Title
The Development of a Psychometric Scale for the Evaluation of the Emotional Predispositions of Pet Dogs

Permalink

Journal

ISSN

Authors
Sheppard, Gill
Mills, Daniel S.

Publication Date
2002

DOI

Copyright Information

Peer reviewed

The Development of a Psychometric Scale for the Evaluation of the Emotional Predispositions of Pet Dogs

Gill Sheppard and Daniel S. Mills
University of Lincoln, United Kingdom

Many pet dogs exhibit problem behaviours which can be corrected through the judicious use of positive and negative reinforcers in training. However, animals often vary in their sensitivity and response to these. It is hypothesised that this variation may offer a biological basis for discriminating between animals that develop certain types of problems (eg fears and phobias) and their response to treatment. The development of a clinical tool that uses owner report to measure individual differences in positive and negative activation in pet dogs is described. The activation scales each measured a single homogeneous construct and effectively differentiated between individuals, demonstrating variation on two dimensions. Test-retest reliability was good and the study provided evidence of validity. Normative data were calculated and may be used in future research that could provide further evidence of validity and in investigations of the underlying structure of canine behaviour disorders. The final scales comprised of 21 items, and so can be administered with little difficulty.

Veterinary behavioural medicine is the diagnosis and treatment of behaviour disorders in animals under the auspices of the veterinary profession. In 1997, the first global conference was held for this emerging field of veterinary science (Mills et al., 1997) and a disparity between alternative approaches to the conceptualisation of behaviour problems was highlighted. It is apparent that in many ways, the discipline has developed along similar lines to the human psychiatric field, but it has the potential to learn from the experiences of this related discipline and recent advances in psychobiology (Sheppard & Mills, 2003). Thus, in contrast to many other streams of comparative psychology, this field is developing through the use of methods and concepts developed in the human field to explore the mind of other animals, more than the reverse.

An area of growing importance, in both the veterinary and human medical fields, is the need for comprehensive theories that link behaviour, affect, motivation and neurobiology (Panksepp, 1998). For example, Depue and Collins (1999) reviewed evidence that links individual differences in positive incentive motivation with variation in dopamine activity and personality. This model also implies that individuals at the extremes of the dimension are more likely to exhibit behaviour, which is inappropriate and problematic in a range of circumstances and so are more likely to present with signs of "clinical mental illness". Psychometric interviews may be developed to quantify such traits, but must be based on an understanding of the underlying psychobiology of the construct to be measured. This has

been used in humans to assess responsiveness to rewarding and aversive experiences (positive and negative activation; e.g. Carver & White, 1994), but it has not previously been developed in veterinary behavioural medicine for the clinical evaluation of patients. It is hypothesised that such an assessment may serve several useful functions in clinical practice. First it may help to predict at risk individuals; for example, individuals with a greater sensitivity to aversion might be expected to be at greater risk of developing fears, phobias and anxiety problems. Second, pet dogs normally undergo some form of training in their life; this involves the judicious use of rewards and/or punishment and at present this choice of reinforcement is often based on the personal predispositions of the trainer. Some claim to only use reward based techniques, whilst others find punishment acceptable (Marschark & Baenninger, 2002). In reality most trainers use a combination of reward and punishment, even if they do not use physical punishment (Mills, 2002). It is hypothesised that an assessment of positive and negative activation, will allow a more objective choice of training style, tailoring the type and magnitude of reinforcement to the subject being trained. This would represent an advance in both routine dog training as well as behaviour therapy programmes in which an animal already has a behaviour problem. Third, in any behaviour modification programme, some animals respond well and others do not for reasons beyond owner compliance. It is hypothesised that differences in emotional sensitivity may underlie at least some of this difference in response, with animals at the extremes of sensitivity perhaps being less likely to make a sustainable response within the realms of normality. An assessment of positive and negative activation may therefore allow greater accuracy in prognostication in any given case. There are currently few prognostic instruments beyond those which have been identified from retrospective epidemiological studies (e.g., Reisner et al. 1994; Takeuchi et al., 2000). On a related note, it has recently been reported that shyness-boldness as assessed from behavioural tests is predictive of performance in working dog trials (Svartberg, 2002). This trait is probably a composite of positive and negative activation and so further elucidation of the structure of the trait, may also allow more specific prediction of performance in normal animals as well.

The psychobiological approach proposed here for the development of the general assessment of pet dogs is fundamentally different to the principals guiding the choice of behaviours in the development of many temperament tests in dogs (e.g., Campbell, 1972; Netto & Planta, 1997; Wilsson & Sundgren, 1998). Sometimes purely descriptive behaviour categories are used and these may lead to the inclusion of behavioural expressions with a variety of motivations, affecting the validity of the subsequent scale (Martin & Bateson, 1986); others focus on predicting the behaviour of animals destined for a specific function in a specific environment. These are inevitably limited in their utility in other contexts. By contrast, a focus on identifying behaviour traits with clear psychobiological foundations in the concepts of positive and negative activation should produce a valid scale which allows behavioural tendencies to be predicted over a wider range of environmental contexts as defined by their reinforcing properties. This is particularly important in veterinary behavioural medicine as pet animals are kept in a variety of environments that frequently change, e.g. a dog may be at home with familiar people one minute and then out in an unfamiliar environment with strange dogs the next.

Broad range predictability is therefore important to the practical utility of a tool for assessing pets.

Although non-human animals are not amenable to direct interview, psychometric profiling using owner/carer interview is possible (Gosling & John, 1999) and it has been argued that comparative studies can also make a serious scientific contribution to the study of personality (Gosling, 2001). The development of an instrument for the assessment of emotional functioning in companion animals could also be used to guide diagnosis, treatment recommendation and prognostication for patients with behaviour problems and help assess their welfare in other contexts.

Internal consistency and unidimensionality are important to the validity of such a scale (Smith & McCarthy, 1995). Internal consistency refers to the mean level of inter-item correlation within a scale whereas unidimensionality indicates that items are measuring a single underlying construct (Clark & Watson, 1995). Measures of internal consistency may conceal the absence of significant correlations between some items if sufficiently strong. Thus, internal consistency alone cannot indicate unidimensionality, and an additional examination of the range of inter-item correlations is required (Clark and Watson, 1995).

Therefore our aim was to develop a psychometric interview of emotional functioning in pet dogs for use primarily in veterinary behavioural medicine, based on current theories of positive and negative activation. A rigorous developmental methodology is an essential prerequisite to the acceptance of a useful clinical tool and so, in keeping with the tradition in clinical psychology (e.g., Watson et al., 1988), we emphasise here first, the procedure adopted during the development of this instrument. We focussed on the constructs of positive and negative activation for two reasons. Firstly because of the practical importance of reward and punishment in dog training as has already been described above, and secondly because these constructs have broad scientific support as fundamental emotional systems which are likely to occur in a range of mammalian species, including dogs. The fundamental importance of the tendency to boldness-shyness as an influence on behaviour in a range of test situations has also recently been reported by Svartberg and Forkman (2002). This does not however preclude the future addition of other scales for other traits, in order to develop the tool described here further for use in both broad and narrow range contexts. The efficacy of the scale in a range of clinical contexts will be described in future publications.

Methods

Item Generation

A literature review identified mood descriptors, personality traits, behavioural tendencies and eliciting stimuli that were associated with positive and negative activation. A broad item pool was generated from this information, based mainly on the following publications: Depue and Collins (1999), Panksepp (1998), Gray (1983, 1994), Watson et al., (1988, 1999). The main objective was that each item should measure either positive or negative activation, but should not reflect both constructs. Items within the broad item pool then had to be assessed for their ability to be readily applied to a typical domestic setting which could be easily and unambiguously identified by the owner.

Items that described behavioural responses to specific situations appeared to have greatest objectivity. However, some behavioural responses may be subject to the effects of training and other environmental constraints, preventing these expressions of positive and negative activation. The in-

clusion of mood descriptors, although much more subjective, could enhance the schedule's ability to measure the proposed constructs. Items were selected by an ultimate consensus from both authors to represent normal "nonproblematic" affect and behaviour as well as the extreme variants that may be associated with disorder, as evidenced by the second author's clinical experience. Thus the range of scores from the assessment instrument should represent the full extent of variation in positive and negative activation.

The questionnaire was administered to a small pilot sample, identifying areas of ambiguity and revision. The prototype questionnaire included 45 items, numbered 1-45 that were hypothesised to measure positive and negative activation (Appendix 1). Demographic data (e.g., age and sex) was also gathered.

Scoring Responses

Scoring methods reflected the hypothesis that high scorers on the dimensions of positive or negative activation would experience the affective state more frequently and intensely than low scorers (Watson et al., 1994). Therefore it was not appropriate to use a scale that referred to frequency alone. Participants were presented with a series of statements and asked to rate their level of agreement with each. Responses were recorded on a 5-point Likert scale, having "strongly agree" (scores 5) and "strongly disagree" (scores 1) at the extremes and "partly agree, partly disagree" at the midpoint. A "not applicable" option was also provided.

Each construct was represented by some positively-worded items and some that were negatively-worded. Thus extreme levels of a construct would be indicated by strong agreement with some items and strong disagreement with others. This format was used to reduce the possibility of a response set (DeVellis, 1991), reflecting acquiescence bias. Scores for items were later reversed as appropriate to ensure that the total scores for each scale would be meaningful.

Participants

Questionnaires were distributed on the university campus (to staff, students and clients of the animal behaviour clinic), at obedience classes, dog shows and at a large retail outlet that specialises in provisions for pets. A stamped addressed envelope was provided with those that were distributed externally. The sample was selected to include a wide range of individuals to reflect the variation in the constructs to be measured, including the norm and the more extreme scores that may be associated with problem behaviours.

Statistical Analyses

All statistical analyses were conducted using Minitab 12.1 (Minitab Ltd).

Missing Data. Complete data sets were required to compute multivariate statistics as the many statistical software packages delete all of the data from questionnaires that have missing responses (Figueredo et al., 2000). Such deletion of data was not desirable as the number of items that were included within the questionnaire determined that a large number of participants were required in order to conduct the statistical analysis. Floyd and Widaman (1995) suggest that a participant to item ratio of at least 4:1 is traditionally recommended for exploratory factor analysis and Kline (1994) recommends that the ratio should be as large as possible. Additionally deletion may introduce bias if, for example, all dogs that were never walked off a lead were excluded from the sample. Inclusion of these items may be necessary for each of the constructs to be fully represented. Thus it was felt necessary to impute values for missing data.

The method of multiple imputation (Longford et al., 2000) was used to complete data records that had missing responses. Items were organized into groups, each consisting of between three and five theoretically similar statements. For example, the items 1, 13, 16 and 41 (appendix 1) represent one of these groupings. The complete data records were examined to determine the distribution of response scores within each of these item groups. This made it possible to calculate the relative frequencies of alternative scores for an item, based on the scores for other items in that grouping. A score was then imputed using random number generation and the probability associated with each of the possible alternative scores. Thus a score was imputed for a missing response, using other data that

the participant had provided and the distribution of response scores in the complete data sets of other participants.

Records were deleted and missing data were not imputed if less than 40 of the 45 items were completed.

Five alternative values were imputed for each missing datum, resulting in five complete data sets that were each subjected to statistical analysis. This facilitated comparisons of results to ensure consistency. Differences between data sets would indicate that the imputation of missing values had influenced the results and be a cause for significant concern in evaluating the validity and utility of the final scale proposed.

Distribution of Item Response Scores. Descriptive statistics were used to identify the distributions of response scores for each item. Items that displayed gross deviations from a normal distribution were initially excluded from further analyses as such items poorly differentiated between individuals. Items were initially excluded if the mean score was greater than 4.0 and the median equalled 5.0 or if the mean was less than 2.0 and the median equalled 1.0.

Scale Development: Positive and Negative Activation. Pearson product moment correlation coefficients were used to construct a correlation matrix of general response scores, which was then subjected to a principal component analysis. The Eigenvalues were plotted against the order of extraction, and visual inspection, using the rules of the scree test (Cattell, 1966) indicated the number of principal components for extraction. A principal components factor analysis with varimax rotation was then used to extract the appropriate number of factors. Varimax rotation identifies orthogonal simple structure (Kline, 1994). This method was selected as the constructs to be measured were hypothesised to be independent and the law of parsimony dictates that simple solutions are preferred.

In order to maximise the specificity of the items, they were retained if their loading on one of the first two factors was greater than 0.35 and their loading on the other factor was relatively weak (i.e., < 0.20). Items that did not fulfil these criteria were deleted.

Scale development allowed the reintroduction of items that were initially excluded due to skewed responses. This increased the coverage of the scales, increasing the likelihood that they would fully reflect the constructs being assessed.

The principal components factor analysis (described above) was repeated after each addition or deletion of items, resulting in further amendments. This continued until all remaining items loaded significantly on one factor and weakly on the other.

The final items of each scale appeared to emphasise some facets of positive and negative activation more than others. This suggested that it may be appropriate to group homogeneous items together, facilitating their representation by a mean score ("facet score"). This approach may prevent scales from reflecting a single facet more than the construct as a whole. Facet scores are also likely to have greater reliability and deviate less from a normal distribution than individual item scores (Reise et al., 2000). Scales were analysed separately, using the multivariate statistical techniques described above and alternative solutions were explored (i.e. the extraction of two, three and four factors). The preferred solution was defined as one where mean item-item correlations within facets were greater than mean itemitem correlations between facets (Clark & Watson, 1995). It also had to provide a meaningful representation of the data.

A final principal components analysis was conducted on the remaining response scores. A scree test (Cattell, 1966) identified the number of principal components to be extracted and a principal components factor analysis with varimax rotation was used for this purpose. These final multivariate statistics were repeated with each of the data sets to facilitate comparison of the results.

The mean score per scale was calculated for each participant's data record, by averaging item or facet scores.

Test-Retest Reliability. The initial questionnaire asked participants to indicate whether they would be willing to complete a follow up questionnaire in the future. One third of those that indicated willingness were selected by removing every third questionnaire. The second questionnaire was dispatched six weeks after the initial response was received. It included the same items as the first schedule but the order had been changed. Participants were asked to ensure that the same individual completed the questionnaire on both occasions.

This approach provided two sets of data for some participants. A Pearson product-moment correlation coefficient was calculated for test-retest item scores to provide a reliability coefficient (Rust & Golombok, 1989). Any items that did not display a significant correlation were removed.

Additionally, Wilcoxon signed rank tests were used to compare the median test-retest scores for each item and significant differences resulted in deletion.

These statistical analyses were repeated with facet scores and scale scores. Cronbach's alpha coefficient was not calculated as an estimate of internal consistency as this is based on mean interitem correlation and so a moderate value could potentially conceal some extreme individual levels of correlation. However, visual inspection of inter-item correlations was conducted to determine whether each fell within an acceptable range (see below). This provided a more appropriate means of maximising unidimensionality whilst avoiding the inclusion of redundant data.

Internal Consistency and Unidimensionality. The mean interitem correlation for each scale was calculated to measure internal consistency. The range of individual interitem correlations was additionally examined to assess the unidimensionality of each scale. Clark and Watson (1995) recommend that "virtually all" interitem correlations within a scale should range from 0.15 to 0.50 to maximise unidimensionality (correlations < 0.15) but avoid redundancy (correlations > 0.50). Interitem correlations within each scale were compared to this range.

Standardisation. Scale scores were standardised, using *z* conversion (Cronbach, 1990). Thus each raw score was expressed in terms of the number of standard deviations above or below the mean score for the scale [i.e., (score – mean)/standard deviation].

Results

Response Rate

In total, 775 questionnaires were distributed and 358 (46.19%) were returned. Table 1 describes the extent of missing data. Items that had fewer responses tended to be those that described a situation that a dog may never experience (e.g., being walked off a lead). The responses of seventeen participants were removed before analyses due to missing data, leaving a total of 343, of which 125 had missing responses.

The second questionnaire was sent to 105 of the original participants, and 78 (74.29%) were returned.

Table 1
The Extent of Missing Data within the Sample.

No. of items completed	No. of dogs	%
0	3	0.84
1 - 4	0	0.00
5 - 9	0	0.00
10 - 14	0	0.00
15 - 19	1	0.28
20 - 24	0	0.00
25 - 29	2	0.56
30 - 34	2	0.84
35 - 39	7	2.23
40 - 44	125	34.92
45	218	60.34
Total	358	100

Imputing Missing Data

In total 209 item responses were imputed to complete the records (1.35% of general data). Multiple imputation resulted in five complete data sets. These will be referred to as "data set 1", "data set 2", etc.

Demographic Data

The demographic data (Tables 2, 3) indicated that the sample included dogs that varied greatly in their age and the length of time that they had been with their current owner. Both sexes appeared to be adequately represented and a large proportion (40.82%) was reported to have a behaviour problem.

Table 2
Sexual status of dogs and frequency of behaviour problems.

Sex		n	%
Male	Neutered	71	20.70
	Entire	113	32.94
Female	Neutered	89	25.95
	Entire	67	19.53
No data		3	0.87
Behaviour problem			
Yes		140	40.82
No		195	56.85
No data		8	2.33

Table 3
Age of Dog and Length of Time with Current Owner.

	Mean	Maximum	Minimum	Standard deviation	No data
Age (years)	4.70	17	0.25	3.89	4
Length of ownership (years)	4.05	16	0.25	3.76	8

Scale Development: Positive and Negative Activation (Data Set 1)

One item [31, "Your dog trembles but makes little attempt to get away when you try to perform any procedure that it dislikes (e.g., giving medication,

grooming, or bathing)] was removed before any analysis was conducted, as it became apparent that it could be ambiguous.

Almost all of the items had response scores that were skewed to a greater or lesser degree. Fifteen items (3, 4, 5, 9, 11, 19, 22, 23, 28, 30, 34, 37, 38, 44, 45) were provisionally excluded as their response scores displayed gross deviations from normal distributions.

The Eigen analysis of the correlation matrix of the remaining items following principal components analysis indicated that the first and second components accounted for 18.9% and 11.8%, respectively, of the total variability in the correlation matrix. Each of the remaining principal components accounted for less than 7% with the sixth principal component and beyond each accounting for less than 4%. The scree plot of the Eigenvalues indicated that a two-factor solution was appropriate for the data. From the load plot of the two factor principal components factor analysis with varimax rotation, seven items did not appear to load strongly on only one of the first two factors (items 6, 8, 12, 13, 17, 32, 35, 42) and so were removed. Many of the items that referred to exploratory behaviour (including response to novel stimuli) were included among these deletions.

Six of the previously deleted items (3, 19, 22, 28, 30, 34) were reintroduced to increase the coverage of factor two and the principal components factor analysis with varimax rotation repeated. This resulted in the deletion of three items (3, 26, 28). Two of these items referred to the dogs' response to novelty (i.e., unfamiliar surroundings and unfamiliar people).

Preliminary reliability analyses identified significant differences in the median test-retest scores of items 21 and 33. These items were removed and the two factor principal components factor analysis with varimax rotation repeated. This resulted in the further deletion of item 43.

The correlation matrix of the remaining items was subject to a principal components analysis. The Eigen analysis indicated that the first two principal components accounted for a much greater proportion of the variation (36.7%) than subsequent principal components. Visual inspection of the scree plot also indicated that there were two main factors (Figure 1).
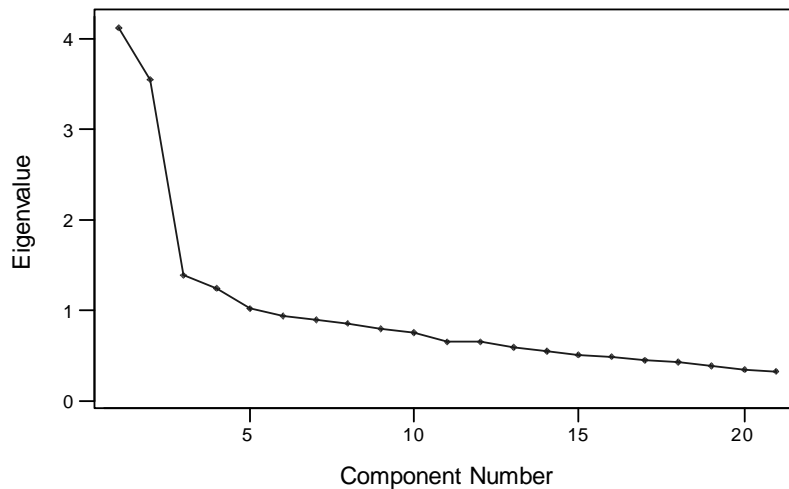
Figure 1. Scree plot of the principal components identified in the analysis of negative and positive activation items.

The two factor factor analysis with varimax rotation confirmed that the remaining items all loaded strongly (> 035) on one factor and weakly (< 0.20) on the other (Table 4). The nature of the remaining items suggested that factor 1 and factor 2 may correspond to negative and positive activation respectively.

Table 4
Rotated Factor Loadings of Positive and Negative Activation Item Scores.

	Activation	
Item	Negative	Positive
Your dog……		
*2 is rarely frightened	0.705	-0.151
7 is easily startled by noises and / or movements	0.681	0.048
15 appears nervous and / or jumpy for several minutes after it has been startled	0.665	-0.045
18 has a specific fear or phobia	0.654	-0.118
*20 appears calm in noisy, crowded places	0.691	0.024
24 is frightened by noises from the television or radio	0.594	0.117
*25 usually appears relaxed	0.569	0.099
*27 adapts quickly to changes in its environment (eg. being cared for by different people, moving house or a family member leaving home)	0.418	-0.124
29 appears afraid of the vacuum cleaner or any other familiar household appliance	0.499	-0.036
*39 appears calm in unfamiliar environments	0.677	0.060
40 appears unsettled by changes to its routine (eg. if it is not fed at the usual time, if it is left alone for longer than usual)	0.382	0.106
1 becomes very excited when it is about to go for a walk (eg. when it sees its lead, or when it hears "walkies", etc.)	-0.016	0.453
10 is very persistent in its efforts to get you to play	-0.030	0.635
*14 shows little interest in its surroundings	-0.150	0.407
16 is easily excited	0.152	0.658
19 tries to escape from the garden	0.146	0.353
22 is full of energy	-0.059	0.787
*30 is lazy	-0.021	0.576
*34 requires a great deal of encouragement to take part in energetic activities	-0.151	0.603
36 persists in being naughty despite being told off for the behaviour	0.167	0.461
41 is very boisterous	-0.019	0.779
Variance	4.136	3.562
% Var	0.197	0.170

Note. * Indicates reversed scored items.

Multivariate statistical analyses of the negative activation items did not identify facets of the domain or provide a useful structure for grouping items.

The principal components analysis on the correlation matrix of positive activation items suggested two, three and four-factor solutions which could be explored. Two and four factor solutions corresponded poorly to the hypothesised construct. The two-factor solution did not appear to adequately represent the different facets whereas a four-factor solution appeared to subdivide one of the facets. The extraction of three-factors using principal components factor analysis with varimax rotation provided the most meaningful representation of the data (Table 5). Additionally, mean item-item correlations within facets ranged from 0.29 to 0.35, whereas mean item-item correlations between facets were lower (0.22 – 0.24).

Table 5
Principal Components Factor Analysis with Varimax Rotation to Extract Facets of the Positive Activation Domain.

Item: Your dog…………….	Energy & interest	Persistence	Excitement
*14 shows little interest in its surroundings	0.561	0.234	-0.168
22 is full of energy	0.689	0.150	0.460
*30 is lazy	0.718	-0.090	0.227
*34 requires a great deal of encouragement to take part in energetic activities	0.770	0.035	0.133
10 is very persistent in its efforts to get you to play	0.369	0.465	0.320
19 tries to escape from the garden	-0.002	0.743	-0.030
36 persists in being naughty despite being told off for the behaviour	0.036	0.737	0.134
41 is very boisterous	0.441	0.491	0.429
1 becomes very excited when it is about to go for a walk (eg. when it sees its lead, or when it hears "walkies", etc.)	0.060	-0.041	0.787
16 is easily excited	0.186	0.334	0.676
Variance	2.268	1.753	1.691
% Var	0.227	0.175	0.169

Note. * Indicates reversed scored items.

The response distributions of facet scores were skewed but not as greatly as many of the item response distributions (Table 6). The apparent deviations from normal distributions did not indicate failure to differentiate between individuals.

Table 6
Descriptive Statistics of Facet Scores (Positive Activation).

Facet	N	Mean	Median	SD	SEM	Min	Max	Q1	Q3
Energy & interest	343	4.26	4.50	0.76	0.04	1.50	5.00	4.00	5.00
Persistence	343	2.76	2.75	0.91	0.05	1.00	5.00	2.00	3.50
Excitement	343	3.99	4.00	0.86	0.05	1.00	5.00	3.50	4.50

Principal components analysis of the facet scores (positive activation) and item scores (negative activation) indicated that the first two components accounted for 43.1% of the total variability in the correlation matrix. The scree plot indicated that a two-factor solution was most appropriate. Each facet had a strong loading on one factor and a minimal loading on the other. This was also apparent in the load plot. These results identified eleven items for inclusion within the negative activation scale and ten items for inclusion in the positive activation scale of the final psychometric instrument (Table 7).

Similar results in the final analysis of facet and item scores were obtained with the alternative imputed data sets. The resulting load plots showed that the items and facets were similarly grouped in each data set. However, two of the data sets produced inverted plots.

Table 7

List of Final Items Included in the Psychometric Scale and Results of Final Principal Components Factor Analysis Using Varimax Rotation of Negative Activation Items and Positive Activation Facets.

Negative activation items: Your dog…………….	Negative	Positive
*2 is rarely frightened	0.715	-0.189
7 is easily startled by noises and / or movements	0.689	0.047
15 appears nervous and / or jumpy for several minutes after it has been startled	0.668	-0.049
18 has a specific fear or phobia	0.668	-0.160
*20 appears calm in noisy, crowded places	0.687	0.056
24 is frightened by noises from the television or radio	0.588	0. 172
*25 usually appears relaxed	0.563	0.147
*27 adapts quickly to changes in its environment (eg. being cared for by different people, moving house or a family member leaving home)	0.426	-0.148
29 appears afraid of the vacuum cleaner or any other familiar household appliance	0.507	-0.039
*39 appears calm in unfamiliar environments	0.675	0.071
40 appears unsettled by changes to its routine (eg. if it is not fed at the usual time, if it is left alone for longer than usual)	0.367	0.177

Positive activation facets

Energy and interest (facet 1)
  -0.120 0.742

*14 Your dog shows little interest in its surroundings

22 Your dog is full of energy

*30 Your dog is lazy

*34 Your dog requires a great deal of encouragement to take part in   energetic activities

Persistence (facet 2)
  0.055 0.812

10 Your dog is very persistent in its efforts to get you to play

19 Your dog tries to escape from the garden

36 Your dog persists in being naughty despite being told off for the behaviour

41 Your dog is very boisterous

Excitement (facet 3)
  0.074 0.767

1 Your dog becomes very excited when it is about to go for a walk (eg. when it sees its lead, or when it hears "walkies", etc.)

16 Your dog is easily excited

	Negative	Positive
Variance	4.062	1.977
% Var	0.290	0.141

Note. * Indicates reversed scored items.

Scale Scores

The distribution of scores for each scale was skewed (Figure 2), but to a lesser extent than many of the individual item scores. The negative activation scale scores were positively skewed whereas the positive activation scale scores were negatively skewed. The response distribution of standardised scores for each scale indicates ability to discriminate between individuals. These standardised scale scores for positive and negative activation were produced whilst maintaining minimal correlation ($r = 0.01$, $p = 0.87$).
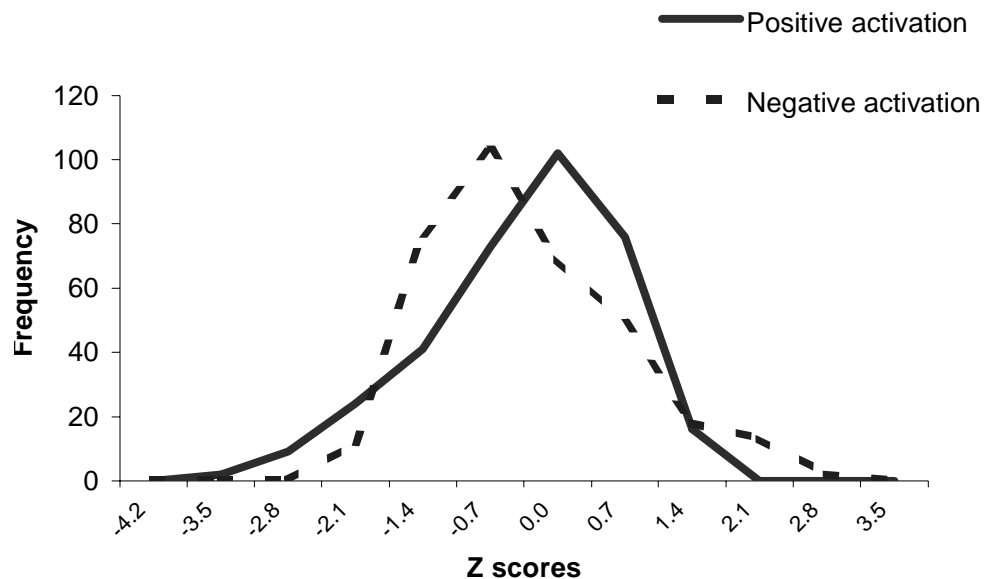

Figure 2. Frequency distributions of standardised scale scores.

Test-Retest Reliability

Each facet of positive activation displayed moderate to high correlations in test-retest scores and there were no significant differences in their median values (Table 8). The positive and negative activation scales both displayed correlations in test-retest scores that were at least moderate, and median scores did not differ significantly.

Internal Consistency and Unidimensionality

Negative Activation. The items of the negative activation scale displayed inter-item correlations that ranged from $r = 0.08$ to $r = 0.49$ with a mean inter-item correlation of 0.29. Only five of the fifty-five inter-item correlations for the negative activation scale fell outside of the optimal desired range (0.15 – 0.50), indicating internal consistency and unidimensionality.

Table 8
Measures of Test-Retest Reliability on Facet and Scale Scores.

	Comparisons of median scores					Correlations per item (Pearsons)	
	N	N missing	N for test	Wilcoxon statistic	p	r	p
Negative Activation							
Scale Scores	78	0	71	1416.0	0.43	0.87	<0.01
Positive Activation							
Facet 1	78	0	53	66.5	0.69	0.80	<0.01
Facet 2	78	0	63	1118.5	0.45	0.70	<0.01
Facet 3	78	0	49	497.0	0.25	0.72	<0.01
Scale Scores	78	0	70	1239.5	0.99	0.75	<0.01

Positive Activation. The items of the positive activation scale displayed inter-item correlations that ranged from 0.05 to 0.54 with a mean inter-item correlation of 0.26. Thirteen of the forty-five correlations fell outside of the optimal desired range (0.15 – 0.50).

When the facets of the positive activation scale were examined separately, the items of facet 1 (energy and interest) displayed correlations that ranged from 0.20 to 0.54 (Mean = 0.35). Two of the six correlations were slightly higher than the optimal desired range (0.15 – 0.50). The items of facet 2 (persistence) displayed correlations that ranged from 0.25 to 0.39 (Mean = 0.31). Facet 3 included only 2 items, displaying a correlation of 0.29. This suggested that the facets of positive activation had internal consistency and unidimensionality. The correlations between the facet scores ranged from 0.40 to 0.45, again falling within the optimal range.

Discussion

This investigation has described two underlying traits that appear to correspond to the dimensions of positive and negative activation commonly reported in humans. Since multiple imputation of missing data appeared to have little impact on the results, the use of all available data sets was justified.

Trait Measures of Positive and Negative Activation

The negative activation scale included items that referred to the frequency of fearful and relaxed states (items 2 and 25), and responses to changing environments (items 27 and 40), unfamiliar environments (39), habituation (items 18, 20, 24, and 29) and the startle response (items 7 and 15). The relevance of each of these to the concept of negative activation is discussed below.

Watson and Tellegen (1985) suggested that fear and relaxation clearly represent the two extremes of the negative affect dimension. The canine scales included references to these and other mood descriptors (e.g., calm, afraid, and nervous) that are regarded as indicators of negative affect in humans (Watson & Tellegen, 1985). The inclusion of items relating to the ability to adapt to environ-

mental change within the negative activation scale is consistent with expectations. Such adaptation requires appropriate responses to novel stimuli and/or signals of the omission of an expected reward and these processes are thought to be controlled by the behavioural inhibition system, which is associated with negative activation (Gray, 1983). Theories of habituation refer to the novel and emotional significance of stimuli (Gati & Ben-Shakhar, 1990), suggesting that individual differences in rates of habituation may reflect sensitivity of the behavioural inhibition system. Individual differences of this nature may be demonstrated by differences in the avoidance of novel or potentially punishing stimuli. The development of such avoidance behaviours in pet dogs has been associated with their early environment (Appleby et al., 2002), suggesting that the sensitivity of the behavioural inhibition system in dogs may, at least in part, be dependent upon early environmental stimuli. Potentiation of the startle response is similarly associated with fear and anxiety (Panksepp, 1998), and may also reflect a low tendency to habituation. Thus it appears that the items that were retained within the final scale may be regarded as measures of negative activation.

The positive activation scale included three facets, referring to energy and interest, persistence and excitement. Watson et al. (1988) used similar mood descriptors ("interested", "active", "determined", and "excited") to measure positive affect in humans. Energised, excited and vigorous behaviours are consistently associated with the motivation systems that are linked to positive activation (Depue & Collins, 1999; Panksepp, 1988). Depue and Collins (1999) also suggest that persistence, effort and resistance to extinction are associated with the behavioural activation system.

Whilst it appears that items in the negative and positive activation scales correspond to facets of the constructs that they were designed to measure, neither construct may be fully represented. The negative activation scale did not identify the hesitance and vigilance that would usually be associated with anxiety (Gray, 1983). Items of this nature loaded on both factors. This might be because owners fail to differentiate between similar behaviours that have different motivations. For example, an animal that attends to the environment may be monitoring signals of reward (positive activation) or signals of punishment (negative activation).

The positive activation scale does not fully reflect the exploratory behaviours that are commonly associated with the underlying motivation system (Depue & Collins, 1999; Gray, 1991; Panksepp, 1998). Items that referred to active exploration (items 17, 32, and 42) loaded on both factors. This may reflect the absence within the scale of specification of the presence or absence of a signal of reward, however, active exploration may also reflect both high positive activation and low negative activation and so not be amenable to assessment in this way.

Generally items that referred to the dog's attachment to the owner (26, 28, 43) did not load strongly on one factor alone. Such attachment may reflect social engagement, which Eysenck and Eysenck (1985) associated with extraversion (positive activation), as well as insecurity, which is more likely to be related to neuroticism (negative activation). Thus it appears that these items are not pure measures of either construct. The significance of attachment in a range of problem behaviours such as separation anxiety, remains controversial (Horwitz, 2002) and so development of an appropriate clinical assessment instrument would be useful. However, given the limitations of owners as accurate observers of behaviour, it

may be that an in-clinic test, such as that described by Prato Previde and colleagues (2001) will be a more fruitful way forward than any test based on interview alone.

Some items were endorsed rarely, resulting in extreme response distributions and poor differentiation between individuals. Such items included those that referred to irritability and aggression. Watson and colleagues (1988) included "hostile" and "irritable" among the items of their PANAS scales, indicating the potential importance of these in measures of activation. Such behaviour was rare in the sample but may occur more commonly than indicated due to the influence of a social desirability response set.

Positive and Negative Activation as Separate Factors Influencing Canine Behaviour

Several studies on dogs have revealed the importance of positive and negative activation in the regulation of behaviour in a broad range of contexts (Brace, 1961; Goddard & Beilharz, 1984, 1985; Svartberg & Forkman, 2002) and some (e.g., Svartberg & Forkman, 2002) have drawn comparison between this and the human supertraits of Extraversion and Neuroticism described by Eysenck (1967). However, there is a tendency in the canine literature to assign these behavioural tendencies to a single dimension, with some form of "boldness" at one end of the scale and "timidity" at the other. This is not consistent with the proposed biological basis of these traits. By contrast, our data suggest that two independent dimensions may be used to describe these tendencies. Draper (1995) in a reanalysis of the data of Hart and Hart (1985) and Hart and Miller (1985), identified three factors which he interpreted as analogous to surgency, agreeableness and openness from the "Big Five" factors commonly described in human personality (Norman, 1967). However, the loadings of the first two factors could just as easily be interpreted as loading on positive and negative activation as separate items. Similarly, Goodloe and Borchelt (1998) identified some factors that may be facets of positive activation and some that may be facets of negative activation. "Friendliness" and "play" factors may be associated with the former whereas "fear or avoidance of strangers" may reflect the latter. Further support for the independent evaluation of these factors comes from Serpell and Hsu (2001) who identified one factor, "energy level", that appears to correspond to a facet of positive activation, and three that refer to fearful behaviour, possibly representing facets of negative activation.

Psychometric Evaluation

Measures of internal consistency and unidimensionality indicated that each of the canine activation scales measured a single homogeneous construct. This suggests that they represent a solid foundation for future investigation.

The positive and negative activation scales in this study accounted for 43.1% of the common variance within the sample. In comparison, trait measures of positive and negative affect in humans using the PANAS scales accounted for 68.7% of common variance (Watson et al., 1988). The BIS / BAS scales that were developed by Carver and White (1994) to measure Gray's constructs of behavioural approach and behavioural inhibition in humans displayed greater similarity

to the results of this study. The factors that they identified accounted for 49% of variance. Whilst this result for the canine negative and positive activation scales could reflect a failure to fully represent all of the facets of each of the domains, it may also be that there are other potentially important sources of variance that are not apparent in investigations of humans. This may include the effects of owners, who may facilitate or inhibit specific aspects of their dogs' behaviour.

Watson and colleagues (1988) found relatively weak correlations between positive and negative affect scales scores, ranging from –0.12 to –0.23. The correlation between canine scale scores for positive and negative activation was much lower (r=0.01). This might indicate that the dimensions of canine activation display a greater degree of independence than those of human affect, supporting the view that positive and negative activation correspond to two unipolar dimensions that are largely independent. Both need to be assessed in any given clinical case as abnormalities in either or both may be important.

Distribution of Response Scores

Individual differences in activation scores reflect inter-dog variation in their perceived environment. There may be real differences in the frequency, intensity and duration of exposure to aversive and appetitive stimuli, however, differences in perception may also reflect variation in underlying neurobiology such that neutral stimuli may be misinterpreted as potentially harmful or rewarding. We believe the latter most probably explains the distribution recorded here, since most items were aimed at being generic in their circumstance.

The theory describing the dimensions of positive and negative activation suggests that scale scores would be expected to correspond to a normal distribution if the sample was representative of the population as a whole. Whilst both scales approached normality, both the positive and the negative activation scale scores had skewed distributions, the former displaying a positive skew and the latter displaying a negative skew. These distributions suggest that the majority of dogs in the sample perceive their environments to contain fewer signals of punishment in relation to signals of reward. This may be a true reflection of the home environment of many dogs and may also indicate that such imbalance in scale scores represents the optimal level of functioning for pet dogs. This may, in part, be the result of selective breeding for a "pet" dog, i.e. animals less sensitive to punishment and more sensitive to rewards are favoured in pet selection. This is consistent with theories of domestication, in which it is frequently proposed that humans may act as a buffer against some aversive stimuli in the wild environment (Price, 1984) and thus reduce selective pressure for escape behaviours. It is also believed, in the case of the dog, that domestication came about as a result of the evolution of a tendency for greater behavioural approach towards human settlements in order to scavenge from their waste sites (Coppinger & Coppinger, 2001). This would not only result in an attenuation of negative activation, but also require accentuation of behavioural activation. The skewed response distributions may, however, also reflect bias within the sample of dog owners. Some owners may think that it would reflect badly on them if their dogs were anxious and lethargic, resulting in a social desirability response set.

It is inevitable that some important behaviours or states will have extreme response distributions. This is particularly likely if items refer to signs of disorders (e.g. aggression and repetitive behaviours), as disorders are relatively rare. The results demonstrated that response distributions were improved by combining correlated items within a single variable and despite the skewed response distributions, the scales did appear to effectively differentiate between individuals.

Test-Retest Reliability and Validity

The scales demonstrated good test-retest reliability. This suggests temporal stability in owners' perceptions of their dogs but does not necessarily imply that those perceptions are accurate. Further investigations of the scales' validity will address this issue.

A rigorous methodology was adopted in the development of these canine positive and negative activation scales and the literature provides a solid biological foundation for the measurement instrument and items included in the scale. Efforts have been made to fully represent the facets and domains of both constructs in order to provide some evidence of content validity (DeVellis, 1991). The multivariate statistical analyses supported the factor structure that has been identified in investigations of positive and negative activation in humans. The psychometric properties indicated that each scale measured a single homogeneous construct and that these constructs appeared to correspond to those used to describe positive and negative activation in humans. However, we acknowledge that in the absence of specific studies aimed at validating the scales against physiological parameters, caution is warranted in assuming that these human and canine scales are indeed truly homologous.

Overall, it appears that the scales hold considerable promise as a method of assessing positive and negative activation in dogs, and investigating the underlying behavioural tendencies of dogs in a much greater variety of contexts than has been possible before with narrower, less biologically based scales and temperament tests.

Limitations of the Scale

The use of a single sample for the development and refinement of scales suggests that the results could be specific to a chance bias that may be present (Smith & McCarthy, 1995). The exclusion of dog-owners that were not sufficiently motivated to complete the questionnaire could bias the results but we consider this to be unlikely, given the range of populations from which responses were received and the distribution of results. However, replication with an independent sample is recommended to indicate the degree to which the results may be generalised. Similarly, these scales have been developed with a population of pet dogs and cannot readily be applied to dogs kept in a non-domestic environment, such as working dogs, laboratory animals, and feral dogs.

Other potential sources of bias may have been reduced by the nature of the questionnaire. This provided descriptions of behavioural tendencies rather than descriptions of specific behaviours. The former may provide more accurate measures of positive and negative activation, as owners may prevent their dogs from

displaying specific behaviours but still recognise the extent to which the behavioural tendency exists.

The measurement instrument relies on owners' subjective assessments of their dogs, but this does not indicate any compromise in terms of validity. Assessments of child temperament in humans are frequently based on questionnaires that have been completed by parents or teachers. No evidence has been found to suggest that these are less valid than more objective measures, such as direct observations (Bates et al., 1994). Some studies of infants have identified a great deal of coherence between data obtained by parental report questionnaires and data obtained in laboratory observations of behaviour (e.g., Carranza et al., 2000) and the robustness of psychometric profiles as predictors of behaviour in a range of contexts and across time have also been established in other nonhuman species (Capitanio, 1999). However, similar correspondence cannot be assumed without investigation. Objective behavioural recording in conjunction with the questionnaire could be used to determine the robustness of the scale (Kazdin, 1995). Hennessey and colleagues (2001) developed a battery of behavioural tests that aimed to assess responses to novelty and threatening situations. These measures were found to be predictive of behaviour problems that were reported six months later (Hennessey et al., 2001). Gray (1994) suggests that behavioural tests investigating sensitivity to signals of reward and punishment may assess activity of the behavioural approach and behavioural inhibition systems. Therefore such tests could be used in conjunction with the questionnaire to identify the effects of social desirability response set and indicate the accuracy of owners' subjective assessments of their dogs.

Conclusion

This paper has described the development of a new tool that uses owner report to measure individual differences in positive and negative activation in pet dogs. The assessment instrument is the first psychometric profile to be developed for use in managing problem behaviour in non-humans that is based on comparative psychological theories that link behaviour, affect, motivation and neurobiology and suggests that these traits are consistent between humans and dogs. The solid biological foundation, combined with the rigour of the developmental process gives confidence to the potential for the instrument to assess general behavioural tendencies that cannot be elucidated with other available canine assessments. The scale may therefore have many uses within veterinary behavioural medicine and the study of canine welfare, analogous to those of the comparable human scales.

References

Appleby, D. L., Bradshaw, J. W. S., & Casey, R. A. (2002). Relationship between aggressive and avoidance behaviour by dogs and their experience in the first six months of life. *Veterinary Record,* **150**, 434-438.

Bates, J. E., Wachs, T. D., & Emde, R. N. (1994). Toward practical uses for biological concepts. In J. E. Bates, and T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 275-306). Washington, DC: American Psychological Association.

Brace, C. L. (1961). *Physique, physiology and behaviour; an attempt to analyze a part of their roles in the canine biogram.* Unpublished PhD thesis, Harvard University.

Campbell, W. E. (1972). A behavior test for puppy selection. *Modern Veterinary Practice,* **53**, 29-33.

Capitanio, J. P. (1999). Personality dimensions in adult male rhesus macaques: prediction of behaviors across time and situation. *American Journal of Primatology*, **47**, 299-320.

Carranza Carnicero, J. A., Perez-Lopez, J., Gonzalez Salinas, M., & Martinez-Fuentes, M. T. (2000). A longitudinal study of temperament in infancy: stability and convergence of measures. *European Journal of Personality*, **14**, 21-37.

Carver, C. S., & White, T. L. (1994). Behavioural inhibition, behavioural activation, and affective responses to impending reward and punishment: the BIS / BAS scales. *Journal of Personality and Social Psychology*, **67**, 319-333.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, **1**, 69-90.

Clark, L. A., & Watson, D. (1995). Constructing validity: basic issues in objective scale development. *Psychological Assessment*, **7**, 309-319.

Coppinger, R., & Coppinger L. (2001). *Dogs – A startling new understanding of canine origin, behaviour and evolution*, New York, Scribner.

Cronbach, L. C. (1990). *Essentials of psychological testing*. Fifth edition. New York: Harper Collins.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioural and Brain Sciences*, **22**, 491-569.

DeVellis, R. F. (1991). *Scale development: theory and applications*. London: Sage.

Draper, T. W. (1995). Canine analogs of human personality factors. *Journal of General Psychology*, **122**, 241-252.

Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C. Thomas.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: a natural science approach*. New York: Plenum.

Figueredo, A. J., McKnight, P. E., McKnight, K. M., & Sidani, S. (2000). Multivariate modeling of missing data within and across assessment waves. *Addiction*, **95**, s361-s380.

Floyd, F. J., & Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, **7**, 286-299.

Gati, I., & Ben-Shakhar, G. (1990). Novelty and significance in orientation and habituation: a feature-matching approach. *Journal of Experimental Psychology: General*, **119**, 251-263.

Goddard, M. E., & Beilharz, R. G. (1984). A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science*, **12**, 253-263.

Goddard, M. E., & Beilharz, R. G. (1985). individual variation in agonistic behaviour in dogs. *Animal Behaviour*, **33**, 1338-1342.

Goodloe, L. P., & Borchelt, P. L. (1998). Companion dog temperament traits. *Journal of Applied Animal Welfare Science*, **1**, 303-338.

Gosling, S. D., & John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science*, **8**, 69-75.

Gosling S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, **127**, 45-86.

Gray, J. A. (1983). Anxiety, personality and the brain. In A. Gale, and J. A. Edwards (Eds.), *Physiological correlates of human behaviour, Vol. III: Individual differences and psychopathology* (pp. 31-43). London: Academic Press.

Gray, J. A. (1991). The neuropsychology of temperament. In J. Strelau, and A. Angleitner (Eds.), *Explorations in temperament: international perspectives on theory and measurement* (pp. 105-128). New York: Plenum.

Gray, J. A. (1994). Personality dimensions and emotion systems. In P. Ekman, and R.J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 329-331). Oxford: Oxford University Press.

Hart, B. L., & Hart, L. A. (1985). Selecting pet dogs on the basis of cluster analysis of breed behavior profiles and gender. *Journal of the American Veterinary Medical Association*, **168**, 1181-1185.

Hart, B. L., & Miller, M. F. (1985). Behavior profiles of dog breeds: A quantitative approach. *Journal of the American Veterinary Medical Association*, **168**, 1175-1180.

Hennessy, M. B., Voith, V. L., Mazzei, S. J., Buttram, J., Miller, D. D., & Linden, F. (2001). Behavior and cortisol levels of dogs in a public animal shelter, and an exploration of the ability of

these measures to predict problem behavior after adoption. *Applied Animal Behaviour Science*, **73**, 217-233.

Horwitz D. F. (2002). Separation-related problems in dogs. In D.F Horwitz, D.S. Mills, and S.E Heath (Eds.), *BSAVA Manual of Canine and Feline Behavioural Medicine* (pp. 154-163). Gloucester, UK: British Small Animal Veterinary Association.

Kazdin, A. E. (1995). Preparing and evaluating research reports. *Psychological Assessment* **7**, 228-237.

Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.

Longford, N. T., Ely, M., Hardy, R., & Wadsworth, M. E .J. (2000). Handling missing data in diaries of alcohol consumption. *Journal of the Royal Statistical Society - Series A*, **163**, 381-402.

Marschark, E. D., & Baenninger, R. (2002). Modification of instinctive herding dog behaviour using reinforcement and punishment. *Anthrozoos*, **15**, 51-68.

Martin, P., & Bateson P. (1986). *Measuring behaviour. An introductory guide.* Cambridge: Cambridge University Press.

Mills D. S. (2002). Learning, training and behaviour modification techniques. In D. F. Horwitz, D. S. Mills, & S. Heath (Eds.). *BSAVA manual of canine and feline behavioural medicine* (pp. 37-48). Gloucester, UK: British Small Animal Veterinary Association.

Mills, D. S., Heath S. E., & Harrington L. J. (1997). *Proceedings of the First International Conference on Veterinary Behavioural Medicine*. Potters Bar, UK: Universities Federation for Animal Welfare.

Netto, W. J., & Planta, D. J. U. (1997). Behavioural testing for aggression in the domestic dog. *Applied Animal Behaviour Science,* **52**, 243-263.

Norman W. T. (1967). *2800 personality trait descriptors: normative operating characteristics for a university population.* Ann Arbor, MI: University of Michigan.

Panksepp, J. (1998). *Affective neuroscience: the foundations of human and animal emotions.* Oxford: Oxford University Press.

Prato Previde, E., Custance, D. M., Sabatini, F., Spiezo, C., & Poli M. (2002). Does attachment charcaterise the dog-human affectional bond? An observational study under controlled conditions using Ainsworth's "Strange Situation Test". In K. L. Overall, D. S. Mills, S. E. Heath & D. Horwitz (Eds.), *Proceedings of the Third International Congress on Veterinary Behavioural Medicine* (pp. 145-147). Wheathampstead, UK: Universities Federation for Animal Welfare.

Price, E. O. (1984). Behavioral aspects of domestication. *Quarterly Review of Biology*, **55**, 1-32.

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, **12**, 287-297.

Reisner, I. R., Hollis, N. E., & Houpt, K. A. (1994). Risk factors for behaviour related euthanasia among dominant-aggressive dogs: 110 cases, (1989-1992). *Journal of the American Veterinary Medical Association*, **205**, 855-863.

Rust, J., & Golombok, S. (1989). *Modern psychometrics: The science of psychological assessment.* London: Routledge.

Serpell, J. A., & Hsu, Y. Y. (2001). Development and validation of a novel method for evaluating behavior and temperament in guide dogs. *Applied Animal Behaviour Science,* **72**, 347-364.

Sheppard, G. and Mills, D. S. (2003). Construct models in veterinary behavioural medicine: lessons from the human experience. *Veterinary Research Communications,* **27**, 175-191.

Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment,* **7**, 300-308.

Svartberg, K. (2002). Shyness-boldness predicts performance in working dogs. *Applied Animal Behaviour Science*, **79**, 157-174.

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

Takeuchi, Y., Houpt, K. A., & Scarlett, J. M. (2000). Evaluation of treatments for separation anxiety in dogs. *Journal of the American Veterinary Medical Association*, **217**, 342-345.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, **98**, 219-235.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, **54**, 1063-1070.

Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology. *Journal of Abnormal Psychology*, **103**, 18-31.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, **76**, 820-838.

Wilsson, E. & Sundgren, P. E. (1998). Behaviour test for eight week old puppies – heritabilities of tested behaviour traits and its correspondence to later behaviour. *Applied Animal Behaviour Science*, **58**, 151-162.

Appendix 1. List of Items in Original Questionnaire

1. Your dog becomes very excited when it is about to go for a walk (eg. when it sees its lead, or when it hears "walkies", etc.)
2. Your dog is rarely frightened
3. Your dog ignores unfamiliar people that you invite into your home
4. Your dog becomes aggressive (i.e. growl, snap or bite) if you try to remove its favourite toy or food
5. Your dog is sluggish in its movements
6. Your dog frequently sniffs its surroundings when being walked
7. Your dog is easily startled by noises and / or movements
8. Your dog approaches most activities with enthusiasm
9. Your dog appears hesitant and / or wary in most of its movements
10. Your dog is very persistent in its efforts to get you to play
11. Your dog is hesitant to approach unfamiliar people that you invite into your house
12. Your dog appears to be very upset when you tell it off
13. Your dog becomes very excited when its dinner is being prepared
14. Your dog shows little interest in its surroundings
15. Your dog appears nervous and / or jumpy for several minutes after it has been startled
16. Your dog is easily excited
17. Your dog chooses to stay close to you when walked off the lead
18. Your dog has a specific fear or phobia
19. Your dog tries to escape from your garden
20. Your dog appears calm in noisy, crowded places
21. Your dog is constantly vigilant, frequently scanning the environment
22. Your dog is full of energy
23. Your dog displays repetitive movements or behaviours (eg. repeatedly licking a specific area of its coat, repeatedly chasing its tail, repeatedly pacing or running a specific path)
24. Your dog is frightened by noises from the television or radio
25. Your dog usually appears relaxed
26. Your dog chooses to stay very close to you in unfamiliar surroundings
27. Your dog adapts quickly to changes in its environment (eg. being cared for by different people, moving house or a family member leaving home)
28. Your dog appears indifferent towards you
29. Your dog appears afraid of the vacuum cleaner or any other familiar household appliance
30. Your dog is lazy
31. Your dog trembles but makes little attempt to get away when you try to perform any procedure that it dislikes (eg. giving medication, grooming or bathing)
32. Your dog is keen to investigate noises, objects or movements in the environment
33. Your dog appears shy around people that it does not know well
34. Your dog requires a great deal of encouragement to take part in energetic activities (eg. running, vigorous games)
35. Your dog is curious about its surroundings
36. Your dog persists in being naughty despite being told off for the behaviour
37. Your dog frequently appears nervous
38. Your dog often appears irritable or grumpy
39. Your dog appears calm in unfamiliar environments
40. Your dog appears unsettled by changes to its routine (eg. if it is not fed at the usual time, if it is left alone for longer than usual)
41. Your dog is very boisterous
42. Your dog is keen to explore new surroundings
43. Your dog tries to stay in close proximity to a person that it knows well

44. Your dog appears to be lethargic
45. Your dog frequently experiences recurrent health problems such as stomach upsets (not relapses or flare ups of an ongoing problem like arthritis)