The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK

Animal Welfare 2010, 19: 75-84
ISSN 0962-7286

# The assessment of emotional expression in dogs using a Free Choice Profiling methodology

J Walker*†, A Dale†, N Waran†, N Clarke†‡, M Farnworth† and F Wemelsfelder§

† Department of Natural Sciences, Unitec, Private Bag 92025, Auckland, New Zealand
‡ The Animal Behaviour and Welfare Group, Department of Clinical Veterinary Science, University of Bristol, Langford, Bristol BS40 5DU, UK
§ Sustainable Livestock Systems, Scottish Agricultural College, Bush Estate, Penicuik EH26 0QE, UK
* Contact for correspondence and requests for reprints: jwalker@unitec.ac.nz

## Abstract

*This study explores the use of Free Choice Profiling (FCP) methodology for the qualitative behaviour assessment of emotional expression in the domestic dog (*Canis familiaris*). Qualitative behaviour assessment is based upon the integration of many pieces of information that in conventional quantitative approaches are recorded separately or not at all. Observers are asked to focus on an animal's interaction with its surrounding environment, and to describe the animal's expressive demeanor, or 'body language'. A specific characteristic of FCP methodology is that it allows observers the freedom to devise their own descriptive terms, and then to use these personal terms to quantitatively score observed subjects. Application of FCP to qualitative behaviour assessment in animals was originally tested for pigs, and more recently for dairy cows, horses, and ponies. The goal of this study was to apply FCP to the domestic dog, and to investigate the inter-observer reliability of assessments of emotional expression in 10 individual Beagles by a group of 18 untrained observers. The data was analysed using Generalised Procrustes Analysis (GPA), a multivariate statistical technique associated with FCP. The observers achieved highly-significant agreement in their assessments of the dogs' expressions, thereby establishing the applicability of this methodology for the first time in the domestic dog.*

**Keywords**: *animal welfare, dog welfare, emotional expression, Free Choice Profiling, Generalised Procrustes Analysis, qualitative behaviour assessment*

## Introduction

Traditionally, establishing the welfare status of an animal is based upon behavioural and physiological measurement and evaluation. However, these methods tend to isolate particular aspects of behaviour for quantification, breaking up the behavioural flow and leading to a loss of 'whole-animal' information which cannot be regained at a later stage (Wemelsfelder & Lawrence 2001). Qualitative assessment of behaviour is based upon the integration of many pieces of information that, in conventional quantitative approaches, are recorded separately, or not at all. This may include incidental behavioural events, subtle details of movement and posture, and aspects of the context in which behaviour occurs (Wemelsfelder *et al* 2001). It focuses not so much on what an animal does, but on how it does it, that is, its dynamic style of interaction with the environment (Stevenson-Hinde *et al* 1980; Stevenson-Hinde 1983; Feaver *et al* 1986; Fagen *et al* 1997; Wemelsfelder 1997). Thus, behaviour is seen as an expressive process that is open to direct observation at any given moment in time, and that can be described using terms such as 'bold', 'shy', 'content', 'frustrated', 'bored' or 'relaxed'.

Scientists frequently use expressive terminologies in studies of animal personality and temperament (eg Stevenson-Hinde *et al* 1980; Stevenson-Hinde 1983; Goddard & Beilharz 1984; Feaver *et al* 1986; Clarke & Boinski 1995; Fagen *et al* 1997; Murphy 1998; Gosling 2001; Serpell & Hsu 2001; Svartberg & Forkman 2002), and in studies involving pain assessment (eg Light *et al* 1993; Hansen 1997; Holton *et al* 1998; Firth & Haldane 1999; Hardie 2000; Kent *et al* 2000; Holton *et al* 2001; Molony *et al* 2002; Wiseman *et al* 2004). However, wary of anthropomorphising, they tend to be reluctant to use terms such as 'anxious', 'relaxed' or 'content' as descriptors in studies of animal behaviour and welfare. Nevertheless, over time, various authors have suggested that qualitative terminologies could potentially be used as indicators of animal experience and emotion (eg Hebb 1946; Wemelsfelder 1997; Weiss *et al* 2006). The question is whether the use of such indicators is scientifically robust (Wemelsfelder 2007).



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

To investigate this question, Wemelsfelder *et al* (2000, 2001) made use of a Free Choice Profiling (FCP) methodology, and found high levels of inter- and intra-observer reliability in observers' spontaneous qualitative assessments of emotional expression in pigs. The distinguishing characteristic of FCP is that it asks observers to generate their own descriptive vocabulary based on direct observation of animals, and thus facilitates the active interpretation by observers of these animals' expressions, rather than providing them with pre-selected descriptive terms, or asking them to infer descriptors retrospectively from quantitative data. FCP was originally developed in food science, and is associated with a multivariate statistical technique called Generalised Procrustes Analysis (GPA), a pattern detection mechanism that does not depend on the use of fixed variables to calculate consensus between groups of assessors (Gower 1975; Oreskovich *et al* 1991). The FCP/GPA approach has been extensively applied and validated in studies of the sensory perception of food qualities (eg Arnold & Williams 1985; Noble & Ebeler 2002). Wemelsfelder *et al* (2000, 2001) were the first to apply this approach for use in animal science. The advantage of this approach is that it allows testing of whether observers perceive animal expressions in similar ways rather than assuming they do, and also that it allows observers to generate descriptors that are most relevant to a particular study rather than being given pre-fixed descriptors. More recently, this approach has been successfully applied to dairy cattle (Rousing & Wemelsfelder 2006), and horses and ponies (Napolitano *et al* 2008; Minero *et al* 2009). These studies also found high levels of inter- and intra-observer reliability, and demonstrated that observers' qualitative assessments showed significant and meaningful correlations with ethogram-based assessments of behaviour.

The aim of this study was to apply FCP methodology for the first time to the domestic dog, and to test the inter-observer reliability of assessing emotional expression in dogs. The application of qualitative behaviour assessment to domestic dogs is not in itself new. Studies investigating pain assessment in dogs, for example, have relied upon qualitative assessments of facial expression and body language to interpret the degree of pain severity (eg Light *et al* 1993; Firth & Haldane 1999; Hardie 2000; Holton *et al* 2001; Wiseman *et al* 2004, 2006). However, these studies were either based on fixed descriptor lists or relied on retrospective judgment, and thus did not facilitate direct investigation of dog behavioural expressions in particular contexts. Such investigation may potentially play a useful role in understanding the dynamics of dog welfare. Dogs play a valuable role in human society, but despite this they can, like all animals, be subject to poor welfare states. The present study asked observers to assess the expressions of individual dogs while in interaction with an unfamiliar human person, both because such a test situation had proven fruitful with individual pigs (Wemelsfelder *et al* 2001), and because interaction with humans is such a crucial part of most dogs' well-being. If successful, the application and testing of new research methodologies, such as FCP, in such situations should contribute to further understanding the behaviour and welfare of the domestic dog.

## Materials and methods

### Animals

The subjects of this study were ten Beagles owned by the Ministry of Agriculture and Forestry for use as customs dogs at Auckland International Airport. Although very similar morphologically, the group of dogs consisted of mixed sex and age (2–8 years). The dogs were individually housed in kennels that comprised of two sections; an indoor section (1 × 1.5 m; length × breadth) and an outdoor section (2.5 × 1.5 m). The outside section also contained a tug rope suspended from the ceiling for the purposes of enrichment. The dogs were maintained on Hills Science Diet, provided once a day (am), each dog receiving between 2–4 cups. Dogs were also provided with a raw hide bone during the day. After morning feeding, the dogs were confined to the indoor section of the kennel whilst daily cleaning of the outside section was carried out. The dogs were then moved to the outdoor section of the kennel whilst the indoor section was cleaned. The dogs were exercised daily as a pack in an outdoor yard (10 × 5 m). Kennel staff were assigned dogs on a daily basis and carried out obedience training with them.

### Video

Between 1300 and 1700h, each dog was brought individually from their kennel into a test pen constructed as part of an outside free run which was small enough (2.5 × 4 m) that the entire area could be filmed. Each individual dog was then allowed to interact with an unfamiliar human situated in the middle of the pen for seven minutes. During the interaction the dog's behaviour was recorded using a digital camcorder (Sony DCR-TRV22E, Tokyo, Japan) mounted at the edge of the pen. Within each session, the human participant, as a general rule, took a passive role in order to standardise the interaction as much as possible. The human participant only interacted with the dog when and if it was solicited. For example, if the dog looked at the human or approached, the human extended a hand towards it. If the dog remained close and initiated further interaction, the human would begin to gently stroke the dog. If the dog stopped interacting with the human, the human would cease to stroke the dog. If the dog became over-excited the human would push the dog off and remain passive. The resulting video footage was edited to create two-minute video clips with a three-minute break in between each clip. Clips were selected according to when the dog was first viewable on the screen for a continuous two-minute period.

### Observer group

The observer group comprised of 18 female undergraduate students, all of whom were currently studying animal behaviour and had experience observing animals, nine of which had specifically studied canine behaviour and had experience observing dogs. None of the observers had previous experience with qualitative behaviour assessment or Free Choice Profiling methodology.



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

## Observer session one

The observers were told the experiment was part of a research project aimed at investigating the reliability of a methodology for assessing the behavioural expressions of animals (behavioural expression was defined as style of interaction, that is, how the animal behaves as opposed to what it does). It was explained that an essential part of the methodology required the observers to generate their own terms with which to score the behavioural expressions of the dogs. Following specific instructions, all observers were simultaneously shown the two-minute clips using a lecture theatre screen. This was followed by three minutes to write down terms which they considered best described the behavioural expression of the dog. With each new clip, observers were free to choose as many or as few terms as they deemed appropriate, and to re-use terms or select new terms for each dog. They were told to refrain from discussing terms throughout the session. Thus, each observer compiled a set of terms describing the expressive repertoire of the 10 dogs. The FCP procedure used in this study has been described in detail by Wemelsfelder *et al* (2001).

## Observer session two

In phase two, each of the observers was provided with a compilation of their own terms that they had generated in phase one, each term was set next to a visual analogue scale (0–125 mm). They were instructed by the experimenter on how to use their personal terminologies as a quantitative measurement tool. The observers were told that only after each two-minute clip had ended they should score each dog, on each of the terms of their personal rating scales by marking a line at an appropriate point between 'minimum' (0 mm) and 'maximum'(125 mm). All the terms had to be scored, if they were considered irrelevant the observer was required to score that expression at the minimum point. The observers watched the same video sequence as in phase one.

## General method of analysis

At the end of phase two, observers had used their personal rating scales to produce a set of scores for all 10 dogs. The score on a certain term for a dog was determined by measuring the distance (mm) between the left 'minimum' point of the scale and the point where the observer's mark crossed the line. Each observers' scores were entered into a data matrix (one for each observer) with each matrix defined by the number of dogs (10) and the number of terms used by each individual observer.

Generalized Procrustes Analysis (GPA) was used to analyse the data. GPA is a multivariate statistical technique that does not rely on fixed variables (Oreskovich *et al* 1991; Noble & Ebeler 2002; Saba & Rosati 2002). It detects the level of consensus between observer scoring patterns on the basis of inter-sample distances specified by each observer. The statistical significance of this consensus is then evaluated against a mean randomised profile, obtained by re-running GPA with randomised observer data sets a hundred times. A one-tailed Student's *t*-test (n = 100) is used to determine whether the true consensus differs significantly from the mean randomised profile; a probability of $P < 0.001$ is generally taken to indicate that the consensus profile was a meaningful feature of the data set rather than a statistical artifact. A detailed description of these GPA procedures can be found in Wemelsfelder *et al* (2000).

### *Statistical procedures*

To find the consensus between data matrices within a data set, GPA assesses each matrix as a multidimensional configuration. The position of the dogs in the multidimensional space is defined by their scores. Columns of zeros are added to individual matrices, so that all observer configurations acquire equal dimensionality. Matching is achieved through a series of iterative transformations (translation, rotation/reflection and scaling) which maintain relative inter-sample relationships within each configuration. A mean of all the transformed individual configurations is taken, thought of as the 'consensus profile', which represents the 'best-possible-fit' of the configurations. The 'goodness-of-fit' of transformed observer configurations is quantified by the Procrustes Statistic, which gives the percentage of variation between observer configurations explained by the consensus profile. Using Principal Co-ordinate Analysis (PCO) of the Procrustes statistic for each pair of observers, the distance between transformed observer configurations and the 'best-of-fit' can be projected visually in an 'observer plot'. PCO estimates the centre of distributions of observers together with a standard deviation, and draws a 95% confidence region. Observers lying outside this region are potentially outliers that is, in some sense, they may differ from the other observers in their assessment of the samples (Gains & Thompson 1990).

GPA thus transforms the 18 different dog-scoring configurations into one multidimensional consensus profile, entirely independently of any interpretation by the experimenter. This consensus profile is defined purely in terms of its geometrical properties, and has as yet no semantic connotations attached to it. A first step towards interpretation is to reduce the number of dimensions of the consensus profile through Principal Component Analysis (PCA). This identifies the principle dimensions of the consensus, and determines how much variation between dogs these dimensions explain. Individual animals are attributed scores on each of these dimensions. These scores are reflected in two-dimensional 'dog-plots' which show the distribution of the 10 dogs along the principle axes of the consensus profile. A standard error ellipse indicates the reliability for each dog's position on the axes. The second step is to calculate how the co-ordinates of the consensus profile correlate with the co-ordinates of each of the 18 original individual matrices; this step confers semantic meaning onto the principle axes of the consensus profile. This information is presented in 18 two-dimensional 'word charts' (one for each observer). On each chart all the terms of a particular observer are correlated with the principle dimensions of the consensus profile; the higher a term correlates with a dimension, the more weight it has as a descriptor for that dimension. The extent to which individual observers concur in their judgment of dogs'



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

**Table 1 Significance of the observers' assessment of behavioural expression in dogs.**

| Factor | Procrustes statistic |
|---|---|
| Consensus profile | 78.79 |
| Mean (± SD) randomised profile | 62.12 (± 0.28) |
| $t_{99}$ | 16.14* |

* $P < 0.001$.

expressions is indicated by the degree of semantic convergence between charts. For example, in one observer's chart, the terms confident/playful may show the highest correlation with consensus dimensions, while in another observer's chart the terms boisterous/bold may correlate best. Even though these are different terms they have similar meanings and suggest similar interpretations of what was seen by observers. If there appears to be satisfactory semantic convergence between observer word charts, a third and final step of interpretation is for the experimenter to summarise this information into one or more labels for the main dimensions of the sample plots. However, this interpretative role of the experimenter is entirely '*post hoc*', and plays no role in the computation of the consensus profile. GPA preserves semantic information as part of the analysis of object-based data sets, independently from the experimenter's interpretation of that information, making it possible to investigate whether observers apply their qualitative vocabulary in similar ways to characterise a group of dogs.

## Results

### Observer plots and their statistical significance

The consensus profile explained a significantly higher percentage of the variation between observer matrices than the mean of 100 randomised profiles (Table 1).

The observer plot also shows good consensus between the observers, with the majority of observers lying within the 95% confidence region (Figure 1). Although there are five outliers (those lying outside the 95% confidence region) they do not belong to one specific group, ie they include both observers who had completed a basic canine behaviour class and those who had not specifically studied canine behaviour. The outliers also did not stop the consensus from being highly significant ($P < 0.001$). Thus, the qualitative assessment of dog expression, made by observers given freedom to create their own terminologies, shows high inter-observer reliability.

### Observer word charts

The word charts for observers 3 and 12 are shown as examples in Figure 2. The axes of these charts reflect the three main dimensions of the consensus profile, showing which of the observers' terms best correlate with each of these dimensions. It should be emphasised that, as explained in *Materials and methods*, these word charts are entirely an outcome of statistical calculations based on the observers' scorings of the dogs, and have not been manipulated in any way by either experimenters or observers.

As can be seen in Figure 2, the word chart of observer 3 characterises dimension 1 as ranging from 'friendly' to 'nervous', while that of observer 12 describes this dimension as 'relaxed' to 'tense'. Dimension 2 was characterised by observer 3 as 'investigative–timid', and by observer 12 as 'active–quiet', while dimension 3 was described as 'excited–calm' and as 'anxious-content' by observers 3 and 12, respectively.

Table 2 provides a reflection of how similar or different these example observer word charts are from the other word charts by listing, for all 18 observers, the two terms which held the highest positive and negative correlations with dimensions 1, 2 and 3. For example, for six of the 18 observers the term 'playful' best describes the positive end of dimension 1, and for 12 of the 18 observers the term 'calm' best describes the negative end of dimension 3. The terms used most frequently to characterise the first dimension of the consensus profile were 'playful', 'happy', and 'confident' versus 'nervous', 'unsure' and 'tense'. The terms used most frequently to characterise the second dimension of the consensus profile were 'alert', 'inquisitive' and 'investigative' versus 'attention-seeking', 'quiet' and 'unsure'. The terms used most frequently to characterise the third dimension of the consensus profile were 'playful', 'nervous' and 'boisterous', versus 'calm', 'relaxed' and 'content'. This shows that a considerable number of observers used the same terms to describe the dimensions. Where observers used different terms, the meanings of the terms was similar (eg 'investigative/explorative/curious', 'nervous/fearful/anxious', or 'calm/gentle/docile'). Observers also produced terms that complemented each other (eg 'confident/alert', 'needy/timid'); although these terms do not have the same meaning they appear to reflect similar aspects of observed behaviour.

Figure 3 shows the strength of correlation for all descriptors of dimensions 1, 2 and 3. The *r*-value of the majority of terms for dimension 1 falls between 0.7–1.0, indicating these terms characterise this dimension well. *R*-values for dimensions 2 and 3 are slightly lower (0.6–1.0 and 0.4–0.9, respectively), but still enable meaningful characterisation of this dimension. These results suggest that generally the observers were able to use their self-generated terms as a coherent and meaningful semantic framework for assessing the dogs' behavioural expression.

### Dog plots

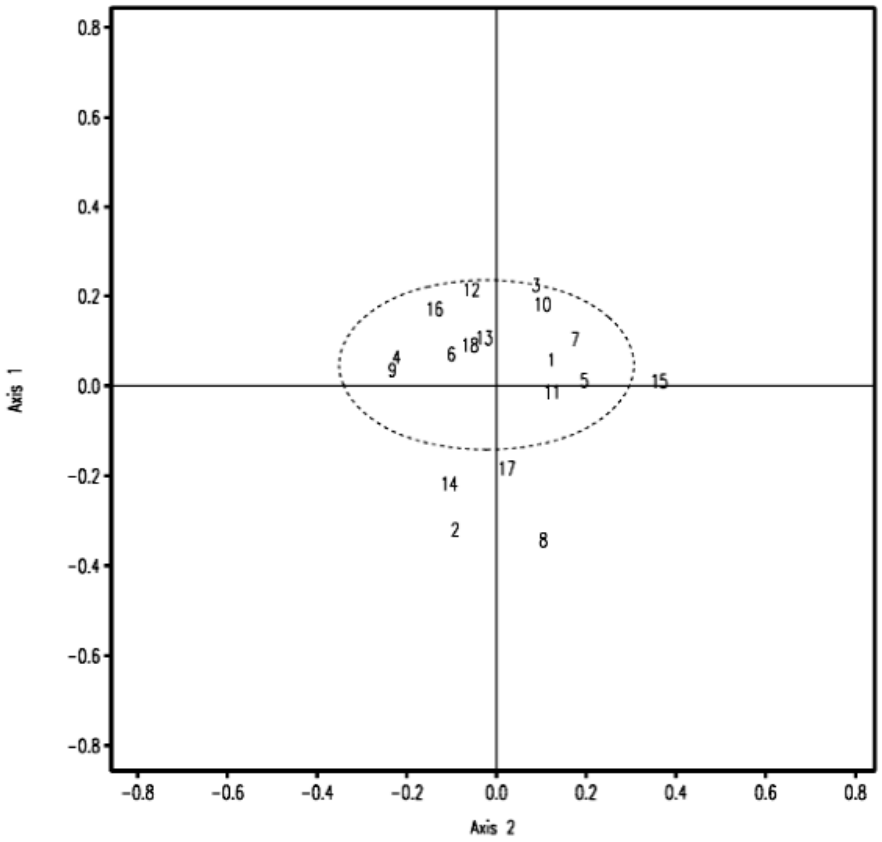
Figure 4 shows the position of the individual dogs on the three main dimensions. The dogs are distributed reasonably evenly over the three dimensions, which suggests that these dimensions effectively characterise observed variances in behavioural expression. The standard errors of the positions of the dogs (as reflected by the small dotted circle in the bottom right hand corner) on the plot are small, suggesting



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

**Figure 1**

Observer plot. Axes reflect PCO scaling values for relative observer distance — numbers represent individual observers. The dotted ellipse represents the 95% confidence region for what may be considered the normal population of observers.




that the dogs' positions on the dimensions are reliable. Dimension 1 explains 41.6%, dimension 2 explains 26.1% and dimension 3 explains 13.2%, giving a total of 80.9% of the variance between the dogs explained.

## Discussion

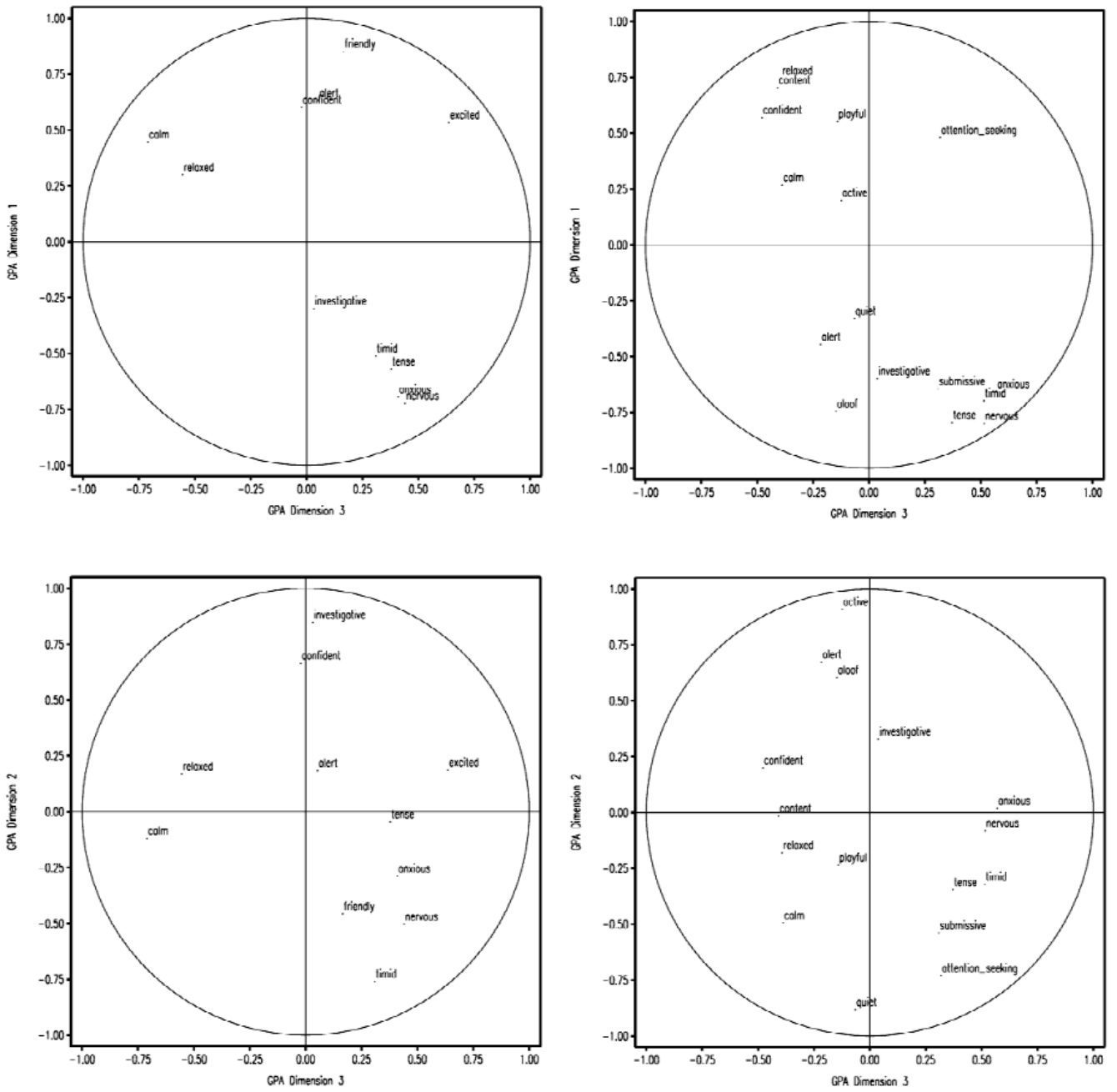
The objective of this study was to investigate the applicability of a novel qualitative behaviour assessment methodology that had previously been tested only with farm animals, to the domestic dog. The results show that, using a Free Choice Profiling methodology (FCP), observers achieved good agreement in their judgment and quantification of emotional expressions in a group of 10 Beagles. The consensus between observers was highly significant ($P < 0.001$) and identified three main dimensions of dog expression that together explained 80.9% of the variation. These dimensions were identified as dimension 1: 'playful/happy/confident' to 'nervous/unsure/tense'; dimension 2: 'alert/inquisitive/investigative' to 'attention-seeking/quiet/unsure', and dimension 3: 'playful/nervous/boisterous' to 'calm/relaxed/content'.

These results provide a first step in testing the scientific application of direct, non-retrospective qualitative judgements of dog behaviour. The dimensions identified in this study contain similar elements as those found in other studies using pre-fixed rating scales, such as, for example, a study by Svartberg and Forkman (2002) which reports 'playfulness', 'curiosity/fearlessness', 'chase-prone', 'sociability' and 'aggressiveness' as dimensions of temperament in dogs. The high agreement achieved in the 'free-choice' procedures in the present study provides support for the use of pre-fixed qualitative rating scales, by demonstrating that observers generally do interpret and quantify dog expressions in similar ways. There is the possibility furthermore that free-choice assessments might encompass a greater level of behavioural and contextual detail than pre-fixed or retrospective assessments, because they allow observers more freedom, scope and versatility in integrating and labelling what they see. Yet this freedom may also seem to bring with it an apparent excess of descriptors, and an ensuing ambiguity in the labelling of dimensions. However, it is important to realise that qualitative methodologies tend not to follow reductionist principles of organisation; descriptors are not meant to designate separate, sharply delineated, causal factors, but complementary, overlapping, mutually-enhancing aspects of the whole organism. Rather than be confused by the multitude of terms, the idea is to perceive the meaning expressed



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

**Figure 2**




Observer word charts. Shown as examples are the word charts of observer 3 (left) and observer 12 (right). The axes of these word charts show the first three principle dimensions of the consensus profile and indicate which of each particular observer terms best correlate with those dimensions.

through them. Only further empirical investigation can tell whether such conceptual richness enhances the clarity of scientific information, or detracts from it.

Our guiding hypothesis is that direct qualitative assessments of behaviour could potentially serve as an effective guide for integrating and interpreting traditional quantitative measures, particularly in cases where behaviour requires interpretation in terms of welfare (Wemelsfelder *et al* 2001; Rousing & Wemelsfelder 2006). In this light, using FCP methodology to investigate the welfare status of dogs appears the logical next step. As a companion animal, the domestic dog's welfare may be compromised by owner-



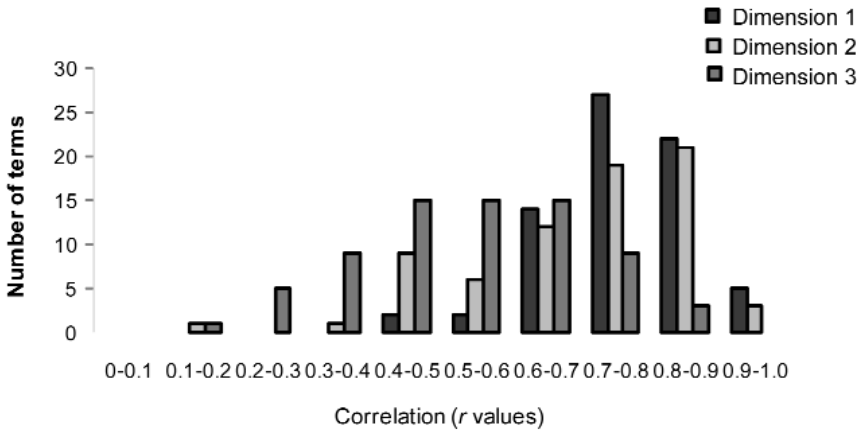
https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

**Table 2 Terms for each of the 18 observers that showed the two highest positive and negative correlations with dimensions 1, 2 and 3 of the consensus profile. Values inside parentheses give number of observers using that term.**

| Dimension | Positive | Negative |
|---|---|---|
| 1 | Playful (6), Happy (4), Confident (3), Friendly (3), Attention-seeking (3), Demanding (2), Excited (2), Affectionate, Comfortable, Alert, Sociable, Affectionate, Attentive, Responsive, Relaxed, Content, Unconcerned, Interactive, Solicitous, Engaged | Nervous (5), Unsure (4), Tense (3), Anxious (3), Cautious (3), Shy (2), Hesitant (2), Weary (2), Aloof (2), Functional, Reluctant, Curious, Distracted, Submissive, Mellow, On edge, Serious, Insecure, Fearful |
| 2 | Alert (5), Inquisitive (4), Investigative (4), Confident (4), Independent (3), Curious (3), Dominant (2), Explorative (2), Aware, Nosey, Restless, Aloof, Responsive, Territorial, Active, Disinterested, Boisterous | Attention-seeking (3), Quiet (3), Unsure (2), Submissive (2), Relaxed (2), Alert (2), Calm, Timid, Insecure, Nervous, Afraid, Stressed, Dependent, Needy, Anxious, Nervous, Weary, Content, Depressed, Passive, Sedate, Docile, Affectionate, Friendly, Cautious, Submissive, Appeasing, Seeking reassurance |
| 3 | Playful (8), Nervous (4), Boisterous (3), Unsure (2), Dominant (2), Attention-seeking (2), Excited (2), Flighty (2), Anxious (2), Alert (2), Exciting, Jumping, Cautious, Hyperactive, Inquisitive, Investigative, Aware | Calm (12), Relaxed (8), Content (2), Tense, Confident, Gentle, Seeking Reassurance, Disinterested, Confident, Docile, Placid, Aloof, Quiet, Dominant, Mellow, Reserved, Functional |

**Figure 3**

Correlation of observer terms with dimensions 1, 2 and 3.



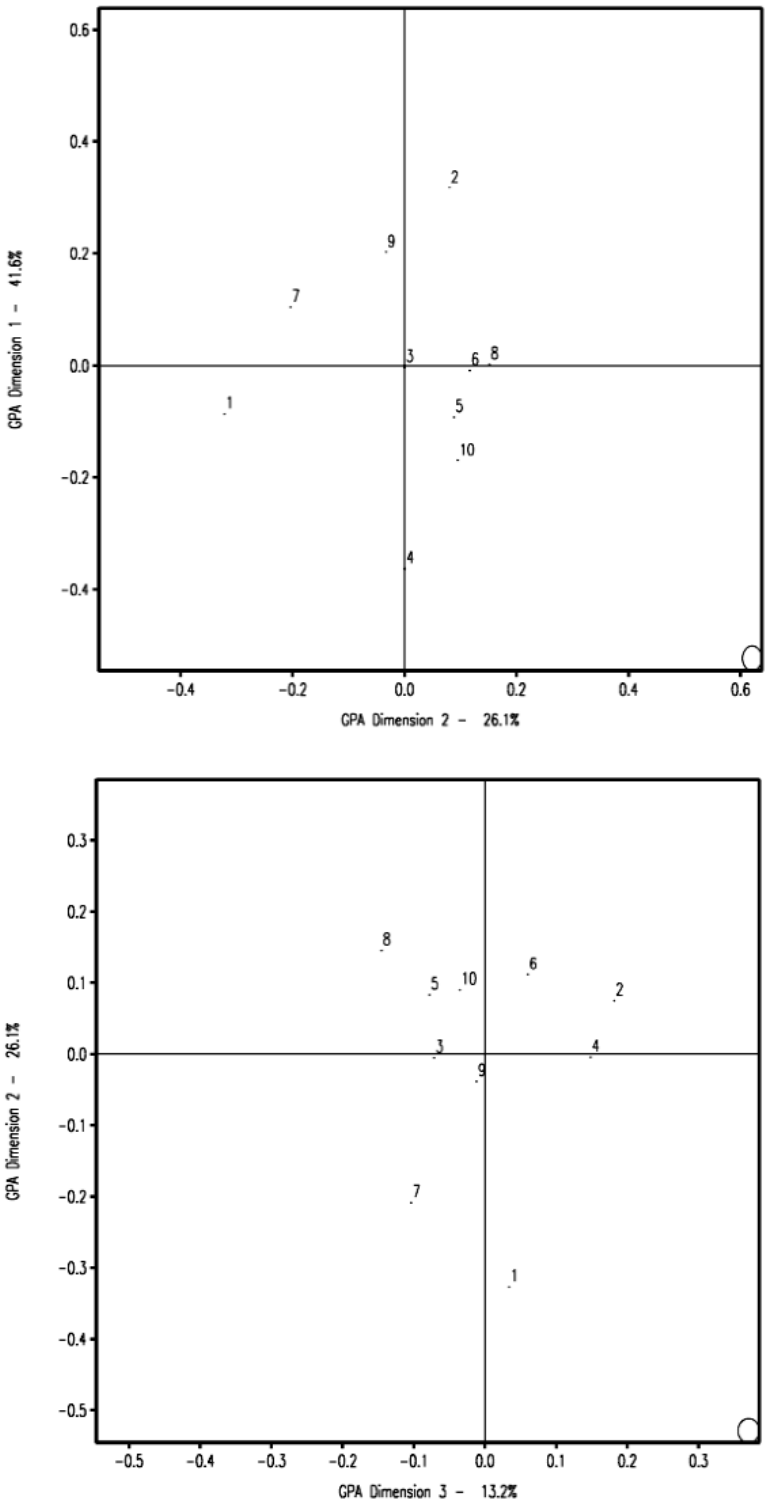

intensified behavioural problems such as separation anxiety, aggression problems or over-feeding resulting in obesity (Hubrecht 1995; Takeuchi *et al* 2001). Behavioural problems are a major cause of abandonment, euthanasia and relinquishment of companion dogs to shelters (Goodloe & Borchelt 1998). The domestic dog is also at great risk of suffering poor welfare within shelters and laboratories (Hubrecht 1995; Sales *et al* 1997; Wells 2004). Despite such problems, the welfare of domestic dogs is relatively understudied. So far, research has focused predominately on the investigation of enrichment ideas to increase the welfare status of confined dogs (eg Wells 2004; Graham *et al* 2005), and on the investigation of stress levels in dogs kept in different types of confinement (Sales *et al* 1997; Beerda *et al* 1999; Hennessy *et al* 2001). This research found, for example, that dogs in animal shelters are significantly stressed during the first three days of their stay, and that poor housing conditions also increase a dog's stress levels, as reflected in increases in so-called 'negative' behaviours (Beerda *et al* 1999). In such studies, qualitative behavioural assessment of a dog's emotional experience may make a useful contribution to establishing an animal's welfare status quickly, reliably and non-invasively.

The Free Choice Profiling approach to qualitative behaviour assessment may also be applied to pain assessment. Studies investigating pain assessment in dogs often rely upon facial expression and body language to interpret the degree of pain severity (eg Light *et al* 1993; Firth & Haldane 1999; Hardie 2000; Holton *et al* 2001; Wiseman *et al* 2004, 2006). Similarly, in human infant patients that cannot communicate, assessment relies on interpreting facial expressions and behaviours which have been categorised to devise pain-scoring systems (Holton *et al* 2001). However, the qualitative rating scales used in such studies tend to be based on pre-selected terms given to observers to assess the animal or human subject (eg how 'uncomfortable' is the cat?). Research has indicated that such pain scales are not always



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

**Figure 4**




Dog plots. Showing the distribution of the 10 dogs along dimensions 1, 2 and 3.

reliable in that different observers can attribute markedly different scores to the same animal (Hansen 1997). Wiseman *et al* (2004, 2006), however, found that qualitative descriptors of behaviour generated by the owners of dogs could serve as useful and reliable indicators of chronic pain in these dogs. Similarly, the research presented here suggests that if observers are allowed to generate their own descriptive terminology during observation, good



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press

agreement in characterising the behavioural expressions of dogs can be achieved. It would, thus, be interesting to further explore the use of an FCP approach to animal and human pain assessment.

This study focused on one breed of dog (Beagle), and the validity of its outcomes with regards to other breeds remains to be tested. Many of the morphological modifications of modern breeds of domestic dog can be explained by changes in the rate of development during domestication from the wolf. These changes have been dominated by paedomorphism and neoteny, so that the adult passes through fewer growth stages and physically resembles a juvenile stage of its ancestor, while at the same time their behaviour also resembles this juvenile state (Goodwin *et al* 1997; Vas *et al* 2005). Such effects may well influence the ability of observers to accurately interpret a dog's emotional expression, because the expression of different breeds in the same situation (eg a wire-haired Fox Terrier in comparison to a Shar Pei or Siberian Husky) is likely to be different. Thus, to further investigate the applicability of this methodology to welfare and pain assessment in the domestic dog, the large variation in morphology within the species must be taken into account.

Finally, several studies suggest that males and females differ significantly in their attitudes towards animals. Females are generally more empathic than males and ascribe higher sentience scores (Paul & Podberscek 2000; Daly & Morton 2006), and in general have more positive views of animals (Herzog 2007). The observer group in the current study consisted entirely of female dog-experienced observers, and it is therefore likely that gender-bias will have affected the reported results. Previous research using FCP methodology has similarly consisted of greater numbers of female than male observers (Wemelsfelder *et al* 2000, 2001; Napolitano *et al* 2008). Thus, future research should make an effort to investigate the effects of gender on the outcomes of qualitative behaviour assessments generated through FCP.

### Animal welfare implications

In conclusion, the results of this study indicate that the qualitative behaviour assessment methodology originally developed and tested for farm animals, also shows good inter-observer reliability when applied to the domestic dog (Beagles). This research should now be extended across a range of dog breeds, and should investigate the usefulness of this methodology in assessing the welfare status of dogs in different environments and treatments. Recently, in scientific literature, there has been growing interest in assessing positive and negative aspects of animal welfare and overall quality of life. The use of qualitative behaviour assessment incorporates both positive and negative aspects of an animal's expressive repertoire, and may well be particularly useful in supporting the interpretation of physical and physiological measurements of behaviour in terms of an animal's emotional experience, thereby providing an effective tool for welfare assessment in dogs and a range of species.

## Acknowledgements

We would like to thank Carol Lockett, Alan Willox and the Ministry of Agriculture and Forestry for their assistance with this research.

## References

**Arnold GM and Williams AA** 1985 The use of generalized Procrustes Techniques in sensory analysis. In: Piggott JR (ed) *Statistical Procedures in Food Research* pp 233-253. Elsevier: London, UK

**Beerda B, Schilder BH, Bernadina W, Van Hoof JARAM, De Vries HW and Mol JA** 1999 Chronic stress in dogs subjected to social and spatial restriction 1. Behavioural responses. *Physiology and Behavior 66*: 233-242

**Clarke SA and Boinski S** 1995 Temperament in non human primates. *American Journal of Primatology 37*: 103-125

**Daly B and Morton LL** 2006 An investigation of human-animal interactions and empathy as related to pet preference, ownership, attachment and attitudes in children *Anthrozoology 19(2)*: 113-127

**Fagen R, Conitz J and Kunib E** 1997 Observing behavioral qualities. *International Journal of Comparative Psychology 10*: 167-179

**Feaver J, Mendl M and Bateson P** 1986 A method for rating the individual distinctiveness of domestic cats. *Animal Behaviour 34*: 1016-1025

**Firth AM and Haldane SL** 1999 Development of a scale to evaluate postoperative pain in dogs. *Journal of the American Veterinary Medicine Association 214(5)*: 651-659

**Gains N and Thomson DMH** 1990 Contextual evaluation of canned lagers using repertory grid method. *International Journal of Food Science and Technology 25*: 699-705

**Goddard ME and Beilharz RG** 1984 A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science 12(3)*: 253-265

**Goodloe LP and Borchelt PL** 1998 Companion dog temperament traits. *Journal of Applied Animal Welfare Science 1(4)*: 303-338

**Goodwin D, Bradshaw JWS and Wickens SM** 1997 Paedomorphosis affects agonistic visual signals of domestic dogs. *Animal Behaviour 53(2)*: 297-304

**Gosling SD** 2001 From mice to men: What can we learn about personality from animal research? *Psychological Bulletin 127*: 45-86

**Gower JC** 1975 Generalized Procrustes Analysis. *Psychometrica 40*: 35-51

**Graham L, Wells DL and Hepper PG** 2005 The influence of olfactory stimulation on the behaviour of dogs housed in a rescue shelter. *Applied Animal Behaviour Science 91*: 143-153

**Hansen B** 1997 Through a glass darkly: Using behavior to assess pain. *Seminars in Veterinary Medicine and Surgery (small animal) 12(2)*: 67-74

**Hardie EM** 2000 Recognition of pain behaviour in animals. In: Hellebrekers LJ (ed) *Animal Pain* pp 51-69. W van der Waas Uitgevery: Utrecht, The Netherlands

**Hebb DO** 1946 Emotion in man and animal: an analysis of the intuitive processes of recognition. *Psychological Review 53*: 88-106

**Hennessy MB, Voith VL, Mazzei SJ, Buttram J, Miller DD and Linden F** 2001 Behaviour and cortisol levels of dogs in a public animal shelter, and an exploration of the ability of these measures to predict problem behaviour after adoption. *Applied Animal Behaviour Science 73(3)*: 217-233

**Herzog HA** Gender difference in human-animal interactions: a review. *Anthrozoology 20(10)*: 7-21

**Holton L, Scott EM, Nolan AM, Reid J, Welsh E and Flaherty D** 1998 Comparison of three methods used for assessment of pain in dogs. *Journal of the American Veterinary Medicine Association 212*: 61-67

**Holton L, Reid J, Scott EM, Pawson P and Nolan A** 2001 Development of a behaviour based scale to measure acute pain in dogs. *The Veterinary Record 148*: 525-531

**Hubrecht R** 1995 The welfare of dogs in human care. In: Serpell J (ed) *The Domestic Dog. Its Evolution, Behaviour and Interactions with People*. Cambridge University Press: Cambridge, UK

**Kent JE, Jackson RE, Molony V and Hosie BD** 2000 Effects of acute pain reduction on chronic inflammatory lesions and behaviour of lambs castrated and tail docked with rubber rings at less than two days of age. *The Veterinary Journal 160*: 33-41

**Light GS, Hardie EM, Young MS, Hellyer PW, Brownie C and Hansen BD** 1993 Pain and anxiety behaviors of dogs during intravenous catheterization after premedication with placebo, Acepromazine or Oxymorphone. *Applied Animal Behaviour Science 37*: 331-343

**Minero M, Tosi MV, Canali E and Wemelsfelder F** 2009 Quantitative and qualitative assessment of the response of foals to the presence of an unfamiliar human. *Applied Animal Behaviour Science 116*: 74-81

**Molony V, Kent JE and McKendrick IJ** 2002 Validation of a method for assessment of an acute pain in lambs. *Applied Animal Behaviour Science 76*: 215-238

**Murphy JA** 1998 Describing categories of temperament in potential Guide Dogs for the Blind. *Applied Animal Behaviour Science 58*: 163-178

**Napolitano F, De Rosa G, Braghieri A, Grasso F, Bordi A and Wemelsfelder F** 2008 The qualitative assessment of responsiveness to environmental challenge in horses and ponies. *Applied Animal Behaviour Science 109(2-4)*: 342-354

**Noble AC and Ebeler SE** 2002 Use of multivariate statistics in understanding wine flavor. *Food Review International 18*: 1-21

**Oreskovich DC, Klein BP and Sutherland JW** 1991 Procrustes Analysis and its applications to Free-Choice and other Sensory Profiling. In: Lawless HT and Klein BP (eds) *Sensory Science: Theory and Applications in Foods* pp 353-393. Marcel Dekker: New York, USA

**Paul ES and Podberscek AL** 2000 Veterinary education and students attitudes towards animal welfare. *Veterinary Record 146*: 269-272

**Rousing T and Wemelsfelder F** 2006 Qualitative Assessment of social behaviour of dairy cows housed in loose housing systems. *Applied Animal Behaviour Science 101(1-2)*: 40-53

**Saba A and Rosati S** 2002 Understanding consumer perception of fermented yoghurt products using Conjoint and Generalised Procrustes Analysis. *Italian Journal of Food Science 14*: 339-350

**Sales G, Hubrecht R, Peyvandi A, Milligan S and Sheild B** 1997 Noise in dog kennelling: Is barking a welfare problem for dogs? *Applied Animal Behaviour Science 52*: 321-329

**Serpell J and Hsu Y** 2001 Development and validation of a novel method for evaluating behaviour and temperament in Guide Dogs. *Applied Animal Behaviour Science 72*: 347-364

**Stevenson-Hinde J** 1983 Individual characteristics: a statement of the problem. In: Hinde RA (ed) *Primate Social Relationships: an Integrated Approach* pp 28-34. Blackwell Scientific Publications: Oxford, UK

**Stevenson-Hinde J, Stillwell-Barnes R and Zunz M** 1980 Subjective assessment of rhesus monkeys over four successive years. *In Primates 21*: 66-83

**Svartberg K and Forkman B** 2002 Personality traits in the domestic dog (*Canis familiaris*). *Applied Animal Behaviour Science 79*: 133-155

**Takeuchi Y, Ogata N, Houpt KA and Scarlett JM** 2001 Differences in background and outcome of three behaviour problems of dogs. *Applied Animal Behaviour Science 70*: 297-308

**Vas J, Topal J, Gacsi M, Miklosi A and Csanyi** 2005 A friend or an enemy? Dogs' reaction to an unfamiliar person showing behavioural cues of threat and friendliness at different times. *Applied Animal Behaviour Science 94*: 99-115

**Weiss A, King JE and Perkins L 2006** Personality and subjective well-being in orangutans (*Pongo pygmaeus and Pongo abelii*). *Journal of Personality and Social Psychology 90(3)*: 501-511

**Wells DL** 2004 A review of environmental enrichment for kennelled dogs, *Canis familiaris*. *Applied Animal Behaviour Science 85*: 307-317

**Wemelsfelder F** 1997 The scientific validity of subjective concepts in models of animal welfare. *Applied Animal Behaviour Science 53*: 75-88

**Wemelsfelder F** 2007 How animals communicate quality of life: the qualitative assessment of animal behaviour. *Animal Welfare 16(S)*: 25-31

**Wemelsfelder F, Hunter TEA, Mendl MT and Lawrence AB** 2000 The spontaneous Qualitative Assessment of behavioural expressions in pigs: First explorations of a novel methodology for integrative animal welfare measurement. *Applied Animal Behaviour Science 67*: 193-215

**Wemelsfelder F and Lawrence AB** 2001 Qualitative Assessment of animal behaviour as an on-farm welfare-monitoring tool. *Acta Agriculturae Scandinavica: Section A, Animal Science 51(30)*: 21-25

**Wemelsfelder F, Hunter TEA, Mendl MT and Lawrence AB** 2001 Assessing the 'whole animal': a Free Choice Profiling approach. *Animal Behavior 62*: 209-220

**Wiseman-Orr ML, Nolan AM, Reid J and Scott EM** 2004 Development of a questionnaire to measure the effects of chronic pain on health-related quality of life in dogs. *American Journal of Veterinary Research 65(8)*: 1077-1084

**Wiseman-Orr ML, Scott EM, Reid J and Nolan AM** 2006 Validation of a structured questionnaire as an instrument to measure chronic pain in dogs on the basis of effects on health-related quality of life. *American Journal of Veterinary Research 67(11)*: 1826-1836



https://doi.org/10.1017/S0962728600001196 Published online by Cambridge University Press