# Shyness–boldness predicts performance in working dogs

## Kenth Svartberg[*]

*Department of Zoology, Stockholm University, S-109 61 Stockholm, Sweden*

Accepted 5 December 2001

## Abstract

This study investigates if there are relationships between personality and performance of dogs (*Canis familiaris*) in working dog trials. Data from 2655 dogs of the two breeds German Shepherd dog (GSD) and Belgian Tervuren (BT) were used. The breeds were chosen because of indications of differences in personality between these breeds, and because both breeds are commonly trained for working dog trials. All dogs were tested in a personality test between 12 and 18 months of age. Using a factor analysis, five factors were extracted: ''Playfulness'', ''Curiosity/Fearlessness'', ''Chase-proneness'', ''Sociability'', and ''Aggressiveness''. Further analyses showed that these factors, with the exception of Aggressiveness, were all related to one higher-order factor, which was interpreted as a shyness–boldness dimension. Because of the risk of confounding variables, the influence of the owners' previous experience was tested. This showed that owner experience was related to performance, as well as to the shyness–boldness score. Therefore, only data from dogs with inexperienced owners were used in the later analyses. According to their success in working dog trials, the dogs could be categorised as low, middle, or high performing. The results show that the shyness–boldness score is related to the level of performance: high-performing dogs have higher scores (i.e. are bolder) compared to low-performing dogs. This difference was significant in Belgian Tervurens of both sexes, and in female German Shepherds. In general, German Shepherds scored higher than Belgian Tervurens, and males scored higher than females. However, in well-performing dogs there were no breed or sex differences. This indicates a threshold effect; to reach high levels in working dog trials the dog, independent of breed or sex, should have a certain level of boldness. These results imply that a lower proportion of dogs of shyer breeds are able to reach higher performance levels, compared to dogs of breeds that in general score higher on the shyness–boldness axis. In German Shepherds, a relationship was also found between personality and age of success; bolder dogs reached success at a younger age. There were no differences in Boldness score between dogs succeeding in different types of working dog trials (tracking, searching, delivering messages, handler protection), suggesting that the personality dimension predisposes trainability in general. The results

[*] Tel.: +46-8-16-40-54; fax: +46-8-16-77-15.
*E-mail address:* kenth.svartberg@zoologi.su.se (K. Svartberg).

might be applied to the selection of breeding dogs in working breeds and in selecting suitable working and service dogs. A test like the one used in this study can give a description of an individual dog's personality, which also can help matching the dog with adequate training.

## 1. Introduction

One of the active fields of research within animal behaviour is the study of individual differences. Research on "animal personality" has revealed behavioural differences, consistent over time or across situations, in a range of species. From this work, several concepts used to describe individual differences in humans have been found to be applicable in the study of animal personality. Examples are the shyness–boldness continuum (Wilson et al., 1994), individual differences in coping styles (Benus et al., 1991), and the concept of "supertraits" like the Big Five (Gosling and John, 1999). These findings indicate the existence of cross-species traits caused by similarities in selection pressures or the sharing of genes due to common ancestors (Buss, 1991; Gosling and John, 1999). The last decades of research in animal personality have resulted in a closer association with the area of human personality psychology. However, one field within personality psychology that has been largely ignored is the relationship between animal personality and performance in learning situations. These relationships have been studied from the 1950s up to today in humans, which has yielded knowledge about relationships between personality and factors affecting learning ability (e.g. Eysenck, 1981). Compared to the amount of studies on humans in this field, only few attempts have been done to find out these relationships in animals. Based on the studies in this area, it seems to be difficult to relate personality and learning ability in a simple way in animals (e.g. acquisition in aversive situations and emotionality in rats: Sing, 1959, 1961; Sing and Eysenck, 1960; Broadhurst and Levine, 1963; Levine and Broadhurst, 1963). In addition to this, the relationship that has been found between personality and learning performance seems to be specific for the test situation and sensitive to experimental circumstances (Garcia-Sevilla, 1984, compared to Broadhurst, 1960; Broadhurst and Levine, 1963). With this background the result from Scott and Fuller's (1965) early study appears remarkable. They studied dogs from five breeds in several aspects during the dogs' first year of life. Using data from dogs of all breeds, as well as hybrids, 40 variables from separate behavioural test situations were analysed (Brace, 1961, in Scott and Fuller, 1965). By factor analysis, one major behavioural factor was found. Variables indicating success from almost all training situations—such as speed and number of errors in mazes, success in motor skills, persistence in spatial orientation, success in string pulling and success in active escape training—together with several other behavioural scores indicating active and confident behaviour, were loaded on the factor. According to Scott and Fuller, this factor indicates the existence of one major trait, influencing behaviour in a range of situations and relates to learning performance in general. The dimension was interpreted as going from "general

good performance, and active confident behaviour'' to ''timidity, or fear, particularly of strange apparatus but also involving some fear of human beings'' (pp. 374–375).

In contrast to the results from studies on rodents, these results indicate the existence of behavioural dimensions that clearly influence general performance in learning situations. The reason for the lack of this relationship in earlier studies in rodents may be because of the focus of the studies. General relationships, if they exist, may not be found in specific situations. First after training and testing in a range of situations, these relationships may be revealed. If this is true, the search after general relationship between personality and learning performance should be directed to areas where several aspects of the animal's trainability are tested. One candidate for this is the tasks working dogs, i.e. police or military dogs, have to manage in order to be functional. These dogs are trained in tracking, searching, and/or handler protection, and are furthermore trained to be obedient in different kinds of situations. To accomplish this, these dogs are trained during long periods in various situations. Therefore, if the dimension found by Brace (1961), in Scott and Fuller (1965), exists outside the laboratory and is relevant in the real world, it would be likely to be reflected in the performance of working dogs. Earlier studies indicate that there are relationships to be found between personality traits and the probability to succeed in the training for working dogs (Wilsson and Sundgren, 1997). However, these results were based on subjective measures of dog personality and the situation for which the dogs were trained for was only vaguely defined.

Svartberg and Forkman (submitted for publication) used data from a personality test and found a higher-order factor in dogs. The factor describes a behavioural dimension, which goes from playful, active, friendly behaviour to fearlessness combined with low interest to play, chase and meet strangers. This dimension, which can be compared to a shyness–boldness axis, is similar to the one reported by Brace (1961), in Scott and Fuller (1965). This, in turn, makes it possible to use Brace's result to create predictions for the boldness dimension about performance in tasks requiring varied training. One such task is working dog trials, a civil counterpart to the tasks military and police dogs are trained for. In this study, a large number of dogs of the two breeds German Shepherd dog (GSD) and Belgian Tervuren (BT) were used. The breeds were chosen because of earlier indications of breed differences in regards to Boldness score, and because dogs from these breeds are common in working dog trials. All dogs were tested in a personality test similar to the one reported in Svartberg and Forkman (submitted for publication), from which the Boldness score was calculated. A prediction derived from Brace's (1961), in Scott and Fuller (1965), results is that dogs that perform well in working dog trials score higher on the boldness dimension compared to less successful dogs. If there is a threshold effect, i.e. there is a certain boldness level required for success in working dog trials, the effect of the personality upon performance should be more pronounced in the breed and in the sex that in general score low on the boldness dimension.

## 2. Methods

### 2.1. Subjects

Data from the two dog breeds GSD and BT were used. The initial sample included 2219 GSD (1178 males and 1041 females) and 436 BT (203 males and 233 females). When

Table 1
Numbers of males and females of both breeds in different categories of performance levels with inexperienced owners

| Category | GSD | | BT | |
|---|---|---|---|---|
| | Males | Females | Males | Females |
| None | 457 | 402 | 121 | 110 |
| Low | 63 | 54 | 8 | 18 |
| Middle | 226 | 215 | 44 | 26 |
| High | 129 | 95 | 18 | 8 |
| Total | 875 | 766 | 191 | 162 |

controlled for owner experience, the sample size was reduced (see Table 1). The dogs were pets living with their owners.

The breeds were used because of their suitability for this kind of comparisons. Both breeds, especially GSD, are commonly trained for performance in working dog trials. Furthermore, differences were found between these breeds with regards to factor score on the shyness–boldness dimension (hereafter called ''Boldness score''). By using these breeds, I could compare the relationship between personality and performance in two breeds that differ in Boldness score.

Behavioural data from two events were used: a personality test and a performance test. These tests are described in detail in the following sections.

## 2.2. The personality test

All dogs had been subjected to a standardised behavioural test (''dog mentality assessment'', DMA), once at age 12–18 months. This test, which can be seen as a personality test for dogs, is used by the Swedish Working Dog Association (SWDA) mainly as a tool in dog breeding, in which the behavioural reactions of the parents and, later on, of the progeny can be compared (Fält, 1997a,b). In the test, the dogs are exposed to several different novel situations and their reactions are described according to a standardised score sheet by official observers, one at every test. During the test a handler accompanies the dog (the owner or another familiar person). A test-leader is responsible for the practical aspects of the test. All persons conducting the test had been trained and certified by SWDA. The data from the personality test were collected during the years 1989–1997 at different testing arenas in Sweden. Scores from the dogs were reported to the SWDA, and further to the Swedish Kennel Club. Permission has kindly been given to use them in this study.

The personality test is an earlier version of the test described in Svartberg and Forkman (submitted for publication). Generally, the tests are very similar; the dogs are exposed to several subtests in which their reactions to strangers, play-invitations, fleeing prey-like object, and several potential fear-eliciting stimuli are described. However, the subtests Distance-play and Play 2 were added in the later version, and are therefore not included here. Another difference is that the subtest Assault was included in this earlier version of the test. However, data from that subtest were not used in this study. There

were slight differences in the assessment of the dogs' behaviour in comparison to the later version. The major difference was in number of variables for each subtest (fewer in the earlier version). The subtests are described in the presentation below. For a more thorough description of the later version, see Svartberg and Forkman (submitted for publication).

The test consists of nine separate subtests, which are carried out outdoors in a specific order. At the start of the test, observer and functionaries are positioned in advance before dog and handler are guided to the first station by the test-leader. In order to standardise the test situation the test-leader instructs the handler what to do and how to act before and during each subtest. The observer or the handler can interrupt the test, for example if the dog shows intense and remaining avoidance reactions.

### 2.3. Behavioural rating

The score sheet contains scales for 23 behavioural variables. The aim with the score sheet is to describe the dogs' behaviour as objectively as possible without inferring subjective opinions (like ''good behaviour'', ''too slow'', ''fearful'', etc.). All variables are scored from 1 to 5 according to intensity of the behavioural reaction, where a low score equals low intensity in the dogs' behavioural reaction. Each step in the scale, for each variable, is described on the score sheet. During every subtest, one to four behavioural reactions are recorded. The variables generally correspond to different phases during each subtest.

### 2.4. The subtests

#### 2.4.1. Social contact
Tests the dog's reaction to a stranger that greets the dog and makes physical contacts to him. The dog's behaviour is described by one behavioural variable, which goes from avoidance of stranger together with threats to intense greeting behaviour towards stranger.

#### 2.4.2. Play
Tests the dog's interest to play with a rag that is thrown between handler and test-leader, and further away from the dog. After that the dog is invited to play tug-of-war with the test-leader. The dog's interest in play, intensity in grabbing and interest in playing tug-of-war are described.

#### 2.4.3. Chase
Tests the dog's reaction to a small, rapidly moving, object that ''flees'' in a zigzag pattern away from the dog. The test is repeated once. The dog's interest in chasing the object and grabbing it in both repetitions are described.

#### 2.4.4. Passive situation
Tests the dog's behaviour during a period when handler and the leashed dog remain in position during 3 min without any disturbances. The dog's activity level during this period is described.

### 2.4.5. Sudden appearance

Tests the dog's reaction to a human-like dummy that is suddenly pulled up in front of the dog. The dog's startle reaction, avoidance and approach behaviour as well as aggressive reactions are described.

### 2.4.6. Metallic noise

Tests the dog's reaction to the noise of a chain with large links that is dragged over a sheet of corrugated metal when the dog is in close proximity. The dog's startle reaction, avoidance and approach behaviour are described.

### 2.4.7. Assault

Tests the dog's reaction to a stranger that hides and assaults the dog when passing the hiding place. The action of movement is interrupted in close proximity of the dog. The dog's startle reaction, approach behaviour as well as aggressive reactions are described.

### 2.4.8. Ghosts

Tests the dog's reaction to persons covered in white sheets ("ghosts") that move slowly, one at a time, towards the dog during several minutes. The dog's avoidance reaction, approach behaviour and aggressive behaviour are described.

### 2.4.9. Gunshot

Tests the dog's reaction to gunshots that are fired during activity (handler plays with the dog) and passivity. The dog's avoidance reaction and anxiety are described.

## 2.5. The performance test

For the measure of performance data from working dog trials were used. These trials are organised by SWDA and are given in four different trial types at four levels. The dog has to be 12 months old at the lowest level and 18 months at the three higher levels. At each level the dog is tested in both obedience and function. Trained and certified judges, two at each trial, judge the dog's performance according to rules provided by SWDA.

## 2.6. Test of obedience

In the test of obedience, the dog's ability to willingly and accurately react to the handler's signals is tested. The test consists of a sequence of subtests that are different at each level. At level 1 the subtests are heel on leash, sending on leash, heel free, heel with drop, recall, retrieve on flat, high jump, and long down; at level 2 heel on leash, sending on leash, heel free, heel with drop, recall, crawling, retrieve on flat, high jump, and long down (with gunshot); at level 3 heel free, recall with stop, sending free, heel with drop, heel with stop, crawling, barking, high jump, retrieve (heavy object), and long down (with gunshot); and the subtests at level 4 are heel free, recall with stop and drop, sending free, crawling, barking, high jump, retrieve (metal object), retrieve (heavy object), ladder climbing, and long down (with gunshot). In general, the complexity and difficulty of the obedience part

increases with each level. Subtests that are considered more advanced are given in the higher classes, and the complexity of the subtests that the dog is tested in at several levels (i.e. heel free or crawling) increases with each level. The dog's score in each subtest is according to the dog's responsiveness to the handler's signals, as well as the speed and accuracy of the performance.

## 2.7. Test of function

In the test of function at a working dog trial, the dog is tested in a general and a specific test. There are four specific function tests given, and the dog can start in one of them at each trial: tracking, searching, delivering messages, and handler protection. These can be seen as separate trial types (e.g. passing from one level to another is within specific function test). The specific tests are described in detail below. The general function test, which is independent of trial type, consists at the two lowest levels of a symbolic delivery of messages between handler and a functionary (50–100 m). The handler sends his dog to the functionary, who sends the dog back. At the second level, a gunshot is fired each time the dog passes midway between handler and functionary. The score is according to the dog's pace and the dog's responsiveness to the handler's and the functionary's signals. At the two highest levels, the general function test consists of a searching test. Four retrievable objects (e.g. gloves, purses or other small objects) are placed in a squared area (50 m × 50 m). The four objects should be found and retrieved within 5 min. The score is according to number of found objects, how the dog deliver the objects to the handler, and the dog's responsiveness to the handler's signals.

### 2.7.1. Tracking

In the tracking test, the dog's ability to find and follow a person-laid track, as well as finding objects in the track is tested. The dog tracks in a long lead (approximately 15 m). The length of the track is from 300 m at the lowest level to 1500 m at the highest level. Similarly, the age of the track increases with level: from 20 to 90 min. The track is laid on varied ground, and contains several angles. On the track, several wooden objects are dropped by the tracklayer. Before the dog can start tracking, he has to find the track within a specified area (is not required at the lowest level). The dog's score is according to the dog's ability to find the track at the start of the test, the number of objects found, and the speed of tracking.

### 2.7.2. Searching

In the searching test, the dog's ability to search and find hidden persons is tested. The dog works free without a lead, and the handler sends the dog in order to find the hidden persons (functionaries unknown to the dog). The search is done within a specified area with varying and broken ground, the size of which increases with each level: from a length of 50 to 300 m. The width at all levels is 100 m (50 m on each side of the trail from which the handler sends the dog). When the dog has found a functionary, he uses one of two ways to "tell" the handler about the finding. Either he remains at the functionary and barks until the handler arrives, or he grabs an object (attached to the collar) with the mouth and runs back to the handler. The handler puts the dog on a lead, and the dog guides the handler to the

functionary. At the lowest level two functionaries are hidden, while there are three functionaries hidden at the higher levels. The score is according to the number of persons found, the dog's behaviour when the dog has found the hidden person, how effective the dog is in searching, and the dog's responsiveness to the handler's signals (i.e. that the dog can be directed by the handler during the search).

### 2.7.3. Messenger dog

In the messenger dog test, the dog's ability to bring messages between known people over long, and partly known, distances (500–2000 m) is tested. During this test two handlers are in action (A and B). Handler A remains at a base station, while handler B walks with the dog from the base station to a second station. During the test the dog is sent between the stations at least twice (at first and second level), bringing a message (a piece of paper) in a small container attached to the collar. The dog that competes at levels three and four is sent over three and four distances, respectively. At all levels except the first handler B walks further away to a new station during the time as the dogs runs to handler A. When the dog is sent back to handler B the dog has to use its scent and track the handler's track between the old and new station to find him. The score is according to the dog's running time between the handlers at the stations, the dog's responsiveness to the handler's signals, and the dog's behaviour during rests (i.e. the dog should be silent).

### 2.7.4. Handler protection

In the handler protection test, the dog's ability to protect its handler is tested. This trial type is not given at the lowest level. Generally, this test consists of a number of situations in which one or two persons threaten the handler. The dog should stop these threatening functionaries and control them. The dog is tested both with and without a muzzle. When the dog is not wearing a muzzle, the functionaries are dressed in padded material to protect themselves from the dog's bites. During this test, the dog is only allowed to bite functionaries when these are trying to flee or attack the handler. The dog is tested in transportation of a functionary, protection of the handler who is assaulted by a functionary, catching a fleeing functionary, controlling of a passive functionary, and warding off an assaulting functionary. The score is according to how effective the dog is to ward of threats, the dog's behaviour during non-threatening situations, and its responsiveness to the handler's signals. Dogs that participate in this trial type are also tested in tracking (at all levels) and searching for hidden persons (at fourth level). These tests are similar, but not as demanding, to the specific tracking and searching tests, respectively.

### 2.8. Scores and criteria in working dog trials

The system for scoring is similar for the two parts of working dog trials. In test of obedience, the score is from 0 to 10 for each subtest. Every subtest has a specific coefficient, which the score is multiplied with. The sum of all multiplied scores makes up the total score for the obedience test.

The specific part of the test of function is either judged as one subtest (search and messenger dog), or divided into several subtests (tracking, and especially handler protection). Each subtest has its coefficient, with which the score (0–10) is multiplied. The overall

function score is the sum of the multiplied scores for each subtest (in specific and general part of the test). The total score from obedience and function parts (which has approximately equal maximums) is added, giving the dog's trial score. To pass on to next level, the dog has to reach a specified overall score, as well as a specified score for the function test.

### 2.9. Calculations of Boldness score and classification according to performance success

#### 2.9.1. The Boldness score

The ratings from the personality test were used to calculate each dog's factor score for the shyness–boldness dimension found in Svartberg and Forkman (submitted for publication). In that study, factor analysis found support for five specific factors (Playfulness, Chase-proneness, Curiosity/Fearlessness, Sociability and Aggressiveness) and one higher-order factor (shyness–boldness) that all specific factors with the exception of the factor Aggressiveness were related to. To calculate each dog's Boldness score, the score on the specific factors were first extracted. The representing variables were chosen according to the result of the factor analysis presented in Svartberg and Forkman (submitted for publication). Because of the slight differences between the personality tests used in that study and in this, there were some differences in the representing variables. However, all four specific factors that were related to the shyness–boldness dimension are measured in a similar way in this earlier version of the personality test.

To calculate score for the specific factor Sociability, the single variable from the subtest Social contact was used (in the current version of DMA, from which the this factor was extracted, the dog's behaviour in the subtest Social contact is measured by three variables). The factor Playfulness is based on three variables from the subtest Play. The factor Curiosity/Fearlessness is based on three variables from subtest Sudden appearance (startle reaction, exploration and avoidance), three variables from subtest Metallic noise (startle reaction, exploration and avoidance) and two from subtest Ghosts (avoidance and approach). The variables measuring the dogs startle reaction and avoidance responses loaded negative on this factor, while the exploration and approach variables were positively loaded. The specific factor Chase-proneness is based on four variables from the subtest Chase.

The dogs' score (1–5) on each variable were standardised (subtracting the mean and dividing by the standard deviation). The standardised value for each of the representative variables was summated, creating factor scores. The secondary factor analysis in Svartberg and Forkman (submitted for publication) showed that these four specific factors were all equally correlated to the higher-order factor (shyness–boldness dimension). Because of this, each of the four relevant specific factors were given equal weight when the Boldness score was calculated. The score on specific factors Playfulness, Chase-proneness, Curiosity/Fearlessness, and Sociability, were standardised and each dog's score was summed. This sum was used as the dog's Boldness score.

#### 2.9.2. Performance categorisation

The dogs were divided into three categories with regards to the highest performance reached. Dogs that had not participated in working trials were grouped together (category

None). This category of dogs is only used as a reference group in this study. Dogs that participated but failed at the first level (i.e. not reached the criteria for passing to the next level), were clumped into one category (Low). Dogs who had reached criteria for passing to the second or third level, but not reached the criteria for the highest level, were included in one category (Middle). Dogs who reached criteria for competing in the highest class were included in the third category (High).

## 2.10. Comparisons and analyses

### 2.10.1. Owner experience

To be able to interpret the relationship between personality and performance, it was important to consider the status of the owner. The skill and experience of the owner could in this case obviously be a confounding variable. Data about the owners' previous dog owning experience were collected with help from the Swedish Kennel Club. The number of earlier owned dogs that reached the middle or high level was used as a measurement of the owners' experience in training working dogs. Only previously owned dogs that were born at least one year before the target dogs were counted (an assumption was that the owner must have had earlier dogs for a while to have gathered experience). Dogs were categorised in three groups according to their owners' experience: (0) owner without previous experience, (1) owner with one previous dog that reached at least the middle level in working dog trials, and (2+) owner with two or more dogs that reached the same level. Dogs owned by the police, the military, or companies were eliminated from the data set. Also dogs with unclear owner circumstances were removed.

### 2.10.2. Level of performance

The Boldness scores for dogs that reached different performance levels, i.e. category Low, Middle or High, were compared. Furthermore, high-performing dogs' Boldness scores were compared with the average score for each breed and sex.

### 2.10.3. Number of attempts to success

This analysis was done in order to investigate whether a dog's Boldness score from the personality test relate to the number of attempts before the dog pass to middle level in working dog trials. Dogs from categories Middle and High were used in this analysis.

### 2.10.4. Age and success

In this comparison, the relationship between Boldness score and time to success in working dog trials was investigated. The measure of time to success was the age when the dog passed to the middle level. Dogs from categories Middle and High were used in this analysis.

### 2.10.5. Success in different trial types

The Boldness score for dogs that had performed well (category High) in different trial types—tracking, searching, delivering messages, and handler protection—were compared. Approximately 30% of the high-performing dogs had reached category High in more than one trial type, which made a separate classification of dogs

necessary. In order to compare high-performing dogs in different trial types, and use exclusive groups, dogs were classified into dogs that reached High in one trial type, say tracking, and all other High dogs with no merits in that trial type (i.e. tracking). In this way, dogs with merits in one trial type could be compared to High dogs with merits exclusively from other trial types. Because of insufficient number of successful BTs in different trial types in group High, only data from GSDs were used in this analysis.

## 2.11. Statistical analysis

All comparisons in this study were analysed using non-parametric statistical tools. For analysis of owner experience and success $\chi^2$-analysis was used. In the analyses of Boldness score in relation to level of performance and owner experience, Kruskal–Wallis ANOVA was used. Pair-wise comparisons with Mann–Whitney $U$-test were used to analyse sex and breed differences, and differences in Boldness scores for dogs that had succeeded in different trial types. In the analysis of number of attempts to success and age at success, Spearman rank correlation was used. The statistical package used was STATISTICA$^{TM}$.

## 3. Results

An analysis of the relationship between owner experience and performance shows an uneven distribution of performance in the three owner categories. The data are skewed towards better performing dogs within the groups of dogs with more experienced owners. The $\chi^2$-analyses show that the distributions for all three groups differ from what would be expected if dogs of different performance categories were equally distributed among owner categories (group 0: $\chi^2 = 31.5$, d.f. $= 2$, $P < 0.001$; group 1: $\chi^2 = 127.4$, d.f. $= 2$, $P < 0.001$; group 2+: $\chi^2 = 29.7$, d.f. $= 2$, $P < 0.001$).

There is also a relationship between owner experience and dog personality. Dogs with more experienced owners have higher Boldness score compared to dogs with less experienced owner (Kruskal–Wallis ANOVA; GSD, males: $H = 36.66$, $P < 0.001$; females: $H = 13.23$, $P = 0.001$; BT, males: $H = 11.4$, $P = 0.003$; females: $H = 9.5$, $P = 0.009$). Because of this relationship it was important to control for the experience of the owner. Therefore, only data from dogs with owners lacking experience of previous success were used in the further analyses (Table 1).

## 3.1. Breed and sex differences in Boldness score

An analysis of each breed shows that the Boldness score for the two breeds differ from each other. GSD score higher on the shyness–boldness dimension compared to BT in both males (Mann–Whitney $U$-test: $Z = 3.37$; $P < 0.001$) and females ($Z = 3.77$, $P < 0.001$). In addition, sex differences were found in both breeds showing that males score higher than females (Mann–Whitney $U$-test; GSD: $Z = 3.64$, $P < 0.001$; BT: $Z = 1.94$, $P = 0.05$).
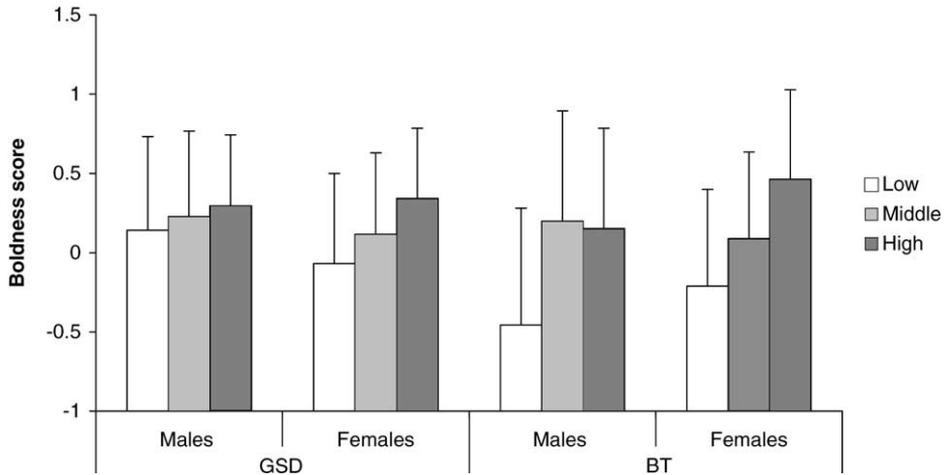
Fig. 1. Boldness scores for German Shepherds (GSD) and Belgian Tervurens (BT) that reached different performance levels (error bars representing standard deviation).

### 3.2. Level of performance and dog personality

The Boldness score was calculated for dogs that reached each performance level— Low, Middle and High. The result of this comparison reveals a step-wise pattern, where Low dogs have the lowest score and High dogs the highest (Fig. 1). A Kruskal–Wallis ANOVA showed that there are significant differences between the three performance categories in GSD females ($H = 21.12$, $P < 0.001$), BT females ($H = 7.38$, $P = 0.025$), and BT males ($H = 6.58$, $P = 0.037$), but not in GSD males ($H = 1.57$, $P = 0.457$).

### 3.3. Comparison between Boldness score in high-performing dogs of different breeds and sex

Because of the general breed and sex differences in regards to Boldness score, it was interesting to analyse if these differences also could be found within the groups of successful dogs. The analyses show that the four groups of High dogs, GSD males and females, BT males and females, do not differ in regards to Boldness score (Kruskal–Wallis ANOVA: $H = 2.60$, $P = 0.458$). The same result was found at level Middle (Kruskal–Wallis ANOVA: $H = 5.47$, $P = 0.141$). This made it possible to merge all High dogs into one group, and all Middle dogs into a second group, and compare these groups' score with the average score for the two breeds. This comparison shows that dogs in the High group have higher Boldness scores than the average for dogs in both breeds and sexes (Mann–Whitney; GSD males: $Z = 4.95$, $P < 0.001$; GSD females: $Z = 7.53$, $P < 0.001$; BT males: $Z = 6.25$, $P < 0.001$; BT females: $Z = 8.14$, $P < 0.001$). The same result was found in the comparison between Middle dog's Boldness score and the two breeds average score (Mann–Whitney; GSD males: $Z = 2.40$, $P < 0.016$; GSD females: $Z = 5.64$,
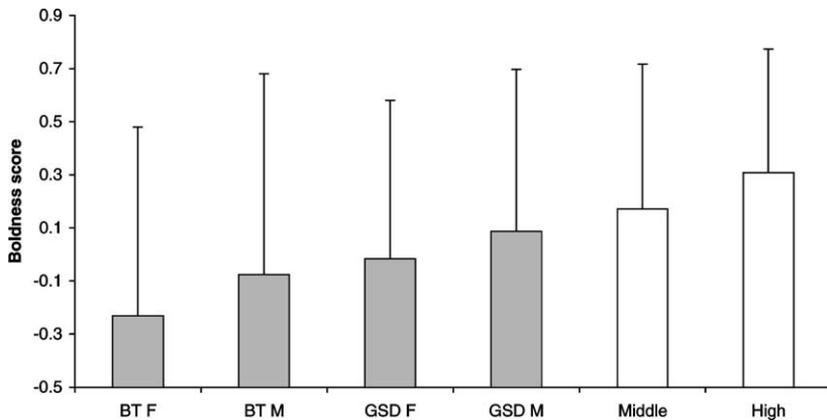
Fig. 2. The average Boldness scores in German Shepherds (GSD) and Belgian Tervurens (BT) males (M) and females (F), as well as the average score for all dogs in the Middle and High groups (error bars representing standard deviation).

$P < 0.001$; BT males: $Z = 4.67$, $P < 0.001$; BT females: $Z = 6.86$, $P < 0.001$). The result is presented in Fig. 2.

### 3.4. Number of attempts to success

In this analysis, the dogs' number of attempts, or number of trials, before entering level Middle was tested against the dogs' Boldness scores. Number of attempts before reaching criteria for level Middle ranged from 1 to 15, with a mean of 1.84 (S.D. $\pm 1.22$) in GSD. In BT, the range was between 1 and 4, and mean 1.61 (S.D. $\pm 0.88$). No correlations were found between number of attempts and Boldness score (Spearman rank correlation; GSD, males: $r_s = 0.01$, $P = 0.999$; females: $r_s = -0.02$, $P = 0.674$; BT, males: $r_s = 0.11$, $P = 0.775$; females: $r_s = -0.01$, $P = 0.990$). There were no sex differences in regards to number of attempts before promotion (Mann–Whitney $U$-test; GSD: $Z = -0.18$, $P = 0.857$; BT: $Z = 1.11$, $P = 0.268$), nor any breed differences (Mann–Whitney $U$-test: $Z = 1.43$, $P = 0.152$).

### 3.5. Age and performance

The dogs' age when reaching level Middle was tested against the dogs' Boldness scores. The dogs' age when reaching criteria for level Middle was on average 714.1 days (S.D. $\pm 293.5$) for GSD and 770.4 days (S.D. $\pm 351.6$) in BT. No significant differences between the breeds was found (males: $Z = -0.46$, $P = 0.647$; females: $Z = -1.00$, $P = 0.318$). Success–age was significantly, and negatively, correlated with Boldness score in GSD males ($r_s = -0.15$, $P = 0.005$) and GSD females ($r_s = -0.17$, $P = 0.002$). In BT, the correlation was in the same direction, but did not reach significant levels (males: $r_s = -0.20$, $P = 0.135$; females: $r_s = -0.25$, $P = 0.152$). There was no difference between sexes in regards to age when reaching criteria for level Middle (Mann–Whitney $U$-test; GSD: $Z = -1.09$, $P = 0.277$; BT: $Z = 0.21$, $P = 0.621$).

*3.6. Boldness scores for dogs succeeding in different trial types*

In this analysis, Boldness scores for dogs that succeeded in different trial types were compared. Males and females were combined into one group in this comparison. This fusion was done based on the similar result in the first analysis of differences between categories in males and females, and because there were no differences in Boldness scores between sexes in this category of dogs (Mann–Whitney $U$-test: $Z = -0.85$, $P = 0.396$). The analyses show that there were no significant differences in Boldness score found between High dogs in different trial types (Mann–Whitney $U$-test: tracking ($n = 212$) versus non-tracking ($n = 12$): $Z = 1.39$, $P = 0.164$; searching (42) versus non-searching (182): $Z = 0.48$, $P = 0.634$; messenger (18) versus non-messenger (206): $Z = -1.19$, $P = 0.235$; protection (19) versus non-protection (205): $Z = -0.78$, $P = 0.434$).

## 4. Discussion

The results in this study show a relationship between personality of dogs and performance in working dog trials, supporting the general prediction. A high score on the shyness–boldness axis, a dimension related to Sociability towards strangers, Playfulness, interest to chase, exploration and fearlessness, correlates with high success in these trials. This relationship is expressed in level of performance; bolder dogs reach higher levels compared to shyer dogs. As suspected, there were breed differences as well as sex differences in regards to Boldness score. German Shepherds score higher compared to Belgian Tervurens, and males score higher than females. However, there are no breed or sex differences in high-performing dogs. This indicates that success at high levels requires a certain Boldness score, which are independent of the dog's breed or sex.

One obvious problem with data like the one used in this study is that there is a risk of different confounding variables. Compared to animals normally used in a controlled experiment, these subjects are exposed to a range of different circumstances from birth to entry on a working dog trial. The dogs live in different kinds of family settings, in different kinds of surroundings, and have owners that differ in several aspects. That such factors may influence performance is exemplified by the finding of a relationship between success and owner experience. The result shows that the owner's earlier experience of working dog trials increases the likelihood that the next dog will be successful. It is likely that there are other similar factors, not controlled for or measured in this study, which influence the probability of success. Because of this, conclusions should be drawn with caution. There are, however, several results that indicate that the shyness–boldness dimension is of real importance for success in working dog trials. One is that the relationship between dog personality and performance were found in a sample controlled for owner experience. It is likely that owner experience is one of the most important factors for success. Another finding that may indicate the relevancy of the shyness–boldness dimension upon performance is that there is a relationship between Boldness score and success–age in German Shepherds, but not between Boldness score and number of attempts necessary to reach success. The fact that low scoring dogs succeed at a higher age is not due to repeated attempts in order to reach promotion; an owner strategy that could otherwise have

explained this relationship. These results instead indicate that dogs with low Boldness score need more and prolonged training to succeed. The relationship between success–age and Boldness score in GSD could also be explained by individual differences in maturation speed: slowly maturing dogs score low on boldness at age 12–18 months as well as taking longer to train, and vice versa. However, the fact that the major result shows a clear relationship between Boldness score and level of performance makes this explanation less plausible. A third point that indicates that the shyness–boldness dimension is important for success is the concordance of the results. The same pattern is found in both German Shepherds and Belgian Tervurens, and in both males and females, even if the same significant levels not are reached in every comparison. This validates the results, and makes it likely that the shyness–boldness dimension is of real importance in applied dog training.

These results are in accordance to what was found by Brace (1961), in Scott and Fuller (1965), which indicates an analogy between the dimension activity–success and the dimension investigated in this study. A summary of the present results—that active, confident and playful dogs are more successful in a performance test (i.e. the working dog trial) compared to fearful dogs with low interest to play and to greet strangers—is nearly identical to Scott and Fuller's description of the dimension activity–success. A result that further supports this is the lack of difference between successful dogs in different trial types. This indicates that this trait is related to trainability in general, and not specifically correlated to a predisposition to be trained in any particular task. If this is correct, the dimension that Brace (1961), in Scott and Fuller (1965), found in laboratory conditions is in action also in applied situations, and plays an important role for the dog's trainability and quality of performance.

Even if there are significant relationships found with two aspects of success—age and level—the trait does not seem to be equally important for the two measures. That only German Shepherds reach significant levels of correlations between Boldness score and success–age indicates that the effect of the trait upon training speed is relatively low. The trait seems to be more important for level of success, which the finding of the same relationship in both breeds shows. Interestingly, there were no sex or breed differences found in well-performing dogs. This is contrasted by the findings that Belgian Tervurens and German Shepherds, as well as males and females, differ in general. These results imply that there is a threshold effect upon successful performance in working dog trials, and that this threshold is the same independent of sex and breed: for a dog to be successful, irrespective of whether it is a Belgian Tervuren or a German Shepherd, male or female, it ought to have a certain boldness. Because of the lower average score in Belgian Tervurens the effect should be more pronounced in this breed compared to German Shepherds. Even if the Boldness score for well-performing dogs differ from what is general in German Shepherds, the greater difference for Belgian Tervurens indicates that there is higher selection within this breed in becoming a working dog. A well-performing German Shepherd ought to be above average in regards to Boldness score, but a Belgian Tervuren, especially if it is a female, have to be extraordinary in the same aspect.

The fact that dog personality and performance are related to owner experience could indicate that the shyness–boldness dimension is unstable and can be influenced by social factors during the first year of living (e.g. O'Farrell, 1997; Podberscek and Serpell, 1997; Roll and Unshelm, 1997). However, there are results that contradict this picture in showing

behavioural stability from early age in dogs (Pfaffenberger et al., 1976; Goddard and Beilharz, 1986; Slabbert and Odendaal, 1999, but see also Beaudet et al., 1994; Wilsson and Sundgren, 1998). The results indicate that the stability in a behaviour dimension found in wolves by Fox (1972) may have its counterpart in dogs, in spite of the differences in living conditions different dogs are exposed to. If this is true, there have to be some other way to explain the relationship between owner experience and dog personality in this study. One plausible explanation is that experienced dog owners are more selective when they buy a new puppy. In order to succeed with the next dog, they may choose a breeder with a history of successful breeding of working dogs. Puppies from these breeders may have a personality that, on average, both influences score on the personality test and predisposes them in the training for working dog trials to a higher degree than puppies from other breeders. There is also a possibility that the choice of the individual puppy in the litter may differ depending on the buyer's earlier experience. Experience of working dog training may help the buyer to choose the most confident, playful and bold puppy. The breeder can also be an active part in such a process; the breeder is aware of the buyer's earlier experience, and help the buyer to choose the puppy that seems to be the most promising for working dog training. Which of these explanations for the relationship between owner experience and dog personality is true, or the most important, is however, impossible to tell by the data in this study. Both the possible owner influences of the dog's personality during young age, and the stability of the traits, are interesting topics that deserve further attention.

## 5. Conclusion

The present study supports earlier results, and shows that in dogs there are general relationships between bold personality and ability to learn and perform well in tasks requiring varied training. Active, playful and bold dogs are more likely to learn more complex behaviour and perform well in situations requiring persistence. Fearless individuals may be less easily distracted and inhibited in novel situations, and the tendency to engage in play and chase may make the opportunity to engage in such activities more rewarding for these dogs. These findings may be useful in the training, selection and breeding of working dogs. For training purposes, it is important to reveal the dog's personality in order to match it with adequate training. Dogs scoring low on shyness–boldness dimension may need more training, especially in order to diminish the risk of distracting and disturbing the dogs performance by fear-elicited responses. In selection of individual potential working dogs and for the purpose of suitable breeding dogs, a test like the one used in this study might be a useful tool. For the purpose of selecting an individual working dog, the relationship seems to be clear—it is an advantage to select bold individuals. When it comes to dog breeding, the uncertainty of the heritability of the shyness–boldness dimension makes it less clear how useful it is for this purpose. However, unpublished and preliminary results (Saetre et al., 2002; Per-Erik Sundgren, personal communication) point towards moderate or high degrees of heritability of all the specific factors that are included in this higher-order dimension. If this is true, a test that reveals the dog's score on these factors can be a very useful tool to select breeding individuals not only in working dog breeds, but also in other breeds, to improve welfare and life quality for dogs and owners.

## Acknowledgements

## References

Beaudet, R., Chalifoux, A., Dallaire, A., 1994. Predictive value of activity level and behavioral evaluation on future dominance in puppies. Appl. Anim. Behav. Sci. 40, 273–284.

Benus, R., Bohus, B., Koolhas, J., van Oortmerssen, G., 1991. Heritable variation for aggression as a reflection of individual coping strategies. Experientia 47, 1008–1019.

Brace, C.L., 1961. Physique, physiology, and behavior: an attempt to analyze a part of their roles in the canine biogram. Ph.D. thesis, Harvard University.

Broadhurst, P.L., 1960. Applications of biometral genetics to the inheritance of behavior. In: Experiments in Personality. Routledge and Kegan Paul, London.

Broadhurst, P.L., Levine, S., 1963. Behavioural consistency in strains of rats selectively bred for emotional elimination. Br. J. Psychol. 54, 121–125.

Buss, D.M., 1991. Evolutionary biology and personality psychology: implications of genetic variability. Pers. Individual Differ. 4, 51–63.

Eysenck, M.W., 1981. Learning, memory and personality. In: Eysenck, H.J. (Ed.), A Model for Personality. Springer, Berlin, pp. 169–209.

Fält, L., 1997a. Anvisningar mentalbeskrivning. Svenska Brukshundklubben.

Fält, L., 1997b. Kompendium mentalitet. Svenska Brukshundklubben.

Fox, M.W., 1972. Socio-ecological implications of individual differences in wolf litters: a developmental and evolutionary perspective. Behaviour 41, 298–313.

Garcia-Sevilla, L., 1984. Extraversion and neuroticism in rats. Pers. Individual Differ. 5, 511–532.

Goddard, M.E., Beilharz, R.G., 1986. Early prediction of adult behaviour in potential guide dogs. Appl. Anim. Behav. Sci. 15, 247–260.

Gosling, S.D., John, O.J., 1999. Personality dimension in nonhuman animals: a cross-species review. Curr. Dir. Psychol. Sci. 8, 69–75.

Levine, S., Broadhurst, P.L., 1963. Genetic and ontogenetic determinants of adult behavior in the rat. J. Comp. Physiol. Psychol. 56, 423–428.

O'Farrell, V., 1997. Owner attitudes and dog behaviour problems. Appl. Anim. Behav. Sci. 52, 205–213.

Pfaffenberger, C.J., Scott, J.P., Fuller, J.L., Ginsburg, B.E., Bielfelt, S.W., 1976. Guide Dogs for the Blind: Their Selection, Development, and Training. Elsevier, Amsterdam.

Podberscek, A.L., Serpell, J.A., 1997. Environmental influences on the expression of aggressive behaviour in English cocker spaniels. Appl. Anim. Behav. Sci. 52, 215–227.

Roll, A., Unshelm, J., 1997. Aggressive conflicts amongst dogs and factors affecting them. Appl. Anim. Behav. Sci. 52, 229–242.

Saetre, P., Strandberg, E., Sundgren, P.-E., Petterson, U., Jazin, E., Bergström, T., 2002. The genetic basis of behavioural patterns in dogs. In: Proceedings of the Workshop on Dog Genetics, Uppsala, Sweden, 22–23 May 2002.

Scott, J.P., Fuller, J.L., 1965. Genetics and the Social Behavior of the Dog. The University of Chicago Press, Chicago.

Sing, S., 1959. Conditioned emotional response in the rat. I. Constitutional and situational determinants. J. Comp. Physiol. Psychol. 52, 574–578.

Sing, S., 1961. Conditioned emotional response in the rat. Effect of stimulant and depressant drugs. J. Physiol. Res. 5, 1–11.

Sing, S., Eysenck, H.J., 1960. Conditioned emotional response in the rat. III. Drug antagonism. J. Genet. Psychol. 63, 275–285.

Slabbert, J.M., Odendaal, J.S.J., 1999. Early prediction of adult police dog efficiency—a longitudinal study. Appl. Anim. Behav. Sci. 64, 269–288.

Svartberg, K., Forkman, B., 2002. Personality traits in the domestic dog (*Canis familaris*). Appl. Anim. Behav. Sci. PII: S0168-1591(02)200121-1.

Wilson, D., Clark, A., Coleman, K., Dearstyne, T., 1994. Shyness and boldness in humans and other animals. Trends Ecol. Evol. 9, 442–446.

Wilsson, E., Sundgren, P.-E., 1997. The use of behaviour test for the selection of dogs for service and breeding. I. Method of testing and evaluating test results in the adult dog, demands on different kinds of service dogs, sex and breed differences. Appl. Anim. Behav. Sci. 53, 279–295.

Wilsson, E., Sundgren, P.-E., 1998. Behaviour test for 8-week old puppies—heritabilities of tested behaviour traits and its correspondence to later behaviour. Appl. Anim. Behav. Sci. 58, 151–162.