Applied Animal Behaviour Science 163 (2015) 1–18

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Review article

Applied personality assessment in domestic dogs: Limitations and caveats

Diana J. Rayment*, Bert De Groef, Richard A. Peters, Linda C. Marston

La Trobe University, Bundoora, VIC 3086, Australia

ARTICLE INFO

Article history:
Accepted 27 November 2014
Available online 9 December 2014

Keywords:
Dog
Shelter
Personality
Behaviour test
Temperament
Coping style
Shyness

ABSTRACT

Modern behavioural tests for domestic dogs are commonly employed across a variety of fields, including the assessment of companion, assistance, working and sporting dogs. While there is increasing scientific attention on the development and evaluation of behavioural tests used to assess personality, a lack of robust assessment of the validity and reliability of many test protocols currently in use 'on the ground' has led to the repeated questioning of their use by both the scientific community and the public. This is particularly prevalent within the welfare sector, where consequences for dogs that are identified as behaviourally unsuitable for rehoming are dire. Despite the large body of work dedicated to canine personality and behavioural test development, relatively few reported protocols have been demonstrated to accurately and reliably predict behavioural traits in dogs, according to accepted validation and reliability assessment protocols used by psychologists. Inconsistencies in the terminology used to describe behaviour, tests and canine personality are common within the literature pertaining to applied personality testing of dogs, making discussion of the usefulness of such tests within the applied environment difficult. Furthermore, highly relevant information gained from the fields of psychology and ethology, such as the limitations of behaviour-only scoring systems, the need to assess both temporal and contextual stability during personality testing, and the effects of arousal, impulsivity and affective state on behaviour, is seemingly overlooked by some test developers, resulting in array of test protocols that suffer from similar, predictable design limitations. This review discusses the most prominent considerations and limitations relevant to the assessment of personality through behavioural tests in domestic dogs, and suggests possible mechanisms through which these limitations may be investigated and addressed. In particular, it is recommended that the results of shelter-based behavioural test batteries are interpreted within the limitations of the context in which they are performed, and test developers consider other methods with which personality may be assessed within the shelter environment.

Abbreviations: DIAS, dog impulsivity assessment survey; DMA, dog mentality assessment; FFM, Five-Factor Model; 5-HIAA, 5-hydroxyindoleacetic acid; HPA, hypothalamo-pituitary-adrenal; SAM, sympatho-adrenal-medullary.

* Corresponding author at: AgriBio, L1Q2, 5 Ring Road, Bundoora, VIC 3083, Australia. Tel.: +61 3 9032 4438; fax: +61 3 9471 0224.
E-mail addresses: d.rayment@latrobe.edu.au (D.J. Rayment), b.degroef@latrobe.edu.au (B. De Groef), richard.peters@latrobe.edu.au (R.A. Peters), linda@marston.com (L.C. Marston).

http://dx.doi.org/10.1016/j.applanim.2014.11.020
0168-1591/

Contents

1. Introduction 2
2. Inconsistencies in terminology 3
2.1. Personality, temperament and trait 3
2.2. Behavioural syndromes 4
2.3. Shyness and the shy–bold continuum 4
3. Prevalent personality models for non-human animals 5
3.1. Coping strategies 5
3.2. Shy–bold continuum 7
3.3. The Five-Factor Model and other questionnaire-based models 8
4. Considerations for assessment of canine personality using behavioural tests 9
4.1. Impulsivity 10
4.2. Judgement bias and affective state 10
4.3. Arousal 11
4.4. Temporal and contextual stability 12
4.5. Behavioural test scoring methods and identification of passively responding individuals 12
5. Where to from here? 13
References 14

1. Introduction

Behavioural tests for domestic dogs are broadly applied in the modern fields of companion dog welfare, assistance and guide dog programmes, pet dog breeding programmes and by modern sports dog clubs. The general aim of these tests is to record the responses of dogs to a variety of stimuli, as test behaviour is thought to provide reliable information about personality and behaviour outside of the test environment. More specifically, working dog tests such as those carried out by guide, service, military and police dog programmes, attempt to assess an individual's aptitude for work-related tasks, and to enable selection of the 'most fit-for-purpose' dogs for breeding or training purposes. Alternatively, tests used by welfare shelters and to a lesser extent by pet dog breeders and enthusiasts, have two broad functions. Firstly, they are a form of risk assessment designed to limit the possibility of rehoming potentially aggressive dogs back into the community (Bennett et al., 2012). Secondly, they attempt to assess canine personality to enable accurate matching of new pet behaviour to owner expectations (Marston and Bennett, 2003). Although these tests are popular within the welfare community, and indeed are firmly embraced, the validity and reliability of behavioural tests used by shelters have been repeatedly questioned by both the scientific community (Taylor and Mills, 2006; Mornement et al., 2009) and the public (Robertson, 2004; Bartholomew, 2012). Primary criticisms of current test protocols include the poor consensus on how tests should be constructed, administered and interpreted, a lack of experience and formal education for some test users and developers, and a lack of robust assessment of the validity and reliability of tests before they are used (Taylor and Mills, 2006; Bräm et al., 2008; Mornement et al., 2010). In an online survey of the Australian public, Mornement et al. (2012) found that while two thirds of respondents agreed or strongly agreed that shelters are trustworthy and sell good-quality adult dogs, approximately one third thought that adult shelter dogs have behavioural problems and a further 38.5% neither agreed nor disagreed with the same statement. This indicates that there is some recognition by interested members of the public that shelters may currently be unable to identify and treat all behaviour problems in dogs that they adopt out. Within the sector, well-justified criticisms of test procedures have eroded trust in the results of behavioural tests (Segurson, 2007; National Canine Research Council, 2013), leading to confusion and conflict (Robertson, 2004), particularly as death is a common consequence for dogs that fail in-shelter tests (Marston et al., 2004; Clark et al., 2012; National Canine Research Council, 2013).

In order to understand the heavy reliance on behavioural tests by many bodies involved with dogs, it is important to appreciate the history of such tests. Behavioural tests were historically performed by breed clubs and working dog interest groups alongside real-world performance and health indicators, as tools for selection of working and breeding animals; one such example is described by Wilsson and Sundgren (1997). They typically entailed 'mock' situations similar to those that the dogs were likely to encounter while working, such as herding or retrieving trials, and specific games developed to assess working ability or work ethic, such as Schutzhund (Ruefenacht et al., 2002; Lindberg et al., 2004; Courreau and Langlois, 2005). Importantly, due to the significant overlap between the situational variables of the test and those encountered in real-world scenarios, these tests were assumed to accurately reflect the inherent working ability of the assessed dogs even though formal validity and reliability evaluations of the tests are rare.

Similarly to early working dog tests, modern behavioural tests, in particular those carried out in welfare shelters, also frequently suffer from a lack of standardisation and formal assessment (Taylor and Mills, 2006; Mornement et al., 2009). However, unlike traditional working dog tests, modern shelter-based behaviour tests typically involve situations quite different from those one would expect a pet dog to encounter outside of the shelter environment. These behaviour assessments usually involve various subtests conducted in a novel space, by a tester that is unfamiliar to the dog (Mornement et al., 2010; Dowling-Guyer et al., 2011; Bennett et al., 2012),

within the shelter environment which is known to be both highly stressful for dogs (Sales et al., 1997; Tuber et al., 1999) and to affect the way they interact with unknown humans (Barrera et al., 2010). As discussed by Gosling (2008) and Stamps and Groothuis (2010) and reviewed at length by Uher (2011a), personality assessment requires across-context reaction norm assessment as well as measures of temporal consistency, making delineation between the effects of personality, learning and motivation on the behavioural responses examined impossible in a one-time assessment performed in a single-context environment.

In a review of the literature, Taylor and Mills (2006) found that very few documented tests were described in enough detail to allow accurate replication and none had been shown to reliable, accurate and valid, a finding replicated by Mornement several years later (2009, 2010). With a small but growing number of notable exceptions, such as Valsecchi et al. (2009, 2011), Sinn et al. (2010), Barnard et al. (2012), Riemer et al. (2014) and Arvelius et al. (2014), it is still the case that relatively few investigators report the predictive and/or concurrent validity of published protocols, or address the shortcomings of single-context tests for the assessment of personality when discussing the limitations of a suggested protocol. This lack of assessment limits the usefulness of these tests for applied purposes as the accuracy and reliability of the protocols is unknown. As highlighted by Gosling (2008) and discussed by Uher (2011b), in order to properly design and evaluate a behavioural test capable of assessing personality, one must first have a sound understanding of the broader field of personality testing in humans and other non-human animals, in order to develop a sound understanding of the limitations of current tests for dogs.

This review will examine inconsistencies in the terminology used to discuss personality, which appear to significantly contribute to confusions surrounding the field of non-human animal personality research. The three most common personality models that have been applied to non-human animals, in particularly dogs, will be considered and several important key considerations for test developers and users will be discussed. Based on these considerations, limitations of current canine behavioural tests, specifically those used by welfare shelters, will be reviewed and possible areas of further research suggested.

2. Inconsistencies in terminology

Psychologists conceptualise behaviour as an expression of the interaction between motivations, innate patterns of behaviour, physiology, learning based on previous experience and circumstantial or contextual factors (Murray, 1938). As such, psychologists have long recognised that the behaviour of an individual in one context is not separate to, or isolated from, its behaviour in other contexts. More recently, this concept is increasingly being recognised as an important consideration by biologists in the study of population ecology and evolution, as traditional models that focus on behavioural variation within a single context, such as the optimality model, consistently fail to account for all behavioural patterns observed in wild populations (Bell, 2007; Wolf and Weissing, 2012). This increasingly wide interest in animal personality research has resulted in a large body of related work conducted by researchers from widely diverse backgrounds. It is perhaps due to this wide variation in approaches that terms which previously had an accepted meaning in the field of human psychology have been redefined by so many researchers, that authors are now required to define their personal understanding of key terms prior to use to avoid misunderstandings (Sih et al., 2004a; Freeman and Gosling, 2010; Toms et al., 2010). It is not surprising, therefore, that a common complaint of researchers attempting to compare personality studies across species and disciplines, is that efforts to directly assess the relationship between personality traits and methodologies across studies are hampered significantly by inconsistencies in terminology and unclear methodological descriptions (Sih et al., 2004b; Jones and Gosling, 2005; Freeman and Gosling, 2010; Toms et al., 2010). The following section will examine the use of several key terms in the personality literature.

2.1. Personality, temperament and trait

Across the non-human animal personality literature, ‘personality’ and ‘temperament’ are often used interchangeably, despite having clear and separate historical definitions in human psychology. ‘Temperament’ describes the stable, inherent and early appearing behavioural tendencies of an individual, while ‘personality’ describes the measureable result of temperament and environment interactions (Rothbart et al., 2000). Thus, one is only able to identify the contribution of temperament to personality measures by determining the heritability of temporally consistent traits. It is interesting to note the widespread and incorrect use of the word ‘temperament’ to describe personality by many involved with canine behavioural assessment (Taylor and Mills, 2006; Valsecchi et al., 2009; Mornement et al., 2010; Barnard et al., 2012). While the terms are commonly used interchangeably, an understanding of the difference between behaviour, temperament and personality is necessary before one can fully grasp the requirements for tests to obtain information about each, and as such, it seems logical for the traditional usage of these terms to be embraced during discussions between those interested in the applied assessment of canine personality, as all three types of tests are utilised within the field.

Probably of greater consequence than the poor use of ‘temperament’ and ‘personality’, is the varied use of the term ‘trait’ when referencing behaviour. In human psychological research, the word trait refers to a dimension or axis upon which consistent individual behavioural differences are identified and maintained within a population (Itoh, 2002). Traits influence the expression of behaviour across different contexts and remain relatively stable over time, distinguishing them from species-specific and context-specific behaviours, such as nest-building methods or mating rituals (Toms et al., 2010). Furthermore, motives, described as the desires or drives that compel an individual to exhibit goal-oriented behaviour (Winter et al., 1998), influence the expression of personality traits. For example, the trait of fearfulness (i.e. a general tendency to

behave fearfully across a variety of situations) can be distinguished from a fear of spiders (i.e. the strong desire to be physically distant from spiders based on anxiety caused by close proximity) by simply exposing an individual to a variety of potentially fear-eliciting stimuli and assessing their response. A fearful or anxious person may or may not be afraid of spiders, but in isolation, his behaviour in response to the presentation of a spider gives, at best, limited information about his fearful tendencies.

In contrast to the psychological definition, biologists define trait more generally as any phenotypic characteristic of a population that shows consistent variation along an axis, including morphological characteristics such as weight and life history characteristics such as infant mortality (Réale et al., 2007). Others define trait more narrowly, proposing that any temporally consistent behavioural variation displayed within a population in a single context should be considered a trait (Sih et al., 2004b). It appears that at least some of this variation in understanding of the term stems from differences in the way 'context' is defined by different researchers. Behavioural and population ecologists generally define 'context' by the apparent function of the behaviour exhibited; for example, all predator avoidance behaviours fall into one context regardless of other factors (Sih et al., 2004b; Bell, 2007). In contrast, other researchers hold the more traditional psychological view that 'context' consists of all external and internal variables which affect the expression of behaviour (Stamps and Groothuis, 2010). This inconsistent use of the terms 'trait' and 'context' inevitably leads to confusion over the definition of 'behavioural syndromes' and highlights the need for close attention to be paid to study design and behavioural scoring methods when evaluating the applicability of previous work.

2.2. Behavioural syndromes

Studies assessing behavioural variation across large subsets of animal populations have indicated that certain traits, including propensity for overt and aggressive responses to threats, social and non-social fearfulness, risk aversion and sociability, tend to co-vary in a somewhat predictable manner (Wilson et al., 1994; Koolhaas et al., 2010). These co-variations in the expression of multiple traits have been termed 'behavioural syndromes' by non-human animal researchers (Sih et al., 2004b). This term is now widely used by ecologists, although some variation in the precise definition still exists. As previously discussed, this inconsistency appears to stem from variation in the definitions of trait and context. Some authors consider consistent co-variation of behaviours across several contexts to be indicative of a behavioural syndrome and understand a behavioural syndrome to be synonymous to a single personality trait (Sih et al., 2004b; Bell, 2007; Dochtermann and Jenkins, 2007; Gosling, 2008). Others, in keeping with traditional definitions of trait and personality from psychology, define behavioural syndromes as co-variations of a sub-set of personality traits, each of which are individually consistent across contexts and therefore only descriptive of one portion of personality (Dalley and Roiser, 2012). This definition equates behavioural syndromes with whole personality descriptions, or so-called 'super traits'. Furthermore, as noted by Beckmann and Biro (2013), an increasing trend towards using single-context tests as proxy measures of behavioural syndromes has emerged, limiting the extent to which studies can be compared even further.

2.3. Shyness and the shy–bold continuum

In human psychological research, the narrow-order trait of shyness has been described as the exhibition of fearful or tense behaviour, along with increased hypothalamic-pituitary-adrenal (HPA) axis activity in response to novel social situations (Cheek and Buss, 1981), or those where an individual feels judged in some way by conspecifics (Crazier, 1979). While shyness may or may not be coupled with low sociability (Bruch et al., 1989; Coplan et al., 2004), it is linked to but distinct from the broader personality traits of introversion and neuroticism (Crazier, 1979; Ebeling-Witte et al., 2007). Early investigations into the consistency and stability of responses to social and non-social novelty in infants and children, suggested that babies that present as extremely shy are temperamentally inhibited and tend to remain so through until at least their 7th year (Kagan et al., 1988; Kagan and Snidman, 1990). Later work by the same authors indicated that while motor arousal and fear reactivity measures showed some co-variation, they were distinct responses indicating that activity level was not always indicative of stress intensity in this sample group (Kagan, 1997). Based on the lack of correlation between motor activity, physiological response and vocalisations observed in these studies, along with the patterns of variation observed in these three variables, Kagan and colleagues (1997) proposed that high and low fear-reactive individuals fall into two distinct groups, rather than representing the extreme ends of a behavioural spectrum. In a later study utilising the same sample group, Schwartz et al. (2003) showed that individuals categorised as highly fearful of novel stimuli as infants, termed 'shy', displayed greater amygdala responses to the presentation of novel stimuli as adults, indicating a biological basis for neophobic responses and aversion to risk that remained consistent despite significant development and learning.

The inhibited-uninhibited categorisation of human infants developed by Kagan and colleagues has been proposed by several authors as the basis of the shy–bold continuum in non-human animals (Coleman and Wilson, 1998; Svartberg and Forkman, 2002; De Meester et al., 2008). In these cases, 'shyness' is often simply defined as the fearful extreme of the shy–bold continuum, a 'super-trait' which is thought to control variation in other personality dimensions such as sociability and playfulness (Svartberg and Forkman, 2002). As such, 'shyness' is often not clearly defined as much of the shy–bold literature focuses specifically on behavioural measures of boldness. However, in studies of boldness in dogs it is common for overt fear behaviours such as fleeing to be used as an indicator for shyness (Svartberg, 2002; Starling et al., 2013). Contrary to other non-human animal researchers, Koolhaas et al. (2010) used the term 'shy' to describe individuals that display high stress responsiveness and a passive behavioural reaction to stress, making little

reference to sociability outside of competitive interactions and non-threat-related motivations.

While the general definition of boldness has been widely accepted as describing an individual with little aversion to risk or novelty, who actively seeks out and engages in social interactions and expresses aggression during competitive encounters (Wilson et al., 1994; Svartberg, 2005; Turcsán et al., 2011; Carter and Feeney, 2012; Starling et al., 2013c; McDermott et al., 2014), the methodology used to assess boldness varies widely between studies. Researchers have often utilised single-context, one-time responses that they consider to be reliable indicators of a bold phenotype, such as reaction to novelty (Drent et al., 2003; Carere and van Oers, 2004; Frost et al., 2007), competitive aggression towards conspecifics (Natoli et al., 2005) or response to predators (Dochtermann and Jenkins, 2007; Biro et al., 2010). These single-context, one-time tests for boldness lend themselves to criticism, as common single-context tests for boldness often fail to show predictive validity for other measures of boldness (Coleman and Wilson, 1998; Beckmann and Biro, 2013; White et al., 2013; reviewed by Carter et al., 2013), leading some authors to suggested that common methods for assessing 'boldness' are not valid (Réale et al., 2007; Toms et al., 2010; Beckmann and Biro, 2013), and may in fact be measuring different traits (Carter et al., 2013). A further source of variability between studies appears to be a poor consensus on the role of novelty in the assessment of the boldness. While some authors consider neophobia an integral component of shy–bold assessments (Toms et al., 2010), others consider the fear of novelty as a distinct motivation and therefore separate to the boldness measures (Beckmann and Biro, 2013).

3. Prevalent personality models for non-human animals

Possibly due to methodological limitations when working with large numbers of animals in novel environments, the most commonly studied personality dimensions in non-human animals are those associated with the stress response. These include all 'coping behaviours', being those behaviours aimed at reducing the aversive impact of potentially fitness-threatening situations or stimuli (Wechsler, 1995). Stress-related behavioural syndromes are evident in a large number of species from different taxa, including fish (Toms et al., 2010), birds (Carere and van Oers, 2004), cats (Lowe and Bradshaw, 2001), dogs (Svartberg and Forkman, 2002), and several species of rodents (Benus et al., 1991; Dochtermann and Jenkins, 2007). Given the wide variety of species in which stress-related behavioural syndromes have been recognised, it has been suggested that the underlying mechanism of behavioural control governing the stress response has been highly preserved throughout evolution (Gosling and John, 1999; Sih et al., 2004a; Koolhaas et al., 2010).

While loosely describing similar patterns, different researchers have used a variety of terminologies to describe coping-related behavioural syndromes, including 'coping styles' (Koolhaas et al., 1999), the 'shy–bold continuum' (Svartberg, 2002; Carere and van Oers, 2004), or more broadly as just 'behavioural syndromes' (Sih et al., 2004a, 2004b). Despite a significant degree of similarity in the patterns described by each author, the terminology and testing methodologies used vary widely between studies. While authors often refer to the results of previous work as either supporting or contradicting their findings, it is uncommon to find objective discussions of the parallels and deviations between models.

In contrast to the behavioural test approach to assessing animal personality, a smaller number of non-human animal researchers have embraced questionnaire-based assessments reliant on subjective ratings or behavioural descriptions by people familiar with the dog/s being assessed. The most common of these is the Five-Factor Model (FFM), developed by psychologists to describe human personality. While the broad applicability and all-inclusiveness of these models are attractive to those interested in canine personality for applied purposes, the stress-inducing nature and novelty inherent to many testing situations limits the suitability of this model for information obtained by typical battery-style tests. However, valid and reliable canine personality questionnaires are valuable tools for assessing concurrent validity during behavioural test design, and as such, are an important consideration for applied behavioural assessments.

The following section will examine the three most common personality models embraced by animal personality researchers: the coping style-emotionality model, the shy–bold continuum and questionnaire-based models (Table 1).

3.1. Coping strategies

There has been comparatively little work carried out in dogs based on the model of coping strategies proposed by Koolhaas et al. (2007, 2010), which describes both behavioural coping style and physiological reactivity to stressful stimuli. However, if this model holds true for domestic canines, work carried out on other species provides a wealth of useful information that is applicable for both pet owners and working dog handlers alike. Early work examining the stress response of laboratory rodents indicated the presence of a personality dimension along which individual 'default' activity level and behavioural response to stressful situations varied, called the coping style axis (Benus et al., 1991). A large body of work in small mammals and birds has stemmed from these early characterisations, providing descriptions of the typical behavioural profiles of individuals at the extreme ends of this spectrum.

Proactive individuals are generally described as behaviourally inflexible and socially competitive. They easily form habits and respond actively when faced with social or environmental challenges, showing high levels of offensive aggression in social conflicts (Sgoifo et al., 1996; de Boer and Koolhaas, 2005; Natoli et al., 2005; Horváth et al., 2007), high levels of active avoidance and escape behaviours when confronted with environmental stressors (Benus et al., 1989, 1990; de Boer and Koolhaas, 2005; Koolhaas et al., 2007) and shorter latencies to explore novel environments or stimuli, with longer adaptation

Table 1
Summary of three common models of non-human animal personality applicable to domestic dogs.

Model	Scope	Assessment method	Current or potential application	Strengths	Limitations
Coping strategies	Responses to aversive stimuli, novelty and social threat	Behavioural and physiological tests	• Response to stress and novelty in unknown dogs (e.g. those in welfare shelters) • Coping response in working or pet dogs	• Separates intensity of stress response from quality of behavioural response • Robust testing mechanism • Large body of existing knowledge developed across a range of species, including physiology of traits	• Very little information specific to domestic dogs • Limited coverage of traits that influence coping strategy, e.g. vocalisations known to vary separately
Shy–bold continuum	Risk aversion, neophobia, sociability and human-directed playfulness	Behavioural tests and owner questionnaires	• Screen for potential in candidate working dogs • Breed- or population-wide behaviour studies	• Existing body of work on domestic dogs • Uses both behavioural tests and owner questionnaires • Traits covered are directly relevant for working dog selection • Broader body of work using other species may also be relevant to dogs	• No clear consensus on the role of novelty and neophobia in model • Some existing work is generalised from single-context tests, so has limited applicability more widely • Canine-specific behavioural tests score stress intensity based on behavioural response; has limitations for passively responding individuals • Behaviour scoring systems using behavioural measures of stress response may fail to identify passively responding individuals
Questionnaire-based models	Entire suite of externally observable traits	Owner questionnaires	• Validation tool for behaviour-based personality tests • Breed or population wide behaviour studies	• Extremely broad scope • Covers long-term trends in behaviour, rather than one-time tests • Large existing body of work specific to domestic dogs	• No clear consensus on personality structure or number of traits • Lexical models not validated for use in dogs • Inconsistency of language use presents methodological limitations • Questionnaire scoring systems using behavioural measures of stress response may fail to identify passively responding individuals

times in the presence of environmental changes (Benus et al., 1987; Sluyter et al., 1996; Bolhuis et al., 2004).

In contrast to the active and impulsive behaviour of proactive responders, reactive individuals show a more tightly coupled stimulus-response pattern of behaviour that resists habit formation (Benus et al., 1987; Bolhuis et al., 2004); they display little aggression during social conflicts (Sgoifo et al., 1996; de Boer and Koolhaas, 2005; Natoli et al., 2005), exhibit inhibited and avoidance behaviours in response to environmental stressors (Benus et al., 1989, 1990; de Boer and Koolhaas, 2005; Koolhaas et al., 2007) and take longer to explore novel environments with quicker adaptation to environmental change than their proactive conspecifics (Benus et al., 1987; Sluyter et al., 1996; Drent et al., 2003).

While the majority of previous studies have focused on the coping style axis and have treated the proactive and reactive groups as homogeneous, work in calves has shown significant within-group variation in activity of the HPA axis, indicating differences in the stress reactivity of individuals who exhibit similar behavioural styles (Van Reenen et al., 2004, 2005). This dimension, later called the emotionality axis by Koolhaas et al. (2007), describes an individual's proneness to anxiety and is somewhat analogous to the concept of trait anxiety (Bar-Haim et al., 2007). Based on the total body of work examining coping strategies of different species, Koolhaas et al. (2007) proposed a revised, two-dimensional model of coping responses, in which coping behaviour patterns and physiological stress reactivity are two independent traits. This revised model shows a number of parallels to the motor arousal and fear reactivity model in human infants, proposed by Kagan (1997). Importantly, this model for coping strategies does not account for individual differences in intelligence, learning or motivations and as such does not describe all personality dimensions that could influence an individual's coping response (Van Reenen et al., 2004; Koolhaas et al., 2007). Interestingly, both Kagan (1997) working with human infants and Van Reenen et al. (2004) working with dairy calves, found that the degree of vocalisation displayed during the stress response was not reflective of either the motor response or the fear response, indicating that a separate dimension related to verbalisations may be present.

The four personality descriptions emerging from the two dimensional model of coping strategies described by Koolhaas et al. (2007) represent the four 'extreme' manifestations of interactions of the two axes. 'Docile' or 'calm' individuals exhibit a low stress reactivity and passive behavioural coping style; 'shy' individuals are stress-prone with a passive coping style; individuals described as 'panicky' are highly stress-prone with an active coping style; and those labelled 'bold' are active responders with a high stress tolerance (Koolhaas et al., 2010). Similar to Réale et al. (2007) and Sih et al. (2004b) and in line with standard psychological protocol for personality assessment, Koolhaas et al. (2010) advised that in order to identify the default behavioural type of an individual, assessments must be conducted using a variety of threatening stimuli or situations, each of which allows both passive and active responses to be equally successful in reducing the perceived threat. Koolhaas et al. (2010) also reiterate that the quality of coping responses varies between individuals and is therefore not a reliable measure of stress intensity, resulting in the need for both behavioural assessment and physiological measures of arousal for any reliable characterisation of default coping strategy to be made.

In a study carried out by Horváth et al. (2007), the coping responses of working police dogs to social stress were measured using a test battery. Three typical response patterns were identified and described as 'fearful', 'aggressive' and 'ambivalent'. Patterns were characterised by varying levels of avoidance and aggressive behaviour, coupled with varying degrees of active handler-greeting behaviours. The group labelled 'fearful' displayed avoidance behaviours, including moving away from the threatening decoy and backing up, along with a moderate level of aggressive behaviour and a small but statistically significant rise in salivary cortisol concentration during the test. In contrast, the aggressive group showed active approach behaviours with little HPA activation and the third 'ambivalent' group displayed passive behaviours such as paw lifting, snout licking and averting gaze when in close proximity to the decoy, along with the largest mean cortisol increase. The authors concluded that the responses observed were indicative of the proactive (aggressive group) and reactive (fearful group) coping styles identified by Svartberg (2002), with a third group of 'ambivalent' dogs which had not previously been described.

The unusually high level of aggression seen in the fearful group and the offensive behaviour noted in the ambivalent group while at a distance from the decoy could indicate a limitation in test design. This protocol required the dogs to be tethered on a relatively short chain and approached closely while being physically threatened (Horváth et al., 2007). As discussed by Koolhaas et al. (2007), tests to determine default coping style must allow for the successful expression of all possible coping behaviours. As this was not the case in this study, the test design may have resulted in a bias towards aggressive responses, regardless of default coping style, as both flight and passive responses were restricted. Also, given that only fully operational police dogs with a known history of low flight responses to threatening situations were used, there is a possibility that the sample included both non-fearful dogs of unknown coping style and fearful, but passively coping dogs. The third group described by Horváth et al. (2007) displayed active behaviours while at a distance from the threatening decoy, along with passive behaviours when approached closely. Given the sample selection criteria, it is possible that this group includes highly stressed but passively coping dogs that would be described as 'shy' by Koolhaas et al. (2007). This, combined with the design limitations of the test could explain the high level of stress reactivity in the 'ambivalent' group, compared to both of the more proactively responding groups. Due to the design limitations of the work, any comparisons of the results to either the coping style or shy–bold models are speculative at best.

3.2. Shy–bold continuum

Of the most prevalent personality models applied to animals, the shy–bold model has received the most attention

from canine researchers. In relation to dogs, the shy–bold dimension describes a behavioural syndrome in which low levels of neophobia and low risk aversion co-occur with high levels of playfulness and sociability directed towards strangers (Svartberg, 2002; Svartberg and Forkman, 2002). The majority of shy–bold work carried out in dogs is based on results of the Swedish Dog Mentality Assessment (DMA), a test battery designed by the Swedish Working Dog Association to describe temperament in working dogs, as an aid in the selection of dogs for breeding, service or working roles (Svenska Brukshundklubben, 2012).

Like similar tests, the DMA scores easily observable, gross behaviours thought to be indicative of personality such as startle responses, flight and aggression. Early work based on analysis of 15,329 DMA results for dogs from 165 breeds, described five narrow-order factors including playfulness, fearlessness, chase-proneness, sociability and aggressiveness and one overarching or broad characteristic that positively correlated with playfulness, sociability and chase-proneness, interpreted as a measure of the shy–bold continuum (Svartberg and Forkman, 2002). A follow-up study in which these factors were compared to owner reports of behaviour gathered 1–2 years after DMA tests were performed, indicated that the traits playfulness, fearlessness and sociability reflected degrees of human-directed play interest, non-social fear and stranger-directed fear and were therefore both valid and temporally stable over a period of 1–2 years (Svartberg, 2005). Using a similar questionnaire, Starling et al. (2013b) found that increasing age and neutering decreased boldness measures and that male dogs were typically more bold than female dogs. Svartberg (2006) observed that trait scores for playfulness, fearlessness, sociability and aggression differed significantly between breeds, but reflected modern selective pressures rather than historical breed functions. To the contrary, using an expanded questionnaire similar to that used by Svartberg (2005), Starling et al. (2013c) found that boldness measures were related to historic breed groups and suggestive of retained personality differences based on breed function, despite modern selective pressures. As discussed by Starling et al. (2013c), this discrepancy perhaps reflects a difference in the sample group as the second study included a larger variety of breed groups and demographic variables.

In a separate study, Svartberg et al. (2005) investigated the test–retest reliability of the six factors identified by Svartberg and Forkman (2002). Whilst test–retest reliability assessment of the measures is the primary aim of repeat testing, Réale et al. (2007) and Stamps and Groothuis (2010) also suggest this 'reaction norm' style test design as a robust mechanism for disentangling the effects of novelty and exploration from boldness measures. This study involved initial testing of 81 dogs and two repeat tests at monthly intervals for the seven highest- and lowest-ranked dogs from each factor, using a novel test area for each round. It also appears that while the handlers remained consistent, more than one 'test leader' was used during repeat tests. All factors demonstrated high rank-order consistency across tests, indicating temporal stability of the traits. However, mean factor scores for curiosity/fearlessness and aggressiveness differed significantly between original and repeat tests. This effect was presumed to be due to a lack of novelty on repeat exposure to the test decreasing the degree of fearfulness and fear-related aggression in neophobic dogs (Svartberg et al., 2005). Based on these results, Svartberg et al. (2005) suggested that the fearlessness and aggression factors were sensitive to novelty measures, while playfulness, chase-proneness and sociability were not. These findings are in contrast with the results of Svartberg (2005) discussed earlier, which indicated that questionnaire-based assessments of stranger-directed aggression, stranger-directed fear and non-social fear were reflected in the DMA test measures 'playfulness', 'chase-proneness' and 'sociability', thus indicating that these measures are sensitive to both social and non-social novelty. The use of a novel area and different test leaders in repeat tests during the test–retest reliability work could explain this discrepancy, as these differences may have been significant enough to influence behaviours reflecting a positive affective state in neophobic individuals (such as 'playfulness'), negating the benefits of reaction-norm testing for these measures.

Notably, the suppressive effect of social fear or neophobia on behavioural responses governed by positive motivations, such as play or social motivation, is generally poorly examined during discussions of the boldness dimension in dogs. In animals it is generally accepted that negative affective states such as fear inhibit the expression of positive affective states (reviewed by Boissy et al., 2007). Similarly, previous work in people has indicated that social fear (termed 'shyness') and social motivation (termed 'sociability' or 'sociality') are related but independent traits in humans (Cheek and Buss, 1981), with shyness moderating the expression of sociability such that individual differences in sociability are impossible to measure in situations which elicit a shy response (Bruch et al., 1989; Coplan et al., 2004). As such, the expression of play behaviours by socially motivated dogs, that are also sensitive to social novelty, is unlikely in a typical behaviour test battery such as the Swedish DMA. As such, further work is required to determine whether boldness measures in dogs govern variation in traits reflecting play and sociable behaviours, or whether this is simply an artefact of test design.

3.3. *The Five-Factor Model and other questionnaire-based models*

The Five-Factor Model (FFM) or Big Five of personality is the most broadly accepted and widely supported model of personality in the field of human trait psychology. The FFM explains personality as an interaction of the traits 'extraversion', 'agreeableness', 'conscientiousness', 'neuroticism', and 'openness to experience' (McCrae and John, 1992). It has remained popular despite repeated criticism (McCrae and John, 1992) and has shown validity across a wide variety of populations (McCrae et al., 1998), ages (Soto et al., 2011), languages (McCrae and John, 1992) and methods of analyses (Goldberg, 1990; Oh et al., 2011). It is therefore not surprising that a number of canine researchers have attempted to examine how canine personality measures align with human FFM traits. Using a variety of methods, researchers have proposed

three- (Draper, 1995), four- (Gosling et al., 2003), five- (Ley et al., 2008), six- (Fratkin et al., 2013) and seven-factor models (Jones and Gosling, 2005) for canine personality structure. Like human psychologists, these researchers have primarily focused on lexical (Ley et al., 2008; Mirkó et al., 2012) or questionnaire-based designs (Gosling et al., 2003; Kubinyi et al., 2009), or a mixture of both (Draper, 1995). However, with the exception of extraversion and neuroticism, the traits identified by canine studies are typically not directly analogous to each other or to the human FFM traits (Fratkin et al., 2013). This lack of consistency appears to stem from two causes. Firstly, as discussed by Ley and Bennett (2007), a number of authors have attempted to describe canine personality using modified human FFM questionnaires, under the assumption that FFM traits exist in dogs and are expressed in the same way by both species. To date, this assumption has not been validated. Secondly, human studies embracing the lexical approach have done so based on the assumptions that common language descriptors for behavioural traits encode real personality constructs (Uher et al., 2013), and that lay-person descriptions of human personality traits are universally recognised and used consistently across varying cultures and languages (McCrae and John, 1992). Due to inevitable methodological limitations when using common language as the basis for all measurement methods, the assumption that lexical analysis can identify the true nature of personality has received scrutiny (Uher et al., 2013; Uher, 2014). Furthermore, while it appears that common language descriptors are used consistently in humans (McCrae et al., 1998; Oh et al., 2011), studies of the ability of lay people, dog owners and experienced canine professionals to recognise, interpret and describe dog behaviour indicate that canine personality and behaviour descriptions are not consistently used or defined. Inconsistencies in use of terminology have been shown in cohorts of lay people describing the emotional or motivational state of dogs (Fidler et al., 1996; Bahlig-Pieren and Turner, 1999), as well as both lay and experienced people using generic adjectives for personality such as 'friendly' (Early et al., 2014; Tami and Gallagher, 2009). As demonstrated by Fratkin et al. (2014), people with varying levels of experience can accurately rate canine behavioural traits with appropriate pre-training. However, as demonstrated by Mariti et al. (2012) and Valsecchi et al. (2011), and in agreement with Fratkin et al. (2014), this appears to be less reliable for subtle behaviours that are not easily observed, and for abstract concepts like 'focus'.

A less common method of examining personality traits in dogs is via meta-analysis of the canine personality literature. Jones and Gosling (2005) compared the methodologies and trait descriptions of 51 existing canine personality studies, finding seven broadly analogous traits that appeared repeatedly across a number of studies, along with a number of narrow traits which were often specific to one study. The resulting trait categories included 'reactivity', 'fearfulness', 'activity', 'sociability', 'responsiveness to training', 'submissiveness', 'aggression' and 'other'. Jones and Gosling (2005) noted that the extreme variety of subjects and methods, as well as the inconsistency in language used to score behaviours and the way in which different authors described traits, made between-study comparisons of personality traits extremely difficult and often impossible. This, along with a lack of objective behavioural descriptions provided in the original studies, limits the usefulness of the trait categories developed by Jones and Gosling (2005) for direct comparisons with human FFM traits, but provides a framework on which further work investigating broad personality traits in dogs can build.

More recently, Fratkin et al. (2013) used meta-analysis of 31 studies to assess the temporal consistency of the traits identified by Jones and Gosling (2005). Fratkin et al. (2013) encountered similar difficulties in classifying previously identified traits based on poor or incomplete behavioural descriptions and subjective factor labels. They subsequently combined the 'reactivity' and 'fearfulness' categories and omitted the 'other' category, along with any traits that appeared to fit into more than one category, retaining the traits 'fearfulness', 'activity', 'sociability', 'responsiveness to training', 'submissiveness' and 'aggression'. Their findings concur with human literature showing that personality becomes more consistent with age (Roberts and DelVecchio, 2000) and that temporal stability decreases with increasing between-test intervals (Stamps and Groothuis, 2010). Fratkin et al. (2013) suggested that these findings indicate that personality dimensions are less stable in dogs than other species; however, neither they nor Jones and Gosling (2005) provided evidence for the content and concurrent validity of the personality categories used, so the results could just as easily indicate that these categorisations were not fully reflective of personality traits. This work highlights the key issues of inconsistency in the terminology used by canine personality researchers and lack of detailed, objective methodological descriptions, which provides a useful starting point for further discussions on the language used to describe canine personality. Clarification of the behaviours and testing methods used by researchers to describe canine personality traits could allow for the development of more robust canine personality questionnaires, based on consistent behavioural variation observed in domestic dogs within particular environments. One such 'bottom-up', behaviour-based approach to personality assessment has been demonstrated in several primate species by Uher and Asendorpf (2008) and Uher et al. (2013a, 2013b). Such questionnaires would prove invaluable during behavioural test design and validation for specific applied purposes, as well as providing a platform on which to study the structure of canine personality for theoretical or comparative purposes.

4. Considerations for assessment of canine personality using behavioural tests

Researchers studying animal personality are faced with a new set of challenges above and beyond those faced by human personality researchers. Information on the internal state of an individual is difficult or impossible to obtain as animals are unable to verbally communicate their motivations, stress levels and affective state, access to an acceptably large sample of individuals is often limited or impossible, and facilities and complex testing protocols

are generally inherently stress-inducing for test animals, reducing the array of traits that can be studied to only those involved in the stress response. All of these factors must be taken into consideration by test developers, so that tests can be designed to control for as many variables as possible and those that cannot be controlled for are noted, so that design limitations are clearly understood and accounted for during interpretation of results. The following is a discussion of a number of key concepts that are both relevant and commonly overlooked by canine personality test developers.

4.1. Impulsivity

The concept of impulsivity as being a general lack of self-control and poor consideration of consequences prior to action is familiar to most people, but there is little scientific consensus on the exact nature or definition of impulsivity. Researchers widely agree that impulsivity is not a distinct personality trait, but rather a number of different factors that interact to produce a common suite of behavioural responses linked to intolerance of delayed rewards, poor behavioural control in unpredictable situations and a lack of consideration for the future or for long-term consequences (Eysenck, 1993; Evenden, 1999; Fineberg et al., 2009). In domestic dogs, a high degree of impulsivity has been linked to high activity levels, poor attention span and human-directed aggression (Peremans et al., 2003; Vas et al., 2007; Wan et al., 2013), characteristics which are generally considered undesirable in both pet (Serpell, 1996; King et al., 2009) and service dogs (Weiss and Greenberg, 1997; Weiss, 2002), and are therefore important to measure in behavioural and personality tests used to assess the suitablity of dogs for these roles. However, impulsivity has received relatively little attention from canine behavioural test developers and no consensus has been reached as yet on the best method by which to assess impulsivity in dogs.

In people, factor analysis of multiple self-reports of impulsivity have indicated that it is likely made up of four discrete factors, related to traits from the FFM. According to Whiteside and Lynam (2001), the factors ‘premeditation’ and ‘perseverance’ relate to the FFM trait ‘conscientiousness’, reflecting variation in one’s ability to consider consequences prior to acting and to remain on task in distracting circumstances. The factor ‘sensation seeking’ is related to the FFM trait ‘extraversion’ and reflects a predisposition towards risk taking behaviours. Lastly, the ‘urgency’ factor is related to the FFM trait ‘neuroticism’, reflecting an inability to control impulses while experiencing a strong negative affective state.

While the four-factor model of impulsivity described by Whiteside and Lynam (2001) is commonly applied in people, impulsivity in non-human animals is most often defined on a very basic level, having only two components, one cognitive and one motor (Broos et al., 2012). Research into impulsivity in animals has historically focused on pharmacological research aimed at obtaining a greater understanding of the mechanisms underlying impulsivity in people (Fineberg et al., 2009). As reviewed by Dalley and Roiser (2012), the general consensus of research in this field is that the serotonergic and dopaminergic systems are intricately related to the expression of both cognitive and motor impulsivity (as assessed through a variety of tests thought to reflect each factor), but the exact mechanisms through which this occurs and how these variations relate to the four-factor model of impulsivity are still unclear.

Due to the social nature of dogs and the wide variation in impulsivity observed across breeds, dogs have been suggested as a superior model for the study of the evolution of impulsivity and inhibition (Vas et al., 2007; Lit et al., 2010). This has led to an increase in work using dogs to investigate the genetic basis of variation in the serotonergic and dopaminergic systems (Hejjas et al., 2007; Liinamo et al., 2007; van den Berg et al., 2008; Kubinyi et al., 2012; Lit et al., 2013; Wan et al., 2013), as well as the physiological basis of impulsive tendencies (Çakiroğlu et al., 2007; Riva et al., 2008; Rosado et al., 2010a, 2010b; Lit et al., 2013). Recently, however, interest in canine impulsivity has moved beyond that as a model for human research, into an appreciation for the role of impulsivity in canine aggression and welfare (Wright et al., 2011, 2012; Miller et al., 2012; Riemer et al., 2013). Wright and colleagues (2011) used an operant-choice, delayed-reinforcement paradigm and measurement of urinary serotonin and dopamine metabolites to validate an owner questionnaire to assess impulsivity in pet dogs, the Dog Impulsivity Assessment Survey (DIAS). Based on this work, overall impulsivity and behavioural regulation as assessed by DIAS were found to negatively correlate with tolerance of delayed rewards, urinary 5-hydroxyindoleacetic acid (5-HIAA):creatinine ratio and corrected 5-HIAA:homovanillic acid ratio, indicating a biological basis to a measureable trait reflecting impulse control in dogs (Wright et al., 2012). DIAS has shown good test–retest reliability at both 6 weeks (Wright et al., 2011) and 6 years (Riemer et al., 2013) and impulse control as measured by the maximum delay to reinforcement achieved by an individual during the operant task has shown significant test–retest reliability (Riemer et al., 2013), indicating that the dimensions assessed via these methods are highly stable over time in pet dogs.

While the relationships between impulsivity and existing personality models are still very unclear, based on the current conclusion that impulsivity is a multi-factorial concept which influences the behavioural expression of several personality traits (Whiteside and Lynam, 2001), impulsivity should perhaps be considered as an innate ‘modifier’ of the expression of temperament rather than a discrete personality trait. Both DIAS and the tolerance of delayed rewards task appear to be sound measures of assessing impulsivity in pet dogs; however, further work is required to assess their usefulness in a shelter situation.

4.2. Judgement bias and affective state

The assessment of emotion in non-human animals is a relatively new and highly complex field of research that generally falls in the domain of the welfare scientist, rather than the behavioural scientist. However, when considering the use of behavioural tests to assess personality traits in dogs chronically exposed to the stressors present in the shelter environment, it is important to consider how affective state (how one feels), cognition (how one

interprets and assimilates information) and behaviour (how one responds) interact, in order to understand the potential impact of environment on the validity and reliability of personality tests. Understanding how changes in affective state influence behaviour is especially important when behaviour test and scoring protocols developed and validated using owned dogs are subsequently applied to dogs housed in welfare shelters and pounds, who are chronically exposed to stressors such as high levels of noise (Sales et al., 1997), social isolation (Hennessy et al., 2002; Wells, 2004; Coppola et al., 2006), loss of routine and repeated exposure to novelty (Tuber et al., 1999). Positive and negative emotional states most likely affect behavioural test results differently based on context (for example a 'pessimistic' dog may be less likely to engage in play under test conditions than an 'optimistic' conspecific, but it may also be less likely to approach or engage a stimulus perceived as threatening), highlighting a need to investigate the interaction between affective state and test results before any specific conclusions are drawn. Furthermore, Burman et al. (2011) found that dogs achieved relatively slower approach and attempt scores in a judgement bias task, when tested immediately after a rewarding event involving food consumption, perhaps indicating a more complex relationship between competing motivations, arousal, satiety and affective state than earlier studies in other species would suggest.

Despite being a very young field of research, the body of work investigating the relationship judgement bias and affective state in domestic dogs is growing rapidly. Recent work in dogs has shown that negative judgement bias, thought to reflect pessimism, is associated with separation-related distress (Mendl et al., 2010) and correlates with a reduced probability of training success (Starling et al., 2013a) in pet dogs. Starling et al. (2014) suggest that population-level analysis of results of an operant task used to assess judgement bias in dogs could provide information on personality-based variations in optimism and risk aversion in domestic dogs. If this is indeed the case, judgement bias has potential to be applied for personality assessments as well as welfare assessments.

4.3. Arousal

As discussed by Koolhaas et al. (2010) and Kagan (1997), the quality of the behavioural response of an individual to stress is not a reliable indicator of the intensity of the stress experienced by an individual. As such, when examining responses to novelty and stress, it is necessary to consider individual variation in propensity to become aroused in response to particular stimuli separately to the behavioural patterns observed. Furthermore, the importance of a solid understanding of the effect of arousal on canine personality and in particular, on behaviours that are considered undesirable by many pet dog owners, is highlighted by the repeated appearance of the descriptors 'excitability', 'fearfulness/nervousness', 'anxiety' and 'aggression' in the literature pertaining to risk factors for relinquishment and poor owner-dog relationships (Jagoe and Serpell, 1996; Serpell, 1996; Kobelt et al., 2003; Bennett and Rohlf, 2007; Shore et al., 2008). While all of these factors have differing affective components, sustained or repeated high levels of arousal contribute substantially to all of them, indicating that reliable assessment of arousal in dogs is crucial to any test aiming to identify successful pet dogs.

In the broadest sense, arousal is defined as a generalised physiological response resulting in increased motor activity levels, increased vigilance and greater sensory processing that prepares an individual for action (Hebb, 1955; Weil et al., 2010). Early discussions of arousal, often called 'drive' by psychologists of the day, centred on the concept of 'wakefulness' and defined arousal as a state of being which allows an individual to respond quickly and appropriately to external stimuli by maintaining the body in a state of readiness (Hebb, 1955). As reviewed by Teigen (1994), Hebb (1955) was perhaps the earliest researcher to propose a relationship between arousal and 'cue function', a term describing goal-directed behaviour based on examination of sensory input, using the 'inverted U' relationship proposed by Yerkes and Dodson (1908) to explain the correlation between stimulus strength and habit formation. Hebb (1955, pp. 249) suggests that arousal "*serves to tone up the cortex with a background supporting action that is completely necessary if the messages proper* (cue functions) *are to have their effect*", going on to elaborate that neither extremely low arousal (deep sleep) or extremely high arousal (emotional disturbance or anxiety) are conducive to optimal performance. Easterbrook (1959) further refined the notion that arousal affects performance through changes in cue utilisation, suggesting that the range of cues utilised by an individual is reduced with increasing arousal and that the optimal level of arousal is that which allows all irrelevant cues to be discarded, without omitting those required for peak performance. This interpretation of the 'inverted U' relationship between arousal and performance has become widely known as the Yerkes-Dodson law (Teigen, 1994; Cohen, 2011) and gives a good indication of the likely effect of stressful environments on behavioural test results.

Whilst arousal is a physiological state rather than a personality trait, a predisposition to becoming easily aroused has been suggested to positively correlate with expression of the traits extraversion and neuroticism in humans (Gray, 1967; Strelau, 1972) and is the basis for the emotionality axis proposed by Koolhaas et al. (2010). Furthermore, chronic under-arousal resulting in increased reactivity to external stimuli has been proposed as a possible biological basis for impulsivity (Eysenck, 1993; Mathias and Stanford, 2003). Alternatively, it has been suggested that the effects of arousal on behaviour and personality are indirect and work by affecting attentional bias (Dickman, 2000), threat perception (Bar-Haim et al., 2007) or inhibition (Bari and Robbins, 2013). As the external expression of personality traits is directly affected by these cognitive processes, the assessment of animal personality and arousal as separate constructs will be fraught with problems until consensus can be reached on the degree of overlap and inter-connectedness between them. In order for this to be achieved, physiological measures of arousal should feature in all tests of animal personality to allow examination of each concept individually and together.

The relationship between arousal and other personality and cognitive constructs becomes increasingly difficult to assess when one recognises that, like impulsivity, despite often being treated as a single continuum by animal researchers, arousal appears to be made up of at least two related factors (Hebb, 1955; Thayer, 1978). Thayer (1978) identified a two-factor 'activation' model containing an energy–tiredness axis, which is reflective of motor activity and vigour, and a tension–placidity axis which reflects anxiety and irritability. The extreme ends of each axis are thought to function in opposition to one another, leading Thayer (1978) to suggest that there may be four separate but interacting pathways through which arousal is controlled, each with a distinct affective element. Experimental evidence supports the existence of the tense-arousal and energetic-arousal model in people, also suggesting complex interactions between the two types of arousal and the attentional bias theory of impulsivity (Dickman, 2000). Schimmack and Grob (2000) extended the integration of affect and arousal by suggesting a three-dimensional variation of the Thayer model, in which pleasure–displeasure, wakefulness–tiredness and tension–relaxation co-vary. As discussed earlier, assessing affective state in animals is an evolving area of research requiring complex testing protocols. It is perhaps due to the design limitations caused by these difficulties that studies on arousal in animals have focused on the physiological and behavioural changes that indicate increased alertness, motor activity and emotional reactivity and therefore suggest the existence of a single, generalised arousal dimension that is both measureable and heritable (Weil et al., 2010; Quinkert et al., 2011).

4.4. Temporal and contextual stability

While the assessment of the temporal stability of responses via test–retest reliability measures is held as the gold standard for assessing personality traits (Taylor and Mills, 2006), in keeping with the psychological definition of trait, it is also important to assess cross-context reaction norms in order to differentiate between true traits and context-specific behaviours (Stamps and Groothuis, 2010). This is achieved by the inclusion of several subtests designed to assess the same trait across different contexts and distinguish between variation caused by situational or motivational factors and that which truly reflects personality differences (Stamps and Groothuis, 2010). Furthermore, when assessing the temporal or contextual stability of a trait, stability refers to the maintenance of the rank order of individuals (i.e. consistent individual position relative to other individuals), rather than consistency in the actual intensity of the response or the expression of specific behaviours (Sih et al., 2004b; Réale et al., 2007). While the failure of many test developers to assess test–retest reliability has been addressed previously (Jones and Gosling, 2005; Mornement et al., 2009, 2010), the need for repeated measures of the same trait across several contexts and for measures of between individual variation, such as rank order comparisons, appear to have largely been ignored by test developers. The environment in which some behavioural tests are used to assess personality in an applied manner, such as welfare facilities, and working or sporting tests, imposes limitations on the number and variation of contexts available in which to assess behaviour. In these cases it is important to interpret all test results in light of the design limitations of such tests, specifically through understanding the effects of the specific context on behavioural reactions. Where possible, it is also prudent to consider methods with which to gather as much information outside of the test environment, in as many contexts as possible, to aid in the interpretation of test results.

4.5. Behavioural test scoring methods and identification of passively responding individuals

A common theme anecdotally acknowledged by those involved with canine behaviour assessment is that individuals vary widely in their ability to make accurate and reliable assessments of canine personality, and that this variance is only partially explained by their level of training and experience with dogs (Tami and Gallagher, 2009). Some individuals are seemingly able to make reasonably accurate assessments of a dog's personality based on a relatively small set of interactions, but protocols developed around these interactions and assessments often do not transfer well when used by others. While inter-rater reliability is not commonly reported in the test literature (Taylor and Mills, 2006; Mornement et al., 2009), reported reliabilities indicate that these inconsistencies appear to be greatest for assessments of passive behaviours (Valsecchi et al., 2009; Sinn et al., 2010; Mornement et al., 2014). One possible explanation for this discrepancy is that 'successful' assessors utilise hard to define measures of arousal or stress when making judgements of behaviour, such as the degree of 'tense energy' displayed, the relative degree of interest or disinterest in environmental stimuli, or subtle facial changes and body language indicators of fear or tension, such as pupil dilation, changes in posture height or tail wagging and changes to the shape of the mouth. These subtle behaviours are often overlooked by owners and dog enthusiasts (Kerswell et al., 2009; Mariti et al., 2012) and inclusion of these indicators into test protocols is problematic, as accurate descriptions are by necessity, detailed and information-heavy, and therefore, are not easily communicated to an uneducated or naïve audience. Furthermore, the degree of relevant experience and education varies widely between assessors in a shelter or welfare facility (Mornement et al., 2010) and resources such as time and space are highly limited (Mornement et al., 2014), so tests must be quick and easy to administer with minimal resources required, in order to be feasible in these environments.

It is perhaps due to these limitations that many test protocols focus largely on gross, easily observed behaviours that are common indicators of fear, such as flight, and obvious signs of aggression such as growling, lunging and trying to bite (Netto and Planta, 1997; Svartberg and Forkman, 2002; Kroll et al., 2004; Lucidi et al., 2005; Saetre et al., 2006; Planta and De Meester, 2007; Bräm et al., 2008; Haverbeke et al., 2009; Mornement et al., 2009; Bergamasco et al., 2010; Sinn et al., 2010; van der Borg et al., 2010; Bennett et al., 2012; Duffy and Serpell, 2012).

This methodological bias towards measurement of active responses is also prevalent in questionnaires designed to assess dog behaviour through owner reports, such as the Canine Behavioural Assessment and Research Questionnaire (C-BARQ) (Serpell and Hsu, 2001; Hsu and Serpell, 2003; Konok et al., 2011). The degree to which this limitation affects test results depends largely on the purpose of the test, be that the detection and assessment of potentially problematic gross behaviours within a particular set of circumstances, or the assessment of canine personality in an attempt to make predictions about future behaviour across contexts and time. Tests which only assess gross, active behaviours and are aimed at assessing personality or predicting the potential for future aggressive or fearful reactions, rely on the incorrect assumption that individuals displaying passive behaviour are not experiencing considerable stress or arousal and will suffer from poor predictability due to these inherent design flaws. Some support for this idea is shown by the degree to which measures of passive behaviour increase the accuracy of tests; battery tests scored primarily on the presence/absence of active responses show lower levels of concurrent validity or predictability validity for behaviours related to aggression and fear (Bennett et al., 2012; Mornement et al., 2014), than those that include measures of subtle, passive behaviours that are thought to be indicative of stress (Barnard et al., 2012; Kuhne et al., 2014). Given the methodological limitations inherent to assessments carried out in shelters and welfare facilities, and to a degree in other areas in which canine personality tests are used, it appears that behaviour-only tests will always suffer from validity and reliability issues due to the trade-off between the high level of detail required to examine all relevant behaviours and making sure tests are both transferable and feasible.

In 2006, Taylor and Mills suggested that physiological measures of arousal in dogs may serve to increase the validity of behaviour tests, as they are objective, easy to measure and can provide insight into an animal's innate or default reactivity to stressful stimuli. A small number of tests have been carried out in dogs and calves to compare the accuracy of behavioural assessments with heart rate and circulating cortisol analysis for assessing the stress response. Vincent and Leahy (1997) found that heart rate profiles in guide dog candidates during training walks reflected their typical arousal levels, as reported by their handlers. Both heart rate profiles and temperament descriptions remained stable over several months, despite the dogs being adolescent and undertaking training during the time between tests, indicating that the dimension being described was temporally and contextually consistent. Van Reenen et al. (2004) observed that behavioural measures of activity level, vocalisations and interactions with novel objects and people were largely independent of each other and not reflective of the overall fearfulness of dairy calves, despite being internally stable over several months after the age of 16 weeks. Beerda et al. (1998) conducted a similar study, by presenting pet dogs with potentially fear-inducing stimuli and found that heart rate and cortisol measures did not correlate with typical behavioural measures of stress. Unlike Van Reenen et al. (2004), Beerda et al. (1998) did not find a correlation between heart rate during stimulus presentation and post-test cortisol measures; however, Beerda et al. (1998) did not correct for the effect of activity level on heart rate, which is a source of significant noise for heart rate measures. Palestrini et al. (2005) found a positive correlation between heart rate and activity level in dogs, but also noted significantly increased heart rate during isolation tests characterised by static behaviour. This was interpreted as an indication of increased stress levels during these times. Most recently, Kuhne et al. (2014) noted that cardiac activity measures, including heart rate and two measures of heart rate variability, correlated negatively with the presence of displacement behaviours and positively with appeasement gestures in pet dogs during static handling tests. Based on these results, it appears that a robust method of correcting heart rate measures for activity level is required before heart rate variability can be considered a valid tool for assessing the stress response. Until this is achieved, it appears that both heart rate and circulating cortisol measures should be used together to provide an accurate physiological measure of arousal levels, and in conjunction with behavioural analysis to determine typical coping responses.

5. Where to from here?

In order for a productive discussion of the application and usefulness of canine personality tests in shelters to take place, some consensus must first be reached on both the use of terminology and methodology in these tests and the aims of those using them. It is also imperative that those using and developing tests have a sound understanding of the limitations inherent to current testing protocols and the environment in which many tests are conducted. While coping behaviours and stress responses have been the most widely studied personality structures in dogs, the core aim of shelter-based pet dog assessments is to correctly describe a dog's entire personality, to enable effective risk identification and management, and to ensure good owner–pet matching (Marston and Bennett, 2003). In order to meet this aim, a 'whole-personality' assessment needs to be developed for use in welfare shelters. Shelter environments are highly stressful for dogs (Hiby et al., 2006; Rooney et al., 2007; Menor-Campos et al., 2011), limiting the available contexts in which behaviour can be assessed to those which include a component of novelty and environmentally- or socially-induced stress. Bearing this in mind, it is not surprising that measures related to fear- and anxiety-based responses show the greatest validity and predictability in existing shelter tests (van der Borg et al., 1991; Dowling-Guyer et al., 2011; Valsecchi et al., 2011; Barnard et al., 2012; Bennett et al., 2012). Furthermore, due to necessary limitations on the detail that can be included in battery test protocols and behavioural scoring systems, current tests often rely on the incorrect assumption that the presence of active behaviours, such as flight or aggression, provides a good measure of arousal in all dogs. As suggested by Sih et al. (2004b), Réale et al. (2007) and Koolhaas et al. (2007), a clear distinction must be made between the quality of a behavioural response, measured by behavioural analysis, and the level of arousal experienced by the animal, best measured via physiological variables such as heart rate

and circulating levels of catecholamines or glucocorticoids. This will both reduce the subjectivity of behavioural test measures and increase the amount of information gathered about an individual, leading to more accurate assessments of an individual's response to test stimuli and providing insight into the effect of stress on the reliability of test results. Further, this increased accuracy will allow 'outliers' to be identified and removed from the sample group, enabling test parameters to be defined more effectively.

Specific tests that are able to separate context-specific reactions, motivations, generalised arousal or emotionality and the normal response of an individual to stressful or threatening situations would be by requirement, highly complex and time consuming, resulting in poor feasibility for use in a shelter environment or companion dog situation. This is perhaps where the mixed assessment model proposed by Marder et al. (2013) for the assessment of shelter dogs, in which both in-shelter assessments by staff and test battery results are combined, has a unique advantage. With further development, this model could allow for the collection of salient information about behaviours that are difficult or impossible to assess via test battery, across a number of contexts and by several individuals. This information could then be collated with the results of combined behaviour and physiological assessments of stress reactivity and response to novelty during battery tests, giving a more robust and complete picture of individual personality. In order for such a model to work efficiently and reliably, a protocol would need to be developed and tested, which identifies all of the concepts to be assessed and the benefits and limitations of each assessment method for each concept, so that conflicts of information can be reviewed and easily dealt with, without compromising on the reliability and validity of the information.

References

Arvelius, P., Eken Asp, H., Fikse, W.F., Strandberg, E., Nilsson, K., 2014. Genetic analysis of a temperament test as a tool to select against everyday life fearfulness in Rough Collie. J. Anim. Sci., http://dx.doi.org/10.2527/jas.2014-8169.

Bahlig-Pieren, Z., Turner, D.C., 1999. Anthropomorphic interpretations and ethological descriptions of dog and cat behavior by lay people. Anthrozoos 12, 205–210, http://dx.doi.org/10.2752/089279399787000075.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M.J., van IJzendoorn, M.H., 2007. Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. Psychol. Bull. 133, 1–24, http://dx.doi.org/10.1037/0033-2909.133.1.1.

Bari, A., Robbins, T.W., 2013. Inhibition and impulsivity: behavioral and neural basis of response control. Prog. Neurobiol. 108, 44–79, http://dx.doi.org/10.1016/j.pneurobio.2013.06.005.

Barnard, S., Siracusa, C., Reisner, I., Valsecchi, P., Serpell, J.A., 2012. Validity of model devices used to assess canine temperament in behavioural tests. Appl. Anim. Behav. Sci. 138, 79–87, http://dx.doi.org/10.1016/j.applanim.2012.02.017.

Barrera, G., Jakovcevic, A., Elgier, A.M., Mustaca, A., Bentosela, M., 2010. Responses of shelter and pet dogs to an unknown human. J. Vet. Behav. 5, 339–344, http://dx.doi.org/10.1016/j.jveb.2010.08.012.

Bartholomew, D., 2012. Officials to debate temperament testing for shelter dogs, cats. Los Angeles Daily News, Los Angeles. http://www.dailynews.com/20121202/officials-to-debate-temperament-testing-for-shelter-dogs-cats (retrieved 5/12/2014).

Beckmann, C., Biro, P.A., 2013. On the validity of a single (boldness) assay in personality research. Ethology 119, 937–947, http://dx.doi.org/10.1111/eth.12137.

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, http://dx.doi.org/10.1016/s0168-1591(97)00145-7.

Bell, A.M., 2007. Future directions in behavioural syndromes research. Proc. R. Soc. B: Biol. Sci. 274, 755–761, http://dx.doi.org/10.1098/rspb.2006.0199.

Bennett, P.C., Rohlf, V.I., 2007. Owner-companion dog interactions: relationships between demographic variables, potentially problematic behaviours, training engagement and shared activities. Appl. Anim. Behav. Sci. 102, 65–84, http://dx.doi.org/10.1016/j.applanim.2006.03.009.

Bennett, S.L., Litster, A., Weng, H.-Y., Walker, S.L., Luescher, A.U., 2012. Investigating behavior assessment instruments to predict aggression in dogs. Appl. Anim. Behav. Sci. 141, 139–148, http://dx.doi.org/10.1016/j.applanim.2012.08.005.

Benus, R.F., Bohus, B., Koolhaas, J.M., van Oortmerssen, G.A., 1989. Behavioural strategies of aggressive and non-aggressive male mice in active shock avoidance. Behav. Process. 20, 1–12, http://dx.doi.org/10.1016/0376-6357(89)90008-9.

Benus, R.F., Bohus, B., Koolhaas, J.M., van Oortmerssen, G.A., 1990. Behavioural strategies of aggressive and non-aggressive male mice in response to inescapable shock. Behav. Process. 21, 127–141, http://dx.doi.org/10.1016/0376-6357(90)90020-g.

Benus, R.F., Bohus, B., Koolhaas, J.M., van Oortmerssen, G.A., 1991. Heritable variation for aggression as a reflection of individual coping strategies. Experientia 47, 1008–1019, http://dx.doi.org/10.1007/bf01923336.

Benus, R.F., Koolhaas, J.M., van Oortmerssen, G.A., 1987. Individual differences in behavioural reaction to a changing environment in mice and rats. Behaviour 100, 105–122, http://dx.doi.org/10.2307/4534578.

Bergamasco, L., Osella, M.C., Savarino, P., Larosa, G., Ozella, L., Manassero, M., Badino, P., Odore, R., Barbero, R., Re, G., 2010. Heart rate variability and saliva cortisol assessment in shelter dog: human–animal interaction effects. Appl. Anim. Behav. Sci. 125, 56–68, http://dx.doi.org/10.1016/j.applanim.2010.03.002.

Biro, P.A., Beckmann, C., Stamps, J.A., 2010. Small within-day increases in temperature affects boldness and alters personality in coral reef fish. Proc. R. Soc. B: Biol. Sci. 277, 71–77, http://dx.doi.org/10.1098/rspb.2009.1346.

Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B., Keeling, L.J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., 2007. Assessment of positive emotions in animals to improve their welfare. Physiol. Behav. 92, 375–397, http://dx.doi.org/10.1016/j.physbeh.2007.02.003.

Bolhuis, J.E., Schouten, W.G., de Leeuw, J.A., Schrama, J.W., Wiegant, V.M., 2004. Individual coping characteristics, rearing conditions and behavioural flexibility in pigs. Behav. Brain Res. 152, 351–360, http://dx.doi.org/10.1016/j.bbr.2003.10.024.

Bräm, M., Doherr, M.G., Lehmann, D., Mills, D., Steiger, A., 2008. Evaluating aggressive behavior in dogs: a comparison of 3 tests. J. Vet. Behav. 3, 152–160, http://dx.doi.org/10.1016/j.jveb.2008.04.001.

Broos, N., Schmaal, L., Wiskerke, J., Kostelijk, L., Lam, T., Stoop, N., Weierink, L., Ham, J., de Geus, E.J., Schoffelmeer, A.N., 2012. The relationship between impulsive choice and impulsive action: a cross-species translational study. PLoS ONE 7, e36781, http://dx.doi.org/10.1371/journal.pone.0036781.

Bruch, M.A., Gorsky, J.M., Collins, T.M., Berger, P.A., 1989. Shyness and sociability reexamined: a multicomponent analysis. J. Pers. Soc. Psychol. 57, 904, http://dx.doi.org/10.1037//0022-3514.57.5.904.

Burman, O., McGowan, R., Mendl, M., Norling, Y., Paul, E., Rehn, T., Keeling, L., 2011. Using judgement bias to measure positive affective state in dogs. Appl. Anim. Behav. Sci. 132, 160–168, http://dx.doi.org/10.1016/j.applanim.2011.04.001.

Çakiroğlu, D., Meral, Y., Sancak, A.A., Çifti, G., 2007. Relationship between the serum concentrations of serotonin and lipids and aggression in dogs. Vet. Rec. 161, 59–61, http://dx.doi.org/10.1136/vr.161.2.59.

Carere, C., van Oers, K., 2004. Shy and bold great tits (*Parus major*): body temperature and breath rate in response to handling stress. Physiol. Behav. 82, 905–912, http://dx.doi.org/10.1016/j.physbeh.2004.07.009.

Carter, A.J., Feeney, W.E., 2012. Taking a comparative approach: analysing personality as a multivariate behavioural response across species. PLoS ONE 7, 1–9, http://dx.doi.org/10.1371/journal.pone.0042440.

Carter, A.J., Feeney, W.E., Marshall, H.H., Cowlishaw, G., Heinsohn, R., 2013. Animal personality: what are behavioural ecologists measuring? Biol. Rev. 88, 465–475, http://dx.doi.org/10.1111/brv.12007.

Cheek, J.M., Buss, A.H., 1981. Shyness and sociability. J. Pers. Soc. Psychol. 41, 330, http://dx.doi.org/10.1037//0022-3514.41.2.330.

Clark, C.C., Gruffydd-Jones, T., Murray, J.K., 2012. Number of cats and dogs in UK welfare organisations. Vet. Rec. 170, 493, http://dx.doi.org/10.1136/vr.100524.

Cohen, R.A., 2011. Yerkes–Dodson law. In: Kreutzer, J.S., DeLuca, J., Capla, B. (Eds.), Encyclopedia of Clinical Neuropsychology. Springer, New York, pp. 2737–2738, http://dx.doi.org/10.1007/978-0-387-79948-3_1340.
Coleman, K., Wilson, D.S., 1998. Shyness and boldness in pumpkinseed sunfish: individual differences are context-specific. Anim. Behav. 56, 927–936, http://dx.doi.org/10.1006/anbe.1998.0852.
Coplan, R.J., Prakash, K., O'Neil, K., Armer, M., 2004. Do you" want" to play? Distinguishing between conflicted shyness and social disinterest in early childhood. Dev. Psychol. 40, 244, http://dx.doi.org/10.1037/0012-1649.40.2.244.
Coppola, C.L., Grandin, T., Enns, R.M., 2006. Human interaction and cortisol: can human contact reduce stress for shelter dogs? Physiol. Behav. 87, 537–541, http://dx.doi.org/10.1016/j.physbeh.2005.12.001.
Courreau, J.-F., Langlois, B., 2005. Genetic parameters and environmental effects which characterise the defence ability of the Belgian shepherd dog. Appl. Anim. Behav. Sci. 91, 233–245, http://dx.doi.org/10.1016/j.applanim.2004.09.003.
Crazier, W., 1979. Shyness as a dimension of personality. Br. J. Soc. Clin. Psychol. 18, 121–128, http://dx.doi.org/10.1111/j.2044-8260.1979.tb00314.x.
Dalley, J.W., Roiser, J.P., 2012. Dopamine, serotonin and impulsivity. Neuroscience 215, 42–58, http://dx.doi.org/10.1016/j.neuroscience.2012.03.065.
de Boer, S.F., Koolhaas, J.M., 2005. 5-HT1A and 5-HT1B receptor agonists and aggression: a pharmacological challenge of the serotonin deficiency hypothesis. Eur. J. Pharmacol. 526, 125–139, http://dx.doi.org/10.1016/j.ejphar.2005.09.065.
De Meester, R.H., De Bacquer, D., Peremans, K., Vermeire, S., Planta, D.J., Coopman, F., Audenaert, K., 2008. A preliminary study on the use of the socially acceptable behavior test as a test for shyness/confidence in the temperament of dogs. J. Vet. Behav. 3, 161–170, http://dx.doi.org/10.1016/j.jveb.2007.10.005.
Dickman, S.J., 2000. Impulsivity, arousal and attention. Pers. Individ. Differ. 28, 563–581, http://dx.doi.org/10.1016/S0191-8869(99)00120-8.
Dochtermann, N.A., Jenkins, S.H., 2007. Behavioural syndromes in Merriam's kangaroo rats (*Dipodomys merriami*): a test of competing hypotheses. Proc. R. Soc. Lond. B: Biol. Sci. 274, 2343–2349, http://dx.doi.org/10.1098/rspb.2007.0622.
Dowling-Guyer, S., Marder, A., D'Arpino, S., 2011. Behavioral traits detected in shelter dogs by a behavior evaluation. Appl. Anim. Behav. Sci. 120, 107–114, http://dx.doi.org/10.1016/j.appanim.2010.12.004.
Draper, T.W., 1995. Canine analogs of human personality factors. J. Gen. Psychol. 122, 241–252, http://dx.doi.org/10.1080/00221309.1995.9921236.
Drent, P.J., Oers, K.V., Noordwijk, A.J.V., 2003. Realized heritability of personalities in the great tit (*Parus major*). Proc. R. Soc. Lond. B: Biol. Sci. 270, 45–51, http://dx.doi.org/10.1098/rspb.2002.2168.
Duffy, D.L., Serpell, J.A., 2012. Predictive validity of a method for evaluating temperament in young guide and service dogs. Appl. Anim. Behav. Sci. 138, 99–109, http://dx.doi.org/10.1016/j.applanim.2012.02.011.
Early, J.B., Arnott, E., Wade, C.M., McGreevy, P.D., 2014. Manual muster: a critical analysis of the use of common terms in Australian working dog manuals. J. Vet. Behav., http://dx.doi.org/10.1016/j.jveb.2014.07.003.
Easterbrook, J.A., 1959. The effect of emotion on cue utilization and the organization of behavior. Psychol. Rev. 66, 183, http://dx.doi.org/10.1037/h0047707.
Ebeling-Witte, S., Frank, M.L., Lester, D., 2007. Shyness, Internet use, and personality. Cyberpsychol. Behav. 10, 713–716, http://dx.doi.org/10.1089/cpb.2007.9964.
Evenden, J.L., 1999. Varieties of impulsivity. Psychopharmacology (Berl) 146, 348–361, http://dx.doi.org/10.1007/PL00005481.
Eysenck, H.J., 1993. The nature of impulsivity. In: McCown, W.G., Johnson, J.L., Shure, M.B. (Eds.), The Impulsive Client: Theory, Research, and Treatment. American Psychological Association, Washington, DC, pp. 57–69, http://dx.doi.org/10.1037/10500-004.
Fidler, M., Light, P., Costall, A., 1996. Describing dog behavior psychologically: pet owners versus non-owners. Anthrozoos 9, 196–200, http://dx.doi.org/10.2752/089279396787001356.
Fineberg, N.A., Potenza, M.N., Chamberlain, S.R., Berlin, H.A., Menzies, L., Bechara, A., Sahakian, B.J., Robbins, T.W., Bullmore, E.T., Hollander, E., 2009. Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. Neuropsychopharmacology 35, 591–604, http://dx.doi.org/10.1038/npp.2009.185.
Fratkin, J.L., Sinn, D.L., Patall, E.A., Gosling, S.D., 2013. Personality consistency in dogs: a meta-analysis. PLOS ONE 8, 1–19, http://dx.doi.org/10.1371/journal.pone.0054907.
Fratkin, J.L., Sinn, D.L., Thomas, S., Hilliard, S., Olson, Z., Gosling, S.D., 2014. Do you see what I see? Can non-experts with minimal training reproduce expert ratings in behavioral assessments of working dogs? Behav. Process., http://dx.doi.org/10.1016/j.beproc.2014.09.028.
Freeman, H.D., Gosling, S.D., 2010. Personality in nonhuman primates: a review and evaluation of past research. Am. J. Primatol. 72, 653–671, http://dx.doi.org/10.1002/ajp.20833.
Frost, A.J., Winrow-Giffen, A., Ashley, P.J., Sneddon, L.U., 2007. Plasticity in animal personality traits: does prior experience alter the degree of boldness? Proc. R. Soc. B: Biol. Sci. 274, 333–339, http://dx.doi.org/10.1098/rspb.2006.3751.
Goldberg, L.R., 1990. An alternative "description of personality": the Big-Five factor structure. J. Pers. Soc. Psychol. 59, 1216–1229, http://dx.doi.org/10.1037//0022-3514.59.6.1216.
Gosling, S.D., John, O.P., 1999. Personality dimensions in nonhuman animals: a cross-species review. Curr. Dir. Psychol. Sci. 8, 69–75, http://dx.doi.org/10.1111/1467-8721.00017.
Gosling, S.D., Kwan, V.S., John, O.P., 2003. A dog's got personality: a cross-species comparative approach to personality judgments in dogs and humans. J. Pers. Soc. Psychol. 85, 1161–1169, http://dx.doi.org/10.1037/0022-3514.85.6.1161.
Gosling, S.D., 2008. Personality in non-human animals. Soc. Pers. Psychol. Compass 2, 985–1001, http://dx.doi.org/10.1111/j.1751-9004.2008.00087.x.
Gray, J.A., 1967. Strength of the nervous system, introversion – extraversion, conditionability and arousal. Behav. Res. Ther. 5, 151–169, http://dx.doi.org/10.1016/0005-7967(67)90031-9.
Haverbeke, A., De Smet, A., Depiereux, E., Giffroy, J.-M., Diederich, C., 2009. Assessing undesired aggression in military working dogs. Appl. Anim. Behav. Sci. 117, 55–62, http://dx.doi.org/10.1016/j.applanim.2008.12.002.
Hebb, D.O., 1955. Drives and the CNS (conceptual nervous system). Psychol. Rev. 62, 243–254.
Hejjas, K., Vas, J., Topal, J., Szantai, E., Ronai, Z., Szekely, A., Kubinyi, E., Horváth, Z., Sasvari-Szekely, M., Miklosi, A., 2007. Association of polymorphisms in the dopamine D4 receptor gene and the activity–impulsivity endophenotype in dogs. Anim. Genet. 38, 629–633, http://dx.doi.org/10.1111/j.1365-2052.2007.01657.x.
Hennessy, M.B., Voith, V.L., Young, T.L., Hawke, J.L., Centrone, J., McDowell, A.L., Linden, F., Davenport, G.M., 2002. Exploring human interaction and diet effects on the behavior of dogs in a public animal shelter. J. Appl. Anim. Welf. Sci. 5, 253–273.
Hiby, E.F., Rooney, N.J., Bradshaw, J.W.S., 2006. Behavioural and physiological responses of dogs entering rehoming kennels. Physiol. Behav. 89, 385–391, http://dx.doi.org/10.1016/j.physbeh.2006.07.012.
Horváth, Z., Igyártó, B.-Z., Magyar, A., Miklósi, Á., 2007. Three different coping styles in police dogs exposed to a short-term challenge. Horm. Behav. 52, 621–630, http://dx.doi.org/10.1016/j.yhbeh.2007.08.001.
Hsu, Y., Serpell, J.A., 2003. Development and validation of a questionnaire for measuring behaviour and temperament traits in pet dogs. J. Am. Vet. Med. Assoc. 223, 1293–1300, http://dx.doi.org/10.2460/javma.2003.223.1293.
Itoh, K., 2002. Personality research with non-human primates: theoretical formulation and methods. Primates 43, 249–261, http://dx.doi.org/10.1007/BF02629652.
Jagoe, A., Serpell, J., 1996. Owner characteristics and interactions and the prevalence of canine behaviour problems. Appl. Anim. Behav. Sci. 47, 31–42, http://dx.doi.org/10.1016/0168-1591(95)01008-4.
Jones, A.C., Gosling, S.D., 2005. Temperament and personality in dogs (*Canis familiaris*): a review and evaluation of past research. Appl. Anim. Behav. Sci. 95, 1–53, http://dx.doi.org/10.1016/j.applanim.2005.04.008.
Kagan, J., 1997. Temperament and the reactions to unfamiliarity. Child Dev. 68, 139–143, http://dx.doi.org/10.2307/1131931.
Kagan, J., Reznick, J.S., Snidman, N., 1988. Biological bases of childhood shyness. Science 240, 167–171, http://dx.doi.org/10.1126/science.3353713.
Kagan, J., Snidman, N., 1990. Temperamental contributions to human development: the biological characteristics of infants influence their initial behavior to unfamiliar contexts, Research & Clinical Center for Child Development: Annual Report. Faculty of Education, Hokkaido University, pp. 59–70.
Kerswell, K.J., Bennett, P., Butler, K.L., Hemsworth, P.H., 2009. Self-reported comprehension ratings of dog behavior by puppy owners. Anthrozoos 22, 183–193, http://dx.doi.org/10.2752/175303709x434202.
King, T., Marston, L.C., Bennett, P.C., 2009. Describing the ideal Australian companion dog. Appl. Anim. Behav. Sci. 120, 84–93, http://dx.doi.org/10.1016/j.applanim.2009.04.011.
Kobelt, A.J., Hemsworth, P.H., Barnett, J.L., Coleman, G.J., 2003. A survey of dog ownership in suburban Australia—conditions

and behaviour problems. Appl. Anim. Behav. Sci. 82, 137–148, http://dx.doi.org/10.1016/S0168-1591(03)00062-5.
Konok, V., Dóka, A., Miklósi, Á., 2011. The behavior of the domestic dog (*Canis familiaris*) during separation from and reunion with the owner: a questionnaire and an experimental study. Appl. Anim. Behav. Sci. 135, 300–308, http://dx.doi.org/10.1016/j.applanim.2011.10.011.
Koolhaas, J., De Boer, S., Buwalda, B., Van Reenen, K., 2007. Individual variation in coping with stress: a multidimensional approach of ultimate and proximate mechanisms. Brain Behav. Evol. 70, 218, http://dx.doi.org/10.1159/000105485.
Koolhaas, J.M., de Boer, S.F., Coppens, C.M., Buwalda, B., 2010. Neuroendocrinology of coping styles: towards understanding the biology of individual variation. Front. Neuroendocrinol. 31, 307–321, http://dx.doi.org/10.1016/j.yfrne.2010.04.001.
Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H., De Jong, I.C., Ruis, M.A.W., Blokhuis, H.J., 1999. Coping styles in animals: current status in behavior and stress-physiology. Neurosci. Biobehav. Rev. 23, 925–935, http://dx.doi.org/10.1016/s0149-7634(99)00026-3.
Kroll, T.L., Houpt, K.A., Erb, H.N., 2004. The use of novel stimuli as indicators of aggressive behavior in dogs. J. Am. Anim. Hosp. Assoc. 40, 13–19.
Kubinyi, E., Turcsán, B., Miklósi, Á., 2009. Dog and owner demographic characteristics and dog personality trait associations. Behav. Process. 81, 392–401, http://dx.doi.org/10.1016/j.beproc.2009.04.004.
Kubinyi, E., Vas, J., Hejjas, K., Ronai, Z., Bruder, I., Turcsan, B., Sasvari-Szekely, M., Miklosi, A., 2012. Polymorphism in the tyrosine hydroxylase (TH) gene is associated with activity–impulsivity in German Shepherd dogs. PLoS ONE 7, 1–6, http://dx.doi.org/10.1371/journal.pone.0030271.
Kuhne, F., Hößler, J.C., Struwe, R., 2014. Behavioral and cardiac responses by dogs to physical human–dog contact. J. Vet. Behav., http://dx.doi.org/10.1016/j.jveb.2014.02.006.
Ley, J., Bennett, P., Coleman, G., 2008. Personality dimensions that emerge in companion canines. Appl. Anim. Behav. Sci. 110, 305–317, http://dx.doi.org/10.1016/j.applanim.2007.04.016.
Ley, J.M., Bennett, P.C., 2007. Understanding personality by understanding companion dogs. Anthrozoos 20, 113–124, http://dx.doi.org/10.2752/175303707x207909.
Liinamo, A.-E., van den Berg, L., Leegwater, P.A.J., Schilder, M.B.H., van Arendonk, J.A.M., van Oost, B.A., 2007. Genetic variation in aggression-related traits in Golden Retriever dogs. Appl. Anim. Behav. Sci. 104, 95–106, http://dx.doi.org/10.1016/j.applanim.2006.04.025.
Lindberg, S., Strandberg, E., Swenson, L., 2004. Genetic analysis of hunting behaviour in Swedish Flatcoated Retrievers. Appl. Anim. Behav. Sci. 88, 289–298, http://dx.doi.org/10.1016/j.applanim.2004.03.007.
Lit, L., Belanger, J.M., Boehm, D., Lybarger, N., Haverbeke, A., Diederich, C., Oberbauer, A.M., 2013. Characterization of a dopamine transporter polymorphism and behavior in Belgian Malinois. BMC Genet. 14, 1–11, http://dx.doi.org/10.1186/1471-2156-14-45.
Lit, L., Schweitzer, J.B., Iosif, A.M., Oberbauer, A.M., 2010. Owner reports of attention, activity, and impulsivity in dogs: a replication study. Behav. Brain Funct. 6, 1–10, http://dx.doi.org/10.1186/1744-9081-6-1.
Lowe, S.E., Bradshaw, J.W., 2001. Ontogeny of individuality in the domestic cat in the home environment. Anim. Behav. 61, 231–237, http://dx.doi.org/10.1006/anbe.2000.1545.
Lucidi, P., Bernabò, N., Panunzi, M., Villa, P.D., Mattioli, M., 2005. Ethotest: a new model to identify (shelter) dogs' skills as service animals or adoptable pets. Appl. Anim. Behav. Sci. 95, 103–122, http://dx.doi.org/10.1016/j.applanim.2005.04.006.
Marder, A.R., Shabelansky, A., Patronek, G.J., Dowling-Guyer, S., D'Arpino, S.S., 2013. Food-related aggression in shelter dogs: a comparison of behavior identified by a behavior evaluation in the shelter and owner reports after adoption. Appl. Anim. Behav. Sci. 148, 150–156, http://dx.doi.org/10.1016/j.applanim.2013.07.007.
Mariti, C., Gazzano, A., Moore, J.L., Baragli, P., Chelli, L., Sighieri, C., 2012. Perception of dogs' stress by their owners. J. Vet. Behav. 7, 213–219, http://dx.doi.org/10.1016/j.jveb.2011.09.004.
Marston, L.C., Bennett, P.C., 2003. Reforging the bond – towards successful canine adoption. Appl. Anim. Behav. Sci. 83, 227–245, http://dx.doi.org/10.1016/S0168-1591(03)00135-7.
Marston, L.C., Bennett, P.C., Coleman, G.J., 2004. What happens to shelter dogs? An analysis of data for 1 year from three Australian shelters. J. Appl. Anim. Welf. Sci. 7, 27–47.
Mathias, C.W., Stanford, M.S., 2003. Impulsiveness and arousal: heart rate under conditions of rest and challenge in healthy males. Pers. Individ. Differ. 35, 355–371, http://dx.doi.org/10.1016/S0191-8869(02)00195-2.
McCrae, R.R., Costa, P.T., Del Pilar, G.H., Rolland, J.-P., Parker, W.D., 1998. Cross-cultural assessment of the Five-Factor Model the revised NEO Personality Inventory. J. Cross Cult. Psychol. 29, 171–188.
McCrae, R.R., John, O.P., 1992. An introduction to the Five-Factor Model and its applications. J. Pers. 60, 175–215.
McDermott, D.R., Chips, M.J., McGuirk, M., Armagost, F., DiRienzo, N., Pruitt, J.N., 2014. Boldness is influenced by sublethal interactions with predators and is associated with successful harem infiltration in Madagascar hissing cockroaches. Behav. Ecol. Sociobiol. 68, 425–435, http://dx.doi.org/10.1007/s00265-013-1657-8.
Mendl, M., Brooks, J., Basse, C., Burman, O., Paul, E., Blackwell, E., Casey, R., 2010. Dogs showing separation-related behaviour exhibit a 'pessimistic' cognitive bias. Curr. Biol. 20, R839–R840, http://dx.doi.org/10.1016/j.cub.2010.08.030.
Menor-Campos, D.J., Molleda-Carbonell, J.M., López-Rodríguez, R., 2011. Effects of exercise and human contact on animal welfare in a dog shelter. Vet. Rec. 169, 388, http://dx.doi.org/10.1136/vr.d4757.
Miller, H.C., DeWall, C.N., Pattison, K., Molet, M., Zentall, T.R., 2012. Too dog tired to avoid danger: self-control depletion in canines increases behavioral approach toward an aggressive threat. Psychon. Bull. Rev. 19, 535–540, http://dx.doi.org/10.3758/s13423-012-0231-0.
Mirkó, E., Kubinyi, E., Gácsi, M., Miklósi, Á., 2012. Preliminary analysis of an adjective-based dog personality questionnaire developed to measure some aspects of personality in the domestic dog (*Canis familiaris*). Appl. Anim. Behav. Sci. 138, 88–98, http://dx.doi.org/10.1016/j.applanim.2012.02.016.
Mornement, K., Toukhsati, S., Coleman, G., Bennett, P.,2009. Reliability, validity and feasibility of existing tests of canine behaviour. In: AIAM Annual Conference on Urban Animal Management. Australian Institute of Animal Management, Victoria, Australia, pp. 11–18.
Mornement, K.M., Coleman, G.J., Toukhsati, S., Bennett, P.C., 2010. A review of behavioral assessment protocols used by Australian animal shelters to determine the adoption suitability of dogs. J. Appl. Anim. Welf. Sci. 13, 314–329, http://dx.doi.org/10.1080/10888705.2010.483856.
Mornement, K., Coleman, G., Toukhsati, S., Bennett, P., 2012. What do current and potential Australian dog owners believe about shelter practices and shelter dogs? Anthrozoos 25, 457–473, http://dx.doi.org/10.2752/175303712x13479798785850.
Mornement, K.M., Coleman, G.J., Toukhsati, S., Bennett, P.C., 2014. Development of the behavioural assessment for re-homing K9's (B.A.R.K.) protocol. Appl. Anim. Behav. Sci. 51, 75–83, http://dx.doi.org/10.1016/j.applanim.2013.11.008.
Murray, H.A., 1938. Introduction. In: Explorations in Personality: A Clinical Exploration of Fifty Men of College Age. Oxford University Press, Oxford, England, pp. 3–35.
National Canine Research Council, 2013. Assumptions about future behavior of shelter dogs seen as food aggressive are unfounded. National Canine Research Council http://www.nationalcanineresearchcouncil.com/blog/assumptions-about-future-behavior-of-shelter-dogs-seen-as-food-aggressive-are-unfounded (accessed 23.10.14).
Natoli, E., Say, L., Cafazzo, S., Bonanni, R., Schmid, M., Pontier, D., 2005. Bold attitude makes male urban feral domestic cats more vulnerable to Feline Immunodeficiency Virus. Neurosci. Biobehav. Rev. 29, 151–157, http://dx.doi.org/10.1016/j.neubiorev.2004.06.011.
Netto, W.J., Planta, D.J.U., 1997. Behavioural testing for aggression in the domestic dog. Appl. Anim. Behav. Sci. 52, 243–263, http://dx.doi.org/10.1016/s0168-1591(96)01126-4.
Oh, I.-S., Wang, G., Mount, M.K., 2011. Validity of observer ratings of the Five-Factor Model of personality traits: a meta-analysis. J. Appl. Psychol. 96, 762, http://dx.doi.org/10.1037/a0021832.
Palestrini, C., Previde, E.P., Spiezio, C., Verga, M., 2005. Heart rate and behavioural responses of dogs in the Ainsworth's Strange Situation: a pilot study. Appl. Anim. Behav. Sci. 94, 75–88, http://dx.doi.org/10.1016/j.applanim.2005.02.005.
Peremans, K., Audenaert, K., Coopman, F., Blanckaert, P., Jacobs, F., Otte, A., Verschooten, F., Bree, H., Heeringen, K., Mertens, J., Slegers, G., Dierckx, R., 2003. Estimates of regional cerebral blood flow and 5-HT2A receptor density in impulsive, aggressive dogs with 99mTc-ECD and 123I-5-I-R91150. Eur. J. Nucl. Med. Mol. Imaging 30, 1538–1546, http://dx.doi.org/10.1007/s00259-003-1250-x.
Planta, D.J., De Meester, R.H., 2007. Validity of the Socially Acceptable Behavior (SAB) test as a measure of aggression in dogs towards non-familiar humans. Vlaams Diergeneeskund. Tijdschr. 76, 359–368.
Quinkert, A.W., Vimal, V., Weil, Z.M., Reeke, G.N., Schiff, N.D., Banavar, J.R., Pfaff, D.W., 2011. Quantitative descriptions of generalized arousal, an elementary function of the vertebrate brain. Proc. Natl. Acad. Sci. U. S. A. 108, 15617–15623, http://dx.doi.org/10.1073/pnas.1101894108.

Réale, D., Reader, S.M., Sol, D., McDougall, P.T., Dingemanse, N.J., 2007. Integrating animal temperament within ecology and evolution. Biol. Rev. 82, 291–318, http://dx.doi.org/10.1111/j.1469-185X.2007.00010.x.

Riemer, S., Mills, D., Wright, H., 2013. Impulsive for life? The nature of long-term impulsivity in domestic dogs. Anim. Cogn., 1–5, http://dx.doi.org/10.1007/s10071-013-0701-4.

Riemer, S., Müller, C., Virányi, Z., Huber, L., Range, F., 2014. The predictive value of early behavioural assessments in pet dogs – a longitudinal study from neonates to adults. PLoS ONE 9, e101237, http://dx.doi.org/10.1371/journal.pone.0101237.

Riva, J., Bondiolotti, G., Michelazzi, M., Verga, M., Carenzi, C., 2008. Anxiety related behavioural disorders and neurotransmitters in dogs. Appl. Anim. Behav. Sci. 114, 168–181, http://dx.doi.org/10.1016/j.applanim.2008.01.020.

Roberts, B.W., DelVecchio, W.F., 2000. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. Psychol. Bull. 126, 3–25, http://dx.doi.org/10.1037/0033-2909.126.1.3.

Robertson, B., 2004. Dog is in the Details: The Many Sides of Temperament Testing. The Bark: Dog is My Co-Pilot. The Bark Inc., California http://thebark.com/content/dog-details

Rooney, N.J., Gaines, S.A., Bradshaw, J.W.S., 2007. Behavioural and glucocorticoid responses of dogs (*Canis familiaris*) to kennelling: investigating mitigation of stress by prior habituation. Physiol. Behav. 92, 847–854, http://dx.doi.org/10.1016/j.physbeh.2007.06.011.

Rosado, B., García-Belenguer, S., León, M., Chacón, G., Villegas, A., Palacio, J., 2010a. Blood concentrations of serotonin, cortisol and dehydroepiandrosterone in aggressive dogs. Appl. Anim. Behav. Sci. 123, 124–130, http://dx.doi.org/10.1016/j.applanim.2010.01.009.

Rosado, B., García-Belenguer, S., Palacio, J., Chacón, G., Villegas, A., Alcalde, A.I., 2010b. Serotonin transporter activity in platelets and canine aggression. Vet. J. 186, 104–105, http://dx.doi.org/10.1016/j.tvjl.2009.07.029.

Rothbart, M.K., Ahadi, S.A., Evans, D.E., 2000. Temperament and personality: origins and outcomes. J. Pers. Soc. Psychol. 78, 122, http://dx.doi.org/10.1037//002.3514.78.1.122.

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, http://dx.doi.org/10.1016/s0168-1591(02)00134-x.

Saetre, P., Strandberg, E., Sundgren, P.E., Pettersson, U., Jazin, E., Bergström, T.F., 2006. The genetic contribution to canine personality. Genes Brain Behav. 5, 240–248, http://dx.doi.org/10.1111/j.1601-183X.2005.00155.x.

Sales, G., Hubrecht, R., Peyvandi, A., Milligan, S., Shield, B., 1997. Noise in dog kennelling: is barking a welfare problem for dogs? Appl. Anim. Behav. Sci. 52, 321–329.

Schimmack, U., Grob, A., 2000. Dimensional models of core affect: a quantitative comparison by means of structural equation modeling. Eur. J. Pers. 14, 325–345, http://dx.doi.org/10.1002/1099-0984(200007/08)14:4%3C325::AID-PER380%3E3.0.CO;2-I.

Schwartz, C.E., Wright, C.I., Shin, L.M., Kagan, J., Rauch, S.L., 2003. Inhibited and uninhibited infants "grown up": adult amygdalar response to novelty. Science 300, 1952–1953, http://dx.doi.org/10.1126/science.1083703.

Segurson, S., 2007. Behavioral Assessment in Animal Shelters Maddie's Fund. http://www.maddiesfund.org/Maddies_Institute/Articles/Behavioral_Assessment_in_Animal_Shelters.html (accessed 23.10.14).

Serpell, J.A., 1996. Evidence for an association between pet behavior and owner attachment levels. Appl. Anim. Behav. Sci. 47, 49–60, http://dx.doi.org/10.1016/0168-1591(95)01010-6.

Serpell, J.A., Hsu, Y., 2001. Development and validation of a novel method for evaluating behavior and temperament in guide dogs. Appl. Anim. Behav. Sci. 72, 347–364, http://dx.doi.org/10.1016/s0168-1591(00)00210-0.

Sgoifo, A., De Boer, S.F., Haller, J., Koolhaas, J.M., 1996. Individual differences in plasma catecholamine and corticosterone stress responses of wild-type rats: relationship with aggression. Physiol. Behav. 60, 1403–1407, http://dx.doi.org/10.1016/s0031-9384(96)00229-6.

Shore, E.R., Burdsal, C., Douglas, D.K., 2008. Pet owners' views of pet behavior problems and willingness to consult experts for assistance. J. Appl. Anim. Welf. Sci. 11, 63–73, http://dx.doi.org/10.1080/10888700701729221.

Sih, A., Bell, A., Johnson, J.C., 2004a. Behavioral syndromes: an ecological and evolutionary overview. Trends Ecol. Evol. 19, 372–378, http://dx.doi.org/10.1016/j.tree.2004.04.009.

Sih, A., Bell, A.M., Johnson, J.C., Ziemba, R.E., 2004b. Behavioral syndromes: an integrative overview. Q. Rev. Biol. 79, 241–277, http://dx.doi.org/10.1086/422893.

Sinn, D.L., Gosling, S.D., Hilliard, S., 2010. Personality and performance in military working dogs: reliability and predictive validity of behavioral tests. Appl. Anim. Behav. Sci. 127, 51–65, http://dx.doi.org/10.1016/j.applanim.2010.08.007.

Sluyter, F., Korte, S.M., Bohus, B., Van Oortmerssen, G.A., 1996. Behavioral stress response of genetically selected aggressive and nonaggressive wild house mice in the shock-probe/defensive burying test. Pharmacol. Biochem. Behav. 54, 113–116, http://dx.doi.org/10.1016/0091-3057(95)02164-7.

Soto, C.J., John, O.P., Gosling, S.D., Potter, J., 2011. Age differences in personality traits from 10 to 65: big five domains and facets in a large cross-sectional sample. J. Pers. Soc. Psychol. 100, 330, http://dx.doi.org/10.1037/a0021717.

Stamps, J., Groothuis, T.G., 2010. The development of animal personality: relevance, concepts and perspectives. Biol. Rev. 85, 301–325, http://dx.doi.org/10.1111/j.1469-185X.;1; 2009.00103.x.

Starling, M.J., Branson, N., Cody, D., McGreevy, P.D., 2013a. Conceptualising the impact of arousal and affective state on training outcomes of operant conditioning. Animals 3, 300–317, http://dx.doi.org/10.3390/ani3020300.

Starling, M.J., Branson, N., Thomson, P.C., McGreevy, P.D., 2013b. Age, sex and reproductive status affect boldness in dogs. Vet. J. 197, 868–872, http://dx.doi.org/10.1016/j.tvjl.2013.05.019.

Starling, M.J., Branson, N., Thomson, P.C., McGreevy, P.D., 2013c. "Boldness" in the domestic dog differs among breeds and breed groups. Behav. Process. 97, 53–62, http://dx.doi.org/10.1016/j.beproc.2013.04.008.

Starling, M.J., Branson, N., Cody, D., Starling, T.R., McGreevy, P.D., 2014. Canine sense and sensibility: tipping points and response latency variability as an optimism index in a canine judgement bias assessment. PLoS ONE 9, e107794, http://dx.doi.org/10.1371/journal.pone.0107794.

Strelau, J., 1972. Nervous system type and extraversion-introversion: a comparison of Eysenck's theory with Pavlov's typology. Polish Psychol. Bull. 1, 17–24.

Svartberg, K., 2002. Shyness–boldness predicts performance in working dogs. Appl. Anim. Behav. Sci. 79, 157–174, http://dx.doi.org/10.1016/s0168-1591(02)00120-x.

Svartberg, K., 2005. A comparison of behaviour in test and in everyday life: evidence of three consistent boldness-related personality traits in dogs. Appl. Anim. Behav. Sci. 91, 103–128, http://dx.doi.org/10.1016/j.applanim.2004.08.030.

Svartberg, K., 2006. Breed-typical behaviour in dogs—historical remnants or recent constructs? Appl. Anim. Behav. Sci. 96, 293–313, http://dx.doi.org/10.1016/j.applanim.2005.06.014.

Svartberg, K., Forkman, B., 2002. Personality traits in the domestic dog (*Canis familiaris*). Appl. Anim. Behav. Sci. 79, 133–155, http://dx.doi.org/10.1016/s0168-1591(02)00121-1.

Svartberg, K., Tapper, I., Temrin, H., Radesäter, T., Thorman, S., 2005. Consistency of personality traits in dogs. Anim. Behav. 69, 283–291, http://dx.doi.org/10.1016/j.anbehav.2004.04.011.

Svenska Brukshundklubben, 2012. Mentalbeskrivning hund. Svenska Brukshundklubben. http://www.brukshundklubben.se/hundar/mentalbeskrivning-mh/ (accessed 14.02.14).

Tami, G., Gallagher, A., 2009. Description of the behaviour of domestic dog (*Canis familiaris*) by experienced and inexperienced people. Appl. Anim. Behav. Sci. 120, 159–169, http://dx.doi.org/10.1016/j.applanim.2009.06.009.

Taylor, K.D., Mills, D.S., 2006. The development and assessment of temperament tests for adult companion dogs. J. Vet. Behav. 1, 94–108, http://dx.doi.org/10.1016/j.jveb.2006.09.002.

Teigen, K.H., 1994. Yerkes–Dodson: a law for all seasons. Theory Psychol. 4, 525–547, http://dx.doi.org/10.1177/0959354394044004.

Thayer, R.E., 1978. Toward a psychological theory of multidimensional activation (arousal). Mot. Emot. 2, 1–34, http://dx.doi.org/10.1007/BF00992729.

Toms, C.N., Echevarria, D.J., Jouandot, D.J., 2010. A methodological review of personality-related studies in fish: focus on the shy–bold axis of behavior. Int. J. Comp. Psychol. 23, 1–25.

Tuber, D.S., Miller, D.D., Caris, K.A., Halter, R., Linden, F., Hennessy, M.B., 1999. Dogs in animal shelters: problems, suggestions, and needed expertise. Psychol. Sci. 10, 379–386, http://dx.doi.org/10.1111/1467-9280.00173.

Turcsán, B., Kubinyi, E., Miklósi, Á., 2011. Trainability and boldness traits differ between dog breed clusters based on conventional breed

categories and genetic relatedness. Appl. Anim. Behav. Sci. 132, 61–70, http://dx.doi.org/10.1016/j.applanim.2011.03.006.
Uher, J., Asendorpf, J.B., 2008. Personality assessment in the Great Apes: comparing ecologically valid behavior measures, behavior ratings, and adjective ratings. J. Res. Pers. 42, 821–838, http://dx.doi.org/10.1016/j.jrp.2007.10.004.
Uher, J., 2011a. Individual behavioral phenotypes: an integrative meta-theoretical framework. Why "behavioral syndromes" are not analogs of "personality". Dev. Psychobiol. 53, 521–548, http://dx.doi.org/10.1002/dev.20544.
Uher, J., 2011b. Personality in non-human primates: what can we learn from human personality psychology? In: Weiss, A., King, J., Murray, L. (Eds.), Personality and Temperament in Nonhuman Primates. Springer, New York, pp. 41–76.
Uher, J., Werner, C.S., Gosselt, K., 2013a. From observations of individual behaviour to social representations of personality: developmental pathways, attribution biases, and limitations of questionnaire methods. J. Res. Pers. 47, 647–667, http://dx.doi.org/10.1016/j.jrp.2013.03.006.
Uher, J., Addessi, E., Visalberghi, E., 2013b. Contextualised behavioural measurements of personality differences obtained in behavioural tests and social observations in adult capuchin monkeys (*Cebus apella*). J. Res. Pers. 47, 427–444, http://dx.doi.org/10.1016/j.jrp.2013.01.013.
Uher, J., 2014. Interpreting "personality" taxonomies: why previous models cannot capture individual-specific experiencing, behaviour, functioning and development, Major taxonomic tasks still lay ahead. Integr. Physiol. Behav. Sci., 1–56, http://dx.doi.org/10.1007/s12124-014-9281-3.
Valsecchi, P., Barnard, S., Stefanini, C., Normando, S., 2009. Validation of a new temperament test as a practical tool for adoptions of sheltered dogs. J. Vet. Behav. 4, 75–76, http://dx.doi.org/10.1016/j.jveb.2008.10.024.
Valsecchi, P., Barnard, S., Stefanini, C., Normando, S., 2011. Temperament test for re-homed dogs validated through direct behavioral observation in shelter and home environment. J. Vet. Behav. 6, 161–177, http://dx.doi.org/10.1016/j.jveb.2011.01.002.
van den Berg, L., Vos-Loohuis, M., Schilder, M.B., van Oost, B.A., Hazewinkel, H.A., Wade, C.M., Karlsson, E.K., Lindblad-Toh, K., Liinamo, A.E., Leegwater, P.A., 2008. Evaluation of the serotonergic genes htr1A, htr1B, htr2A, and slc6A4 in aggressive behavior of Golden Retriever dogs. Behav. Genet. 38, 55–66, http://dx.doi.org/10.1007/s10519-007-9179-7.
van der Borg, J.A.M., Beerda, B., Ooms, M., de Souza, A.S., van Hagen, M., Kemp, B., 2010. Evaluation of behaviour testing for human directed aggression in dogs. Appl. Anim. Behav. Sci. 128, 78–90, http://dx.doi.org/10.1016/j.applanim.2010.09.016.
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, http://dx.doi.org/10.1016/s0168-1591(05)80047-4.
Van Reenen, C.G., Engel, B., Ruis-Heutinck, L.F.M., Van der Werf, T.N., Buist, W.G., Jones, R.B., Blokhuis, H.J., 2004. Behavioural reactivity of heifer calves in potentially alarming test situations: a multivariate and correlational analysis. Appl. Anim. Behav. Sci. 85, 11–30, http://dx.doi.org/10.1016/j.applanim.2003.09.007.
Van Reenen, C.G., O'Connell, N.E., Van der Werf, J.T., Korte, S.M., Hopster, H., Jones, R.B., Blokhuis, H.J., 2005. Responses of calves to acute stress: individual consistency and relations between behavioral and physiological measures. Physiol. Behav. 85, 557–570, http://dx.doi.org/10.1016/j.physbeh.2005.06.015.
Vas, J., Topál, J., Péch, É., Miklósi, Á., 2007. Measuring attention deficit and activity in dogs: a new application and validation of a human ADHD questionnaire. Appl. Anim. Behav. Sci. 103, 105–117, http://dx.doi.org/10.1016/j.applanim.2006.03.017.
Vincent, I.C., Leahy, R.A., 1997. Real-time non-invasive measurement of heart rate in working dogs: a technique with potential applications in the objective assessment of welfare problems. Vet. J. 153, 179–183, http://dx.doi.org/10.1016/S1090-0233(97)80038-X.
Wan, M., Hejjas, K., Ronai, Z., Elek, Z., Sasvari-Szekely, M., Champagne, F.A., Miklosi, A., Kubinyi, E., 2013. DRD4 and TH gene polymorphisms are associated with activity, impulsivity and inattention in Siberian Husky dogs. Anim. Genet. 44, 717–727, http://dx.doi.org/10.1111/age.12058.
Wechsler, B., 1995. Coping and coping strategies: a behavioural view. Appl. Anim. Behav. Sci. 43, 123–134, http://dx.doi.org/10.1016/0168-1591(95)00557-9.
Weil, Z.M., Zhang, Q., Hornung, A., Blizard, D., Pfaff, D.W., 2010. Impact of generalized brain arousal on sexual behavior. Proc. Natl. Acad. Sci. U. S. A. 107, 2265–2270, http://dx.doi.org/10.1073/pnas.0914014107.
Weiss, E., 2002. Selecting shelter dogs for service dog training. J. Appl. Anim. Welf. Sci. 5, 43–62, http://dx.doi.org/10.1207/s15327604jaws0501_4.
Weiss, E., Greenberg, G., 1997. Service dog selection tests: effectiveness for dogs from animal shelters. Appl. Anim. Behav. Sci. 53, 297–308, http://dx.doi.org/10.1016/s0168-1591(96)01176-8.
Wells, D.L., 2004. A review of environmental enrichment for kennelled dogs, Canis familiaris. Appl. Anim. Behav. Sci. 85, 307–317, http://dx.doi.org/10.1016/j.applanim.2003.11.005.
White, J.R., Meekan, M.G., McCormick, M.I., Ferrari, M.C.O., 2013. A comparison of measures of boldness and their relationships to survival in young fish. PLOS ONE 8, 1–11, http://dx.doi.org/10.1371/journal.pone.0068900.
Whiteside, S.P., Lynam, D.R., 2001. The Five Factor model and impulsivity: using a structural model of personality to understand impulsivity. Pers. Individ. Differ. 30, 669–689, http://dx.doi.org/10.1016/S0191-8869(00)00064-7.
Wilson, D.S., Clark, A.B., Coleman, K., Dearstyne, T., 1994. Shyness and boldness in humans and other animals. Trends Ecol. Evol. 9, 442–446, http://dx.doi.org/10.1016/0169-5347(94)90134-1.
Wilsson, E., Sundgren, P.-E., 1997. 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, http://dx.doi.org/10.1016/s0168-1591(96)01174-4.
Winter, D.G., John, O.P., Stewart, A.J., Klohnen, E.C., Duncan, L.E., 1998. Traits and motives: toward an integration of two traditions in personality research. Psychol. Rev. 105, 230, http://dx.doi.org/10.1037//0033-295X.105.2.230.
Wolf, M., Weissing, F.J., 2012. Animal personalities: consequences for ecology and evolution. Trends Ecol. Evol. 27, 452–461, http://dx.doi.org/10.1016/j.tree.2012.05.001.
Wright, H.F., Mills, D., Pollux, P.M.J., 2011. Development and validation of a psychometric tool for assessing impulsivity in the domestic dog (*Canis familiaris*). Int. J. Comp. Psychol. 24, 210–225.
Wright, H.F., Mills, D.S., Pollux, P.M.J., 2012. Behavioural and physiological correlates of impulsivity in the domestic dog (*Canis familiaris*). Physiol. Behav. 105, 676–682, http://dx.doi.org/10.1016/j.physbeh.2011.09.019.
Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. 18, 459–482, http://dx.doi.org/10.1002/cne.920180503.