Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Applied Animal Behaviour Science 97 (2006) 51–72
APPLIED ANIMAL BEHAVIOUR SCIENCE
www.elsevier.com/locate/applanim

Behavioural testing in dogs: A review of methodology in search for standardisation☆

Claire Diederich *, Jean-Marie Giffroy

Department of Anatomy and Ethology of Domestic Animals, University of Namur, 6, rue Muzet, 5000 Namur, Belgium

Available online 9 January 2006

Abstract

As a consequence of their living close to humans as pets, for working purposes or as laboratory animals, dogs give evidence of behavioural variability, stemming from their innate capacities as well as from environmental influences. This paper reviews the behavioural tests used for dogs—tests which serve as an evaluation tool and those which serve as a means of classifying individual animals. In search of a consensus and standardisation, some material and methodological aspects of behavioural testing in dogs were collected. Behavioural test parameters that were taken into account were the terminology of the temperament concept, the test quality requirements and their implementation in the literature, the characteristics of the dog tested (source, breed, age, sex), the characteristics of the social and environmental stimuli used to elicit canine behaviour, the characteristics of the behavioural variables collected and the characteristics of the physical and physiological concomitant data obtained while assessing the behaviour. This review brings to light a lack of consensus regarding all these parameters. The procedures of testing are often particular to the investigator and thus unique. We emphasised this statement by comparing six research studies using a ball, carried out over 40 years. In view of all these differences in methodology, standardisation is suggested through the creation of a reference manual.

Keywords: Dogs; Canine; Animal behaviour; Behaviour test

☆ This paper is part of the special issue entitled International Society for Applied Etholgy Special Issue—A Selection of Papers from the 38th International Congress of the ISAE, Helsinki, Finland, August 2004, Guest Edited by Victoria Sandilands and Carol Petherick.

* Corresponding author. Tel.: +32 81 74 05 52; fax: +32 81 72 44 20.
E-mail address: claire.diederich@fundp.ac.be (C. Diederich).

0168-1591/$ – see front matter
doi:10.1016/j.applanim.2005.11.018

1. Introduction

Over several decades, a great number of behavioural tests have been developed for puppies and for adult dogs. Their areas of application are numerous and varied: selection for breeding purposes (Wilsson and Sundgren, 1997a), behavioural development (Stanley, 1970; Fox, 1971), effects of genetics (Scott and Fuller, 1965; Murphree and Dykman, 1965) and environmental factors (Fox and Stelzner, 1967) on behaviour, learning abilities (Wyrwicka, 1959; Pongracz et al., 2005), prediction of working aptitude (Slabbert and Odendaal, 1999), likelihood of being adopted from a shelter (Ledger and Baxter, 1997), identification of behavioural problems (Netto and Planta, 1997), as an animal model of human disease (Head et al., 1995), and welfare (Beerda et al., 1997).

The behaviour of dogs can be observed either in their usual environment (i.e. at home, in kennels or while working) or, in contrast, when they are in the standardised conditions of a research laboratory. Animal behavioural tests are defined as standardised experimental situations where stimuli serve to elicit behaviour that is statistically compared with that of other individuals placed in the same situations, in order to classify the subject tested (Serpell and Hsu, 2001).

The need for greater standardisation in tests designed to evaluate dog behaviour has been highlighted (Goodloe, 1996; Murphy, 1998). This review explores the wide variety of dog tests conducted according to the following points: (1) the terminology of the temperament concept in dog research; (2) the tests' quality requirements and the way they are implemented; (3) the characteristics of the dogs tested (source, breed, age, sex); (4) the characteristics of the social stimuli and the environmental stimuli used in behavioural tests; (5) the characteristics of the behavioural variables collected; (6) the characteristics of physiological and physical measures that are taken in parallel with behaviour. Based on this review, we suggest directions for the future as regards standardisation in dog behavioural testing.

This paper is organised into sections independent of the type of behavioural testing (i.e. temperament testing or welfare assessment) because the same methodological aspects might be shared by several types of behavioural testing. For example, the behavioural response to a short and strong acoustic stimulus is analysed in papers focusing on selection of service dogs (starting pistol, Wilsson and Sundgren, 1997a), on animal welfare (fog horn, Beerda et al., 1997), on dogs' personality traits (starter's gun, Svartberg and Forkman, 2002), on a particular trait of temperament such as nervousness (weight dropped into a bucket, Murphree and Dykman, 1965), and on selection of breeding pet dogs (large bore blank pistol, Ruefenacht et al., 2002).

2. Terminology of the temperament concept in dog research

The concept of temperament is introduced to explain individual or breed behavioural differences observed when testing animals, including dogs (Goodloe, 1996). When temperament is defined (e.g. Lawrence et al., 1991; Schneider et al., 1991), it implies that these differences are: (1) present at an early age; (2) elicited in a set of situations; (3) (relatively) stable over time. Parameters that can influence behaviour and help us to

understand the complexity underlying temperament are, for example, genetics, the prenatal environment, the mother–young relationship, early experience and the learning capacities that can intervene at any time in an animal's life (Stur, 1987).

Temperament is made up of several behavioural traits, also called temperament categories (Murphy, 1998), behavioural profiles (Hart and Miller, 1985; Goodloe and Borchelt, 1998), personality traits (Svartberg and Forkman, 2002) or factors, as they are usually identified from factorial analyses. Each factor is comprised of behavioural variables that are related to each other (correlations). The factor is labelled by examining clusters of behaviours because each factor is thought to represent one ''functional unity'' of temperament (Royce, 1955). Three levels of differences are observed that make the comparison between studies laborious: (1) the origin of dogs tested (e.g. laboratory colony of dogs, Cattell and Korth, 1973; guide dogs for the blind, Goddard and Beilharz, 1984; pet dogs, Sheppard and Mills, 2002); (2) the origin of the data analysed (e.g. owners' ratings, Goodloe and Borchelt, 1998; behavioural test, Svartberg and Forkman, 2002); (3) the behavioural content of the factors (cluster of behaviours from factorial analysis or standard general description of a particular trait, Murphy, 1998). It is outside of the scope of this paper to discuss the relevance of how each trait is labelled but we want to draw attention to the disparity that exists within the field. This may be understood because each study has collected its own set of behavioural data and the behavioural clustering depends on these data. For example, the factor 'aggressiveness' in a personality test of German Shepherds and Belgian Tervueren (Svartberg, 2002) is not comparable to the factor 'aggressiveness' from six breeds of laboratory dogs in Royce (1955), nor to the personality trait 'aggression' when assessing sheltered dogs (Ledger and Baxter, 1997) or to the four factors related to aggression of companion dogs identified by Goodloe and Borchelt (1998).

There does not seem to be differences in definition between 'temperament' and 'personality' (Ledger and Baxter, 1997) or between 'temperament' and 'character' (Ruefenacht et al., 2002), as these authors use both terms interchangeably. Sometimes, the study of temperament is restricted to the assessment of a particular aptitude and/or performance of working dogs in specific tasks related to their employment, e.g. guide dogs (Serpell and Hsu, 2001) or military working dogs (Burghardt, 2003).

At this point, the understanding of terminology problem worsens when one of the dog behaviour traits aimed at describing the dog's temperament is itself labelled ''temperament''. In Wilsson and Sundgren (1997b), temperament was defined as ''the degree of liveliness, dogs with high temperament are more responsive to all types of stimuli'' and was evaluated during tests. In Slabbert and Odendaal (1999), temperament was scaled from 0 to 10 in a startle test. Ruefenacht et al. (2002) analysed a behaviour test where temperament was one of eight traits to be evaluated and was defined as ''physical flexibility and intensity of reaction to different environmental stimuli'' (Seiferle and Leonhardt, 1984; Ruefenacht et al., 2002). In this particular case, the evaluation of temperament was first verbal and subjective then transformed into numerical scores.

To conclude this section in a context of standardisation, we suggest that the word temperament should only be employed when characterising the dog behaviour as a whole. Particular traits, which make up temperament should be labelled in an objective way that can meet both scientific concerns and applied applications.

3. Test quality requirements and their implementation in literature

3.1. Definitions

There are four quality requirements for any test (Martin and Bateson, 1986; Pichot, 1991). Firstly, the administration and the notation of the test must be standardised, with the only variable being the animal tested. Secondly, a test must be reliable. If it is applied twice, results need to be significantly correlated. Thirdly, the scoring of the test must be sensitive. The individual behavioural differences have to be translated on a precise and objective behavioural scale. The fourth quality requirement is the validity of the test: does the test accurately measure what the investigator really wishes to measure?

3.2. Implementation of test quality requirements in literature

The implementation of these test quality requirements does not appear to be systematic in the literature. However, with regard to standardisation, this condition is generally fulfilled with an appropriate experimental protocol and if it is not the case, the test is conducted in ways that reduce the errors of measurement (i.e. by providing test administration's guidelines, Goodloe, 1996).

The test–retest consistency must be verified to make sure that a test is reliable (Pichot, 1991). The same test is applied twice to the same subject after a certain interval of time and correlation coefficients are calculated on the similarities or differences between the behaviour's frequency, duration, intensity, etc. On the one hand, if the measurement of behaviour is constant in time, the correlations will be significant. Unfortunately, this check does not take place systematically in the literature, and when implemented the interval of time is empirical and ranges from 30 min (Hoffmann et al., 1995) to 6 months (Netto and Planta, 1997). On the other hand, it is possible that the behaviours observed on these two occasions differ because learning has occurred, as was the case for ''sniffer'' dogs whose behavioural capacities improved in the second part of the experimental period (Gazit and Terkel, 2003). Therefore, the data accrued from the test–retest does not have any bearing on the tests reliability.

If several observers rate the behavioural responses of the same animals in the same place or different animals in different places, it is also necessary to check the inter-observer reliability because the raters may be an additional factor of variance in the responses collected (Ruefenacht et al., 2002). As an example, Goodloe and Borchelt (1998) assessed an inter-owner reliability in pairing observations with two members of a family evaluating the same dog.

The question of sensitivity is never approached as such in the literature. However, preliminary observations as well as pre-tests often precede the experiment itself (for example, Beaudet, 1990). This procedure aims to determine the best method of behavioural notation.

The requirement for test validity is evaluated with regard to internal (content and construct) and external (predictive) validity (Bonboir, 1972). Content validity describes how well the different variables of the behavioural test are representative of the desired behavioural trait to be assessed (Goodloe, 1996). To reach this requirement, the test is evaluated by behavioural scientists and people working in the dog's field and/or it is based

on what is known about the behavioural trait. For example, assessing the general activity of an animal during a test will reflect its tendency to be emotive (James, 1951; Fox, 1972). However, although King et al. (2003) chose stimuli likely to evoke fear in some species, they were not able to validate behavioural variables as measures of fear in dogs. They concluded that they may have highlighted appropriate measures of some aspects of fear (such as responses to novelty and startling stimuli) but that their validity needed to be further investigated.

Construct validity refers to the extent to which a scale measures what it claims to measure. For example, in Hewson et al. (1998), three behavioural changes in canine compulsive disorder (frequency, duration and number of eliciting contexts) were used to validate two rating scales of this behavioural problem.

If a test measures a particular characteristic, it should be correlated with any criterion that measures the same or a connected characteristic. In the case of predictive validity, the criterion is an indicator of a predicted performance (Laveault and Gregoire, 2002). For example, to set up an identification tool for puppies with good working capacities, results from behavioural tests during the animals' development are correlated with the degree of success in a later performance, such as detecting substances or protecting military quarters (Hilliard and Burghardt, 2001). The predictive ability of puppy testing is observed in varying proportions depending on the future desirable characteristics. It ranges from no prediction for military search dogs' aptitude (Rooney et al., 2003), 55% of prediction of environmental fear responses in pet dogs (Diederich, 1999) to 91.7% of accurate prediction of police dogs aptitude (Slabbert and Odendaal, 1999).

To conclude, the creation of a new behavioural test requires accurate preparation to get a useful measurement tool. The appreciation of the differences and the interpretation of the results of a test will depend on its standardisation, its reliability, its sensitivity and its validity. The implementation of each of these requirements is not clearly visible in the current literature even if reliability and validity are the main objectives of some recent studies (see Hewson et al., 1998; Serpell and Hsu, 2001). It seems clear that time and energy would be saved if the investigators would have at their disposal behavioural tests whose quality requirements have already been checked.

4. Characteristics of the dogs tested

Four characteristics have been analysed in the behavioural testing of dogs: the source, breed, age and sex of the tested animals.

4.1. Source of dogs

Besides using laboratory animals in behavioural research (e.g. Beerda et al., 1999a), the tendency for some years has been to recruit pet dogs belonging to volunteer owners by means of private or university veterinary hospitals, dog training schools or animal shelters. The owners or guardians are then invited to complete a questionnaire about their observations of their animal's behaviour (Serpell and Hsu, 2001; Stephen and Ledger, 2003) or to present the dog for behavioural testing in a controlled test setting (Seksel et al.,

1999; Szetei et al., 2003) or to accept that their dog be tested in its familiar surrounding (Mills and Ledger, 2001; Cronin et al., 2003). Subjects are also located through canine associations, such as the Swedish Dog Training Centre (Wilsson and Sundgren, 1997a), the South African Service Dog Breeding Centre (Slabbert and Odendaal, 1999), the Swiss German Shepherd Dog Breeding Club (Ruefenacht et al., 2002) or the Swedish Working Dog Association (Svartberg, 2002). This procedure then allows researchers to work with large dog samples from the same or from different breeds.

4.2. Dog breeds

Dog breeds that have been subjected to behavioural testing are presented in Table 1, according to FCI's (Fédération Cynologique Internationale) grouping. These are the tested

Table 1
Dog breeds[a] tested behaviourally were separated into groups according to Fédération Cynologique Internationale (FCI): Group 1, Sheepdogs and Cattledogs (referred to as "Sheepdogs", only these 10 breeds have been tested out of the 40 possible breeds in this group); Group 2, Pinschers, Schnauzers, Mastiffs, Swiss mountain and Cattledogs ("Pinshers", 8/45 breeds); Group 3, Terriers (8/30 breeds); Group 4, Dachshunds (1/3 breeds); Group 5, Primitive type dogs and Spitzes ("Primitive dogs", 4/44 breeds); Group 6, Scent hounds and related breeds ("Scent hounds", 2/76 breeds); Group 7, Pointing dogs ("Pointers", 9/37 breeds); Group 8, Retrievers, Water dogs and Flushing dogs ("Retrievers", 4/19 breeds); Group 9, Companion and Toy dogs ("Companion dogs", 5/27 breeds); Group 10, Sighthounds (4/12 breeds)

Sheepdogs	Pinschers	Terriers	Dachshunds	Primitive dogs
German Shepherd	*Boxer*	*Fox Terrier*	Standard Dachshund	*Basenji*
Australian Shepherd	Miniature Schnauzer	American Staffordshire Terrier		Siberian Husky
Border Collie	Dogo Argentino	Staffordshire Bull Terrier		Shiba Inu
Shetland Shepherd	Standard Pinsher	Bedlington Terrier		Samoyed
Welh Corgi	Rottweiler	Bull Terrier		
Belgian Tervueren	Fila Brasileiro	Australian Terrier		
Belgian Malinois	Doberman Pinscher	Welsh Terrier		
Belgian Cattledog	Shar Pei	West Highland White Terrier		
Old English Sheepdog				
Rough Collie				

Scent hounds	Pointers	Retrievers	Companion dogs	Sighthounds
Beagle	*English Pointer*	*Golden Retriever*	*Miniature Poodle*	*Greyhound*
Dalmatian	Weimaraner	*Labrador Retriever*	Standard Poodle	Borzoi
	Red Irish Setter	English Cocker Spaniel	Chihuahua	Whippet
	English Setter	English Springer Spaniel	Boston Terrier	Irish Wolfhound
	German Longhaired Pointer		Lhasa Apso	
	German Shorthaired Pointer			
	German Wirehaired Pointer			
	Wirehaired Pointing Griffon			
	Large Munsterlander			

The most numerous are shown in italics.
[a] There were also crossbreeds and mongrels used, which are not listed here.

breeds among all the breeds available within each group. It makes no sense to discuss the number of dogs tested per breed because it depends on their availability and on the objectives of the authors. For example, Wilsson and Sundgren (1998) analysed the predictive value of behavioural tests on 1235 German Shepherd puppies as potential service dogs, whereas, Mills and Ledger (2001) recruited six young adult German Shepherds – and other breed representatives – to study the effect of a chemotherapeutic agent on their learning abilities. Thus it does not make sense to discuss the number of breeds tested per FCI's group because in some cases, only one dog per breed has been tested and can not be considered as representative of its breed (for example, one Samoyed in Szetei et al., 2003).

Some authors do not attach particular importance to the breed being tested because they are trying to bring out general reactions in the canine species or else the dogs tested are intended to be representative of pet dogs. Their work deals with subjects belonging to very different breeds, crossbreeds and mongrels. For example, Blackshaw et al. (1990) studied the aversive nature of miscellaneous auditory and sparkling stimuli with a Chihuahua, Dachshund, German Shepherd, Borzoi and numerous crossbreeds. On the other hand, some researchers attach major importance to the choice of breed because they want to describe or exploit their specific behavioural features, or else they try to select individuals with particular behavioural capacities (see examples in papers on working and service dogs quoted elsewhere in this review).

4.3. Age of the subjects

As regards the age of the animals, Table 2 reveals two predominant testing periods, from 1 to 11 weeks of age and from 1 to 10 years, with a central period where the dogs are infrequently tested (3–11 months). The first period (that is, puppy testing) is taken advantage of to study behavioural development and environmental factors that influence behaviour. In the 1-year or older period, some dogs are young adults and are tested for behaviour characteristics of the general canine species while older dogs provide evidence in the study of the ageing phenomenon. Between 3 and 11 months, dogs are maturing to adulthood and are probably least tested because they are neither physically nor behaviourally fully developed.

In puppy testing, some authors keen to follow their behavioural development retest the same animals at various periods of their lives (Fox and Spencer, 1969; Wright, 1980) or they attempt to predict what characteristics the dog will have in adulthood, and thus what task it will be most suited to, for example, to become a guide dog for the blind (Pfaffenberger and Scott, 1976; Goddard and Beilharz, 1984), a police dog (Verga, 1983; Fält, 1984), a military search dog (Rooney et al., 2003) or a pet (Campbell, 1975). The "Tests of Campbell" (Campbell, 1975) include five tests that assess the degree of socialisation of a puppy to humans and the human–puppy dominance relationship. The tests allow a potential owner to choose the animal that will correspond to his expectations. However, although these tests are well-known to the general public and quite often implemented, their predictive value has been demonstrated neither by their author nor by Young (1988). In contrast, Beaudet et al. (1994) demonstrated a predictive value of these tests between 7 and 16 weeks of age, by introducing into the analysis an independent supplementary measure (total number of squares crossed in the test arena).

Table 2
Age when dogs were assessed using behavioural tests, according to the literature over the last 50 years

Authors	Weeks											Months									Years									
	1	2	3	4	5	6	7	8	9	10	11	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10
James (1952)	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wyrwicka (1959)	–	–	–	–	x	x	x	x	x	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Scott and Fuller (1965)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	–	–	–	–	–	–	–	–	–
Fox and Stelzner (1967)	x	x	x	–	x	–	–	x	–	–	–	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Fox and Spencer (1969)	–	–	–	–	x	–	–	x	–	–	–	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Plutchik (1971)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	x	x	x	–	–	–
Fox (1972)	x	x	x	x	x	x	x	x	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Campbell (1975)	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Pfaffenberger and Scott (1976)	–	–	–	–	–	–	–	x	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Gurski et al. (1980)	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wright (1980)	–	–	–	–	x	–	–	x	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Verga (1983)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–	–
Toman (1984)	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wilsson (1984)	–	–	x	x	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Goddard and Beilharz (1985)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–
Goddard and Beilharz (1986)	–	–	–	x	x	x	x	x	x	x		x	x	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–
Hepper (1986)	–	–	–	x	x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Lore and Eisenberg (1986)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	x	–	–	–	–
van der Borg et al. (1991)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	x	x	x	x	x	x	x	x	–
Head and Milgram (1992)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	x	x	x	x	x	x
Beaudet et al. (1994)	–	–	–	–	–	–	x	–	–	–	–		x	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Hoffmann et al. (1995)	–	–	–	–	–	–	x		–	–	–	–		x	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–
Head et al. (1997)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	x	x	x	x	x	x	x	x	x
Weiss and Greenberg (1997)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	–	–	–	–	–	–	–	–
Wilsson and Sundgren (1997a)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	–	–	–	–	–	–	–	–
Schmutz and Schmutz (1998)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–
Wilsson and Sundgren (1998)	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–	x	x	–	–	–	–	–	–	–	–
Diederich (1999)	x	x	x	x	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–
Seksel et al. (1999)	–	–	–	–	–	–	–	–	–	–	–	x	x	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–
Slabbert and Odendaal (1999)	–	–	–	–	–	–	–	x	–	–	–	x	x	–	x	–	–	x	–	–	–	–	–	–	–	–	–	–	–	–
Mills and Ledger (2001)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	x	x	x	x	x	–	–	–
Svartberg (2002)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	–	–	–	–	–	–	–	–
Ruefenacht et al. (2002)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	x	x	–	–	–	–	–	–	–	–
Rooney et al. (2003)	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	–	–	–

(x): age of tests, (–): no tests.

Behavioural tests in adult dogs are particularly concerned with the selection of shelter dogs to be reintroduced as pets (van der Borg et al., 1991; Ledger and Baxter, 1997), or as service animals (Weiss and Greenberg, 1997), or the selection of pets or working dogs for breeding purposes (Netto and Planta, 1997; Wilsson and Sundgren, 1997a). Adult dogs are also tested to study the sensory function of the canine species (visual acuity: Neitz et al., 1989; olfactory function: Ezeh et al., 1992; for a review, see Myers (1991)), the effects of chemotherapeutic agents on behaviour (Head and Milgram, 1992; Mills and Ledger, 2001), and the effects of ageing in dogs as animal models of human diseases (Head et al., 1995, 1997).

4.4. Sex and sexual status of the dogs

Differences between the subjects tested may be observed according to the sex or the sexual status of the animals, i.e. Head and Milgram (1992) observed an increase in locomotor activity in female beagles and a decrease in frequency of urination in male beagles and mixed-breed dogs after a single oral dose of L-deprenyl. Podberscek and Serpell (1996) showed that males, females, and neutered English Cocker Spaniels were significantly different in types of aggression observed in 13 every day situations. Wilsson and Sundgren (1997a) found that, in two breeds of working dogs, German Shepherd and Labrador males are different from their female counterparts in courage, prey drive and defence drive. Lore and Eisenberg (1986) showed that male pet dogs reluctantly approach and make less body contact with an unfamiliar man than female dogs.

To conclude, it is useful for the purpose of standardisation that data from large samples of dogs from the same breed, but especially from different breeds, are available, thanks to non-laboratory sources. Access to such variety could help overcome the effect of breed in the search for behavioural responses expressed in canine species in general. To what extent the characteristics of dog age (particularly in predictive puppy testing for selection of adult characteristics), dog sex and sexual status have on standardising tests needs more investigation.

5. Characteristics of social stimuli in behavioural tests

Social stimuli are employed to measure dogs' abilities to accept the close proximity of and to relate to conspecifics (intraspecific socialisation) or to any other living species (interspecific socialisation) (Beaver, 1994). In particular cases, social stimuli from other species are used to elicit predatory behaviour.

In assessing intraspecific socialisation, stimulus dogs may be young (Fox, 1971) or adults (Goddard and Beilharz, 1985), presented in a cage (Plutchik, 1971) or moving freely (Fox and Stelzner, 1967). A kennel mate may also be removed and the effects of its departure on behaviour of the remaining dog are observed (Tuber et al., 1996). Stimuli representing dogs are also employed but less often: a picture of a dog (Fox, 1971), mirror reflection (Fox and Spencer, 1969; Wright, 1983), or a stuffed dog (Fox and Spencer, 1969; Plutchik, 1971). These models serve to elicit a social interaction but they carry less social information than a living dog.

To assess interspecific socialisation, the stimuli are mainly human beings, mostly unknown to the dog being tested. Other stimuli representing human beings are less frequently employed: a rag doll (Goddard and Beilharz, 1986), a mechanical doll representing a 2- or 3-year-old child (van der Borg et al., 1991), some life-size paper figures of people (Wilsson and Sundgren, 1997a), a human-like dummy (Netto and Planta, 1997; Svartberg and Forkman, 2002) and a human picture with the eyes staring straight ahead (Wilsson and Sundgren, 1997a). These models are employed when human safety needs to be ensured where the test aims to elicit a fear and/or aggressive response in the dog tested.

The human beings are often referred to as manipulators, either keeping still, standing, sitting or lying in front of the animal (van der Borg et al., 1991; Filiatre et al., 1991), or

moving, walking towards the dog in a threatening way or going away abruptly then returning (Murphree and Dykman, 1965; Svartberg and Forkman, 2002) or leaving the dog alone after a period of time together (Newton and Lucas, 1982). In addition, the people may handle the dog in different ways: they may lift the dog or place/roll it on its back (Campbell, 1975), or punish the dog by smacking its head with a roll of paper (Stanley and Elliot, 1962), provoke pain by pinching an ear (Pfaffenberger and Scott, 1976, to test body sensitivity) or the skin between the toes (Weiss and Greenberg, 1997, to detect excessive submission or aggression) or the skin of the groin (Netto and Planta, 1997, to score aggressive behaviour). Also the manipulator may command the dog with orders such as ''come'', ''sit'', ''stay'', ''drop'' and ''wait'' (Seksel et al., 1999) to evaluate its obedience.

Other species are also used as interspecific social stimuli. Puppies may be reared with cats and then presented to dog littermates to study their socialisation to their own species (Fox, 1971). Puppies' social behaviour to a cat in a wire cage was rated in Seksel et al. (1999); rats presented in cages have been used to assess dogs' exploratory behaviour (Fox and Spencer, 1969), and dogs may also be tested with quails (small prey, Murphree and Dykman, 1965) or with sheep (large prey, Arons and Shoemaker, 1992) to test hunting behaviour.

To conclude, it is clear that no standard test of reference is available to test socialisation in dogs. There is a need to have a standard socialisation test, which would be to test a dog with (a) a friendly dog and (b) a friendly human as the two social stimuli. With such test, the normal responses observed in the canine species could be used to study the influence of other factors such as breed, age and reproductive status of the stimulus dog and of the tested dog, and the age, sex, appearance and body language of the human stimulus.

6. Characteristics of environmental stimuli in behavioural tests

A list of environmental stimuli for testing dogs is presented in Table 3. It will never be complete because it grows with the inventiveness of the researchers. In this review, the environmental stimuli have been classified into two classes: objects and places.

Among the 'objects' stimuli, six subgroups have been identified: mobile, immobile, auditory, visual, odorous, or mixed (mobile and auditory or mobile and visual). The list is organised according to the most potent stimulus effect. For example, a bone is a mixed stimulus (immobile and odorous) but above all, an odorous stimulus for the dog. The subject tested may sometimes establish physical contact with some object stimuli. The 'places' stimuli are outside or inside locations to which a dog is led in order to observe its behaviour in that place. It is not possible to describe all the stimuli that a dog can perceive outside, during a walk or a drive, for example. On the other hand, when tests take place inside, a known or unknown arena or open-field is frequently quoted in the literature. The size depends on the choice of the authors. For example, in puppy testing, Stanley and Elliot (1962) tested 3–4-week-old Basenjis in a 13.3 m^2 arena, Wright (1983) tested 5.5–8.5-week-old German Shepherds and Beagles in a 6.25 m^2 arena, and Beaudet (1990) tested 7–16-week-old puppies from six breeds including Toy Poodles and German Shepherds in a 3.35 m^2 arena. If this test is presented to elicit emotional reactivity or exploratory behaviour in puppies, the size of the arena should be adapted to the size of the animal

Table 3
Review of the environmental stimuli for testing dogs

Stimuli	Type	Description
Objects	Mobile	Ball (variable size and colour) (i.e. Plutchik, 1971)
		Umbrella (King et al., 2003)
		Mechanical or rubber snake (Mahut, 1958)
		Rag or cloth shaken in front of the dog (Fox, 1972)
		Crumpled sheet of newspaper (Goddard and Beilharz, 1986)
		Leather lead trailed on the ground (Pfaffenberger and Scott, 1976)
		Remote control car (King et al., 2003)
	Immobile	Carnival mask (Mahut, 1958)
		Elevated plus maze (King et al., 2003)
	Auditory (strong and prolonged noise)	Clock alarm (Toman, 1984)
		Siren (Plutchik, 1971)
		Alarm (Blackshaw et al., 1990)
		Horn (Pfaffenberger and Scott, 1976)
		Doorbell (Pfaffenberger and Scott, 1976)
		Whistle (Goddard and Beilharz, 1986)
		Vacuum cleaner (Vincent and Michell, 1992)
	Auditory (strong and short noise)	Starting pistol (Svartberg and Forkman, 2002)
		Bursting balloon (Plutchik, 1971)
		Metal buckets that fall (Wilsson and Sundgren, 1997a)
		Plastic bottle that falls (Toman, 1984)
		Large links chain on a metal sheet (Svartberg and Forkman, 2002)
		Fog horn (Beerda et al., 1998)
	Visual	Flashing bright light (Blackshaw et al., 1990)
		Artificial day–night transition (Plutchik, 1971)
		Light/dark box (King et al., 2003)
		Illuminated translucent panels (Neitz et al., 1989)
	Odorous	C4 explosive (Gazit and Terkel, 2003)
		Food: bone, meat, fish, ... (Scott and Fuller, 1965; Fox, 1972; Wright, 1980; Toman, 1984)
		Coffee (Gazit and Terkel, 2003)
		Human odour (Schoon, 1996)
		Footprints (Mackenzie and Schultz, 1987)
		Benzaldehyde and eugenol samples (Ezeh et al., 1992)
		N-amylacetate and *d*-limonen samples (Hirano et al., 2000)
		Cloth bedding from the bitchs pen (Fox and Stelzner, 1967)
	Mixed	Musical top (Mahut, 1958)
		Toy-car with flashing light (Goddard and Beilharz, 1984)
Place	Outside	Arena with novel stimuli (Pfaffenberger and Scott, 1976)
		To go for a walk (Goddard and Beilharz, 1986)
		To go for a drive (van der Borg et al., 1991)
		Busy shopping area (Murphy, 1998)
		Open area (Ruefenacht et al., 2002)
		Wooded area (Svartberg and Forkman, 2002)
	Inside	Test room (Netto and Planta, 1997)
		Test arena (Fox and Spencer, 1969)
		Open-field with complex and noncomplex sides (Wright, 1980)
		Kennel (Beerda et al., 1998)
		Owner's home (Mahut, 1958)

tested, in order for the results to be relative to others. In addition, all the 'objects' stimuli can be presented to a dog in conjunction with a 'place' stimulus, i.e. inside as well as outside.

To conclude, no standard environmental stimuli and no standard place exist to test a dog, except in particular aptitude testing (i.e. explosive detection, Gazit and Terkel, 2003). The perfecting of a test is an investigation per se. There is a need for standard environmental and place stimulus tests that researchers could refer to and in most cases use. For example, in the case of testing fear in dogs, there is not yet any consensus regarding the way fear can be induced, recognised, collected and analysed (King et al., 2003).

7. Characteristics of behavioural variables

7.1. What behavioural variables should be collected?

To develop this point, the activity in an open-field has been chosen as an example. According to Head et al. (1997), open-field activity is age-sensitive in dogs, but breed and test conditions are also essential factors. In an attempt to find a consensus of the behavioural variables to collect, we put together four studies testing exploratory behaviour in puppies (Table 4). The subjects tested were of similar ages, and were either Beagles, German Shepherds, mongrels or a combination of these, the size of the arena was 6 or 13.3 m^2, the squares on the floor were, respectively, 30.5 and 60 cm, and the number of objects to explore were similar (3–6). The behavioural data would be expected to be the same, but this is not the case. Regarding the exploration of the objects, for example, the behavioural variables were: time spent interacting with objects (Fox and Stelzner, 1967), frequency and duration of interaction with objects (Fox and Spencer, 1969), number of explorations of the objects (Wright, 1983) and number of objects visited (Wilsson and Sundgren, 1998). A possible explanation for these differences may lie in the methods to collect these variables: the number of observers, the direct or video recorded notation, the length and frequency of the trials are factors that affect the type and the number of behavioural variables to be collected.

7.2. How to note behavioural variables?

The dog's behavioural responses can be measured by either objectively noting frequencies and durations, or scoring them with a number grading the intensity of responses (Svartberg and Forkman, 2002), or subjectively noting them according to the observer's experience (Gosling, 2001; Ruefenacht et al., 2002). However, subjectively scoring is difficult to implement consistently by people other than the author. For example, in Pfaffenberger and Scott (1976), the puppy's fear is assessed during training as the animal ''seems at ease in pen'', ''moves about in puppy pen freely'', ''looks pen and tester over calmly'', ''is friendly with testers'', ''responds to tester's encouragement'' and ''is willing to do what tester wants it to do and shows pleasure in doing it''. Every statement is graded on a 0–6-point scale. Toman (1984) notes the reactions to an unknown noise: surprise, panic or attempts to escape. The problem of subjectively scoring can be eliminated if the

Table 4
Subjects, age of testing, behavioural variables and collection methods in exploratory behaviour testing in puppies

Subjects	Age (weeks)	Behavioural variables	Method to collect variables
8 Beagles, 9 mongrels (Fox and Stelzner, 1967)	5, 8, 12	Time spent interacting with three objects; number of 30.5-cm squares crossed	13 m^2 arena; two observers (time); one observer (squares); direct notation; 10-min trial per week
35 Beagles, 22 mongrels (Fox and Spencer, 1969)	5, 8, 12, 16	Time to enter the arena, out of the start box; number of vocalisations; number of 30.5-cm squares crossed; frequency of interaction with six objects; duration of interaction with six objects	6 m^2 arena; four observers; direct notation; 15-min trial per week
18 German Shepherds, seven Beagles (Wright, 1983)	5.5, 8.5	Latence (s) to enter the complex side of the arena; time (s) spent in the complex side; number of explorations of four objects; number of crossings from one arena's side to the other; number of 30-cm squares crossed in both arena's sides	6.25 m^2 arena; four observers; videotapes; 10-min trial per week
630 German Shepherds (Wilsson and Sundgren, 1998)	8	Time (s) spent inside a central circle with human stimulus; number of 60-cm squares crossed; number of four objects visited	13 m^2 arena; not indicated; one trial of 5 min

behavioural variables are exactly described and are scored according to their intensity, every level of intensity also being accurately described, reflecting a continuum in the response (Svartberg and Forkman, 2002). The objective (quantitative) and subjective (qualitative) approaches can be complementary, as demonstrated in pigs (Wemelsfelder et al., 2003).

Regarding scoring behaviours in either decreasing or increasing value, guidelines do not seem to exist. In some cases, the preferred behaviour gets the higher score. For example, the canine behavioural responses (social behaviour, reaction to novel stimulus, reaction to handling and obedience to commands) were rated in Seksel et al. (1999) on a 100-points scale where the higher score means a ''good'' behaviour and the lower score, an ''undesirable'' one. In other cases, the preferred behaviour gets the lower score. For example, aggressiveness was estimated in Netto and Planta (1997) on a five-point scale from (1) no aggression to (5) to bite and/or to attack with the intention of biting. In another example, the overall impression of the trainer about a dog's temperament for guiding work was scored on a 0–3 scale where 0 meant excellent (Murphy, 1995).

7.3. When to observe behavioural variables?

In the majority of the cases, the behavioural responses of the dog are observed after the presentation of the stimulus. The duration of the test can vary by as much as 200%. For example, the duration of the exploratory behaviour test in an open-field ranged from 5 to 15 min (see Table 4).

In the case of fear testing, it is suggested to record the initial reaction of the dog in the first 10 s after the stimulus presentation in Mahut (1958) or the immediate reaction of the dogs in Slabbert and Odendaal (1999) because it has been observed that the dog no longer responds or loses interest in the stimulus after this. If the dog may explore the startling stimulus, latency to approach and time spent near the stimulus are also counted (King et al., 2003). In the case of identification of aggressive dogs, the occurrence of the expected behaviour (aggression) is sufficient (Netto and Planta, 1997), regardless of the time it was observed during the test.

To conclude, the behavioural variables collected in testing are dependent upon the test conditions and the protocol developed by the researchers. Again, it is suggested to standardise these conditions to prepare foundations for further studies.

8. Characteristics of concomitant physiological and physical measures

Some physiological changes may be concomitant with behavioural reactions. That was the case of nervous Pointers that were distinguishable from the normal line by a lower cardiac resting rate (Newton et al., 1978) and while being petted (Newton and Lucas, 1982), by a bilateral hearing deficit in 74% of them (Klein et al., 1988) and by a lower IGF-I (Insulin-like Growth Factor I) serum rate (Uhde et al., 1992). In the context of predatory behaviour, Arons and Shoemaker (1992) established a correlation between its stages (approach, catch, kill and eat) and the distribution of catecholamines and beta-endorphin in various cerebral zones of three breeds of dogs (Border Collie, Shar Planinetz and Siberian Husky).

When confronted with disturbing stimuli, in addition to behavioural responses, physiological responses are observed due to the activation of the autonomic nervous system and of the hypothalamic-pituitary-adrenal cortex axis (Broom and Johnson, 1993). Behavioural responses have been identified as relevant indicators of acute stress after the presentation of a noise of moderate intensity (Beerda et al., 1997), whereas, it is not clear if salivary cortisol concentration or heart rate is the more reliable parameter in response to an auditory stimulus (Beerda et al., 1998) or in response to the rapid opening of a brightly coloured umbrella (King et al., 2003). Furthermore, behavioural observations do not always move in parallel with the physiological measures (responses to a vacuum cleaner in Vincent and Michell, 1992; responses to wearing muzzles in Cronin et al., 2003). It is suggested that the differences in protocols used by the researchers may be partially responsible for the variety of behavioural responses reported (Beerda et al., 1997; King et al., 2003). Concerning the assessment of chronic stress effects resulting from social and spatial restriction, behaviours and salivary (and/or urinary) cortisol measures may be useful parameters (Beerda et al., 1999a,b).

Furthermore, dogs with certain physical attributes may be distinguishable by their behavioural responses. Mugford (1984) described the low threshold aggression or rage syndrome in unicolour English Cocker Spaniel lines that is characterised by sudden and unpredictable episodes of aggression. Since then this type of aggression has been associated with dominance aggression (Podberscek and Serpell, 1996) and the aggression, whatever the type in English Cocker Spaniels, is correlated with the unicolour coat and with environmental factors such as the age of the owner, what the dog is to be used for (guide dog, working dog, pet, etc.), and the occurrence of an illness at an early age (Podberscek and Serpell, 1997). In German Shepherd guard dogs, Mackenzie et al. (1985) found a negative correlation between some working abilities and hip dysplasia. The nervous Pointer line is distinguishable from the normal line by a lighter weight and a lower size/weight ratio (Uhde et al., 1992).

To conclude, the existence of physiological responses associated to the behavioural responses speaks in favour of standardisation of the testing procedures to identify the more reliable physiological parameters.

9. Directions for the future: is some standardisation possible?

Using behaviour as a tool of measurement requires a long building process where questions have to be answered in order to obtain results that are relevant and useful (Laveault and Gregoire, 2002). Some of these questions are: What is this test used for? What will this test have to measure? What kind of data should be collected? Has the data/test been pre-tested? Is this test validated and reliable? If this test has to be widely distributed, do guidelines exist that permit its correct use? While guides exist in other scientific fields (i.e. international standards ISO 5667-1 in water quality analysis: sampling (Anonymous, 1996), analytical quality control ISO 13530 (Anonymous, 1998) and measurement uncertainty estimation ISO 21748 (Anonymous, 2004), canine behavioural testing suffers from an apparent lack of standardisation in the way it is implemented, despite its numerous fields of use. We suggest that if some testing procedures could be

Table 5
Illustration of the diversity as regards the aims, the age of the tested animals, the stimuli, the collected data and their interpretation with a selection of six research studies using a ball

	Age	Stimuli	Data	Interpretation
a	7–120 months	Balloon inflated to approx. 30 cm diameter	No response; curiosity; brief investigation; contact; approach/avoid; wariness; avoidance	Fear
b	16–20 weeks	Ball (5 cm diameter) rolled across the arena	Orienting response; contacting response (weak or strong); number of contacts	Depression; panic response; lack of stimulation
c	1–7 years	Ball (8 cm diameter), on the floor	Immediate and overall response (approach, withdrawal or indifference); contact time; activity (grid-crossing); defecations and urinations	Timidity; approach; emotionality
d	12–18 months	Beach ball (30 cm diameter) thrown away	Nature of contact: no contact, sniff, bite, carry; avoidance: from no sign to strong withdrawal	Fearfulness; emotionality; nervousness
e	7 weeks	Tennis ball rolled away for 2 m	Not running after; run, not picking up; run, pick up, not bringing back; run, pick up and bring back	Fearfulness
f	8 weeks	Ball (13 cm diameter) in the centre of the pen	Immediate reaction (scored from 1 = withdraw from the ball to 5 = play with the ball)	Willingness to retrieve; cooperativeness; fear

Aims: (a) Emotional behaviour (Mahut, 1958); (b) factors modifying the post-isolation syndrome (Fuller and Clark, 1966); (c) approach and withdrawal behaviours (Plutchik, 1971); (d) factor analysis of fearfulness (Goddard and Beilharz, 1984); (e) puppy tests as predictors of adult dog behaviour (Hoffmann et al., 1995); (f) prediction of adult behaviour (Wilsson and Sundgren, 1998).

approved, then their implementation could be accelerated and the risks of misunderstanding and subjectivity could be reduced as well as minimising the difficulty of comparing results and differing opinions. Our suggestion is particularly relevant in light of a review of stress responses in dogs (Beerda et al., 1997) whose authors suggested that a proportion of the observed behavioural variance may be due to differences in stressor properties and in the protocols used by the different researchers.

To illustrate the diversity observed in the literature as regards the aims, the stimuli, the collected data and the interpretation of behavioural tests, six research studies were selected because each of them used a ball as a stimulus. We wanted to show how this stimulus was used and how the authors interpreted the dog's responses (Table 5). We observed that the aims and interpretations would often be regarded under one heading. In fact, while interpreting their results, the authors commented on (among other things) fear, panic, curiosity, shyness, and emotionality. All these terms characterise exploratory behaviour when faced with an unknown object, the attraction of novelty or simply the emotion fear. Furthermore, no consensus appeared as regards to the size of the ball, which ranged from 5 to 30 cm. At least three methods were used to introduce the stimulus: it was put down in front of the dog, or it was rolled towards the dog or away from it. Finally, although a relatively similar behavioural reaction was measured from one piece of research to another, the data was collected and noted according to the preferences of each individual author. To conclude, if we planned to test dogs with a ball, chosen as a novel environmental stimulus, we must start with validating the test, because the procedure varied according to each investigator.

10. Conclusions

For research on canine behaviour to develop in an optimal way, it is suggested that there be a search for uniformity in the material and methods of testing. For example, there is not yet any consensus regarding the way fear can be induced, recognised, gathered and analysed in dogs (King et al., 2003). This review has highlighted some methodological aspects of dog behavioural testing where standardisation could find a place. As the analysis of behaviour has always had elements of reference (i.e. Martin and Bateson, 1986; Siegel and Castellan, 1988), the Manual of Dog Testing Techniques from Charles et al. (1950) could be updated, putting forward standardised, reliable, sensitive and validated tests. This updated manual would enable researchers and authors to clearly define behavioural reference responses, which could be correlated with data from physiological, cellular and molecular analysis at a later date, as suggested in Wayne and Ostrander (1999). For example, the manual would present a standardised interspecific socialisation test with a human as stimulus. The characteristics of the human (i.e. sex, size, clothing, movements or attitudes) could be described as well as the related responses of the dog.

The advantages of a standardised methodology to test dog behaviour are the ease to conduct tests, to score behaviours, to interpret and discuss results, the opportunity to highlight effects of parameters i.e. breed, sex or age of the subject, the possibility to refer to standard responses and last but not least, to save time. It might be argued that science would lose its productiveness if restricted to standardised procedures, consequently reducing the

knowledge we might gain from dog testing. It does not seem, through this review on behavioural tests' diversity over several decades, that a standard methodology has arisen in the way dogs are tested. On the contrary; we have demonstrated that procedures of testing are unique because they are particular to the investigators. However, behavioural tests could be constructed that would be useful tools for helping science to progress.

Acknowledgments

This review was conducted with the assistance of the staff of the Laboratory of Anatomy and Ethology of Domestic Animals, University of Namur (Belgium). The authors would like to thank Dr. V. Sandilands, Guest Editor, ISAE Special Issue, of Scottish Agricultural College, UK, for helpful advice and discussions throughout the reviewing process of this manuscript, as well as the three anonymous referees for their comments.

References

Anonymous, 1996. Water quality – Sampling – Part 1: Guidance on the Design of Sampling Programmes. ISO 5667-1. International Organisation for Standardisation, Geneva.

Anonymous, 1998. Water Quality—Guide to Analytical Quality Control for Water Analysis. ISO 13530. International Organisation for Standardisation, Geneva.

Anonymous, 2004. Guidance for the Use of Repeatability, Reproducibility and Trueness Estimates in Measurement Uncertainty Estimation. ISO 21748. International Organisation for Standardisation, Geneva.

Arons, C.D., Shoemaker, W.J., 1992. The distribution of catecholamines and beta-endorphin in the brains of three behaviourally distinct breeds of dogs and their F1 hybrids. Brain Res. 594 (1), 31–39.

Beaudet, R., 1990. Mise au point d'un test d'évaluation du tempérament applicable à la sélection de chiens de compagnie. Thèse de doctorat en anatomie et physiologie vétérinaires. Université de Montréal, Canada, p. 128.

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

Beaver, B., 1994. The Veterinarian's Encyclopedia of Animal Behaviour, 1st ed. Iowa Stare University Press, Ames, Iowa, pp. 251–252.

Beerda, B., Schilder, M.B.H., van Hooff, J.A.R.A.M., de Vries, H.W., 1997. Manifestations of chronic and acute stress in dogs. Appl. Anim. Behav. Sci. 52, 307–319.

Beerda, B., Schilder, M.B.H., van Hooff, J.A.R.A.M., de Vries, H.W., Mol, J.A., 1998. Behavioural, saliva cortisol and heart rate responses to different types of stimuli in dogs. Appl. Anim. Behav. Sci. 58, 365–381.

Beerda, B., Schilder, M.B.H., van Hoff, J.A.R.A.M., de Vries, H.W., Mol, J.A., 1999a. Chronic stress in dogs subjected to social and spatial restriction. I. Behavioural responses. Physiol. Behav. 66 (2), 233–242.

Beerda, B., Schilder, M.B.H., Bernardina, W., van Hoof, J.A.R.A.M., de Vries, H.W., Mol, J.A., 1999b. Chronic stress in dogs subjected to social and spatial restriction. II. Hormonal and immunological responses. Physiol. Behav. 66 (2), 243–254.

Blackshaw, J.K., Cook, G.E., Harding, P., Day, C., Bates, W., Rose, J., Bramham, D., 1990. Aversive responses of dogs to ultrasonic, sonic and flashing light units. Appl. Anim. Behav. Sci. 25, 1–8.

Bonboir, A., 1972. La méthode des tests en pédagogie. Coll. Sup. ''L'éducateur'', Presses Universitaires de France, Paris, pp. 38–52.

Broom, D.M., Johnson, K.G., 1993. Systems regulating body and brain. In: Chapman, Hall Animal Behaviour Series, Stress and Animal Welfare, London, pp. 8–34.

Burghardt, W.F., 2003. Behavioural considerations in the management of working dogs. Vet. Clin. Small Anim. 33, 417–446.

Campbell, W.E., 1975. Behaviour Problems in Dogs. American Veterinary Publications, California, pp. 137–144.

Cattell, R.B., Korth, B., 1973. The isolation of temperament dimensions in dogs. Behav. Biol. 9, 15–30.
Charles, M., Fuller, J.L., Easler, C., Marston, M.V., Shade, C., Smith, M., Scott, J.-P., Williams, E., 1950. Manual of Dog Testing Technics. Roscoe B. Jackson Memorial Library, Bar Harbor, Maine, p. 72.
Cronin, G.M., Hemsworth, P.H., Barnett, J.L., Jongman, E.C., Newman, E.A., McCauley, I., 2003. An anti-barking muzzle for dogs and its short-term effects on behaviour and saliva cortisol concentrations. Appl. Anim. Behav. Sci. 3, 215–226.
Diederich, C., 1999. Recherche de la persistance de différences comportementales individuelles chez le chien, de la naissance à l'âge de 15 mois. PhD Thesis in Veterinary Sciences. Presses Universitaires de Namur. p. 332. ISBN 2-87037-285-X.
Ezeh, P.I., Myers, L.J., Hanrahan, L.A., Kemppainen, R.J., Cummins, K.A., 1992. Effects of steroids on the olfactory function of the dog. Physiol. Behav. 51, 1183–1187.
Fält, L., 1984. Inheritance of behaviour in the dog. In: Anderson, R.S. (Ed.), Nutrition and Behaviour in Dogs and Cats. Pergamon Press, New York, pp. 183–187.
Filiatre, J.C., Millot, J.L., Eckerlin, A., 1991. Behavioural variability of olfactory exploration of the pet dog in relation to human adults. Appl. Anim. Behav. Sci. 30, 341–350.
Fox, M.W., 1971. Integrative Development of Brain and Behaviour in the Dog. The University of Chicago Press, Chicago, pp. 40–70, 256–264.
Fox, M.W., 1972. Understanding Your Dog. C.M., G. Inc., New York, p. 59.
Fox, M.W., Spencer, J.W., 1969. Exploratory behaviour in the dog: experiential or age dependent? Dev. Psychol. 2 (2), 68–74.
Fox, M.W., Stelzner, D., 1967. The effects of early experience on the development of inter and intraspecies social relationships in the dog. Anim. Behav. 15, 377–386.
Fuller, J.L., Clark, L.D., 1966. Genetic and treatment factors modifying the postisolation syndrome in dogs. J. Comp. Physiol.-Psychol. 61 (2), 251–257.
Gazit, I., Terkel, J., 2003. Explosives detection by sniffer dogs following strenuous physical activity. Appl. Anim. Behav. Sci. 81, 149–161.
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., 1985. Individual variation in agonistic behaviour in dogs. Anim. Behav. 33, 1338–1342.
Goddard, M.E., Beilharz, R.G., 1986. Early prediction of adult behaviour in potential guide dogs. Appl. Anim. Behav. Sci. 15, 247–260.
Goodloe, L.P., 1996. Issues in description and measurement of temperament in companion dogs. In: Voith, V.L., Borchelt, P.L. (Eds.), Readings in Companion Animal Behaviour. Veterinary Learning Systems, Trenton (NJ), pp. 32–39.
Goodloe, L.P., Borchelt, P.L., 1998. Companion dog temperament traits. J. Appl. Anim. Welf. Sci. 1 (4), 303–338.
Gosling, S.D., 2001. From mice to men: what can we learn about personality from animal research? Psychol. Bull. 127 (1), 45–86.
Gurski, J.C., Davis, K., Scott, J.P., 1980. Interaction of separation discomfort with contact comfort and discomfort in the dog. Dev. Psychobiol. 13 (5), 463–467.
Hart, B.L., Miller, M.F., 1985. Behavioural profiles of dog breeds. J. Am. Vet. Med. Assoc. 186 (11), 1175–1180.
Head, E., Milgram, N.W., 1992. Changes in spontaneous behaviour in the dog following oral administration of L-deprenyl. Pharmacol. Biochem. Behav. 43, 749–757.
Head, E., Mehta, R., Hartley, J., Kameka, M., Cummings, B.J., Cotman, C.W., Ruehl, W.W., Milgram, N.W., 1995. Spatial learning and memory as a function of age in the dog. Behav. Neurosci. 109 (5), 851–858.
Head, E., Callahan, H., Cummings, B.J., Cotman, C.W., Ruehl, W.W., Muggenberg, B.A., Milgram, N.W., 1997. Open-field activity and human interaction as a function of age and breed in dogs. Physiol. Behav. 62 (5), 963–971.
Hepper, P.G., 1986. Sibling recognition in the domestic dog. Anim. Behav. 34 (1), 288–289.
Hewson, C.J., Luescher, U.A., Ball, R.O., 1998. Measuring change in the behavioural severity of canine compulsive disorder: the construct validity of categories of change derived from two rating scales. Appl. Anim. Behav. Sci. 60, 55–68.

Hilliard, S., Burghardt, W.F., 2001. Development and validation of behavioural testing instruments for longitudinal study of military working puppies. In: Proceedings of the International Working Dog Breeding Conference. International Working Dog Breeding Association, San Antonio, TX.

Hirano, Y., Oosawa, T., Tonosaki, K., 2000. Electroencephalographic olfactometry (EEGO) analysis of odour responses in dogs. Res. Vet. Sci. 69, 263–265.

Hoffmann, G., Blackshaw, J.K., Smith, G.A., 1995. Puppy tests: intercorrelations and test–retest reliability. In: Proceedings of the 7th International Conference of Human-Animal Interactions, Geneva, September 6–9.

James, W.T., 1951. Social organisation among dogs of different temperaments, terriers and beagles, reared together. J. Comp. Physiol. Psychol. 44, 71–77.

James, W.T., 1952. Observations on behaviour of new-born puppies: method of measurement and types of behaviour involved. J. Gen. Psychol. 80, 65–73.

King, T., Hemsworth, P.H., Coleman, G.J., 2003. Fear of novel and startling stimuli in domestic dogs. Appl. Anim. Behav. Sci. 82, 45–64.

Klein, E., Steinberg, S.A., Weiss, S.R.B., Matthews, D.M., Uhde, T., 1988. The relationship between genetic deafness and fear-related behaviours in nervous pointer dogs. Physiol. Behav. 43, 307–312.

Laveault, D., Gregoire, J., 2002. Introduction aux théories des tests en psychologie et en sciences de l'éducation, 2nd ed. De Boeck Université, Bruxelles, pp. 10–15, 163–197.

Lawrence, A.B., Terlouw, E.M.C., Illius, A.W., 1991. Individual differences in behavioural responses of pigs exposed to non-social and social challenges. Appl. Anim. Behav. Sci. 30, 73–86.

Ledger, R.A., Baxter, M.R., 1997. The development of a validated test to assess the temperament of dogs in a rescue shelter. In: Proceedings of the First International Conference on Veterinary Behavioural Medicine, Birmingham, UK, pp. 87–92.

Lore, R.K., Eisenberg, F.B., 1986. Avoidance reactions of domestic dogs to unfamiliar male and female in a kennel setting. Appl. Anim. Behav. Sci. 15, 261–266.

Mackenzie, S.A., Oltenacu, E.A.B., Leighton, E., 1985. Heritability estimate for temperament scores in German Shepherd dogs and its genetic correlation with Hip Dysplasia. Behav. Genet. 15 (5), 475–482.

Mackenzie, S.A., Schultz, J.A., 1987. Frequency of back-tracking in the tracking dog. Appl. Anim. Behav. Sci. 17, 353–359.

Mahut, H., 1958. Breed differences in the dog's emotional behaviour. Can. J. Psychol. 12 (1), 35–44.

Martin, P., Bateson, P., 1986. Measuring Behaviour. An Introductory Guide. Cambridge University Press, p. 200, Reprinted 1987.

Mills, D., Ledger, R., 2001. The effects of oral selegiline hydrochloride on learning and training in the dog: a psychobiological interpretation. Prog. Neuro-Psychopharmacol. Biol. Psychiat. 25, 1597–1613.

Mugford, R.A., 1984. Aggressive behaviour in the English Cocker Spaniel. Vet. Annu. 24, 310–314.

Murphree, O.D., Dykman, R.A., 1965. Litter patterns in the offspring of nervous and stable dogs. 1. Behavioural tests. J. Nerv. Ment. 111 (3), 321–332.

Murphy, J.A., 1995. Assessment of the temperament of potential guide dogs. Anthrozoös 7 (4), 224–228.

Murphy, J.A., 1998. Describing categories of temperament in potential guide dogs for the blind. Appl. Anim. Behav. Sci. 58, 163–178.

Myers, L.J., 1991. Use of innate behaviours to evaluate sensory function in the dog. Vet. Clin. North Am., Small Anim. Pract. 21 (2), 389–399.

Neitz, J., Geist, T., Jacobs, G.H., 1989. Colour vision in the dog. Visual Neurosci. 3, 119–125.

Netto, W.J., Planta, D.J.U., 1997. Behavioural testing for aggression in the domestic dog. Appl. Anim. Behav. Sci. 52, 243–263.

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

Newton, J.E.O., Lucas, L.A., 1982. Differential heart-rate responses to person in nervous and normal pointer dogs. Behav. Genet. 12 (4), 379–393.

Pfaffenberger, C.J., Scott, J.P., 1976. Early rearing and testing. In: Pfaffenberger, C., et al. (Eds.), Guide Dogs for the Blind: Their Selection, Development and Training. Elsevier, Amsterdam, pp. 13–37.

Pichot, P., 1991. Les tests mentaux. 13ème édition. Coll. Que sais-je? Presses Universitaires de France, pp. 5–17.

Plutchik, R., 1971. Individual and breed differences in approach and withdrawal in dogs. Behaviour 40, 302–311.

Podberscek, A.L., Serpell, J.A., 1996. The English Cocker Spaniel: preliminary findings on aggressive behaviour. Appl. Anim. Behav. Sci. 47, 75–89.

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.

Pongracz, P., Miklosi, A., Vida, V., Csanyi, V., 2005. The pet dogs ability for learning from a human demonstrator in a detour task is independent from the breed and age. Appl. Anim. Behav. Sci. 90, 309–323.

Rooney, N.J., Gaines, S.A., Bradshaw, J.W.S., 2003. How predictive are puppy tests? Evidence from a puppy walking programme for military search dogs. In: Proceedings of the 37th International Congress of the ISAE, Abano Terme, Italy, June 24–28, p. 41.

Royce, J.R., 1955. A factorial study of emotionality in the dog. Psychol. Monogr. 69 (22), 1–27.

Ruefenacht, S., Gebhardt-Henrich, S., Miyake, T., Gaillard, C., 2002. A behaviour test on German Shepherd dogs: heritability of seven different traits. Appl. Anim. Behav. Sci. 79, 113–132.

Schmutz, S.M., Schmutz, J.K., 1998. Heritability estimates of behaviours associated with hunting in dogs. J. Hered. 89 (3), 233–237.

Schneider, M.L., Moore, C.F., Suomi, S.J., Champoux, M., 1991. Laboratory assessment of temperament and environmental enrichment in Rhesus Monkey infants (*Macaca mulatta*). Am. J. Primatol. 25, 137–155.

Schoon, G.A.A., 1996. Scent identification lineups by dogs (*Cani familiaris*): experimental design and forensic application. Appl. Anim. Behav. Sci. 49, 257–267.

Scott, J.P., Fuller, J.L., 1965. Genetics and the Social Behaviour of the Dog. University of Chicago Press, Chicago, p. 468.

Seiferle, E., Leonhardt, E., 1984. In: Erni, S., Druck, A.G., (Eds.), Wesensgrundlagen und Wesensprüfung des Hundes, 2nd ed. Utznacherstr. 3, CH-8722 Kaltbrunn, Switzerland.

Seksel, K., Mazurski, E.J., Taylor, A., 1999. Puppy socialisation programs: short and long term behavioural effects. Appl. Anim. Behav. Sci. 62, 335–349.

Serpell, J.A., Hsu, Y., 2001. Development and validation of a novel method for evaluation behaviour and temperament in guide dogs. Appl. Anim. Behav. Sci. 72, 347–364.

Sheppard, G., Mills, D.S., 2002. The development of a psychometric scale for the evaluation of the emotional predispositions of pet dogs. Int. J. Comp. Psychol. 15, 201–222.

Siegel, S., Castellan Jr., N.J., 1988. Nonparametric Statistics for the Behavioural Sciences, 2nd ed. McGraw-Hill Book Company, New York, p. 399.

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.

Stanley, W.C., 1970. Feeding behaviour and learning in neonatal dogs. In: Bosma, J.F. (Ed.), Proceedings of the Second Symposium on Oral Sensation and Perception, Springfield, IL, pp. 242–290.

Stanley, W.C., Elliot, O., 1962. Differential human handling as reinforcing events and as treatments influencing later social behaviour in basenji puppies. Psychol. Rep. 10, 775–788.

Stephen, J.M., Ledger, R.A., 2003. Owners are reliable observers of their own dog's behaviour. In: Proceedings of the 37th International Congress of the ISAE, Abano Terme, Italy, June 24–28, p.190.

Stur, I., 1987. Genetic aspects of temperament and behaviour in dogs. In: Edney, T.A.B. (Ed.), Proceedings of the Canine development throughout life. Waltham Symposium No. 8 28 (11), J. Sm. Anim. P 957–964.

Svartberg, K., 2002. Shyness-boldness predicts performance in working dogs. Appl. Anim. Behav. Sci. 79, 157–174.

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

Szetei, V., Miklosi, A., Topal, J., Csanyi, V., 2003. When dogs seem to lose their nose: an investigation on the use of visual and olfactory cues in communicative context between dog and owner. Appl. Anim. Behav. Sci. 83, 141–152.

Toman, J.R., 1984. Informationen für züchter. Der Dachshund, pp. 261–264.

Tuber, D.S., Hennessy, M.B., Sanders, S., Miller, J.A., 1996. Behavioural and glucocorticoid responses of adult domestic dogs (*Canis familiaris*) to companionship and social separation. J. Comp. Psychol. 110 (1), 103–108.

Uhde, T.W., Malloy, L.C., Slate, S.O., 1992. Fearful behaviour, body size, and serum IGF-I levels in nervous and normal pointer dogs. Pharm. Biochem. Behav. 43 (1), 263–269.

van der Borg, J.A.M., Netto, W.J., Planta, D.J.U., 1991. Behavioural testing of dogs in animal shelters to predict problem behaviour. Appl. Anim. Behav. Sci. 32, 237–251.

Verga, M., 1983. Relative influence of genetic and environmental factors on the behaviour of German Shepherd dogs and Rottweilers. The human–pet relationship. Proceedings of the Symposium at the Institute for Interdisciplinary Research on the Human–Pet Relationship (IEMT), Vienna, 27–28 October, Aust. Acad. Sci. 89–97.

Vincent, I.C., Michell, A.R., 1992. Comparison of cortisol concentrations in saliva and plasma of dogs. Res. Vet. Sci. 53, 342–345.

Wayne, R.K., Ostrander, E.A., 1999. Origin, genetic diversity, and genome structure of the domestic dog. BioEssays 21 (3), 247–257.

Weiss, E., Greenberg, G., 1997. Service dog selection tests: effectiveness for dogs from animal shelters. Appl. Anim. Behav. Sci. 53, 297–308.

Wemelsfelder, F., Batchelor, C., Jarvis, S., Farish, M., Calvert, S., 2003. The relationship between qualitative and quantitative assessments of pig behaviour. In: Proceedings of the 37th International Congress of the ISAE, Abano Terme, Italy, June 24–28, p. 42.

Wilsson, E., 1984. The social interaction between mother and offspring during weaning in German Shepherd dogs: individual differences between mothers and their effects on offspring. Appl. Anim. Behav. Sci. 13, 101–112.

Wilsson, E., Sundgren, P.E., 1997a. 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. Appl. Anim. Behav. Sci. 53, 279–295.

Wilsson, E., Sundgren, P.E., 1997b. The use of a behaviour test for the 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.

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.

Wright, J.C., 1980. The development of social structure during the primary socialisation period in German Shepherds. Dev. Psychobiol. 13 (1), 17–24.

Wright, J.C., 1983. The effects of differential rearing on exploratory behaviour in puppies. Appl. Anim. Ethol. 10, 27–34.

Wyrwicka, W., 1959. Studies on detour behaviour. Behaviour 14, 240–264.

Young, M.S., 1988. Puppy selection and evaluation. Dogs: Companions or Nuisances? Public Seminar. vol. 22. Werribee Vet. Clin. Ctr, Princes Highway, pp. 8–15.