rating and behaviour coding (for a review see Gosling, 2001). In a recent study Konok et al. (2011) constructed a questionnaire to collect information from owners about their dog’s separation behaviour and in parallel coded the dogs’ separation-related behaviour in a behavioral test. Their results show that owners provided a realistic picture about their dog’s reactions in a separation situation.

To our knowledge, there has been few empirical comparisons of the two methods regarding the personality of individual dogs (e.g. Svartberg 2005). By definition personality refers to the characteristics of individuals which describe and account for consistent patterns of behaving, feeling and thinking (Jones and Gosling, 2005). If both behaviour coding and subjective rating as methods are reliable to the same extent, the rating and coding measures of

the traits reflecting the same underlying factors must converge (Vazire et al., 2007). For instance, if the owner rates the dog as highly noise sensitive, the dog should avoid an object producing a loud noise in a behaviour test. To study convergence between behaviour coding and subjective rating we evaluated the correlations between a behavioural test (FIDO Personality test, Bröder et al., submitted) and a questionnaire (Dog Personality Questionnaire, Mirkó et al., 2012).

A number of researchers argue that during a test battery the behaviour of an animal in reaction to a particular stimulus can be under the influence of various factors like the presence of other dogs around the test scene etc. and the inner state of the target animal can also change during the course of the test situations (e.g. Jones and Gosling, 2005). The behaviour of the target animal in a test battery thus might not accurately reflect a particular aspect of a dog's character (e.g. Jones and Gosling, 2005). If a dog, for example, does not typically bark a lot, but in a test scene keeps barking at its conspecifics, the animal's performance in a behaviour test may substantially differ from its manners at home. In this case it might happen that no association can be detected between owner's report and behaviour coding not because the former assesses behaviour in a qualitative while the latter does so in a quantitative way, but because of the fact that behaviour assessments based on test batteries lack the aggregated character of questionnaire studies (Vazire et al., 2007). In order to investigate correspondence between the dogs' behaviour in the test battery and owners' report with the exclusion of any potential divergences resulting from the qualitative vs quantitative quality of the two methods, we carried out a second experiment. Independent observers unfamiliar with each dog were asked to watch the video clips of the FIDO Personality test (Bröder et al., submitted), then rate the animals by filling in the same questionnaire the owner had completed. In this way the data sets we gained only differed in the amount of knowledge the assessors had about each dog.

Few studies have examined the influence of the assessors' previous experience on subjective ratings of dogs based on video recordings of test batteries (Tami and Gallagher, 2009). To investigate the potential impact of prior experience, the other aim of our second study was to compare the assessments of three rater groups each having different levels of hands-on experience with dogs. The three groups comprised dog trainers, independent dog owners and people who had never kept a dog.

Our investigations follow the direction suggested by Jones and Gosling (2005), namely that future research on the personality of individual dogs (*Canis familiaris*) should focus on the comparison of different methods.

2. Methods

2.1 Study 1: The comparison of the questionnaire study and behaviour coding

2.1.2 Subjects

We collected 100 dogs, carried out the FIDO Personality test (Brüder et al., submitted) with all of them and asked the owners to fill in our Dog Personality Questionnaire (Mirkó et al., 2012). In this way we obtained both the video recordings of the test battery and the completed questionnaires of the 100 dogs. The protocol of the FIDO Personality test is described in Appendix II. The dogs and their owners were recruited on voluntary basis to participate in our Family Dog Research Programme. The sample comprised 27 breeds registered in the FCI (Fédération Cynologique Internationale) breed groups and six mongrels, for further details see Appendix I. The sample consisted of 42 males, 48 females, 4 neutered males and 6 neutered

females. The mean age was 3 years ($SD=2.00$) Thirty-seven dogs took part in organized obedience training, sixty-three did not.

2.1.3. Procedure

The behavioural tests were carried out in Budapest, in the park next to the building of Eötvös Lóránd University, at a holiday camp for dogs and their owners in Dombrád, Hungary, at a festival organized for owners of beagles in Tatabánya, Hungary and in Debrecen, Hungary. The behavioural tests were conducted by three female students, including the first author (E.M.), and a researcher working for the Department of Ethology at Eötvös Lóránd University. Prior to conducting the test batteries, all the experimenters were trained on over 20 dogs how to carry out the FIDO test until the analysis of the video recordings showed that all of them conducted it meticulously according to the protocol. The 20 dogs the experimenters practised on were not included in the present study. The subjects were tested individually with no other dogs nearby. Three participants, the dog, the owner and the experimenter took part in the tests. The behaviour of the dogs was recorded from the side by a camerawoman from a distance of about 8 metres.

Collecting ratings by means of the questionnaire

Before the behaviour test, each owner was asked to fill in the Dog Personality Questionnaire (DPQ) (see Mirkó et al., 2012). On the basis of the questionnaire Mirkó et al. (2012) established four personality factors which included the following adjectives; Stranger-directed Sociability: brave, not afraid of noises, not mistrustful, initiative; Activity: not lazy, likes fetching balls, likes games of fight, overactive; Aggressiveness: likely to bite a human,

jealous of dogs, hysterical, shows a tendency to bark, retaliative; Trainability: learns things easily, likes pleasing the owner, controllable, not disorganised. Statistical analysis showed good internal and external consistency of these factors and correlation analysis confirmed that they were unrelated dimensions of dog personality (Mirkó et al., 2012).

Behaviour coding via video recordings

The videotapes of the test batteries were coded with Solomon Coder beta 30 June 2011 (Copy right András Péter <http://solomoncoder.com>). The behaviour coding was done by the first author. Twenty per cent of the videos (N=20) were also coded by a second observer. Inter-observer reliability was determined by computing Cohen's Kappa coefficients between the two observers. The reliability can be considered very good as Cohen's Kappa ranged from 0.78-1.0. Table 1 shows the coded behaviour variables hypothesized to measure the same constructs as the adjective items (see above) loading high on the personality dimensions identified by our DPQ.

2.1.4 Hypothesised correspondences between subjective rating by the owner and behavioural coding

As has been mentioned above, in Study 1 we wanted to establish association between subjective rating and behavioural coding. Based on preliminary experience and literature survey (see for example, Svartberg 2005), we made some specific predictions with regard to these associations.

Stranger-directed Sociability consists of the questionnaire items: 'brave', 'not afraid of noises', 'not mistrustful with unfamiliar humans' and 'initiative'. We assumed correlation

between the dog's reaction (moves the head backwards; steps back; runs away) to the suddenly unfolding umbrella opened by the experimenter and the owner's report regarding the traits 'brave' and 'not afraid of noises', because the sudden appearance of the umbrella (*Sudden appearance test*) was hypothesised to provide a strong visual stimulus supposed to frighten a dog and its opening was also accompanied by a noise (Svartberg, 2005). The traits 'initiative' and 'not mistrustful with unfamiliar humans' were assumed to show association with the latency of approaching the experimenter ('stranger') in the *Greeting test* and the duration of playing with the experimenter. Both of these traits and test battery variables were supposed to be related to dogs' contact seeking behaviour towards unfamiliar humans (see „social impulsivity” in Héjjas et al., 2009).

The personality factor *Activity* consists of the following questionnaire items: not lazy, likes fetching balls, likes games of fight, overactive. On the basis of previous findings (e.g. Vas et al., 2007) our hypothesis was that the trait 'overactive' and 'not lazy' will correlate with the activity level of the dog observed in the *Spontaneous activity test*. We also supposed positive association between the trait 'likes fetching balls' and the duration of 'following the ball' and 'keeping the ball in the mouth' (*Ball-playing with the owner*).

The personality factor *Aggressiveness* consists of 'likely to bite a human', 'jealous of dogs', 'hysterical', 'shows a tendency to bark' and 'retaliative'. The questionnaire variables 'jealous of dogs' and 'retaliative' were excluded from the analysis, since the FIDO test did not involve any tests on the basis of which these traits could be rated.

The questionnaire items 'likely to bite a human' was assumed to correlate with the responses of the dogs in all the three tests in which we examined the animals' potential aggressive behaviour in the following three subtests: *Getting DNA sample*, *Bone take-away test* and *Threatening approach*. When getting the DNA sample, the experimenter puts her hand into the mouth of the dog and wipes its lower molars. This process may trigger

ambiguity and/or fear in dogs. Since a number of authors reported convergence between fear and aggression (e.g. Guy et al., 2001; Overall, 2001; O'Sullivan et al., 2008) we also expected positive association between the owners' report and the dogs' reactions.

Food guarding behaviour can also manifest in aggressive tendencies in a dog when the experimenter is trying to get the bone in the *Bone take-away test* (Coppinger and Coppinger, 2001). Finally, based on the results of Vas et al. (2008), the *Threatening approach* test was also thought to evoke aggressive behaviour in the subjects.

We presumed convergence between the trait 'hysterical' and the dogs' behaviour during getting the DNA sample. We assumed that in the case of those dogs that were scored high on this trait the process of getting the sample will last longer, because moving the wiper in the animals' mouth may cause them distress. Finally, on the basis of the study by Pongrácz (2010), the trait 'shows a tendency to bark' was expected to correlate with the dogs' behaviour in all the three subtests: *Bone take-away test*, *Threatening approach* and *Getting DNA sample*.

Trainability was characterised by the phrases 'learns things easily', 'pleases the owner', 'controllable' and 'not disorganised'. We presumed that if a dog is scored high by the owner on the trait 'learns things easily', it may have a short latency of obtaining the piece of meat in the *Problem-solving test*. We also supposed positive association between the trait 'pleases the owner' and the duration of playing interaction in the *Ball-playing with the owner test*, since in the test situation it was the owner's task to initiate play with the dog. Thus those dogs that were scored high on this trait were expected to play with the owner for a long period of time. Previous results suggest that the owner's success in laying down a dog on the side is related to the ability of the owner to enforce his or her will on the dog. (Netto and Planta, 1997; Rooney and Bradshaw, 2002, 2003). For this reason, the trait 'controllable' was hypothesised to correlate with the latency of the owner's success to make the dog lie down

(*Lying down on the side test*). Finally, as regards the questionnaire item ‘not disorganised’ positive association was assumed with the duration of watching the swinging sausage in the *Attention test*.

2.2 Study 2: The comparison of the owners’ subjective ratings and the video-based assessments using the DPQ

2.2.1 Subjects

As has been described above, we had both the video recordings of the FIDO Personality test and the DPQ completed by the owners of 100 dogs. Of these, we used the data of 33 beagles to compare the owners’ subjective ratings and the video-based subjective ratings of naïve observers. We chose to use this breed, because it represented the biggest pure-breed subgroup in our sample and it was considered important that raters should assess dogs belonging to the same breed-group so that any bias regarding breed typical stereotypes could be prevented. The sample comprised 18 males, 13 females and 2 neutered females. The mean age was 3.8 years ($SD=2.28$). Eight dogs took part in organized obedience training, twenty-five did not.

2.2.2 Procedure

The recorded personality tests of the thirty-three beagles were shown to three groups of observers with three different levels of hands-on experience with dogs. The first group of raters comprised 33 dog trainers whom we called ‘trainers of non-included dogs’ the second one 33 owners of family dogs, whom we called ‘owners of non-included dogs’, while in group three there were 33 observers who had never kept a dog before, they were termed ‘non-

owners'. Working with as many observers as possible reduces measurement error because the errors due to the systematic idiosyncracies of an observer can be eliminated (Block, 1961) and the effect of pseudoreplication can be avoided (Kroodsma et al., 2001). For this reason, all the raters saw only one dog's FIDO Personality test and on the basis of that filled in the Dog Personality Questionnaire. Each dog was assigned randomly to each rater. None of the raters knew any of the dogs and any of the owners and they did not know each other either so that no information could be exchanged about any of the target dogs or the behaviour test itself. Each rater saw the video clip of the behaviour test in a room alone and only once, then filled in the questionnaire right after seeing the recording. In this way, together with the owners' reports, 4 different assessments of each dog were made.

2.3 Data analysis

We used SPSS 13.00 for the statistical analyses.

Study 1. The distribution of the questionnaire variables and that of the behaviour coding was examined with the Shapiro-Wilk test. Since some of the items of the behaviour variables were not normally distributed, Spearman correlation was applied to investigate the associations between the corresponding variables of the personality dimensions and the behaviour coding (see above).

Study 2. The questionnaire variables 'jealous of dogs' and "retaliative' on the factor *Aggressiveness* were excluded from the analysis, because the FIDO test did not involve any tests on the basis of which these traits could be rated. To compare owners' reports and the assessments of each rater group with regard to each factor, the data was analysed with one-way ANOVA and *post hoc* test was done with Dunnett's pairwise comparison. Spearman

correlation was used to investigate the associations between the four sets of assessments given by owners, trainers of non-included dogs, owners of non-included dogs and non-owners.

3. Results

3.1 The association between the owners' subjective rating and behaviour coding

The four factors established by Mirkó et al. (2012) were examined one by one according to the predicted associations described above.

In the case of *Stranger-directed Sociability* the owners' rating on the variable 'not afraid of noises' did not correlate with any of the variables describing the dogs' reactions to the suddenly opening umbrella.

Correlation was found between the questionnaire variable 'initiative' and 'latency of approaching the experimenter' ($r = -0.25$, $p < 0.01$) and the 'duration of playing with the experimenter' ($r = 0.24$, $p < 0.01$). The questionnaire variable 'not mistrustful with unfamiliar humans' correlated negatively with the personality test variable 'latency of approaching the experimenter' ($r = -0.27$, $p < 0.01$).

In the case of the factor *Activity* no correlation was found between the owner's report and the behaviour of dogs displayed in the Spontaneous activity phase of the test battery. Relatively strong correlation was revealed between the questionnaire variable 'likes fetching balls' and the 'duration of following the ball' ($r = 0.47$, $p < 0.001$). The personality test variable 'duration of keeping the ball in the mouth' correlated with the questionnaire variables 'likes tug of war' ($r = 0.25$, $p < 0.01$) as well as with 'likes fetching balls' ($r = 0.39$, $p < 0.001$). The personality test variable 'duration of playing with the owner' correlated with the

questionnaire variables 'likes tug of war' ($r = 0.27, p < 0.01$) and with 'likes fetching balls' ($r = 0.43, p < 0.001$).

As regards the factor *Aggressiveness*, the questionnaire variable 'shows a tendency to bark' correlated with the dogs' aggressive reactions (growl, bite, attack) in the bone take-away subtest ($r = 0.25, p < 0.01$) and with the 'latency of getting the DNA sample' ($r = 0.30, p < 0.001$). No further correlations could be highlighted.

With the factor *Trainability*, no correlation could be detected between any of the questionnaire variables and the hypothetically corresponding test battery variables.

3.2 The comparison of the owners' ratings and the video-based ratings done by humans with different levels of experience with dogs

First, in the case of each factor, we compared the degrees of ratings done by the owners and by the three rater groups with different levels of hands-on experience with dogs. We found no difference in the case of *Stranger-directed Sociability*, *Activity* and *Trainability* ($F(3,128)=1.17, p=0.17$; $F(3,128)=0.99, p=0.4$; $F(3,128)=1.6, p=0.18$ respectively). However, there was a significant variability in judging *Aggressiveness* ($F(3,128)=2.86, p=0.04$). Dunnett's post hoc comparisons showed that all the three video-based rater groups characterised the dogs to have more aggressive tendencies ($p<0.05$) than the dogs' owners (Figure 1).

In the case of *Stranger-directed Sociability*, *Activity* and *Aggressiveness* the owner's report correlated with the assessments of all the three rater groups and the assessments of the rater groups correlated with each other as well. As regards *Trainability*, the owners' report did not correlate with the assessments of any of the rating groups, but the assessments of all the rating groups correlated with each other. For the correlation matrix see Table 3.

4. Discussion

4.1 Association between owners' subjective rating and behaviour coding

The first aim of this article was to compare subjective rating and behaviour coding focusing on the personality dimensions of individual dogs.

In the personality factor *Stranger-directed Sociability*, the trait 'initiative' was defined as „Initiates contact with unfamiliar humans, when he fails to make the owner tender him/her, he/she tries to do the same with a nearby unfamiliar human.” The dogs' behaviour in the FIDO test regarding contact seeking with the unfamiliar experimenter is in line with the owners' report, since the higher a dog was rated on the questionnaire item 'initiative', the shorter the latency of approaching the experimenter was in the Greeting test. In a similar vein, those dogs who were scored high on 'initiative' also played more with the experimenter. These findings, in line with the investigations of Svartberg and Forkman (2002) suggest that the latency of approaching an unfamiliar human and the time ratio a dog spends playing with an unfamiliar experimenter in a test situation can provide reliable information about a dog's behavioural profile concerning its attitude towards unacquainted humans in a non-threatening situation.

Despite our hypothesis, the dogs' reaction to the suddenly opening umbrella was not found to be associated with the questionnaire items 'not afraid of noises' and 'brave'. The lack of association could be due to the fact that the visual stimulus the umbrella provided was far more dominant than the loudness of the noise the opening umbrella made. In the DPQ (Mirkó et al., 2012) 'brave' was defined as 'He/She does not get frightened easily by different objects, animals or humans, is confident even in new environments.' This may explain why there was a lack of association. The opening of the umbrella was a quick and unexpected

event even confident animals might get frightened by. In other words, as Vazire et al. (2007) put it, behaviour codings are very sensitive to the situational variation in an animal's behaviour, while a human observer, for example, an owner is able to discount the situational influences on a dog's behaviour when rating the animal.

The factor *Activity* consists of the variables 'overactive' (He/She is quite lively, moving about in all circumstances, not preferring sitting or lying for long), 'not lazy' (Often lies down), 'likes games of fight' and 'likes fetching balls'. Despite our hypothesis (based on Vas et al., 2007) and more recent positive findings by the means of the FIDO test (Kubinyi et al., 2012) we did not find activity-related correlation between the DPQ and behaviour scoring. The lack of association could be due to the fact that in the Spontaneous activity phase the dogs were held on their leash for 60 seconds at an unfamiliar test scene. Diverio et al. (2008) pointed out that in novel places dogs are more cautious and less self-confident than in their own territory. Besides that, during the test, the animals were only observed for a relatively short time and not 'for long', as can be read in the definition of the questionnaire item.

As regards the questionnaire item 'likes fetching balls' defined as „Likes ball games very much”, moderate correlation was found with both test battery variables 'the duration of running after the ball' and 'the duration of keeping the ball in the mouth'. Thus dogs' preference for ball games can be predicted on the basis of their performance during a ball game in a test situation. The same seems to be applied to playing tug of war. In the definition of the questionnaire variable 'likes games of fight' playing tug of war was given as an example. The fact that it correlated with the duration of playing tug of war with the owner indicates that dogs' propensity for this type of game can reliably be measured by a test battery. Moreover, in contrast to the findings of Tóth et al. (2008), the association highlighted between the trait 'likes fetching balls' and the 'duration of play with the owner' as well as between 'likes games of fight' and 'ball in mouth' hints that a dog's tendency to play ball

games and its inclination to play tug of war might also be related. This suggests the presence of a „playfulness” factor (see Svartberg and Forkman, 2002) in the personality of the dog.

In the case of *Aggressiveness*, ‘hysterical’ and ‘likely to bite a human’ did not correlate with any of the predicted behavioural variables, while ‘shows a tendency to bark’ was found to be associated positively with the dogs’ reactions to the experimenter’s attempt to take the bone away and also with the long latency of getting the DNA sample. The former finding is in line with other studies suggesting that barking is often related to aggressive inner state (Netto and Planta, 1997; Horváth et al., 2007; Tami and Gallagher, 2009; van den Berg et al., 2010). When taking the DNA sample, the experimenter puts her hand into the mouth of the dog to wipe the lower molars and the dogs try to divert their heads from the experimenter’s hand. This process may evoke ambiguity and fear in dogs (who may be fearful in other social contexts too). Thus the positive correlation between the latency of getting the DNA sample and the dogs’ tendency to bark appears to be in accordance with previous studies finding that despair and fear related behaviours are also accompanied by barking (Molnár et al., 2006; Pongrácz et al., 2010).

It is important to note, however, that an observed behaviour (e.g. growling or biting) can only be regarded a characteristic trait if it is consistent and shows reliable stability over time (Vas et al., 2005). Owing to uncontrolled changes in the social environment, aggression related actions are not consistent (Goddard and Beilharz, 1985; Svartberg et al., 2005). Aggressiveness measured in a test battery shows low association with owner’s report (Svartberg, 2005), and only shows relatively high correspondence with the biting history of dogs in specific test series of aggression (Netto and Planta, 1997; Planta and De Meester, 2007). In addition, territorial and intra-specific aggression might not be reliably measured by testing (Christensen et al., 2007).

The factor *Trainability* consists of the questionnaire variables ‘not disorganised’, ‘learns things easily’, ‘fond of pleasing the owner’ and ‘controllable’. Neither of our hypotheses was confirmed (see above). In the case of the trait ‘disorganised’ and the dog’s persistence to concentrate on the swinging sausage the lack of correlation might be due to the fact that this test might not be apt for detecting whether a dog is disorganised or not. At the beginning of this test the dogs attempt to catch the swinging sausage, but the experimenter lifts it up to prevent it from being eaten. After the unsuccessful attempts the animals appear to lose interest in the sausage and start looking at the owner and/or the experimenter or at any other stimulus coming from the nearby environment. The short duration of the dogs’ concentration on the food thus does not seem to reflect that the animals are disorganised but rather that they lost interest. Further, Topál et al. (1997) found that dogs’ poor problem-solving performance was less dependent on their cognitive abilities than on their dependent relationship with the owner. We could assume that the dogs’ relationship with the owner masked any association between the questionnaire items of *Trainability* and their performance in the test.

The lack of positive association between the duration of playing with the owner and the dogs’ scores on the questionnaire item ‘fond of pleasing the owner’ might also root in the fact that the test battery and the questionnaire variable measure two different things. The definition of ‘fond of pleasing the owner’ was: “Repeats an activity the owner is very pleased about several times and expects praise for that.” However, playing tug of war game with the dog may not be among the everyday activities the owner is pleased about and praises the dog for. Moreover, Tóth et al. (2008) found that the behaviour of family dogs in a playing situation is more influenced by their motivation to play than by the identity of the play partner and their tendencies for cooperative behaviour.

Previous studies suggest that the owner's success in laying down a dog on the side is related to the owner's assertiveness over the dog (Netto and Planta, 1997; Rooney and Bradshaw, 2002, 2003). The definition of our questionnaire item 'controllable' was „Returns right away when being called”. This item content thus implied that those owners whose dogs were scored high on this trait were supposed to be more assertive over their pet, thus the latency of rolling the pet over on the side would be short in the FIDO test. The lack of associations suggests that following the owner's commands regarding learned activities is not associated with a dog's willingness to take up a body posture in which he/she will be in a defenceless position.

Though the correlations found are statistically significant, they seem to be relatively low (for comparable level of correlations see also Gosling et al. 2003). A possible explanation for that can be that subjective ratings focus on the aggregated character of animal behaviour. In contrast, in the case of behaviour coding, the occurrence of discrete, well-defined behaviour units without reference to their function are recorded. The low r-values thus might reflect the qualitative vs quantitative character of the two methods.

Another possible explanation for the apparently low correlations may be that behaviour can vary due to, for instance, pre-test condition, daily fluctuations, or seasonal effects. For a researcher aiming to assess the personality traits of an individual dog, this variance represents noise. Questionnaire studies reduce this variability, since an owner can discount situational influences on behaviour when making his/her ratings. As opposed to that, an observer making behaviour codings does not take the potential situational factors (e.g. unfamiliar test scene) into account (Vazire, 2007). Data collected via questionnaires are thus aggregated across all the times, while behaviour codings lack the aggregated character of observation. That results in a greater sensitivity to the situational variation in a dog's behaviour over time (Vazire, 2007). The situational variance can obscure the cross-situational

component of behaviour which may result in low r-values when trying to find associations between trait ratings and behaviour codings. Svartberg (2005) also investigated association between owner's report using the CBARQ (Canine Behavioral Assessment and Research Questionnaire) and the DMA (Dog Mentality Assessment) behaviour test. Similarly to our findings, in Svartberg's study, the values of the significant correlations were also low, and the average value of the significant correlations was even lower than in the present study.

Since both questionnaire studies and behaviour codings have their advantages, the personality of an individual dog might be best investigated by the co-application of both of them. As has been mentioned above, Svartberg (2005) aimed to establish association between owner's report using the CBARQ (Serpell & Hsu, 2003) and the DMA behaviour test. As opposed to his expectations, chase-proneness was found to be related to play interest and non-social fear in everyday life, and not, as expected, to predatory behaviour (Svartberg, 2005). The present analysis highlighted in the case of which trait the data collected by the FIDO Personality test and the Dog Personality Questionnaire converge. The lack of the correlations between the hypothesised variables provide cause for concern for future researchers to be wary of using one or the other method of personality assessment uncritically and make them work on making both methods more reliable.

4.2 The comparison of the owners' ratings and the video-based ratings done by humans with different levels of experience with dogs

The second aim of this study was to investigate correspondence between owners' ratings and independent observers' ratings on the dogs' behaviour in FIDO Personality test (Brüder et al., submitted).

Our findings are in harmony with previously published results. Studies carried out on cats (Feaver et al., 1986) and pigs (Wemelsfelder et al., 2001) revealed a high level of correspondence in describing animals' modes of behaviour by human observers. In relation to dogs, Simpson (1997) found that humans were able to interpret the meaning of a number of communicative signals given by dogs. Bahlig-Pieren and Turner (1999) reported humans' ability to recognize facial expression of fear and curiosity in dogs with an agreement of over 80%. It is important to note, however, that hands-on experience with a particular species can improve the interpretation and prediction of an animal's behaviour (Serpell, 1986; Bahlig-Pieren and Turner, 1999; Irvine, 2001). Using Free Choice Profiling (FCP) methodology for the qualitative behaviour assessment of emotional expression in dogs Walker et al. (2010) also found good consensus between the observers despite of one group of judges who had more extensive experience with dogs.

In the second experiment the correlation coefficients were substantially higher than in the first one. This suggests that the three independent observer groups' subjective assessment of the target dogs' character may reflect the owners' experience more than behaviour coding in the case of *Stranger-directed Sociability*, *Activity* and *Aggressiveness*.

The reason for this, as discussed in the case of the first experiment as well, might lie in the fact that measuring behaviour in a qualitative vs quantitative way may affect the strength of the correlations. The fact that none of the independent observers were familiar with any of the target dogs in the video clips, consequently, could only do their ratings on the basis of the test battery, suggests that trait ratings are heavily affected by the applied research tool. Our findings in the second experiment thus seem to support what we discussed in the case of study one: trait rating and behaviour coding represent two different approaches to behaviour research and their co-application may offer a deeper insight into the subject under focus.

As regards *Trainability*, the owners' report did not correlate with the assessments of any of the rating groups, but the assessments of all the rating groups correlated with each other. This might indicate that on the basis of FIDO Personality test the traits comprising the personality dimension *Trainability* cannot be measured reliably.

Besides the above, we also compared the degrees of ratings done by the owners and by the three rater groups with different levels of hands-on experience with dogs with regard to all the four factors. This was done in order to detect possible population-level biases in the evaluations.

Regarding *Stranger-directed Sociability*, *Activity* and *Trainability* we found no difference in the degrees of ratings provided by the owners and by the three rater groups with different levels of hands-on experience with dogs. In the case of *Aggressiveness* the trainers of non-included dogs, the owners of non-included dogs and the group of non-owners characterised the target animals more aggressive than their owners did.

In the case of *Aggressiveness* a possible explanation why the owners reported a lower degree of their dog's aggressive tendencies than the video-based rater groups could root in the nature of the test situation where the video-taped animals were exposed to aggression-eliciting stimuli which are much less frequently present in everyday life (Svartberg, 2005). Furthermore, as Svartberg (2005) also suggests, *Aggressiveness* is more difficult for owners to assess than any other aspects of dog behaviour.

5. Conclusions

Our findings suggest that there is a need and room for testing whether behaviour traits measured by the means of questionnaires and behaviour tests converge.

The fact that the assessments of all the three rater groups correlated with owners' report in the case of *Stranger-directed Sociability, Activity and Aggressiveness* hints that the FIDO Personality test might be suitable to measure these aspects of the dogs' behaviour, since the amount of information the test situations provided could serve as a basis for judging these factors with strong correspondence with the owners' experience. The lack of associations between the hypothetically corresponding questionnaire items or factors suggest that there is further need for collaboration between researchers developing behaviour tests and questionnaire studies. Ideally, this research should be done in parallel.

Acknowledgements

This research was supported by HAS-ELTE Comparative Ethology Research Group (MTA 01 031) and by the EU FP7 ICT-215554 LIREC. The authors are grateful to Enikő Kubinyi and Ildikó Brúder for helping with the data collection and for making the video recordings of the FIDO behaviour test available for this study. We are also very grateful to an anonymous referee, who gave us a number of valuable comments on the previous version of this manuscript.

References

- Bahlig-Pieren, Z., Turner, D.C., 1999. Anthropomorphic interpretations and ethological descriptions of dog and cat behavior by lay people. *Anthrozoös* 12, 205-210.
- Block, J., 1961. The Q-sort method in personality assessment and psychiatric research. Springfield, IL. Charles C. Thomas.

Brúder, I., Kubinyi, E., De Meester, R., Pluijmakers, J., Miklósi, Á., Submitted. Introducing a simple behaviour test battery for family dogs (FIDO) and a case study for cross-laboratory comparison.

Christensen, E., Scarlett, J., Campagna, M., Houpt, K.A., 2007. Aggressive behaviour in adopted dogs that passed a temperament test. *Appl. Anim. Behav. Sci.* 106, 85–95.

Coppinger, R., Coppinger, L. 2001. *Dogs: A startling new understanding of canine origin, behaviour and evolution.* New York: Scribner

Diverio, S., Tami, G., Barone, A., 2008. Prevalence of aggression and fear-related behavioral problems in a sample of Argentine Dogs in Italy. *J. Vet. Behav.: Clin. Appl. Res.* 3, 74-86.

Feaver, J., Mendl, M., Bateson, P., 1986. A method for rating the individual distinctiveness of domestic cats. *Anim. Behav.* 34, 1016-1025.

Goddard, M.E., Beilharz, R.G., 1985. Individual variation in agonistic behaviour in dogs. *Anim. Behav.* 33, 1338-1342.

Gosling, S. D., 2001. From mice to men: What can we learn about personality from animal research? *Psychol. Bull.* 127, 45-86.

Gosling, S. D., Kwan, V. S. Y., John, O. P., 2003. A dog's got personality: a cross-species comparative approach to evaluating personality judgments. *J. Pers. Soc. Psychol.* 85, 1161-1169.

Guy, N.C., Luescher, U.A., Dohoo, S.E., Spangler, E., Miller, J.B., Dohoo, I.R., Bate, L.A. 2001.a Demographic and aggressive characteristics of dogs in a general veterinary caseload. *Appl. Anim. Behav. Sci.* 74, 15-28.

Guy, N.C., Luescher, U.A., Dohoo, S.E., Spangler, E., Miller, J.B., Dohoo, I.R., Bate, L.A. 2001.b Risk factors for dog bites to owners in a general veterinary caseload. *Appl. Anim. Behav. Sci.* 74, 29-42.

- Guy, N.C., Luescher, U.A., Dohoo, S.E., Spangler, E., Miller, J.B., Dohoo, I.R., Bate, L.A. 2001.c A case series of biting dogs – characteristics of the dogs, their behaviour, and their victims. *Appl. Anim. Behav. Sci.* 74, 43-57.
- Héjjas, K., Kubinyi, E., Rónai, Zs., Székely, A., Vas, J., Miklósi, Á., Sasvári-Székely, M., Kereszturi, E. 2009. Molecular and behavioral analysis of the intron 2 repeat polymorphism in canine dopamine D4 receptor gene. *Genes, Brain, Behaviour.* 8, 330-336.
- Horváth, Zs., Igyártó, B.Z., Magyar, A., Miklósi, Á. 2007. Three different coping styles in police dogs exposed to a short-term challenge. *Hormones and Behavior.* 52, 621–630.
- Irvine, L., 2001. The power of play. *Anthrozoös.* 14, 151-160.
- Jones, A.C., Gosling, S.D., 2005. Temperament and personality in dogs (*Canis familiaris*): A review and evaluation of past research. *Appl. Anim. Behav. Sci.* 95, 1–53.
- Konok, V., Dóka, A., Miklósi, Á. 2011. The behavior of the domestic dog (*Canis familiaris*) during separation from and reunion with the owner: A questionnaire and an experimental study. *Appl. Anim. Behav. Sci.* 135, 300-308.
- Kroodsma, D. E., Byers, B. E., Goodale, E., Johnson, S., Liu, W. 2001. Pseudoreplication in playback experiments, revisited a decade later. *Anim. Behav.* 61, 1029-1033.
- Kubinyi, E., 2009. Dog and owner demographic characteristics and dog personality trait associations. *Behav. Proc.* 81, 392-401.
- Kubinyi, E., Vas, J., Héjjas, K., Ronai, Zs., Brúder, I., Turcsán, B., Sasvári-Székely, M., Miklósi, Á. 2012. Polymorphism in the tyrosine hydroxylase (TH) gene is associated with activity-impulsivity in german shepherd dogs. *PLoS ONE*, 7, e30271.
- Lehner, P.N., 1996. *Handbook of ethological methods.* Cambridge University Press.
- Ley, J., Bennett, P., Coleman, G., 2008. Personality dimensions that emerge in companion canines. *Appl. Anim. Behav. Sci.* 110, 305-317.

- Martin, P., Bateson, P., 2007. *Measuring Behaviour: An introductory guide*. Cambridge University Press, Cambridge
- Mirkó, E., Kubinyi, E., Gácsi, M., Miklósi, Á. 2012. Preliminary analysis of an adjective-based dog personality questionnaire developed to measure some aspects of personality in the domestic dog (*Canis familiaris*). *Appl. Anim. Behav. Sci.* 138, 88-98.
- Molnár, Cs., Pongrácz, P., Dóka, A., Miklósi, Á. 2006. Can humans discriminate between dogs on the base of the acoustic parameters of barks? *Behav. Processes.* 73, 76-83.
- Netto, W.J., Planta, D.J.U., 1997. Behavioural testing for aggression in the domestic dog. *Appl. Anim. Behav. Sci.* 52, 243-263.
- O'Sullivan, E.N., Jones, B.R., O'Sullivan, K., Hanlon, A.J., 2008. The management and behavioural history of 100 dogs reported for biting a person. *Appl. Anim. Behav. Sci.* 114, 149-158.
- Overall, K.L., Love M. 2001. Dog bites to humans – demography, epidemiology, injury and risk. *J. Am. Vet. Med. Assoc.* 12, 1923-1933.
- Planta, J.U.D., De Meester, R.H.W.M, 2007. Validity of the Socially Acceptable Behavior (SAB) test as a measure of aggression in dogs towards non-familiar humans. *Vlaams Diergeneeskundig Tijdschrift* 76, 359-368.
- Pongrácz, P., Molnár, Cs., Miklósi, Á. 2010. Barking in family dogs: An ethological approach. *The Vet. J.* 183, 141-147.
- Rooney, N.J., Bradshaw, J.W.S., 2002. An experimental study of the effects of play upon the dog-human relationship. *Appl. Anim. Behav. Sci.* 75, 161-176.
- Rooney, N.J., Bradshaw, J.W., 2003. Links between play and dominance and attachment dimensions of dog-human relationships. *J. Appl. Anim. Welfare Sci.* 6, 67-94.
- Serpell, J., 1986. *In the Company of Animals: A Study of Human-Animal Relationships*. Cambridge University Press, Cambridge.

Simpson, B.S., 1997. Canine communication. *Vet. Clin. North Am. Small Anim. Pract.* 27, 445-464.

Stevenson-Hinde, J., Zunz, M., 1978. Subjective assessment of rhesus monkeys. *Primates*. 19, 473-482.

Svartberg, K., Forkman, B., 2002. Personality traits in the domestic dog (*Canis familiaris*). *App. Anim., Behav. Sci.* 79, 133-155.

Svartberg, K., Tapper, I., Temrin, H., Radesater, T., Thorman, S., 2005. Consistency of personality traits in dogs. *Anim. Behav.* 69, 283-291.

Svartberg, K., 2005. A comparison of behaviour in test and in everyday life: evidence of three consistent boldness-related personality traits in dogs. *Appl. Anim. Behav. Sci.* 91, 103-128.

Tami G., Gallagher A., 2009. Description of the behaviour of domestic dog (*Canis familiaris*) by experienced and inexperienced people. *Appl. Anim. Behav. Sci.* 120, 159-169.

Topál, J., Miklósi, Á., Csányi, V. 1997. Dog-human relationship affects problem solving behavior in the dog. *Anthrozoös*. 10, 214-224.

Tóth, L., Gácsi, M., Topál, J., Miklósi, Á. 2008. Playing styles and possible causative factors in dogs' behaviour when playing with humans. *Appl. Anim. Behav. Sci.* 114, 473-484.

Uher, J., Asendorpf, J.B., 2008. Personality assessment in the Great Apes: Comparing ecologically valid behaviour measures, behaviour ratings, and adjective ratings. *Journal of Research in Personality*, 42, 821-838.

van den Berg, S.M., Heuven, H.C.M., van den Berg, L., Duffy, D.L., Serpell, J.A., 2010. Evaluation of the C-BARQ as a measure of stranger-directed aggression in three common dog breeds. *Appl. Anim. Behav. Sci.* 124, 136-141.

Vas, J., Topál, J., Gácsi, M., Miklósi, Á., Csányi, V. 2005. A friend or an enemy? Dogs' reaction to an unfamiliar person showing behavioural cues of threat and friendliness at different times. *Appl. Anim. Behav. Sci.* 94, 99-115.

- Vas, J., Topál, J., Péch, É., Miklósi Á. 2007. Measuring attention deficit and activity in dogs: A new application and validation of a human ADHD questionnaire. *Appl. Anim. Behav. Sci.* 103, 105-117.
- Vas, J., Topál, J., Győri, B., Miklósi, Á. 2008. Consistency of dogs' reactions to threatening cues of an unfamiliar person. *Appl. Anim. Behav. Sci.* 112, 331–344.
- Vazire, S., Gosling, S.D., Dickey, A.S., Schaprio, S.J., 2007. Measuring personality in nonhuman animals. In R.W. Robins, R.C. Fraley, R. Krueger (Eds.), *Handbook of research methods in personality psychology*. New York, NY: Guilford, pp. 190-206.
- Walker, J., Dale, A., Waran, N., Clarke, N., Farnworth, M., Wemelsfelder, F., 2010. The assessment of emotional expression in dogs using a Free Choice Profiling methodology. *Anim. Welf.* 19, 75-84.
- Weinstein, T.A.R., Capitanio, J.P. and Gosling, S.D., 2008. Personality in Animals. In John, O.P., Robins, R.W., and Pervin, L.A. (Eds.), *Handbook of Personality Theory and Research*. New York: Guilford, pp. 328-248.
- Wemelsfelder, F., Hunter, E.A., Mendl, M.T., Lawrence, A.B., 2001. Assessing the 'whole animal': a Free-Choice-Profiling approach. *Anim. Behav.* 62, 209-220.

Tables and legends

Table 1. The table shows the personality factors obtained by Mirkó et al. (2012) and respective items (left column) and the corresponding subtests of the FIDO personality test with the coded behaviour variables. (-) and (+) indicates that the specific variable has a negative/positive loading on the factor. For the statistical analysis duration data were re-calculated as % of time.

<i>Personality Factors</i>	<i>The tests and the scoring of the coded variables</i>
<i>Stranger-directed Sociability</i> Brave (+) Afraid of noises (-) Mistrustful (-) Initiative (+)	Test 2. Greeting Latency of approaching the experimenter 0: not approaching; 1: approaches when being called; 2: approaches in 3 seconds after the experimenter stood in front of the dog; 3: approaches immediately Test 6. Playing with the experimenter The duration of playing with the experimenter (s) Test 13. Sudden appearance 0: no reaction; 1: moves the head backwards; 2: steps back; 3: runs away
<i>Activity</i> Lazy (-) Likes fetching balls (+) Likes games of fight (+) Overactive (+)	Test 1. Spontaneous activity The duration of moving the legs (s) Test 8. Playing with the owner The duration of playing with the owner (s) Test 9. Ball-playing with the owner The duration of running after the ball (s) The duration of keeping the ball in the mouth (s)
<i>Aggressiveness</i> Likely to bite a human (+) Hysterical (+) Shows a tendency to bark (+)	Test 3. Getting DNA sample Latency of getting the DNA sample 0: sample cannot be taken; 1: 10-60s; 2: <10s; 3: immediately Test 11. Bone take-away test 0: no aggressive reaction; 1: growling; 2: biting the artificial hand; 3: trying to attack the experimenter Test 12. Threatening approach 0: interrupting eye contact and/or moving away; 1: growling; 2: barking; 3: straining the leash in order to attack
<i>Trainability</i> Learns things easily (+) Pleases the owner (+) Controllable (+) Disorganised (-)	Test 4. Attention test The duration of orienting the swinging sausage (s) Test 8. Playing with the owner The duration of playing with the owner (s) Test 10. Problem-solving test Latency of obtaining the piece of meat 0: meat is not obtained; 1: 10-60s; 2: <10s; 3: immediately Test 14. Laying down to the side

	Latency of making the dog lie down 0: dog does not lie down; 1: 10-60s; 2: <10s; 3: immediately
--	--

Table 2. Summary table for Spearman's correlations (N=100) of the FIDO test variables and the questionnaire (DPQ) items loading on the factors Stranger-directed Sociability, Activity, Aggressiveness and Trainability (for details see Results). We tested only for the predicted associations and all the other cells in the table were left blank intentionally. Significant correlations are shown in bold. (-) indicates that the variable scores were reversed for analysis

Test battery variables	Variables of the personality factors			
	<u><i>Stranger-directed Sociability</i></u>			
	Afraid of noises (-)	Brave (+)	Mistrustful (-)	Initiative (+)
Umbrella	0.13 (p=0.22)	-0.19 (p=0.06)		
Appr E lat			-0.27 (p<0.01)	-0.25 (p<0.01)
Play E %			-0.06 (p=0.54)	0.24 (p<0.01)
	<u><i>Activity</i></u>			
	Lazy (-)	Overactive (+)	Fetching balls (+)	Likes games of fight (+)
Moving legs %	0.05 (p=0.60)	0.00 (p=0.98)		
Follow ball			0.47 (p<0.001)	
Ball in mouth			0.39 (p<0.001)	0.25 (p<0.01)
Play owner %			0.43 (p<0.001)	0.27 (p<0.01)
	<u><i>Aggressiveness</i></u>			
	Bites (+)	Barks (+)	Hysterical (+)	
Bone	0.03 p=0.77	0.25 (p<0.01)		
Threat	0.14 p=0.18	0.00 (p=0.98)		
DNA lat	0.04 p=0.67	0.30 (p<0.001)	-0.19 (p=0.07)	
	<u><i>Trainability</i></u>			
	Disorganised (-)	Pleases the owner (+)	Learns easily (+)	Controllable (+)
Pendulum	-0.04 (p=0.73)			
Play owner %		0.05 (p=0.63)		
Problem lat			-0.14 (p=0.16)	
Lay lat				-0.09 (p=0.41)

Table 3. Spearman's Correlations between different raters (Owner, Trainers of non-included dogs, Owners of non-included dogs, Non-owners) and owners' scores ($p < 0.05$). Raters provided their scores for the personality traits Stranger-directed Sociability, Activity, Aggressiveness, and Trainability after watching a short video record of the FIDO test (see text and Appendix II). Within a group of raters each person viewed a different dog on video. Significant correlations are shown in bold.

Stranger-directed Sociability			
	Owner	Owners of non-included dogs	Non-owners
Trainers of non-included dogs	0.66 (p<0.01)	0.59 (p<0.001)	0.57 (p<0.001)
Owners of non-included dogs	0.72 (p<0.01)	-	0.79 (p<0.001)
Non-owners	0.50 (p<0.01)	-	-
Activity			
	Owner	Owners of non-included dogs	Non-owners
Trainers of non-included dogs	0.83 (p<0.001)	0.58 (p<0.001)	0.59 (p<0.001)
Owners of non-included dogs	0.60 (p<0.001)	-	0.97 (p<0.001)
Non-owners	0.57 (p<0.001)	-	-
Aggressiveness			
	Owner	Owners of non-included dogs	Non-owners
Trainers of non-included dogs	0.60 (p<0.001)	0.85 (p<0.001)	0.79 (p<0.001)
Owners of non-included dogs	0.48 (p<0.01)	-	0.95 (p<0.001)
Non-owners	0.41 (p=0.02)	-	-
Trainability			
	Owner	Owners of non-included dogs	Non-owners
Trainers of non-included dogs	0.18 (p=0.31)	0.38 (p=0.03)	0.51 (p<0.001)
Owners of non-included dogs	0.15 (p=0.40)	-	0.86 (p<0.001)
Non-owners	0.25 (p=0.16)	-	-

Figure legend

Figure 1. The means of the questionnaire scores (+SE) of the personality factors (Stranger-directed Sociability; Activity; Aggressiveness; Trainability) given by different types of raters (Owner, Trainers of non-included dogs, Owners of non-included dogs, Non-owners) watching a single dog out of N=33 (post hoc test, $p < 0.05$). * denotes that the group of owners differs significantly from all the other groups in the case of Aggressiveness

