ELSEVIER

Applied Animal Behaviour Science 58 (1998) 151–162

APPLIED ANIMAL BEHAVIOUR SCIENCE

Behaviour test for eight-week old puppies—heritabilities of tested behaviour traits and its correspondence to later behaviour

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

[a] *Department of Zoology, Stockholm University,S-106 91 Stockholm, Sweden*
[b] *Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden*

Accepted 25 June 1997

Abstract

In order to test if adult behaviour could be predicted at eight weeks of age, 630 German shepherd puppies were tested. All dogs were also tested at 450–600 days of age according to regimen used to select service dogs. Significant gender differences were found in 4 of the 10 score groups of the puppy test. There were also significant correlations between the puppy test score groups. Correspondence of puppy test results to performance at adult age was negligible and the puppy test was therefore not found useful in predicting adult suitability for service dog work. Heritability was medium high or high for behaviour characteristics of the score groups in the puppy test. Maternal effects on the puppy test results were found when comparing estimations based on sire and dam variances. It also suggests that maternal effects are more likely to be seen in juvenile than in adult behaviour.

Keywords: Dog; Behavioural test; Temperament; Genetics; Ontogeny

1. Introduction

Puppy behavioural tests for different purposes have been presented. Scott and Fuller (1965) in their work at Bar Harbor performed a large number of tests on puppies and young dogs, from birth up to the age of one year. The purpose of these tests were mainly to describe behavioural, developmental and genetic differences in different breeds of dogs. One important finding in this report is that some characteristics show a consider-

* Corresponding author.

0168-1591/98/$19.00
PII S0168-1591(97)00093-2

able consistency along the maternal line indicating maternal effects to be important. The tests in Scott and Fuller's study were not used to predict the future usefulness of the dogs tested.

In a later study Scott and Bielfelt (1976) described a number of repetitive tests performed on puppies 8–12 weeks of age with the purpose to predict their future usefulness as guide dogs. The tests were found to be acceptably predictive, already after the first tests at eight weeks of age, suggesting that repetitive testing was not necessary. However the heritability of the tested characteristics was found to be low and effects of maternal environment to be high.

Goddard and Beilharz (1984) evaluated tests for guide dogs in Australia and found high heritability, especially for fearfulness. Goddard and Beilharz suggested that fearfulness, being one of the most common reason for rejection of dogs, could be evaluated as early as at 12 weeks of age. In a later report (1986) this statement was somewhat modified and it was suggested that the predictive value increased if dogs were tested later in life.

A number of tests on young dogs have also been performed in order to evaluate the effects of different experimental conditions. Murphree and Dykman (1965) described a test for puppies used to separate dogs from stable and nervous lines. In a later report Newton et al. (1978) compared behaviour tests with physiological parameters and concluded that the distinction between nervous and stable dogs could be made with 100% certainty by using physiological tests and by 95% certainty by behavioural tests. Martinek et al. (1978) reported differences in the ability of neonate puppies to orient from a cold surface to a warmer and that this ability corresponded to habituation tests at four months of age. Fox (1972) and Campell (1975) have also described behavioural tests for puppies used to select puppies from a litter. Neither of these tests were evaluated or proven to have any predictive value.

The purpose of this study was to investigate if a puppy test could be used to predict adult usefulness. The goal was also to estimate the hereditary effects on the behaviour of puppies. The puppy test results are compared to the results at adult age, a test previously shown to be useful when selecting dogs for work and breeding (Wilsson and Sundgren, 1997a).

2. Material and methods

A total of 1235 eight-week old German shepherd puppies, all bred at the Swedish Dog Training Centre (SDTC) were tested during the years 1978–1983. Most of these dogs were also tested as adults, at 450–600 days of age, according to regimens previously described by Wilsson and Sundgren (1997a). After excluding dogs with incomplete records, the material consisted of 867 dogs with complete records from both tests. In the preliminary analysis in year 2, puppies weighed 25% less than puppies in year 1 due to changes in feeding regime. The analysis was therefore carried out on the remaining 630 dogs tested in year 3–6. The estimations of heritability were made on all litters after sires with at least two litters and with at least two puppies in each litter during year 3–6 which leaves 277 males and 277 females after 84 sires and 109 dams.

The mothers were recalled from the hosts two weeks prior to estimated time of whelping. After parturition the bitches were fed four times per day. The puppies were offered solid standard commercial dog food twice a day from the age of three weeks with one additional meal per day for each week until four meals a day were fed from the sixth week on. The mothers were continually housed with the puppies until the puppies were tested and sent to puppy walkers where they remained until recalled for testing at 450–600 days of age.

Most of the characteristics tested at adult age have not matured at eight weeks of age. The test performed on adult dogs also include situations where the dog is threatened or exposed to fear eliciting stimuli, situations not appropriate for puppies. Therefore, the intention was not to adapt the puppy test to the adult test. The puppy test focused instead on characteristics with considerably variability among eight-week old puppies, i.e., sociability, independence, fearfulness, competitiveness, general activity and explorative behaviour. Some of the test situations were adapted from earlier tests (Scott and Fuller, 1965; Fox and Stelzner, 1966; Fox, 1972; Scott and Bielfelt, 1976; Fuller and Clark, 1968). The total test time per puppy did not exceed 30 min. The tests were conducted by four persons, at least two present at each test occasion.

2.1. Description of puppy test

The test is divided into parts I and II.

2.1.1. Part I: vocalisation when isolated

One puppy is randomly picked out and placed in an empty breeding pen (approx. 2.5×3.0 m). The puppy is left alone in the pen and the door to the pen is closed. The test is carried out in order to test the emotional distress when isolated.

2.1.1.1. Yelp. The time until the puppy whines or yelps is noted. If the puppy does not vocalise within 5 min the score time 300 s is noted.

2.1.1.2. Shriek. The time until the puppy has given at least three distinctive shrieks is noted. If the puppy does not show the described reaction within 5 min the score time 300 s is noted.

2.1.1.3. Contact I. When the previous test is completed the test leader (TL), not previously encountered by the puppy, enters the pen without paying attention to the puppy. The puppy's reaction is scored from 1 to 5. Score 5 is given puppies overly willing to make contact, yelping and jumping up towards the person. Score 4 is given puppies spontaneously making contact without hesitation, without jumping and yelping. Score 3 is given puppies spontaneously making contact however after some hesitation. Puppies initially withdrawing but later allowing TL to make contact is given score 2 and those not allowing TL to make contact is given score 1.

After the score is determined the TL shows the puppy a tennis ball. When the puppy's attention is gained, the ball is rolled over the floor. The retrieving test is divided

into two parts, fetch and retrieve. Willingness to retrieve objects in general is considered a good measurement of cooperativeness and have previously been shown to have a high predictive value (Scott and Bielfelt, 1976) and it was also one of the tests that showed the highest heritability in the same study.

2.1.1.4. Fetch. The time it takes for the puppy to take the ball into its mouth, from time of becoming attentive to the ball, is noted. If this does not happened within 3 min the score time of 180 s is noted.

2.1.1.5. Retrieve. The puppy's willingness to retrieve the ball is scored from 1 to 5. Score 5 is given puppies that retrieve the ball spontaneously at least three successive times. Score 4 is given puppies retrieving once or twice. Score 3 is given puppies carrying the ball without bringing it back to the TL. Score 2 is given puppies running after the ball however without taking it in the mouth and score 1 to puppies not showing any interest in the ball.

2.1.1.6. Large ball. A larger ball (13 cm diameter) is placed in the centre of the pen. The puppy's immediate reaction to the ball is scored from 1 to 5. Score 5 is given puppies that spontaneously starts to play with the ball, score 4 to puppies starting to play with the ball after it is rolled over the floor. Score 3 is given puppies investigating or playing with the ball only after it is rolled over the floor. Score 2 is given puppies running after the ball however without investigating it before or after it has been rolled. Score 1 is given puppies that withdraw from the ball. This test situation is similar to the previous retrieving test however larger objects may in some puppies elicit a fear response.

2.1.1.7. Tug of war. The TL then draws the puppy's attention to a cotton rag. If the puppy grasps the rag, TL does not let go. The engagement of the puppy is evaluated and scored from 1 to 5. Score 5 is given puppies that highly engage and shake the rag, i.e., 'death shake'. Score 4 is given puppies getting highly engaged but not showing the 'death shake'. Score 3 is given puppies biting and chewing but lets go of the rag as soon as the TL starts to pull. Score 2 is given puppies investigating the rag without taking it in the mouth. Score 1 is given puppies not paying any attention to the rag. The purpose of the test was to test the possibility to evaluate the puppy's future interest in objects, a characteristic often used when training adult dogs.

2.1.2. Part II: arena test

All puppies in a litter are subjected to all of the tests in part I before part II, the arena test, is started. All puppies are moved to an unfamiliar building. The arena is partitioned of area 3.6×3.6 m where the floor is marked in squares, 0.6×0.6 m. In the middle of the arena a circle 2 m in diameter is marked. Spread out on the floor of the arena are four objects; a piece of wood $3 \times 7 \times 20$ cm and the three objects used in part I (the tennis ball, the large ball and the cotton rag) (Section 2.1.1). For 5 min, the TL sits quietly and passively on a chair in the middle of the circle. The purpose of the test is to test the dog's explorative behaviour with a passive person present. The three score groups reflect different aspects of the explorative tendencies of the dog.

2.1.2.1. Contact II. Time (in seconds) spent inside the circle during the 5 min is noted.

2.1.2.2. Activity. Number of squares the puppy moves over during the 5 min.

2.1.2.3. Objects visited. Number of objects visited during the test.

2.2. Statistical analysis

Statistical analyses were made using a standard statistical computer program, Stat View v. 4.0. The linear models for Table 1 was calculated by the use of GLM procedure of Data Desk v. 4.0 and the exact model was:

$$y_{ijkl} = \mu + G_i + P_j + L_k + W_l + e_{ijkl}$$

where y_{ijkl} = test result; μ = general mean; G_i = fixed effect of gender, discontinuous; P_j = fixed effect of parity, continuous; L_k = fixed effect of litter size, continuous; W_l = fixed effect of weight at 50 days, continuous; e_{ijkl} = error term.

Heritabilities of puppy scores were estimated from intra class correlations between sibs within groups of full and half sibs. Calculations were made in accordance to formulas given by Becker (1985) for treatment of unbalanced data, i.e., variable sizes of progeny groups for both sires and dams. Estimates were made separately on sire components, dam components and combined sire–dam components of variance. Before calculations, the data was corrected for effects of year and sex.

All comparisons to adult behaviour are made by comparing the puppy test scores to partial index values derived from scores achieved in the adult test. A detailed description of the adult test and of the methods for calculating the partial index value is presented by Wilsson and Sundgren (1997a). The partial index value can be explained as a predictive value of the dog's suitability for service dog work. High partial index values are desirable for all evaluated characteristics. The index value is the sum of all partial index values given each individual.

Table 1
LS-means in the puppy test results for each sex respectively. Differences between sexes and level of significance (GLM procedure)

Characteristic	Mean		
	Males ($n = 320$)	Females ($n = 310$)	Difference (m − f)
Yelp	44.3	50.8	−6.5
Shriek	222.4	226.6	−4.2
Contact I	4.8	4.8	0.0
Fetch	61.6	49.4	12.2 *
Retrieve	4.1	4.2	−0.1
Large ball	4.4	4.6	−0.2
Tug of war	3.3	3.3	0
Activity	84.8	92.3	−7.5 *
Contact II	113.8	94.0	19.8 * * *
Objects visited	6.9	8.1	−1.1 * * *

Level of significance: * $P < 0.05$; * * $P < 0.01$; * * * $P < 0.001$.

3. Results

3.1. Effects of year and weight

As shown by Fig. 1, mean puppy weight at 50 days of age decreased during the second year, due to changes in feeding regimen. Data from year 1 and 2 therefore was excluded and data from year 3–6 was analysed on the remaining 630 puppies (320 males and 310 females).

3.2. Effects of sex and age

The mean test age was 56.15 days with a standard deviation of 1.95. Test age was not found to have any significant effect on test results with one exception; the regression on contact II for females (regression coefficient $= -5.66$; $P < 0.001$, one day in test age corresponded to a decrease in time spent in circle of 5.66 s).

The mean value for acquired test scores are shown for each sex separately in Table 1. Significant differences between sexes were found in 4 of the 10 score groups. Female puppies showed significantly shorter fetch times, were more active, spent less time within the circle and visited more objects during the arena test compared to male puppies.

3.3. Correlation analysis

Table 2 shows the correlation between the puppy score groups. For both sexes a strong correlation was found between the two score groups in the isolation test (yelp and shriek). Strong correlation was also found between score groups where puppies manipulated objects (fetch, retrieve, reaction to large ball, tug of war). There was also a strong correlation between the different score groups in the arena test (activity, contact II, objects visited).

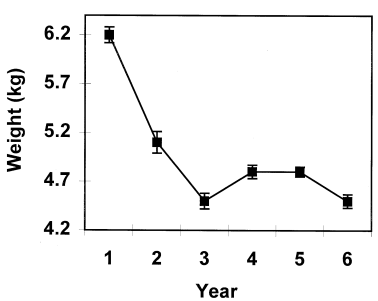

Fig. 1. Mean weight at 50 days of age for puppies tested in year 1 to 6. Mean ± standard error.

Table 2
Correlation between characteristics tested in 320 male (m) and 310 female (f) puppies. Test of significance by Fisher's r to z

		Yelp	Shriek	Contact I	Fetch	Retrieve	Large ball	Tug of war	Activity	Contact II
Shriek	(m)	0.35***	1.00							
	(f)	0.37***								
Contact I	(m)	−0.19***	−0.11*	1.00						
	(f)	−0.19***	−0.08							
Fetch	(m)	−0.04	−0.01	0.03	1.00					
	(f)	−0.12*	−0.08	0.06						
Retrieve	(m)	0.03	−0.02	0.01	−0.83***	1.00				
	(f)	0.10	−0.02	0.05	−0.80***					
Large ball	(m)	0.07	−0.05	0.10	−0.48***	0.52***	1.00			
	(f)	−0.02	0.05	0.16**	−0.38***	0.44***				
Tug of war	(m)	0.01	−0.11	0.06	−0.30***	0.33***	0.32***	1.00		
	(f)	0.01	−0.07	0.05	−0.40***	0.36***	0.18***			
Activity	(m)	−0.01	−0.15**	−0.04	−0.19***	0.12*	0.16**	0.14*	1.00	
	(f)	−0.13*	−0.06	0.10	−0.08	0.13*	0.17**	0.13*		
Contact II	(m)	−0.12*	−0.07	0.10	0.08	−0.07	−0.06	−0.06	−0.27***	1.00
	(f)	−0.09	−0.13*	0.09	0.18**	−0.11*	−0.21***	−0.24***	−0.19***	
Objects visited	(m)	0.08	−0.03	−0.03	−0.13*	0.11	0.10	0.09	0.31***	−0.26***
	(f)	0.02	−0.01	0.03	−0.09	0.09	0.15**	0.13*	0.22***	−0.05

Level of significance: * $P < 0.05$; ** $P < 0.01$ and *** $P < 0.001$.

Table 3
Coefficients of regression for puppy test results on characteristics tested at mature age. Coefficient of regression not significant (ns) are not shown. The calculation is based on both sexes ($n = 630$, year 3–6)

	Yelp	Shriek	Contact I	Fetch	Retrieve	Large ball	Tug of War	Activity II	Contact II	Objects visited
Courage	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Sharpness	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Defense drive	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.024**
Prey drive	ns	ns	ns	−0.010*	0.047**	ns	ns	ns	ns	ns
Nerve stability	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Temperament	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Hardness	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Ability to cooperate	ns	ns	−0.166*	ns	ns	ns	ns	ns	ns	ns
Affability	−0.001*	ns	ns	ns	ns	ns	ns	ns	ns	ns
Index value	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Level of significance: * $P < 0.05$; ** $P < 0.01$ and *** $P < 0.001$.

Table 4
Heritabilities ± standard errors estimated on the puppy test results for 277 male and 277 female puppies tested year 3–6.

	Sire	Dam	Combined
Yelp	0.20 ± 0.24	0.25 ± 0.25	0.22 ± 0.08
Shriek	−0.24 ± 0.33	0.33 ± 0.73	0.24 ± 0.09
Contact I	0.05 ± 0.26	0.36 ± 0.28	0.21 ± 0.08
Fetch	0.22 ± 0.23	0.19 ± 0.24	0.21 ± 0.08
Retrieve	0.23 ± 0.23	0.17 ± 0.23	0.20 ± 0.08
Large ball	0.27 ± 0.25	0.27 ± 0.25	0.27 ± 0.09
Tug of war	0.21 ± 0.34	0.76 ± 0.36	0.48 ± 0.11
Activity	−0.60 ± 0.51	1.67 ± 0.60	0.53 ± 0.13
Contact II	0.04 ± 0.35	0.80 ± 0.38	0.42 ± 0.10
Objects visited	−0.28 ± 0.34	0.82 ± 0.40	0.27 ± 0.09

3.4. Correspondence of puppy test to adult behaviour test result

In Table 3, the regression of the puppy test score groups on the nine characteristics tested at adult age and the resulting index value are shown. Only three significant regressions were found at the $P < 0.05$ level and two at the $P < 0.01$ level, which is exactly what would be expected by pure chance. Different statistical methods used in order to weigh information from each score group of the puppy test failed to give any significant correspondence to performance at adult age.

3.5. Heritability of puppy test score groups

Estimates of heritability based on data corrected for year and sex differences are presented in Table 4. The highest estimates (combined values) were found on the score groups tug of war, activity, and contact II. Other estimates of heritability ranges from 0.20–0.27. Higher estimates on dam than on sire were found in the score groups shriek, contact I, tug of war, activity, contact II and objects visited.

4. Discussion

Behavioural sex differences in dogs have so far scarcely been reported. However, Scott and Fuller (1965) found males to be dominant over females in competitive situations at the age of eight weeks. Behavioural sex differences in adult dogs have also previously been reported by Wilsson and Sundgren (1997a). The results presented in Table 1 show a significant difference between the sexes in 4 out of 10 score groups. Female puppies were quicker to grab (fetch), were more active, spent less time in the circle in the arena test (contact II) and visited more objects. Based on the results one can say that female puppies were more active and independent than male puppies.

The correlation matrix (Table 2) shows highly significant correlation between the results in the isolation test (yelp and shriek) and between yelp and contact I. There was also a strong correlation between all score groups where puppies manipulated objects

(fetch, retrieve, large ball, and tug of war). Significant correlation between the score groups in the arena test were also found, which is not surprising since active puppies automatically spend less time in the circle (contact II) and visit more objects.

Estimates of heritabilities were found to be medium high or high. Mackenzie et al. (1986) (also cited by Willis (1995)) presented unpublished results from preliminary data from the tests evaluated in this paper. These calculations were however made on results achieved in year 1–4, when puppy weights changed. Since weight was found to have a great impact on the test results it is most likely that these changes in conditions also had a significant effect on the results presented by Mackenzie et al. (1986).

Heritability estimates based on sire and dam variances separately, in some cases show unexpected values. This can be explained by the rather small material where the distribution of variance components between sires and dams is strongly influenced by few observations. This may also explain why negative heritabilities in some cases are found. Normally higher heritability based on dam variance is interpreted as effects of dams or common litter environment. This tendency is shown for several of the estimates presented in Table 4. Negative estimates of heritability should however not be interpreted as large maternal effects. One should however be careful not to jump to conclusions regarding effects of dam or common litter environment unless there are substantial reasons to do so based on other observations.

Puppies are maturing rapidly at the age of eight weeks. If the test results of puppies are affected by level of maturation we would expect this to have a large impact on the results. One can also assume that there is a higher inter than intra litter variation in maturation. If so the degree of maturation would be a factor common to the litter and hence increase estimations of heritability based on dam. Much of the maternal effects seen in Table 4 may in fact be a consequence of high differences between litters in maturation.

This assumption is supported by the fact that despite low deviations in test age, there was a highly significant effect of test age on contact II, however only in females. The fact that test age had a significant effect on females but not in males imply different changes in behaviour due to maturation in the sexes. If this holds true we would expect the highest difference in test result between the sexes in the score group being the most affected by test age. This was certainly the case since contact II was the only score group significantly affected by test age and it was also the score group with the highest difference between the sexes (Table 1). As seen in Table 4, contact II was also the score group most effected by maternal effects.

If the behaviour of puppies is affected by degree of maturation to such an extent this would certainly diminish the predictive value of the test. Considering this, one would expect the predictive value of a test on young dogs to increase if dogs are tested later when the differences in behaviour is less affected by degree of maturation.

Medium high to high estimates of heritabilities found in puppies in this report and on adult dogs (Wilsson and Sundgren, 1997b) show that behaviour is strongly affected by hereditary factors. The lack of correspondence between puppy test results and test results at adult age could potentially be explained if juvenile behaviour is governed by different genes than behaviour at adult age. Another interpretation would be that completely different behaviour systems were tested in puppies than in adults. Although a test for

puppies has to be different from a test designed for adult dogs we would expect some correspondence between the tests if the adult phenotype can be predicted at the age of eight weeks. Scott and Bielfelt (1976) found high predictive values on tests performed at the age of eight weeks for tests evaluating interaction confidence in humans during easy obedience tasks. It is possible that the results would have been different in this study if some test situations of that type had been used. The differences in results between this study and that by Scott and Bielfelt may also be explained by the fact that their studies included several different breeds.

The negligible predictive value of the puppy test presented in this study is however similar to the results presented by Young [1] who found no correspondence between the results of Campbell's puppy test and the behaviour at adult age in 327 dogs tested at seven weeks and again at three years of age.

Goddard and Beilharz (1986) have shown that fearfulness, a characteristic often disqualifying guide dogs, may be predicted as early as at 12 weeks of age, but that a better prediction is achieved if dogs are tested at the age of six months. It is likely that the predictive value of the puppy test evaluated in this study may have increased if puppies had been tested at older age. It is also possible that the predictive value of the test would have increased if it was focused more on the fear reactions in puppies.

There are a number of papers presenting results showing maternal effects on behaviour (the results from this study, Scott and Fuller, 1965; Scott and Bielfelt, 1976). However all of these tests were conducted on puppies or comparatively young animals. Most reports dealing with behaviour tests on older dogs, more than one year old, have shown small or negligible maternal effects (Newton et al., 1978; Bartlett, 1976; Goddard and Beilharz, 1982; Wilsson and Sundgren, 1997b). Maternal effects therefore is most likely to be found on behaviour characteristics tested at juvenile age. As previously suggested, it is possible that some of the maternal effects seen when comparing heritabilities estimated on sire and dam are achieved due to inter-litter differences in maturation.

Since puppies are tested during a period when behaviour changes rapidly it makes correspondence to adult behaviour less likely to be found. From this point of view dogs should be tested later, when their behaviour has matured and when changes per time is less, for the results to predict the adult phenotype. The same general conclusions have previously been made by Goddard and Beilharz (1986) in their evaluations of guide dog tests in Australia.

The major conclusion of this paper is that individual variation in behaviour observed among puppies to a large extent can be explained both by hereditary factors and by effects of common litter environment. The results also implies that adult behaviour can not be predicted as early as at eight weeks of age. Breeding programs aimed to improve behaviour in dogs may not be based on information collected on tests performed as early as at eight weeks of age.

[1] Young, M., 1986. Alderens påvirkning på adferdstest av hunder. (The effect of age on behavioural tests in dogs). Lecture held in Geilo Norway, 4 and 5 October 1986. In: Owren, T. (Ed.), Dyr Og Mennesker I Fokus, in Norwegian (unpublished data).

References

Bartlett, C.R., 1976. Heritabilities and genetic correlation between hip dysplasia and temperament traits of seeing-eye Dogs. Masters Thesis, Rutgers University, New Brunswick, NJ.

Becker, W.A., 1985. Manual of Quantitative Genetics. Academic Enterprises, Washington State University, p. 190.

Campell, W.L., 1975. A behavior test for puppy selection. In: Campell, W.L. (Ed.), Behavior Problems in Dogs. Am. Vet. Publ., Santa Barbara, pp. 137–144.

Fox, M.W., 1972. Understanding Your Dog. Coward McCAnn & Geoghegan, New York.

Fox, M.W., Stelzner, D., 1966. Behavioural effects of differential early experience in the dog. Anim. Behav. 14, 273–281.

Fuller, J.L., Clark, C.L., 1968. Genotype and behavioral vulnerability to isolation in dogs. J. Comp. Physiol. Psychol. 66, 151–156.

Goddard, M.E., 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., Beilharz, R.G., 1984. A factor analysis of fearfulness in potential guide dogs. Appl. Anim. Behav. Sci. 12, 253–265.

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

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

Martinek, Z., Lat, J., Pribik, V., 1978. Individual behavioural differences in puppies during the neonatal period. Physiol. Bohemoslov. 27 (3), 260.

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

Newton, J.E.O., Dykman, R.A., Chapin, J.L., 1978. The prediction of abnormal behavior from autonomic indices in dogs. J. Nerv. Ment. Dis. 166 (9), 635–641.

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

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

Willis, M.B., 1995. Genetics of dog behaviour with particular reference to working dogs. In: Serpell, James (Ed.), The Domestic Dog, It's Evolution, Behaviour and Interactions with People. Cambridge Univ. Press, pp. 51–64.

Wilsson, E., Sundgren, P.E., 1997a. The use of a behaviour test for a 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., 1997b. The use of a behaviour test for selection of dogs for service and breeding. II. Heritability for tested parameters and effect of selection based on service dog characteristics. Appl. Anim. Behav. Sci. 54, 235–241.