Validation of a temperament test for domestic cats

Janice M. Siegford[*] *, Sally O. Walshaw*[†]*, Petra Brunner*[††] *and Adroaldo J. Zanella*[*]

[*] Animal Behavior and Welfare Group, Department of Animal Science, Michigan State University, East Lansing, USA
[†] Dept. of Biomedical Sciences, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, PE Canada
[††] Sltana Pharma KG, Konstanz, Germany

Abstract

Cats are popular companion animals, particularly in Europe and North America, and appear in correspondingly large numbers in animal shelters. Temperament tests are not widely used to assess cats before adoption from shelters. However, cats exhibit a wide range of temperaments as do the families adopting them and ensuring compatibility between the two could increase the rate of successful placement. Scores on a feline temperament profile (FTP), which measures a cat's responses to standardized interactions with an unfamiliar person, were compared between cats and over time and related to responses of cats to familiar and unfamiliar persons and to basal salivary cortisol levels. Cats showed significant differences in FTP scores ($p<0.001$). Ranking cats according to FTP scores resulted in three distinct groups of cats. Over eight months, changes in FTP scores were minor, with cats scoring somewhat more acceptably and less questionably following adoption. Acceptable scores on pre-adoption FTPs were positively correlated with 1) positive responses to familiar caretakers in housing rooms ($p=0.01$) and 2) average percent of time spent near either unfamiliar men or women in open field tests in novel rooms ($p=0.01$ in both instances). Thus, cats displaying general positive responses to humans did so in both familiar and test environments and with familiar and unfamiliar persons. No correlation was seen between FTP scores and basal salivary cortisol levels ($p>0.05$), though there were significant differences in cortisol levels between cats ($p=0.04$). The data indicate that the FTP was relatively stable over time for adult cats, and test scores correlated well with ethological observations of cats' interactions with humans. The FTP could provide an accurate, consistent assessment of cat temperament, leading to more successful placement of cats.

Keywords: *assessment, feline, open-field test, personality, salivary cortisol*

Address for correspondence: J. M. Siegford, 1287C Anthony Hall, Animal Behavior and Welfare Group, Department of Animal Science, Michigan State University, East Lansing, MI, USA. Ph: 517-432-8212; fax: 517- 353-1699; e-mail: siegford@msu.edu.

There is a growing need to understand factors that contribute towards successful adaptation of cats to the human environment. Animal shelters across the US take in an estimated 6 to 8 million dogs and cats each year, and in many shelters approximately 70% of these cats are euthanized (Human Society of the United States, National Council on Pet Population Study and Policy). For example, in 1998 shelters in the state of Michigan euthanized more than 33,000 cats (Michigan Department of Agriculture). Among the top reasons cats are likely to be relinquished to shelters is because they do not match the expectations or lifestyle of the adopting family (Salman et al. 1998; New et al. 2000).

Temperament testing has long been performed on dogs in order to assist their placement with families, particularly at shelters placing animals for adoption. Temperament, individuality, and personality are terms often used interchangeably in studies of animals to describe the sum total of behavioral attributes which characterize an individual and set it apart from others (Mendl and Harcourt 2000). While the merits of different forms of canine temperament testing may be debated, temperament tests and physiological measures can be used successfully to predict problem behaviors in dogs following adoption (van der Borg, Netto and Planta 1991; Hennessy et al. 2001). No such practice is used routinely for cats, however, despite the fact that several checklists or easily scored ethograms have been developed that could be used for this purpose (Lee et al. 1983; McCune 1995; Kessler and Turner 1997).

The absence of temperament assessment of cats often leads to the placement of cats on the basis of appearance, age, and sex of the animal. While these characteristics may be important features of owner expectation as they select a cat, the temperament of the cat and its suitability for its adoptive family should be considered. Cat owners, or persons who have intensely observed the behavior of an individual cat, develop a sense that each cat has unique behavioral characteristics (Feaver, Mendl and Bateson 1986; Mendl and Harcourt 2000). It becomes important, therefore, to match cats and owners because discrepancies between an animal's actual behavior and the owner's vision of ideal behavior can affect owner attachment to the animal. For example, owners that reported their cats were affectionate to them also reported themselves as more affectionate to their cats (Turner and Stammbach-Geering 1990). Thus, owner attachment based on desired behaviors probably affects the likelihood that the owner will relinquish the animal (Serpell 1996). It would therefore appear that shelters should provide prospective owners with an assessment of a cat's temperament to be used with its physical characteristics in order to find a cat that meets owner

expectations on all levels. Additionally, cat temperament tests could be used by veterinarians to establish baseline temperament profiles of feline patients which could be reassessed on subsequent visits.

Methods have been developed to reliably assess individual differences in cats in an objective as well as a subjective manner (Feaver, Mendl and Bateson 1986; McCune 1995; Kessler and Turner 1997; Turner 2000a). A variety of studies have examined the ontogeny of cat behavior and individuality and their consistency over time (Martin 1986; Reisner et al. 1994; Bradshaw and Cook 1996; Durr and Smith 1997; Lowe and Bradshaw 2001), response of cats to stressful situations (Kessler and Turner 1997, 1999a, 1999b), cat–human interactions under a variety of conditions (Mertens and Turner 1988; Mertens 1991; Podberscek, Blackshaw and Beattie 1991; McCune 1995; Turner 2000a), and strength of owner attachment (Karsh and Turner 1988). The results of this work suggest that individual personalities of cats are one of the most significant factors influencing cats' behavior towards people (Mertens and Turner 1988), and that the friendliness of cats to humans depends both on genetic factors and socialization at an appropriate age (Karsh and Turner 1988; Reisner et al. 1994; McCune 1995; Turner 2000b). Additionally, these studies showed that many aspects of cat individuality remain stable over time, particularly once the animal is 4–5 months old (e.g., Lowe and Bradshaw 2001). In sum, previous studies have revealed salient features of cat individuality and the basis for their affinity for humans, and demonstrated that many traits relevant to cat–human interaction remain stable over time and across situations (for reviews of individuality in the cat and cat–human relationships, see Mendl and Harcourt 2000 and Turner 2000b, respectively).

However, the techniques used to assess cats in many of the previous studies would not be practical for assessing the temperaments of cats in shelters or in veterinarians' offices. Ethological observations, while objective, are often time consuming and require personnel trained in behavioral observation and statistics in order to accurately record the necessary data and to analyze it for use. Subjective assessments, on the other hand, require the assessor to have spent enough time with the cat to form an intimate knowledge of its behavior under a variety of circumstances, which is unlikely to occur in a shelter situation.

A standardized temperament test has the advantage of requiring less training by the user, less time per animal and if a scoring system is used, easier analysis of the results. Additionally, simpler behavioral criteria may be more relevant to owner–cat attachment (Serpell 1996). Thus, an effective feline temperament test could be used in shelters across the country by

existing staff. However, feline temperament tests generally have not been scientifically examined either by administering the test repeatedly to the same group of cats across time and varying situations (but see McCune 1995; Kessler and Turner 1997 1999a, 1999b for exceptions) or by comparing ethological observations to test scores to assess whether these tests measure relevant and stable aspects of temperament, and, in particular, features relating to cat–human compatibility.

There were three objectives of the present study. The first objective was to assess an existing feline temperament test (Lee et al. 1983) by comparing the results of the test to ethological observations of the same cats in the presence of familiar and unfamiliar humans in familiar and unfamiliar locations, in order to validate the accuracy of the test. Secondly, we examined the results of the test over time and under changing circumstances, i.e., before and after adoption, to determine if scores remained consistent under such varying circumstances. Finally, we explored the relationship between scores on the test and salivary cortisol levels at baseline to determine if a relationship existed between a cat's performance on the test and its stress level.

Methods

Animals and housing

This study was approved by the Michigan State University All-University Committee on Animal Use and Care. Twenty female, domestic shorthair cats, ten months of age at the start of the study, were examined over an eight-month period. All were specific-pathogen free cats, which were bred and raised at the same research animal production facility and shared a similar genetic background. Additionally, the breeding facility mandated that all cats receive the same amount and type of handling (including petting and holding) each day to ensure uniform socialization from infancy onward. The cats were housed in groups of ten in two identical rooms at the laboratory animal facility at Michigan State University (MSU). Each room contained two food and water dishes, two litter boxes, various toys, and several carriers, tubs, and buckets of varying sizes allowing cats to hide. Following the experiment, cats were adopted by members of the community and returned to MSU three and six months later for follow-up tests.

Experimental procedures

Cats underwent evaluation using a standardized temperament test (Lee et al. 1983), the feline temperament profile (FTP), twice before and 3 and 6 months after adoption (with the exception of one cat, L069, who was not adopted). This test was originally designed to assess the suitability of cats for placement

in nursing homes by evaluating general levels of sociability, aggressiveness, and adaptability to new situations. The test consisted of ten different phases during which the investigator proceeded from calling the cat at a distance of a few meters to more proximate interactions, such as holding the cat while petting it or pulling on the cat's tail. A list of five to seven possible responses by the cat was present under the description of each phase along with space for observations. Each response in the list was described as "acceptable" or "questionable." Occurrence (1) or absence (0) of each listed response by each cat was noted during each phase of the test. For responses that could occur during multiple phases of the test, such as eye contact, occurrence and absence scores from all phases were combined into one score. For example, eye contact was listed as a possible response in five phases of the test; therefore a cat could have a score of 0 to 5 for eye contact on the test. Some of the listed responses never occurred during the experiment and were not used in the scoring system or examined in the analysis.[1] Using this scoring system, the maximum acceptable score on the FTP was 38 and the maximum questionable score a cat could receive was 16. A list of observed responses and their categorization as acceptable or questionable is given in Table 1. The order in which the cats were tested was randomized on each occasion. All FTP tests were administered by an investigator, the "tester," and scored by a second investigator, the "observer." Both tester and observer recorded scores. However, as these scores had a high degree of inter-observer reliability and as there were occasional cases where the tester could not see the cat as clearly as the observer, scores from the observer were used in generating the FTP. Overall, acceptable and questionable scores for each test were obtained by combining scores for items of each category together. To rank cats, mean acceptable and mean questionable scores using all four FTP tests were generated for each cat. Cats were then ranked by acceptable scores from largest to smallest scores, then again by questionable scores from smallest to largest. A mean rank for each cat was then generated by averaging the acceptable and questionable rankings.

Prior to adoption, interactions between the cats and their caretakers at MSU were monitored in the rooms that housed the cats during three daily visits (28 minutes each) using time-lapse video recording. Proximity to caretaker was measured every 30 seconds during the visit. Frequency of the following behaviors was recorded during the visit: approaches person, touches person, and retreats. Proximity scores were added to those of approach and touch frequency and these values were averaged over the three exposures and termed "positive responses." The frequency of retreats was considered a "negative response" and the frequency of this behavior was averaged over the three visits.

Table 1. Summary of scoring of the Feline Temperament Profile.

Cat Responses	Scored as:
Responses to Tester	
Eye contact	Acceptable
Approaches/circles around tester	Acceptable
Sniffs hand	Acceptable
Rubs/bumps head against tester	Acceptable
Rolls	Acceptable
Vocalization: meow/purr/chirrups	Acceptable
Call cat, cat approaches tester	Acceptable
Retreats/withdraws	Questionable
Vocalization: hisses/growls	Questionable
Challenge 1: Pull Tail	
Turns around	Acceptable
No reaction	Acceptable
Struggles	Questionable
Growls/strikes/hisses	Questionable
Challenge 2: With Toy	
Watches toy	Acceptable
Chases toy	Acceptable
Allows stroking	Acceptable
Challenge 3: Drop Object	
Turns and relaxes	Acceptable
Runs to investigate	Acceptable
Ignores noise	Acceptable

Responses listed on the original FTP that never occurred during our study included: strikes hand, threatens to strike hand, bites or attempts to bite, jumps up on lap, ignores toy, attends to something else in the room rather than the toy, does not hear object dropped, startles and runs to hide, startles then shows aggressive posture when item is dropped. All but one of the unobserved responses fell into the "questionable" response category.

Prior to adoption, cats were also tested in an open field situation to determine their response to two unfamiliar persons in a novel place (which would be similar to a visit to the veterinarian or shelter). The open field test was conducted three times for each cat (n=19) with one test per day and a week between tests. A novel room (3 m x 3.7 m) the cats had never entered before was used as the open field area. The room was divided into 30 (0.6 m x 0.6 m) areas using tape on the floor (designated A–E across the shorter wall and 1–6 across the longer wall, giving each cell in the grid a unique label, e.g., A1). A door was present in the middle of one of the longer walls

(E3–4). During testing, a man sat on the floor at one end of the room in the center of one of the shorter walls, and a woman sat on the floor at the other end of the room (C1 and C6, respectively). Different unfamiliar men and women were used in each of the three open field tests for each cat. Cats were carried into the room in a carrier, which was placed in the center of the floor and opened. Time started when the cat left the carrier (some cats had to be removed from the carrier) and each test lasted five minutes. The number of line crossings was recorded (one in and one out of a cell equaled one total crossing) as was amount of time spent in each cell, which was then converted into percent of total time in each cell. For analysis, cells near the door were combined into one category. Cells in corners were similarly combined. Cells within a one-cell radius surrounding the woman were combined into one category and cells surrounding the man were combined into a second category, to determine if cats responded differently to unfamiliar persons of different sex.

Cortisol measurements

Basal cortisol profiles for each cat were obtained in a non-invasive manner using saliva samples collected from unrestrained cats twice a day, at 07:00 and 19:00, on four different days. Cats were conditioned to come to the researcher by tapping on an open can of cat food with a spoon. Cats were then presented with cotton swabs and allowed to chew on the swabs until they were thoroughly moistened. Cats were rewarded for chewing on the swabs by giving them some of the canned food. Swabs were centrifuged to remove saliva, which was stored at -20°C until processing. An enzyme-linked immunoabsorbent assay (ELISA) was used to measure cortisol as described by Cooper et al. (1989). Briefly, polystyrene microtitre plates were coated with the secondary antibody (goat anti-rabbit IgG at 1:8000), washed, then incubated overnight with primary antibody (rabbit anti-bovine cortisol-3- carboxymethyloxime) diluted 1:4000. Duplicate samples of saliva (in a range of dilutions from 1:4 to 1:100), controls, reagents, and triplicate standards (1-1000 pg) were added to the plates over ice followed by addition of horseradish peroxidase-labelled cortisol (1:30000) to prescribed wells. Plates were incubated for two hours then washed and 3,3',5,5'-tetramethylbenzidine in dimethyl sulfoxide (10mg/ml), made fresh with hydrogen peroxide, added to the plates. The color reaction was stopped by addition of sulfuric acid, then optical density was read in a plate reader (Bio Rad) equipped with a 450 nm filter. Inter-assay variability was 15.0% and intra-assay variability was 4.3%.

Analyses

To ensure that the responses listed on the FTP measured only positive or negative reactions of the cat to the tester, the relationships between acceptable and questionable scores using average scores from all four FTP were assessed using a Spearman rank correlation. We hypothesized that acceptable and questionable scores would be negatively correlated.

Comparisons of acceptable scores on the FTP between cats across all tests and over time for all cats were performed using a one-way analysis of variance (ANOVA), with Bonferroni multiple comparisons tests where needed. A Kruskal-Wallis one-way ANOVA was used to examine the differences in questionable scores between cats across all tests and over time for all cats because the data for these scores did not meet the assumptions of normality. Where needed, multiple comparisons on these data were performed using a Kruskal-Wallis Multiple-Comparison Z-Value test with a Bonferroni correction factor. Cats were ranked 1–20 by highest to lowest acceptable scores averaged across all four FTP tests, then ranked 1–20 by lowest to highest questionable scores averaged across all four FTP tests. These two ranks were then averaged to create an overall rank based on FTP responses. A cluster analysis was performed (SAS Institute, Cary, NC) by cat on mean scores from all four FTPs for each response, to assess the validity of grouping cats by FTP scores.

A Spearman rank correlation was used to assess each of the following relationships: acceptable scores on pre-adoption FTPs with positive responses to a familiar caretaker, questionable FTP scores with negative responses to caretakers, acceptable scores on pre-adoption FTPs with mean percent of time spent in proximity to unfamiliar men or women in the open field test, questionable scores on pre-adoption FTPs with mean time spent in the corners or at the door in the open field test, and the mean number of cell crossings made in all open field tests with acceptable or questionable pre-adoption FTP scores.

Circadian cortisol rhythms were square root transformed to meet the assumptions of normality, then analyzed using a mixed general linear model (proc mixed) with repeated measures (SAS Institute, Cary, NC), with day and time of sample as independent factors and cat as a random effect. Least squares means of cortisol levels were compared and adjusted using the Tukey-Kramer method to control for Type-I error rate. Results following transformation were similar to those obtained using non-transformed data in a similar analysis; therefore results from the transformed analysis are reported in the text. Basal cortisol levels were compared between cats using a Kruskal-Wallis one-way ANOVA to determine if sig-

nificant differences existed between cats. A Kruskal-Wallis Multiple-Comparison Z-Value test with a Bonferroni correction factor was then used to determine which cats had differing basal cortisol levels. Correlations between scores on the pre-adoption FTPs and basal cortisol levels were analyzed using Spearman rank correlation. Correlations between basal cortisol levels and groups of cats that resulted from FTP ranking were analyzed using a one-way ANOVA.

For all analyses, an alpha of less than 0.05 was considered significant. All data are given as mean ± standard error of the mean (SEM). Unless otherwise mentioned, analyses were performed using NCSS (Number Cruncher Statistical Systems, Kaysville, UT).

Results

Behavioral measures and correlations between measures

Mean acceptable scores on all FTPs were negatively correlated with mean questionable scores on the test (Figure 1; n=20, rho=-0.85, p<0.001). Comparison of acceptable scores received on all FTPs revealed significant differences in scores between cats (Table 2; $F_{(19,60)}$=5.95, p<0.001). No difference in questionable scores was observed between cats, likely due to the generally low scores received in this category (Table 2; n=19, χ^2=25.1, p=0.16). However, several cats stand out as having higher questionable scores compared to their peers, while several others have lower scores (Table 2). Ranking of cats (Table 2) was based on the mean ranking received for acceptable scores (higher rank for larger acceptable scores) and ranking received for questionable scores (lower rank for larger questionable scores). Cluster analysis revealed the lowest ranked cats were clearly separated from other cats, and some separation of cats with the highest ranks from those ranked in the middle also existed (Figure 2).

Comparison of FTP scores over time revealed a non-significant increase in acceptable scores (Figure 3; $F_{(3,76)}$=1.29, p=0.28). Questionable scores changed significantly over time (Figure 3; $F_{(3,76)}$=6.01, p=0.001); six months after adoption, cats had lower questionable scores when compared with scores on the second pre-adoption test or scores three months post-adoption.

The correlation between acceptable scores on pre-adoption FTPs and the positive responses of cats to familiar caretakers in their housing rooms was positive and significant (Figure 4; n=20, rho=0.51, p=0.02). There was no correlation between questionable scores on pre-adoption FTPs and negative responses of cats to caretakers (n=20, rho=0.2, p=0.39). A sig-

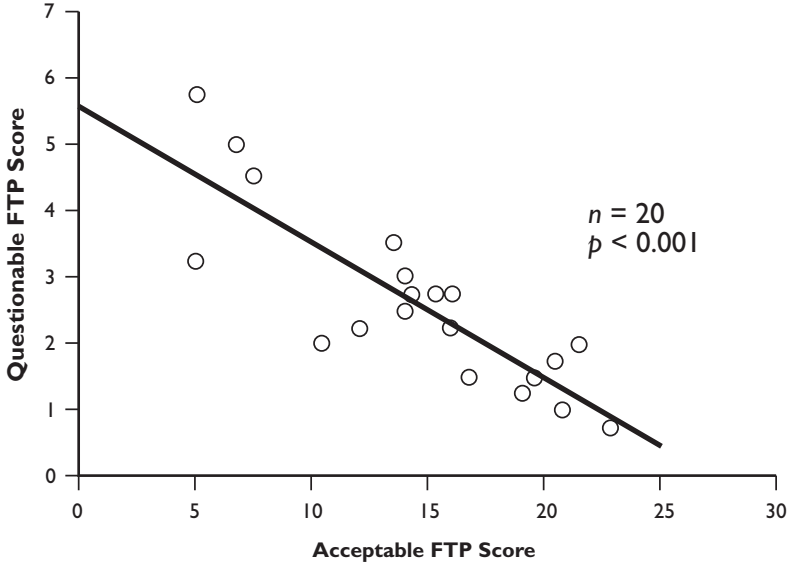

Figure 1. Relationship between average acceptable and questionable average FTP scores. Note different scales on the x and y axes. A trend line has been added to the data to highlight the significant relationship of the data sets in this and subsequent scatter plots.

nificant positive correlation was observed between acceptable scores on the pre-adoption FTPs and mean percent of time the cats spent near the unfamiliar male in the novel room during the open field test (Figure 5a; n=19, *rho*=0.60, p=0.01). A similar, significant positive correlation was observed between acceptable scores on pre-adoption FTPs and mean percent of time the cats spent near the unfamiliar female in the novel room during the open field test (Figure 5b; n=19, *rho*=0.57, p=0.01). A significant positive correlation was observed between questionable scores on pre-adoption FTPs and percent of time cats spent in corners of the novel room during open field tests (Figure 5c; n=19, *rho*=0.51, p=0.03), while no significant correlation was observed between questionable scores on pre-adoption FTPs and mean percent time spent near the door in open field tests (Figure 5d; n=19, *rho*=0.19, p=0.44). The number of cell crossings made in open field tests was positively correlated with acceptable scores on pre-adoption FTPs (Figure 6; n=19, *rho*=0.54, p=0.02), but was not related to questionable scores on pre-adoption FTPs (n=19, *rho*=-0.32, p=0.18).

Cortisol measurements

Salivary cortisol measures from unrestrained cats on each of four consecutive sampling days were not significantly different between sample times of 7:00 and 19:00 ($F_{(1,130)}$=0.06, p=0.81), nor did the cortisol levels vary sig-

Table 2. Average scores of cats from all temperament tests.

Cat No.	Acceptable Score	Questionable Score	Rank	Grouping
L097	22.80[a]	0.75	1	High
L099	20.80[a,d]	1.00	2	High
L041	21.50[a,d]	2.00	3	High
L089	19.50[a,c,d]	1.50	3	High
L307	19.00[a,c,d]	1.25	3	High
L101	20.50[a,d]	1.75	6	High
L093	16.80	1.50	7	Middle
L371	16.00	2.25	8	Middle
L007	16.00	2.75	9	Middle
K227	15.30	2.75	10	Middle
L077	14.00	2.50	10	Middle
L057	10.50[b,c,d]	2.00	12	Middle
L037	14.30	2.75	13	Middle
L091	12.00	2.25	13	Middle
L079	14.00	3.00	15	Middle
L065	13.50	3.50	16	Middle
L075	7.50[b,c]	4.50*	17	Low
L039	5.00[b]	3.25	17	Low
K285[†]	6.75[b]	5.00*	19	Low
L069[‡]	5.00[b]	5.75*	20	Low

Different superscripted letters represent significant differences between scores in that category. Scores with the same superscripted letters were not significantly different from one another and scores without any superscripts were not significantly different from any other scores in that category. Asterisks denote negative scores that appeared higher than those of other cats despite a lack of significant difference. [†]Cat K285 was returned one month after adoption. [‡]Cat L069 was never adopted.

nificantly between sample days ($F_{(3,130)}$ =0.23, p=0.87). There was, however, a significant interaction between time and day the cortisol samples were taken ($F_{(3,130)}$=2.92, p=0.04), owing to a significant difference in cortisol levels between samples from the morning and evening of the first day ($t_{(130)}$=2.55, p=0.012) and a nearly significant difference between cortisol levels in samples from the mornings of the first and third days ($t_{(130)}$=1.97, p=0.051). Overall, comparisons of basal cortisol levels between cats revealed significant individual differences (Figure 7; n=20, χ^2=31.1, p=0.04). However, after correcting for Type-I error rate, multiple comparisons revealed no individuals significantly different from one another.

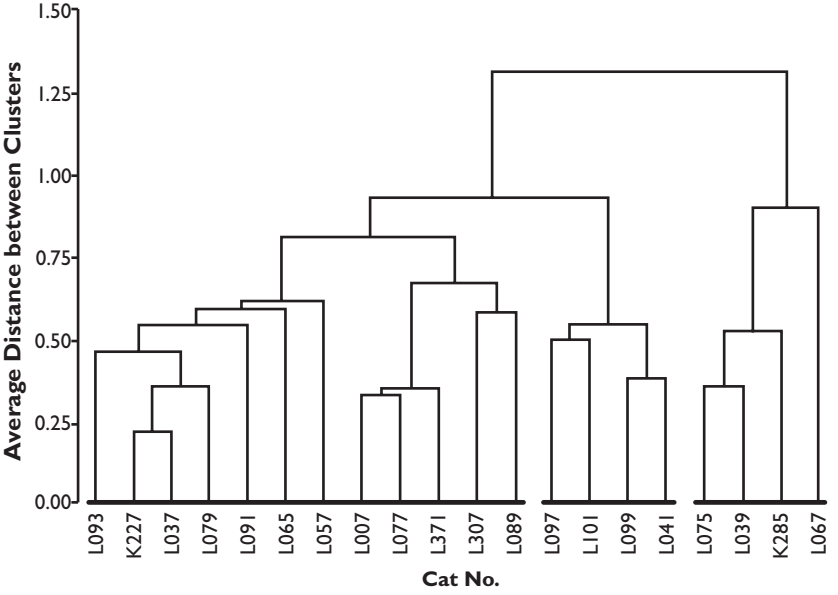

Figure 2. Cluster analysis of cats based on an average of scores over all four FTPs for each response during the FTP. Heavy, horizontal lines directly above cat identification numbers indicate clustering of cats together in groups. See Table 2 for comparison of this grouping with overall acceptable and questionable score ranking.

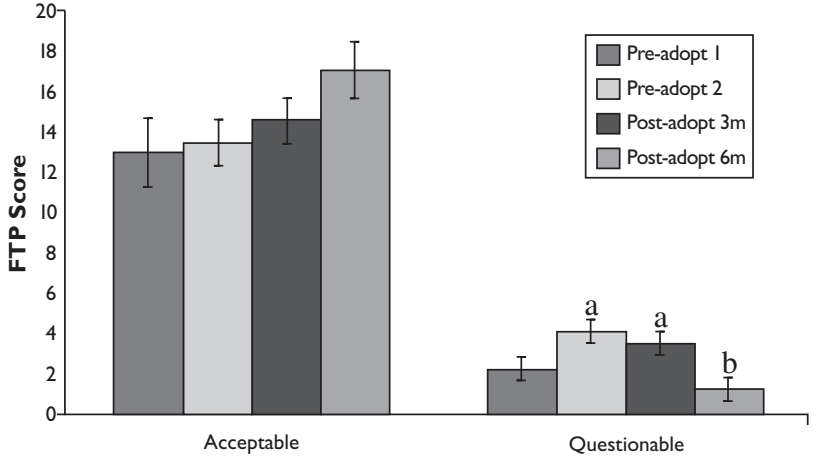

Figure 3. Temporal changes in acceptable and questionable FTP scores. Pre-adopt 1 = pre-adoption FTP in September; pre-adopt 2 = pre-adoption FTP in October; post-adopt 3m = FTP three months post-adoption in December-January; post-adopt 6m = FTP six months post-adoption in April.

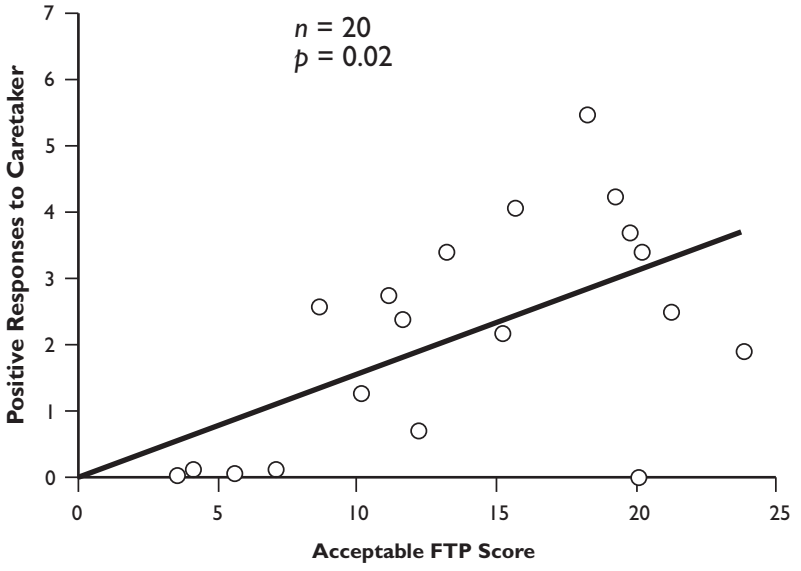

Figure 4. The relationship between acceptable FTP scores and positive responses of cats to familiar caretakers.

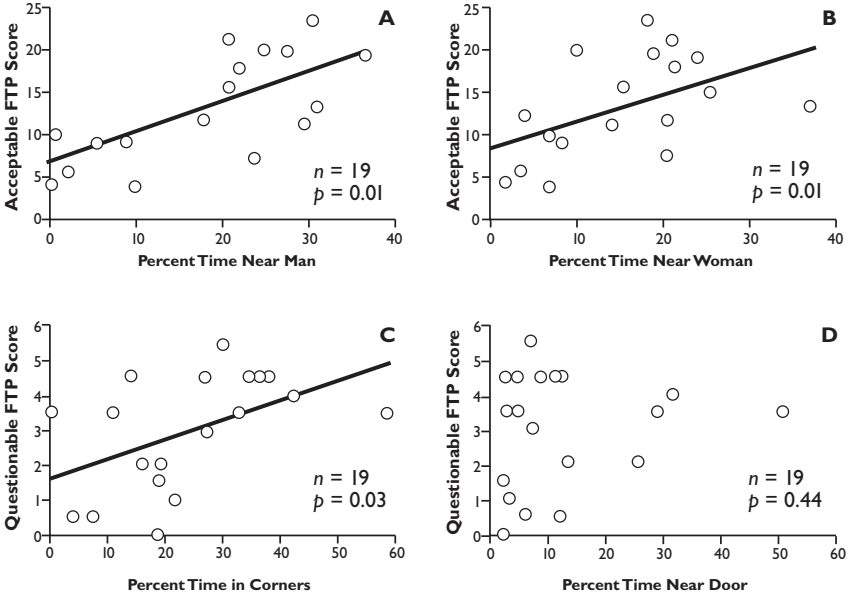

Figure 5. Relationships between A) acceptable FTP scores prior to adoption and average percent time cats spent near the unfamiliar man across all open field tests, B) acceptable FTP scores prior to adoption and percent time cats spent near the unfamiliar woman across all open field tests, C) Questionable pre-adoption FTP scores and percent time spent in corners during the open field tests, and D) questionable pre-adoption FTP scores and percent time spent near the door during the open field tests.

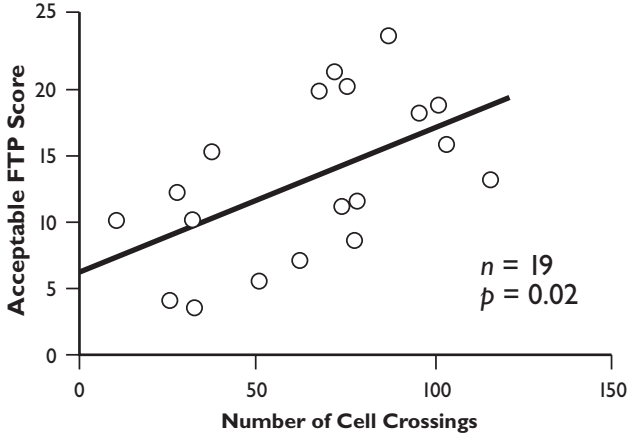

Figure 6. The relationship between the average number of cell crossings cats made across all open field tests and acceptable FTP scores prior to adoption.

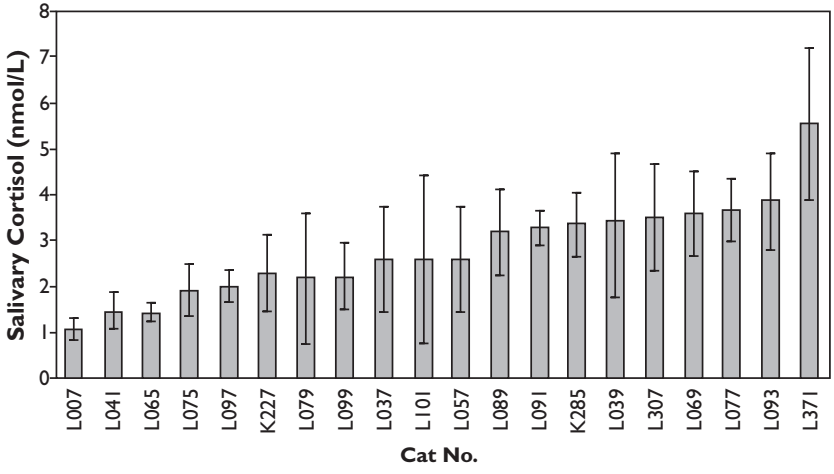

Figure 7. Mean salivary basal cortisol levels in individual cats prior to adoption, displayed from lowest to highest levels. Error bars represent ± SEM.

Correlations between temperament test scores and basal cortisol levels

No relationship was observed between basal cortisol levels from saliva and acceptable or questionable scores on pre-adoption FTPs (n=19, t=0.70, p=0.49 and n=19, t=0.11, p=0.91, respectively). No relationship was observed between basal cortisol levels and the groups of cats that resulted from the FTP rankings ($F_{(2,17)}$=0.39, p=0.68).

Discussion

Feaver, Mendl and Bateson (1986) and Turner (2000a) used subjective measures of cat individuality, which, while they were strongly correlated with ethological observations and were quick and easy to perform, required that the evaluator know the cat well enough to rate many aspects of the cat's behavior and individuality. This is not practical in a shelter setting or veterinarian's office where staff may have limited time and opportunity to observe each cat, though experienced personnel are likely practiced at assessing animals quickly. Additionally in these situations, the cats themselves may have limited opportunity to perform a wide range of behaviors and are likely to be stressed in such situations, providing staff with little information on which to base their judgments.

Other tests, such as the "cat stress test" developed by Kessler and Turner (1997), assess the general levels of stress a cat is experiencing in a given situation rather than temperament. While this test evaluates a cat's reaction to situations that may be stressful, such as housing or density (Kessler and Turner 1997, 1999a, 1999b), and perhaps could evaluate responses to handling or presence of humans, it does not provide an indication of a cat's affinity for humans.

The ideal temperament test, then, would combine ease of use, quickness, consistency, and reliability. However, few tests of this type exist for cats (Lee et al. 1983; McCune 1995) and these have either not been scientifically validated or not assessed for consistency and reliability across time, or correlated with other measures. The temperament test used in the current study was developed by Lee et al. (1983) for use in evaluating cats for placement in nursing homes. The test is very similar to those used to evaluate dogs for adoption from shelters, and it evaluates cats' responses to a variety of increasingly interactive and challenging situations in a novel environment: ranging from a novel person calling the cat to this same person involving the cat in play, and the cat's reaction to a startling noise. This type of test, which allows for an objective tester rather than one with extensive knowledge of the cat, has the advantage of being simple to perform and score. As we used the test, we found some of the responses listed under various phases of the test did not occur (Table 1). However, as our cats were from a homogenous genetic and social background, these responses may be observed when a more heterogeneous population of cats is examined and therefore should be retained in the FTP until such testing has been conducted. After more extensive validation in practical settings, it may be possible to simplify the FTP to make it more economical for shelter use.

Results from the current study indicate that the FTP performs consistently over time, even when the cat's circumstances change dramatically following adoption. Acceptable scores on the FTP correlated positively with the

cats' responses to familiar caretakers and unfamiliar humans in housing rooms and novel (i.e., the open field test) environments, suggesting that the test also accurately measures affinity for humans. All of the cats studied received equal amounts of socialization as kittens, therefore it will be important to compare FTP scores from cats that received varying amounts of socialization to ensure the test is capable of measuring the full range of variability demonstrated by cats in their affinity for humans. Additionally, the FTP itself should be performed on cats in familiar environments to evaluate scores under comfortable circumstances. This would allow for comparison with scores from the novel situations used in the current study, mimicking potential changes in scores that could occur between shelter and home environments.

Examination of changes in scores over time shows that they remain fairly stable, providing a consistent measure over eight months and changing circumstances. Observed decreases in questionable scores suggest that cats might develop more affinity to their adopters over time or with age, though the decrease was not enough to markedly change the rankings of cats between the first and last test, particularly of those cats that fell into the lowest ranking group.

A negative correlation between acceptable and questionable scores on the FTP indicates that the test is capable of discriminating between cats that react favorably to humans from those that do not. If acceptable and questionable scores had been positively correlated, the test might have been measuring active versus passive responses of cats rather than temperament. If the scores were uncorrelated, however, the test would not have much predictive value because a cat with a high acceptable score could be as likely to have a high as a low questionable score. Using FTP scores, the cats in the present study could be ranked by larger acceptable scores and smaller questionable scores (Table 2). This system resulted in three natural groups of cats, with lowest ranking cats, in particular, separated from the other groups by ranking (Table 2), which compared favorably with the statistical results of the cluster analysis (Figure 2). Cats in the high group had both high acceptable and low questionable scores, and as such are cats that could be placed with a family, a novice cat owner, or a person desiring a sociable pet. Cats in the middle group had mid range scores in both the acceptable and questionable categories, suggesting that they might require a bit more experience and understanding from an owner (and perhaps only older children in the household) in order to form an attachment. Cats in the low group had both low acceptable scores and the highest questionable scores. These cats would require an experienced cat owner or someone that does not expect a sociable, attention-seeking pet. For example, in the present study, cat K285 was adopted by a single woman living alone, however, the woman did not

mention that her grandchildren visited frequently. A month following adoption the cat was returned to MSU because she was not social with the grandchildren, as could have been predicted by her FTP scores.

Lowe and Bradshaw (2001) showed that features such as investigative behavior and boldness were stable features of cats (between four months and two years of age) studied in the home. However, other work also indicates that many aspects of cat behavior and temperament continue to change as the animals mature (Reisner et al. 1994), making a temperament test of kittens unpredictive of adult behavior, as is the case with puppies (Wilsson and Sundgren 1998). However, knowledge of early socialization of cats towards humans during the sensitive phase in the first few months of life or friendliness of the father may be able to suggest the affinity a kitten will have for humans when it matures (Karsh and Turner 1988; McCune 1995; Turner 2000b).

Use of the feline temperament test is not intended to condemn those cats which fall into the lowest ranking group; rather it offers a realistic assessment of their temperament in order to match cats with the expectations of adopters. The goal is to use this tool to achieve successful adoptions of cats as measured by low return rates and high owner satisfaction, as satisfaction and retention of adopted cats have been related to the individuality, compatibility, and behavior of the cat (Salman et al. 1998; Neidhart and Boyd 2002). Similar studies have demonstrated that dogs' performances on temperament tests can be related to problem behavior after adoption (van der Borg, Netto and Planata 1991; Ledger and Baxter 1997), giving shelters the ability to use such tests to place dogs more appropriately.

Cortisol levels reflect activity of the hypothalamic-pituitary-adrenal axis, which is a stress responsive system, and measures of salivary cortisol have been validated in many species, including dogs, as a non-invasive alternative to measuring cortisol in plasma (Cooper et al. 1989; Beerda et al. 1996). In keeping with previous findings that feline adrenal function does not have a circadian rhythm (Johnston and Mather 1979), we found no difference in cortisol levels between samples taken at 07:00 versus 19:00.

The absence of a relationship between cortisol levels in saliva and performance on the temperament test may suggest no correlation between intrinsic cortisol levels and aspects of cat temperament important in cat–human interactions. A more relevant comparison, however, may be found between individual salivary cortisol levels (as in Figure 7) and scores from a cat stress test (Kessler and Turner 1997). Alternatively, use of saliva as a means of non-invasively sampling cat cortisol levels may not provide accurate information on physiological stress, as little cortisol is likely excreted in saliva and what is present may be rapidly metabolized (Carlstead et al.

1992; Carlstead, Brown and Strawn 1993; Graham and Brown 1996; Schatz and Palme 2001). However, Beerda et al. (1996) correlated cortisol levels of saliva with those of plasma in dogs, despite the fact that these levels were ten-fold lower than those in plasma, suggesting that with validation, use of saliva to measure cortisol levels in cats may also be possible.

Conclusion

Currently, cats are often selected for adoption on superficial characteristics, such as coat color, while the cat's behavior is a frequently cited reason for relinquishment of cats to shelters where many are subsequently euthanized. While many organizations involved in cat adoption realize the need to assess a cat's temperament in order to place the cat in a compatible home, feline temperament tests either have not been validated or made available for this purpose. This study demonstrates that temperament testing in cats can provide an accurate and consistent measure of a cat's sociability, aggressiveness and adaptability, and could prove a useful tool for shelter staff, veterinarians and others needing to assess the temperament of a cat.

Acknowledgements

This project was funded by a grant from the Companion Animal Fund at the College of Veterinary Medicine at Michigan State University awarded to Dr. S. O. Walshaw and Dr. P. Brunner. Christine Heinz provided much appreciated assistance with data collection. Our thanks to Dr. C. Bollinger and Dr. R. Walshaw for sharing their cats with us for this study.

Notes

1. Responses listed on the original FTP that never occurred during our study included: strikes hand, threatens to strike hand, bites or attempts to bite, jumps up on lap, ignores toy, attends to something else in the room rather than the toy, does not hear object dropped, startles and runs to hide, startles then shows aggressive posture when item is dropped. The majority of these unobserved responses fell into the "questionable" response category. As our cats were from a homogenous genetic and social background, these responses may be observed when a more heterogeneous population of cats is examined and should be retained in the FTP until such testing has been conducted.

References

Beerda, B., Schilder, M. B. H., Janssen, N. S. C. R. M. and Mol, J. A. 1996. The use of salivary cortisol, urinary cortisol, and catecholamine measurements for a noninvasive assessment of stress response in dogs. *Hormones and Behavior* 30: 272–279.

Bradshaw, J. W. S. and Cook, S. E. 1996. Patterns of pet cat behaviour at feeding occasions. *Applied Animal Behaviour Science* 47: 61–74.

Carlstead, K., Brown, J. L., Monfort, S. L., Killens R. and Wildt, D. E. 1992. Urinary monitoring of adrenal responses to psychological stressors in domestic and nondomestic felids. *Zoo Biology* 11: 165–176.

Carlstead, K., Brown, J. L. and Strawn, W. 1993. Behavioral and physiological correlates of stress in laboratory cats. *Applied Animal Behaviour Science* 38: 143–158.

Cooper, T. R., Trunkfield, H. R., Zanella, A. J. and Booth, W. D. 1989. An enzyme-linked immunoabsorbent assay for cortisol in the saliva of man and domestic farm animals. *Journal of Endocrinology* 123: R13–16.

Durr, R. and Smith, C. 1997. Individual differences and their relation to social structure in domestic cats. *Journal of Comparative Psychology* 4: 412–418.

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

Graham, L. H. and Brown, J. L. 1996. Cortisol metabolism in the domestic cat and implications for non-invasive monitoring of adrenocortical function in endangered felids. *Zoo Biology* 15: 71–82.

Hennessy, M. B., Voith, V. L., Mazzei, S. J., Buttram, J., Miller, D. D. and 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.

Johnston, S. D. and Mather, E. C. 1979. Feline plasma cortisol (hydrocortisone) measured by radioimmunoassay. *American Journal of Veterinary Research* 40: 190–192.

Karsh, E. B. and Turner, D. C. 1988. The human–cat relationship. In *The Domestic Cat: The Biology of its Behaviour,* 1st ed., 159–177, eds. D. C. Turner and P. Bateson. Cambridge: Cambridge University Press.

Kessler, M. R. and Turner, D. C. 1997. Stress and adaptation of cats (*Felis silvestris catus*) housed singly, in pairs and in groups in boarding catteries. *Animal Welfare* 6: 243–254.

Kessler, M. R. and Turner, D. C. 1999a. Socialization and stress in cats (*Felis silvestris catus*) housed singly and in groups in animal shelters. *Animal Welfare* 8: 15–26.

Kessler, M. R. and Turner, D. C. 1999b. Effects of density and cage size on stress in domestic cats (*Felis silvestris catus*) housed in animal shelters and boarding catteries. *Animal Welfare* 8: 259–267.

Ledger, R. A. and Baxter, M. R. 1997. The development of a validated test to assess the temperament of dogs in a rescue shelter. In *Proceedings of the First International Conference on Veterinary Behavioral Medicine*, 87–92, eds. D. S. Mills, S. E. Heath and L. J. Harrington. Wheathampsted, UK: Universities Federation for Animal Welfare.

Lee, R. L., Zeglen, M. E., Ryan, T. and Hines, L. M. 1983. Guidelines: animals in nursing homes. *California Veterinarian* 3: 22a–26a.

Lowe, S. E. and Bradshaw, J. W. S. 2001. Ontogeny of individuality in the domestic cat in the home environment. *Animal Behaviour* 61: 231–237.

Martin, P. 1986. An experimental study of weaning in the domestic cat. *Behaviour* 99: 221–249.

McCune, S. 1995. The impact of paternity and early socialization on the development of cats' behaviour to people and novel objects. *Applied Animal Behaviour Science* 45: 109–124.

Mendl, M. and Harcourt, R. 2000. Individuality in the domestic cat: origins, development and stability. In *The Domestic Cat: The Biology of its Behaviour*, 2nd ed., 47–64, eds. D. C. Turner and P. Bateson. Cambridge: Cambridge University Press.

Mertens, C. 1991. Human–cat interactions in the home setting. *Anthrozoös* 4: 214–231.

Mertens, C. and Turner, D. C. 1988. Experimental analysis of human–cat interactions during first encounters. *Anthrozoös* 2: 83–97.

Neidhart, L. and Boyd, R. 2002. Companion animal adoption study. *Journal of Applied Animal Welfare Science* 5: 175–192.

New, J. C., Jr., Salman, M. D., King, M., Scarlett, J. M., Kass, P. H. and Hutchison, J. M. 2000. Characteristics of shelter-relinquished animals and their owners compared with animals and their owners in U.S. pet-owning households. *Journal of Applied Animal Welfare Science* 3: 179–200.

Podberscek, A. L., Blackshaw, J. K. and Beattie, A. W. 1991. The behaviour of laboratory colony cats and their reactions to a familiar and unfamiliar person. *Applied Animal Behaviour Science* 31: 119-130.

Reisner, I. R., Houpt, K. A., Erb, H. N. and Quimby, F. W. 1994. Friendliness to humans and defensive aggression in cats: the influence of handling and paternity. *Physiology and Behaviour* 55: 1119–1124.

Salman, M. D., New, J. G., Jr., Scarlett, J. M., Kass, P. H., Ruch-Gallie, R. and Hetts, S. 1998. Human and animal factors related to the relinquishment of dogs and cats in 12 selected animal shelters in the United States. *Journal of Applied Animal Welfare Science* 1: 207–226.

Schatz, S. and Palme, R. 2001. Measurement of faecal cortisol metabolites in cats and dogs: a non-invasive method for measuring adrenocortical function. *Veterinary Research Communications* 25: 271–287.

Serpell, J. A. 1996. Evidence for an association between pet behavior and owner attachment levels. *Applied Animal Behaviour Science* 47: 49–60.

Turner, D. C. 2000a. Human–cat interactions: relationships with, and breed differences between, non-pedigree, Persian and Siamese cats. In *Companion Animals and Us: Exploring the Relationships between People and Pets,* 257–271, eds. A. L. Podberscek, E. S. Paul and J. A. Serpell. Cambridge: Cambridge University Press.

Turner, D. C. 2000b. The human–cat relationship. In *The Domestic Cat: The Biology of its Behaviour*. 2nd ed., 191–206, eds. D. C. Turner and P. Bateson. Cambridge: Cambridge University Press.

Turner, D. C. and Stammbach-Geering, K. 1990. Owner assessment and the ethology of human–cat relationships. In *Pets, Benefits and Practice*, 25–30. ed. I. Burger. London: BVA Publications.

van der Borg, J. A. M., Netto, W. J. and Planta, D. J. U. 1991. Behavioural testing of dogs in animal shelters to predict problem behaviour. *Applied Animal Behaviour Science* 32: 237–251.

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