ELSEVIER

Applied Animal Behaviour Science 53 (1997) 279–295

APPLIED ANIMAL BEHAVIOUR SCIENCE

The use of a 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

Erik Wilsson [a,*], Per-Erik Sundgren [b]

[a] *Department of Zoology, University of Stockholm, S-106 91 Stockholm, Sweden*

[b] *Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden*

Accepted 14 October 1996

Abstract

The behaviour test results of 1310 German shepherds and 797 Labrador retrievers, 450–600 days of age, were evaluated. The purpose was to investigate whether the behaviour tests, previously used at the Swedish Dog Training Centre, could be used to select dogs for different kinds of work and for breeding. Ten behavioural characteristics were scored based on the dogs' reactions in seven different test situations. All tests were conducted by one experienced person.

Marked differences in mental characteristics were found between breeds and sexes, but particularly between various categories of service dogs. Regardless of differences in the behaviour profiles of these service categories, there were marked similarities between different categories of service dogs compared with dogs found to be unsuitable for training as service dogs. To interpret the data, an index value was created, based on the test results for each individual dog, and was found to be an excellent instrument for selecting dogs for different types of work.

For both breeds the factor analysis resulted in four factors. In comparing the different characteristics, the same pattern was found in both breeds, with the exception of the characteristic prey drive, which seems to be irrelevant for Labrador retrievers. The conclusion is that a subjective evaluation of complex behaviour parameters can be used as a tool for selecting dogs suitable as service dogs. The results also show that the use and correct interpretation of behaviour

* Corresponding author. Present address: Apt 3, 41 Hawthorne St, Cambridge, MA 02138, USA.

0168-1591/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved.
PII S0168-1591(96)01174-4

tests can be enhanced by adjusting the results for each breed and planned service category. © 1997 Elsevier Science B.V.

Keywords: Dog; German shepherd; Labrador retriever; Behaviour tests; Breed differences; Sex differences; Selection; Service dogs; Guide dogs; Police dogs; Breeding

1. Introduction

Behaviour testing has been used for 60 years to aid the selection of service dogs for various types of work and for breeding (Humphrey, 1934; Reuterwall and Ryman, 1973; Scott and Bielfelt, 1976; Pfleiderer-Högner, 1979; Goddard and Beilharz, 1982, , 1984; Mackenzie et al., 1984). The hereditary aspects of working dogs abilities have been reviewed by Mackenzie et al. (1986) and Willis (1989, 1995). Most of these studies were carried out in order to select dogs for one type of work, and most of them were carried out on only one breed. Very little has been reported regarding evaluating behaviour tests used to select dogs from different breeds and for different types of work, i.e. police dogs, narcotic dogs.

In other studies behaviour tests were used exclusively to study the effects of varied study conditions or hereditary variations in different types of dogs (Murphree and Dykman, 1965; Scott and Fuller, 1965). In most cases, however, these tests were performed on a comparatively small number of individuals, and the value of the tested parameters with regard to the usefulness of the dogs for various types of work has not been tested. Reuterwall and Ryman (1973) evaluated the results of dogs tested at the Swedish Army School (later renamed the Swedish Dog Training Centre, SDTC) during the years 1966–69. The testing procedure and the scoring system was somewhat different than those used in this study.

The aim of this study was to analyse the results for German shepherds and Labrador retrievers tested at SDTC, and to compare those results with their future capacity as service dogs. We also wished to compare the demands on different types of service dogs and to make it easier to interpret the test results from one individual when selecting dogs for different types of work and for breeding.

This is the first of four reports. The second report deals with hereditary aspects of the tests evaluated here, and the third and fourth reports evaluate a puppy test performed on 867 eight-week-old German shepherd puppies. Reports three and four also compare the results of the puppy test and body weight versus the test results on mature animals described here.

2. Materials and methods

The former SDTC has been one of Europe's largest centre for training and breeding service dogs since the late 1930s. The institution has raised and trained most of the service dogs used in Sweden either as police dogs, guide dogs, protection dogs, or narcotic dogs. In 1960–90 the breeding facility produced 300–500 pups per year, mostly German shepherds. Periodically, pups were also purchased from private breeders.

Regardless of whether the pups were raised or purchased they were placed in private homes at the age of eight weeks and later recalled to the SDTC for testing.

Only dogs in the age range 450–600 days were included in this study because earlier studies (unpublished data) have shown that test age influences the test results more than previously assumed. After the exclusion of dogs with incomplete test records, or who were tested outside the age range 450–600 days, the samples included 1310 German shepherds (730 males, 580 females) and 797 Labrador retrievers (343 males, 454 females). 1002 (76%) of the German shepherds and 329 (41%) of the Labrador retrievers were bred by the SDTC, and the rest were purchased from private breeders at the age of eight weeks. All dogs were exposed to the same handling routines after this age, and were tested over the period 1983–91.

All dogs that did not fulfil the demands of a service dog were disqualified and were sold as companion animals or donated to the puppy walkers who had previously had cared for them. About 50% of the dogs selected for training were disqualified during the training period. After completion of the training program all dogs were finally run through a battery of working tests performed by the buyer of the dogs (the police force, guide dog association). In this work only the dogs that passed this final test are referred to as service dogs.

2.1. Description of the test situations

The dogs were tested after an acclimatisation period of at least two weeks at the SDTC. During this time the dogs were handled daily by the same person who had also handled the dog during the test. The dog was exposed to a number of test situations and the test leader conducted a subjective evaluation of 10 different characteristics of the dog. Tables 1–7 show the classification of the evaluated characteristics. Each dog was tested once only.

2.1.1. Test situation 1: Approachability and tendency to compete for objects

The dog's initial reaction to the test leader (TL) is observed; the test leader up to this point is unfamiliar to the dog. The approach is observed where the test leader remains passive and again later when the TL actively tries to make contact with the dog. Thereafter, the handler takes the dog through a short session of leash exercises, and the dog–handler relationship is noted. Next the handler and later the test leader attempt to engage the dog in the game of tug of war using a rag.

2.1.2. Test situation 2: Startling test 1

The dog is held on a leash by the TL, and the handler runs away from the dog, while calling it, into a wooded area where a number of life-size paper figures of people are placed in an area measuring 25×25 m. As soon as the handler is out of view he stops, and the dog is turned loose. When the dog is in the middle of the paper figures another paper figure is suddenly pulled up in front of it. In most cases the dog perceives this as a threatening situation, while at the same time it is highly motivated to reach the handler.

2.1.3. Test situation 3: Startling test 2

The handler walks with the dog along a path. After about 10 m a life-size rag doll is placed about 1 m in front of the dog. During the test the handler is prepared, if needed,

to let go of the leash to enable the dog to leave the area. Special attention is given to the dog's reaction after the encounter with the doll, whether it has the courage to approach the doll, and how it acts in that situation.

2.1.4. Test situation 4: Reaction to a loud noise

The handler walks with the dog along a path. Leaning against a tree is a ladder from which a number of metal buckets and other objects are suspended. As the dog passes the ladder the objects are released and fall, creating loud noises. The dog's reaction to the noise, and particularly how it acts immediately afterwards, is noted.

2.1.5. Test situation 5: Reaction to a successively approaching threat

A paper figure, the top half of a person, is mounted on a wooden sled. The eyes of the figure are emphasised and appear to be staring straight ahead. The handler stands with the dog by his or her side. Suddenly the sled is pulled into view to emerge about 15 m in front of the dog. The sled is then pulled slowly towards the dog. Most dogs experience this situation as threatening.

2.1.6. Test situation 6: Attack on the handler

This test is performed only on German shepherds, and only if the TL requires more information in order to judge its sharpness or defensive behaviour. The test is performed as a staged attack on the handler by the TL, who wears a protective sleeve. The handler walks with the dog along the wall of a building when the TL suddenly rushes out and pretends to attack the handler and the dog.

2.1.7. Test situation 7: Reaction to gunfire

The dog's reaction to gunfire is tested by the repeated firing of a 9 mm starting pistol, first while the handler is playing with the dog and later with the dog sitting by the handler's side. The person fires the pistol while standing about 5–10 m from the dog.

2.2. Evaluated characteristics

The following characteristics were evaluated and scored in the seven test situations.

A. *Courage*: the ability to overcome fear. Courage can only be interpreted in situations where the dog actually became frightened. Score 1–9.

B. *Sharpness*: the tendency to react with aggression. i.e. the tendency to use aggression in order to reach a certain goal. Score 1–6.

C. *Defence drive*: the tendency to defend itself or its handler. In most cases defence is combined with aggression. However a dog may show defensive tendencies without being aggressive, and aggressive behaviour is not always combined with defensive tendencies. Score 1–9.

D. *Prey drive*: the willingness to engage in competitive games, e.g. "tug of war", also termed competitive drive or social competitive drive. In this study the tendency to defend the object ("prey") is also included in the character. Score 1–6.

E. *Nerve stability*: the appropriateness of the dog's reaction to a certain situation. This includes the dog's ability to adapt to various types of situations, to concentrate when highly aroused or in a situation of conflict, as well as its ability to relax and to overcome a frightening situation. Score 1–8.

F. *Reaction to gunfire*: this characteristic is scored from one to three, describing reactions from no sign of fear to one of lasting fear. Score 1–3.

G. *Temperament or energy level*: the degree of liveliness. Dogs with high temperament are more responsive to all types of stimuli. Score 1–9.

H. *Hardness*: the lack of a lasting effect of a pleasant or frightening experience. A dog with a low score is very easily affected by corrections and/or a frightening experience, whereas a dog with high scores is hard to affect. For training purposes a dog with a medium score is preferable. Score 1–6.

I. *Ability to cooperate*: the tendency to be influenced by the handler without being given a direct command or sign. Another often used term is willingness to please. Score 1–6.

J. *Affability*: the dog's willingness to make contact with people. Score 1–12.

The dogs were scored according to the subjective judgement of the TLs assigned to assess each of these characteristics. The evaluation of each individual characteristic was based on the dog's reaction to several test situations. The scores describe the characteristic on scales of increasing intensity. The dogs were tested just once, and the test took about one hour.

2.3. Statistics: Calculation of the index value (I)

The index value was based on the information gathered from the nine characteristics. This was achieved by calculating a partial index for each score within each of the nine characteristics. The characteristic "Reaction to gunfire" was not included in the calculation of the index value because all forms of gun shyness disqualify a dog for use as a service dog. High partial index values for one particular characteristic were given when dogs obtained scores found to be desirable; low partial index values indicate that the dog was given an undesirable score. The mean partial index for all dogs for one particular characteristic is 1.0. For dogs with desirable characteristics the mean partial index should be higher than 1.0.

For each characteristic, the partial index (I_x) for each score was calculated using the formula

$$I_x = O_x/E_x$$

where O_x is the observed number of dogs selected as service dogs in category x of a certain characteristic. For German shepherds, only police dogs and protection dogs were included, and for Labrador retrievers only guide dogs. E_x is the expected number of service dogs in a certain characteristic category, calculated as

$$E_x = n_x * N_y/N$$

where n_x is the number of individuals obtaining the score x, N is the total number of individuals tested, and N_y is the number of service dogs in the total number of dogs

tested. In those cases where $I_x > 1$ and $E_x < 6$, I_x was set to 1.0 in order to reduce the effect on I_x of very small groups of dogs. The index value was calculated as $I = \Sigma(I_{x1-x9})$; i.e. the sum of all the partial index values achieved by each dog.

Because different breeds and sexes are used for different types of service work, the partial index values were calculated separately for male and female German shepherds, and for male and female Labrador retrievers. The calculated index value is therefore based on different interpretation templates depending on sex and breed.

Interrelations between the traits studied were analysed by means of principal component analysis as a factor extraction method. Factors were extracted until 75% of the original variance was explained by the factors retained. The computer program StatView 4.0 (Abacus Concepts) was used for the calculations.

Breed and sex differences were calculated from differences in the expected cell means for each trait using the standard MANOVA procedures of the computer program Data Desk (1995). Separate analyses were made to calculate the differences between the breeds, and those between the sexes both within and between breeds. The results presented here are based either on scores or on partial index values calculated from the results.

3. Results

One critique of the scoring system used in the test was the small range of scores for certain score groups, several of which contained less than 1% of the tested dogs.

3.1. Differences between breeds

Table 1 shows that there were significant differences between breeds when comparing males and females separately in three of ten tested characteristics. When comparing breeds, disregarding sex, there were significant differences in eight of ten characteristics. The German shepherds scored significantly higher both for sharpness and defence drive,

Table 1
Differences in test scores between breeds within each sex and in both sexes (German shepherds minus Labrador retrievers)

	Males	Females	Males and females
Courage	−0.08 ns	−0.07 ns	−0.06 *
Sharpness	0.06 ns	0.08 **	0.07 *
Prey drive	−0.03 ns	−0.08 *	−0.04 ns
Defence drive	0.70 ***	0.18 **	0.50 ***
Nerve stability	−0.27 ***	−0.39 ***	−0.32 ***
Temperament	−0.26 **	0.06 ns	−0.10 ns
Hardness	−0.12 ***	−0.02 ns	−0.05 *
Ability to cooperate	0.06 ns	−0.28 ***	−0.12 **
Affability	−0.27 **	−0.10 ns	−0.18 *
Reaction to gunfire	−0.16 ***	−0.10 **	−0.13 ***

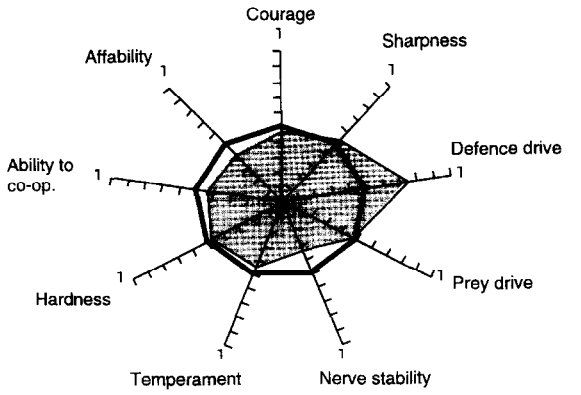

Fig. 1. Comparison of test results for all German shepherds (grey area, $n = 1310$) and all Labrador retrievers (thick lines, $n = 797$), based on test scores. The means for Labrador retrievers have been set to 0.5 for each characteristic and the score range to 1.0.

while the Labradors scored significantly higher for courage, nerve stability, hardness and reacted less to gun fire. They were also more cooperative and more affable than the German shepherds (Table 1 and Fig. 1)

3.2. *Differences between sexes*

A comparison of Figs. 2 and 3 shows that differences between sexes vary between the two breeds. For both breeds there were significant differences between sexes in six of the ten test situations (Table 2). In both breeds the males scored significantly higher than the females for courage, prey drive and defence drive. In German shepherds the

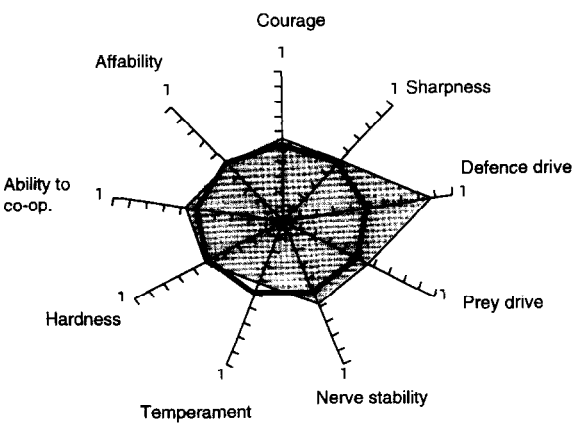

Fig. 2. Comparison of test results for male German shepherds (grey area, $n = 730$) and female German shepherds (thick lines, $n = 580$), based on test scores. The means for females have been set to 0.5 for each characteristic and the score range to 1.0.

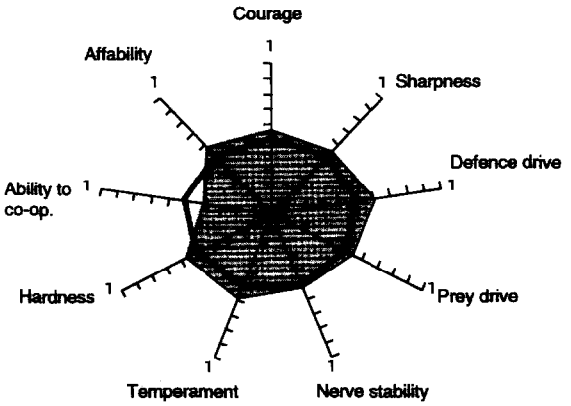

Fig. 3. Comparison of test results for male Labrador retrievers (grey area, $n = 343$) and female Labrador retrievers (thick lines, $n = 454$), based on test scores. The means for females have been set to 0.5 for each characteristic and the score range to 1.0.

males scored significantly higher than females for nerve stability, but there was no significant difference between sexes for the Labradors. Male German shepherds achieved significantly higher scores for ability to cooperate compared with females of the same breed, while in Labradors the opposite was found: the females scored significantly higher than the males (Table 2).

3.3. Factor analysis

The correlations of all evaluated characteristics and factor analysis were first calculated between the sexes of both breeds. The differences in results between sexes were small when the analysis was performed between sexes of each breed separately.

Table 3(a,b) shows that the correlation for several of the tested characteristics was relatively high for all dogs in each breed. In both breeds the factor analysis gave four factors with varied proportionate contribution of variance (Table 4). Table 5(a,b) shows

Table 2
Differences in scores between sexes within the two breeds (males minus females) and in both breeds

	German shepherds	Labrador retrievers	German shepherds and Labrador retrievers
Courage	0.10 **	0.11 **	0.10 ***
Sharpness	0.05 ns	0.07 *	0.07 **
Prey drive	0.13 ***	0.08 *	0.10 ***
Defence drive	0.73 ***	0.22 ***	0.59 ***
Nerve stability	0.15 ***	0.02 ns	0.06 *
Temperament	−0.15 *	0.17 ns	−0.04 ns
Hardness	0.05 ns	0.15 ***	0.08 ***
Ability to cooperate	0.14 **	−0.21 ***	0.00 ns
Affability	0.00 ns	0.17 ns	0.04 ns
Reaction to gunfire	−0.04 ns	0.01 ns	−0.04 ns

Table 3
Correlation of characteristics for (a) German shephers and (b) Labrador retrievers

	Courage	Sharpness	Prey drive	Defence drive	Nerve stability	Temperament	Hardness	Cooperativeness	Affability
(a) *German shepherds* (1310 animals: 730 males, 580 females)									
Courage	1								
Sharpness	0.215	1							
Prey drive	0.346	0.304	1						
Defence drive	0.448	0.439	0.490	1					
Nerve stability	0.497	0.106	0.185	0.296	1				
Temperament	0.131	0.151	0.380	0.179	−0.034	1			
Hardness	0.585	0.276	0.401	0.457	0.433	0.217	1		
Ability to cooperate	0.162	−0.009	−0.190	0.037	0.356	−0.479	0.100	1	
Affability	0.101	0.085	0.240	−0.024	0.204	−0.062	0.100	0.222	1
(b) *Labrador retrievers* (797 animals: 343 males, 454 females)									
Courage	1								
Sharpness	0.125	1							
Prey drive	0.356	0.269	1						
Defence drive	0.316	0.488	0.379	1					
Nerve stability	0.421	0.150	0.185	0.138	1				
Temperament	0.060	0.170	0.353	0.164	−0.035	1			
Hardness	0.533	0.212	0.433	0.397	0.369	0.199	1		
Ability to cooperate	0.201	−0.089	0.019	−0.023	0.236	−0.457	0.039	1	
Affability	0.141	0.009	0.128	0.053	0.076	−0.046	0.120	0.226	1

Table 4
Factor analysis: proportional variance contribution of the different factors (orthogonal)

	German shepherds	Labrador retrievers
Factor 1	0.358	0.334
Factor 2	0.242	0.247
Factor 3	0.159	0.247
Factor 4	0.241	0.171

Table 5
Factor analysis for (a) German shepherds and (b) Labrador retrievers. Factor loading for the four factors. Orthogonal transformation, solution Varimax

	Factor			
	1	2	3	4
(a) *German shepherds*				
Courage	**0.824**	−0.012	0.012	0.151
Sharpness	0.024	0.003	−0.119	**0.894**
Prey drive	0.441	0.329	0.003	**0.501**
Defence drive	0.479	0.027	−0.116	**0.673**
Nerve stability	**0.760**	−0.273	−0.180	−0.032
Temperament	0.202	**0.870**	0.042	0.114
Hardness	**0.770**	0.105	0.045	0.244
Ability to cooperate	0.250	**−0.806**	0.177	0.019
Affability	0.075	−0.084	**0.973**	0.032
(b) *Labrador retrievers*				
Courage	**0.409**	−0.026	−0.097	−0.027
Sharpness	−0.173	−0.118	**0.660**	−0.080
Prey drive	0.157	0.231	0.064	0.255
Defence drive	0.037	−0.073	**0.534**	−0.009
Nerve stability	**0.428**	−0.156	−0.116	−0.254
Temperament	0.071	**0.576**	−0.118	0.101
Hardness	**0.357**	0.106	−0.032	−0.003
Ability to cooperate	0.081	**−0.478**	−0.034	0.176
Affability	−0.123	−0.006	0.032	**0.891**

Table 6
Factor analysis. Variable complexity (orthogonal) for the tested characteristics

	German shepherds	Labrador retrievers
Courage	1.068	1.081
Sharpness	1.038	1.014
Prey drive	**2.725**	**3.538**
Defence drive	1.953	1.261
Nerve stability	1.381	1.273
Temperament	1.148	1.081
Hardness	1.247	1.350
Ability to cooperate	1.296	1.475
Affability	1.029	1.037
Mean	1.432	1.457

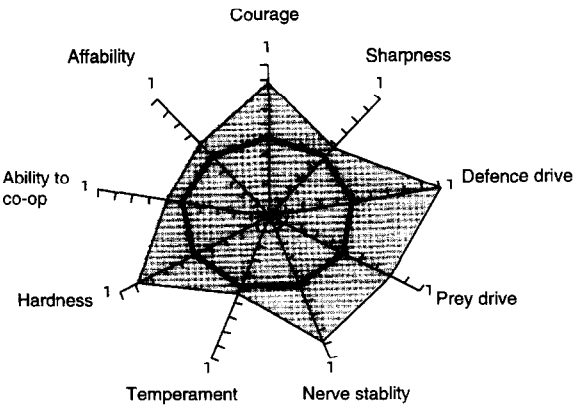

Fig. 4. Comparison of partial index values for 126 police dogs (grey area) and all German shepherds (thick lines, $n = 1310$). The means for all German shepherds have been set to 0.5 for each characteristic and the partial index range to 1.0.

that factor 1 in both breeds comprises the characteristics courage, nerve stability and hardness, and factor 2 comprises the characteristics temperament and ability to cooperate. The characteristic affability alone makes up factor 3 (German shepherd) or factor 4 (Labrador retriever). For German shepherds, factor 4 comprises the characteristics sharpness, prey drive and defence drive. For Labrador retrievers the corresponding factor 3 comprises the characteristics sharpness and defence drive, while prey drive does not contribute to any of the four factors. Table 6 shows that the characteristics prey drive in both breeds, and the defence drive in German shepherds are relatively complex. The ideal value of complexity is 1.0, i.e. each characteristic contributes to only one of the factors.

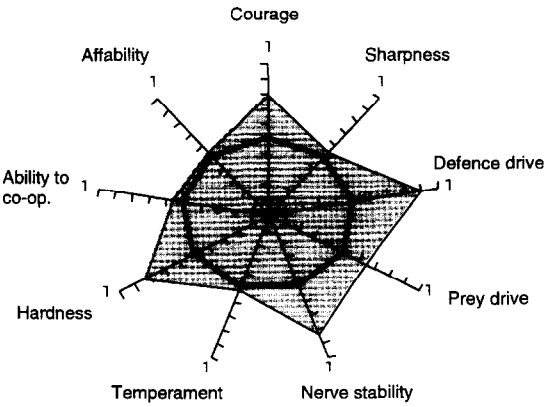

Fig. 5. Comparison of partial index values for 53 protection dogs (grey area) and all German shepherds (thick lines, $n = 1310$). The means for all German shepherds have been set to 0.5 for each characteristic and the partial index range to 1.0.

The primary correlations between factors 1 and 4 (German shepherd) and factors 1 and 3 (Labrador retriever) were 0.391 and 0.287, respectively. Since factor 3 for Labrador retrievers is the same as factor 4 for German shepherds, the complex of courage, nerve stability and hardness correlate with the complex sharpness, prey drive and defence drive in both breeds.

3.4. Test results and age

The mean test age for German shepherds were 526.5 ± 38.4 days and for Labrador retrievers 534.5 ± 33.2 (mean ± SD). Within these age ranges the effects of the test age were low and the changes in the index value was only +0.002 (ns) units per day for male German shepherds and +0.001 (ns) units per day for female German shepherds.

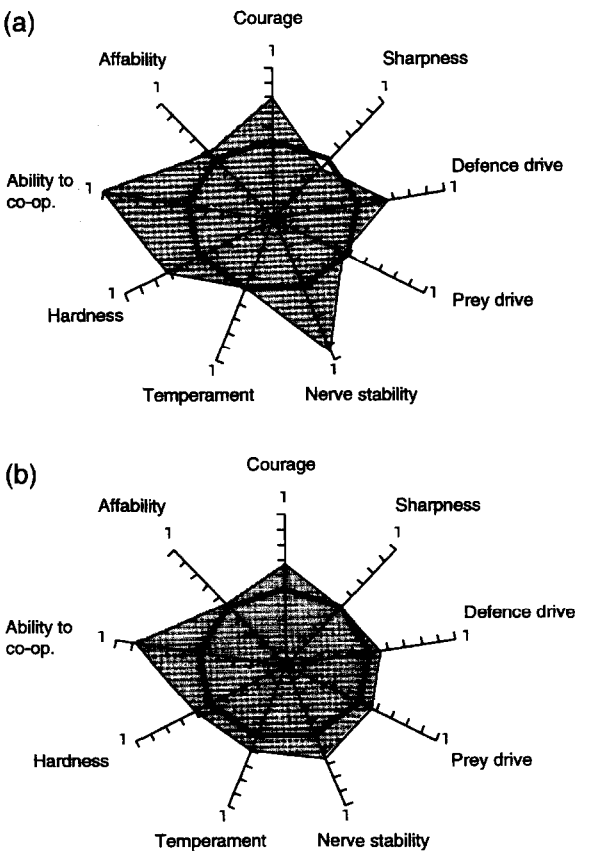

Fig. 6. (a) Comparison of partial index values for 18 German shepherds (grey area) selected as guide dogs, and all German shepherds (thick lines, $n = 1310$). The means for all German shepherds have been set to 0.5 for each characteristic and the partial index range to 1.0. (b) Comparison of partial index values for 75 Labrador retrievers (grey area) selected as guide dogs, and all Labrador retrievers (thick lines, $n = 779$). The means for all Labrador retrievers have been set to 0.5 for each characteristic and the partial index range to 1.0.

The corresponding values for Labrador retrievers were +0.006 (ns) and +0.005 (ns). Based on this, the index values were not corrected for age in any of the calculations.

3.5. Differences in test results for different types of service dogs

The different partial index values of different types of service dogs compared to the average value within the breed are shown in Figs. 4–9. Fig. 4 shows that police dogs scored markedly higher for the characteristics courage, hardness, prey drive and defence drive, and showed better nerve stability than average for the breed. The temperament of the police dogs, however, did not differ significantly from the average results for all German shepherds. Figs. 4 and 5 show that requirement nerve stability was the same for protection dogs and police dogs. The results for German shepherds sold as guide dogs differed from those of all dogs tested in a different way (Fig. 6a). German shepherds sold as guide dogs have higher partial index values for the characteristics ability to cooperate, courage, and nerve stability than the average dog within the breed. Compared with police dogs (Fig. 4), they had lower partial index values for sharpness, defence drive and prey drive, but a higher partial index value for ability to cooperate. Otherwise, there are small variations from the mean value for all German shepherds.

Table 7
Index values for various categories of (a) German shepherds, and (b) Labrador retrievers

	n	Index value		
		Mean	Min.	Max.
(a) *German shepherds* ($n = 1310$)				
Breeding	64	13.57	5.30	18.80
Police dogs	126	13.37	8.70	19.80
Protection dogs	53	12.11	5.40	16.20
Search dogs (narcotics, rot detection)	17	12.69	7.90	18.20
Guard dogs	9	8.52	4.10	13.60
Guide dogs	18	11.90	7.60	18.80
Other service dogs	34	10.18	5.20	15.20
Rejected, sold as companion animals	89	9.00	3.10	17.50
Rejected, given away	699	7.43	2.20	19.20
Rejected, euthanized for medical reasons	50	7.64	2.90	16.60
Rejected, euthanized for behavioural reasons	147	6.87	1.40	17.20
Others	4	12.35	10.30	17.50
All dogs	1310	8.77	1.40	19.80
(b) *Labrador retrievers* ($n = 797$)				
Breeding	41	10.89	5.90	14.90
Search dogs (narcotics, rot detection)	87	10.94	4.80	17.00
Guide dogs	75	10.73	6.60	15.00
Other service dogs	1	10.70	10.70	10.70
Rejected, sold as companion	67	8.17	2.20	13.30
Rejected, given away	463	7.76	2.00	17.10
Rejected, euthanized for medical reasons	18	7.72	4.30	11.60
Rejected, euthanized for behavioural reasons	42	7.06	2.90	13.00
Others	3	11.40	10.40	13.30
All dogs	797	8.56	2.00	17.10

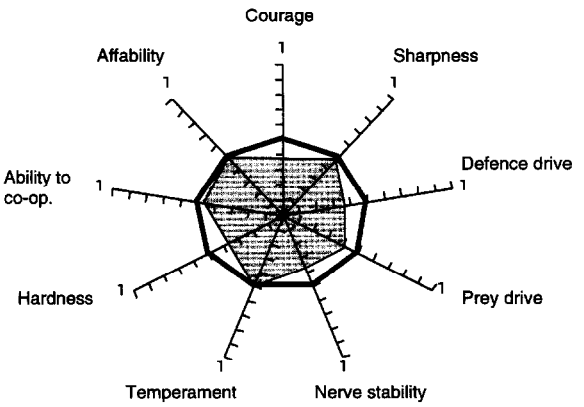

Fig. 7. Comparison of partial index values for 699 German shepherds (grey area) rejected as service dogs and all German shepherds (thick lines, $n = 1310$). The means for all German shepherds have been set to 0.5 for each characteristic and the partial index range to 1.0.

From a comparison of Labrador retrievers selected as guide dogs and the average for the breed, the greatest difference was found in the characteristic ability to cooperate. This was also evident in German shepherds. In contrast with German shepherds, the Labrador retrievers selected as guide dogs did not differ as much with regard to nerve stability and courage compared with the average dog within the breed due to the fact that Labrador retrievers generally scored higher than German shepherds for these characteristics (Table 7).

Dogs rejected as service dogs were donated to private owners if suitable as companion animals, and the dogs rejected as being unsuitable as companion animals were euthanized. Fig. 7 shows that donated animals or those sold as companion dogs

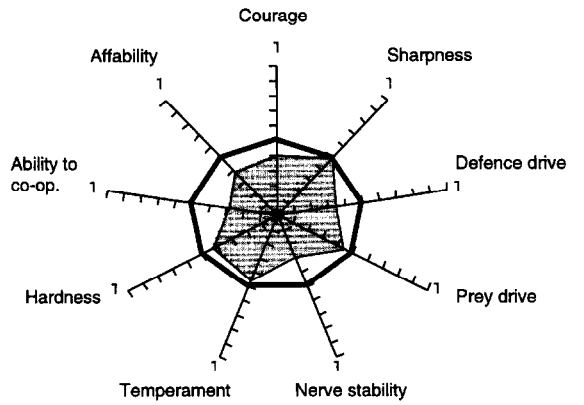

Fig. 8. Comparison of partial index values for 147 German shepherds (grey area) euthanized for behavioural reasons, and all German shepherds (thick lines, $n = 1310$). The means for all German shepherds have been set to 0.5 for each characteristic and the partial index range to 1.0.

generally achieved lower partial index values than the average for the breed. 147 German shepherds euthanized for behavioural reasons differed even more from the average for the breed (Fig. 8). The euthanized dogs showed the greatest deficiencies in nerve stability, defence drive, affability and ability to cooperate.

4. Discussion

The results show marked differences in the mentalities of the two breeds. German shepherds, which are mostly used as police and protection dogs, scored higher for sharpness but even more for defence drive than the Labrador retrievers. Labrador retrievers scored higher for nerve stability, reacted less strongly to gunfire and were more cooperative than the German shepherds. These differences alone make Labrador retrievers more suitable as guide dogs, and German shepherds more suitable as police or protection dogs. These differences in the behaviour of the two breeds were previously shown by Scott and Fuller (1965) to be due to the different ways in which the two breeds were originally used. The breed difference in this study can be explained the same way. Labrador retrievers were originally used as hunting dogs, working closely with their handlers and able to stand gunfire at close range, while German shepherds were used for herding and guarding livestock, but for the last few decades have been used as protection dogs. Because the German shepherds from the SDTC originated from a closed breeding colony, and because the majority of German shepherds in this study were bred by SDTC (1002 out of 1310), it is possible that this breed comparison would have turned out differently if only privately bred German shepherds had been used. Differences in the characteristics of purpose-bred guide dogs and privately bred dogs have also been reported by Goddard and Beilharz (1982, /83).

Tables 1 and 2, and Figs. 2 and 3 also show that the inter-breed differences vary somewhat between sexes. This is especially evident for the characteristic ability to cooperate (Table 1); male Labrador retrievers were less able to cooperate than male German shepherds, whereas female Labrador retrievers were more able to cooperate than female German shepherds. Male German shepherds scored lower for temperament than male Labrador retrievers, while female German shepherds scored higher for temperament than female Labrador retrievers (Table 2, Figs. 2 and 3). This can be explained by the greater male sex characteristics often shown by male Labrador retrievers. This difference in behavioural sexual dimorphism between breeds is best illustrated by comparing Figs. 2 and 3. Male Labradors are usually perceived by trainers to be only suitable for work where the ability to work independently is valued. The lack of ability to cooperate with the handler is considered to disappear if the dogs are castrated before they are one year old, which would suggest that this negative side of male Labrador retrievers is hormonally regulated.

Goddard and Beilharz (1984) point out that factor analysis should have a predictive value, and suggest that the calculated factors from a factor analysis could be used as a basis for summarising the assessed parameters into new ones. Since the terminology used by dog handlers to describe characteristic dog behaviour often varies, the results of a factor analysis could be helpful in establishing a standard terminology.

The four factors obtained from the factor analysis (Table 5a,b) could therefore form a basis for the definition of four altogether new characteristics. Factor 1, comprising courage, nerve stability and hardness, could be summarised into mental stability. Factor 2, comprising temperament and ability to cooperate, could be combined under cooperation or willingness to please. Factor 3 (German shepherds) and factor 4 (Labrador retrievers) comprise only affability, indicating that the way in which a dog relates to humans is a characteristic separate from other behaviour characteristics. Factor 4 (German shepherds) comprises sharpness, defence drive and prey drive, could be described as ardour. Contained in the corresponding factor for Labrador retrievers (factor 3) are the characteristics sharpness and defence drive; the characteristic prey drive does not contribute to any of the factors. Neither sharpness nor defence drive are desirable in Labrador retrievers, which are mostly used as guide dogs. The characteristics most used in the training of German shepherds are defence drive, sharpness, and above all their prey drive. This is shown in Figs. 4 and 5, where German shepherds selected as police and protection dogs have higher partial index values for defence drive and prey drive than the average.

A comparison of the index values for service dogs and the dogs that were rejected for various reasons, show that this system of creating one index value from added partial index values for each dog can be one way of simplifying and increasing the reliability of the selection of dogs for various types of service work.

After the index value had been evaluated, the results of this calculation were routinely entered into the database used by SDTC. The test results for each dog were then also entered, and an index value was automatically calculated. This index value could then be used to evaluate both individual dogs and for progeny testing.

All dogs selected for training to become service dogs, in spite of the varied requirements for the different kinds of service dogs, achieved higher index values than the rejected animals (Fig. 8). The most obvious explanation for this is that various categories of service dogs share more mental similarities than differences. For German shepherds, the partial index value was calculated only to evaluate dogs selected as police and protection dogs, whereas for Labrador retrievers it was used to evaluate those selected as guide dogs. The different characteristics of the breed therefore determine how the index value should be calculated for each breed.

The tests were performed identically for the two breeds in spite of inter-breed variations and different intended uses, but the results were interpreted differently. One of the conclusions of this study is therefore that the same behaviour test can be used to select different types of service dogs, even when from different breeds. Another conclusion is that behaviour tests used to select working dogs should first provide important information about the individual. The interpretation of the results can be adjusted thereafter, according to known weaknesses within the selected population as well as for the intended use of the dog.

The most obvious criticism of this test is that with regard to the Labrador retrievers it does not evaluate the various types of motivations most pronounced in the breed, nor the actual use they are bred for, considering the use of the breed as a gun dog. The characteristic prey drive in German shepherds is interpreted as an engagement in objects and the tendency to establish ownership of the object. In Labrador retrievers this could

be expressed as a drive to search for objects, which was not evaluated in this test. For Labrador retrievers the test should therefore be extended to focus on the dogs' search or retrieving drive in order to establish their suitability as search dogs (i.e. for detecting narcotics, bombs, or rot or mould).

In any behaviour testing, where the results will be used for various kinds of selection, it is of utmost importance that the test results are objective as well as reliable. This study, however, is based on an experienced test leader's subjective observations of a dog's reaction during and immediately after a test situation. This allows an entire behaviour complex to be evaluated, rather than just the dog's immediate reaction to a test situation.

The characteristic nerve stability, which is judged by trainers as the most important characteristic, would be very difficult to evaluate by an objective description of reaction. The results of this study show that this, and other subjectively evaluated complex behavioural parameters, can be evaluated with a high degree of reliability. However, the reliability of the results is inversely proportional to the number of people doing the observations. The reliability of the test could perhaps be improved by repetitive testing procedures and independent judgements by two or more test leaders. The behaviour test could also easily be improved by changing the scoring system to combine existing groups with comparatively few animals, as well as by dividing groups containing large numbers of dogs into two or three groups with different scores.

References

Data Desk, 1995. Data description Inc. Ithaca, NY 14852, USA.

Goddard, M.E. and Beilharz, R.G., 1982. Genetic and environmental factors affecting the suitability of dogs as guide dogs for the blind. Theor. Appl. Genet., 62: 97–102.

Goddard, M.E. and Beilharz, R.G., 1982/83. Genetics of traits which determine the suitability of dogs as guide dogs for the blind. Appl. Anim. Ethol. 9 (1982/83): 299–315.

Goddard, M.E. and Beilharz, R.G., 1984. A factor analysis of fearfulness in potential guide dogs. Appl. Anim. Behav. Sci., 12: 253–65.

Humphrey, E.S., 1934. Mental tests for Shepherd dogs. J. Hereditas, 25: 128–36.

Mackenzie, S.A, Oltenacu, E.A.B and Leighton, E., 1984. Heritability estimate for temperament scores in German shepherd dogs and its genetic correlation with hip dysplasia. Behav. Genet., 15: 475–482.

Mackenzie, S.A., Oltenacu, E.A.B. and Houpt, K.A., 1986. Canine behavioral genetics: A review. Appl. Anim. Behav. Sci., 15: 365–393.

Murphree, O.D and Dykman, R.A., 1965. Litter patterns in the offspring of nervous and stable dogs I: Behavioral tests. J. Nerv. Ment. Dis., 141: 321–332.

Pfleiderer-Högner, M., 1979. Möglichkeiten der Zuchtwertschätzung beim Deutschen Schäferhund anhand der Schutzhundprufung I. Doctoral Thesis, Ludwig-Maxmilians Universität, Munich, Germany.

Reuterwall, C. and Ryman, N., 1973. An estimate of the magnitude of additive genetic variation of some mental characters in Alsatian dogs. J. Hereditas, 73: 277–84.

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

Scott, J.P. and Bielfelt, S.W., 1976. Analysing the puppy-testing program, in: C.J. Pfaffenberger et al., Eds., Guide Dogs for the Blind: Their Selection, Development, and Training. Elsevier, Amsterdam, pp. 39–76.

Willis, M.B., 1989. Genetics of the Dog. Witherby, London.

Willis, M.B., 1995. Genetics of dog behaviour, with particular reference to working dogs, in: J. Serpell, Ed., The Domestic Dog. Cambridge University Press, Cambridge.