The Dissertation Committee for Amanda Claire Jones certifies that this is the approved version of the following dissertation:

DEVELOPMENT AND VALIDATION OF A DOG PERSONALITY QUESTIONNAIRE

Committee:

Samuel D. Gosling, Supervisor

Robert A. Josephs

James W. Pennebaker

Timothy J. Loving

J. Gregory Hixon

DEVELOPMENT AND VALIDATION

OF A DOG PERSONALITY QUESTIONNAIRE

by

Amanda Claire Jones, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2008

Dedication

To Peanut, Jackie, Baya, Rufus, and all the dogs whose life paths have yet to join with my own. You all do have your own unique personalities, but that fact is not all you have to share and to teach.

Acknowledgements

I am not much for mixing sentiment and work, but for once it seems appropriate.

First and, in the context of academic pursuits, foremost, thanks to Sam Gosling. Though I have not always liked it, I am thankful for your guiding, encouraging, and pushing me through the past five years. The specific lessons I've learned from you are too many to innumerate.

Second, my heart-felt thanks to Bob Josephs. You've been a lot of fun to have as an advisor. I've gotten to toy with new ideas one week, and scrap them entirely the next. And your support through life's little and not-so-little crises has been invaluable.

Third, I am especially thankful to my son Abram. You've given me a new perspective on life and motivation to finish my dissertation. I've learned more from the past 23 months as a mother than I ever learned in school. I'm almost afraid of how much remains to be learned. I have a new understanding of my own parents' loving support.

And finally, I'd like to acknowledge the unfaltering support of my friends, particularly Miki and Conor. As with any truly loving and supportive friendships, if I were to try to describe all the ways I have benefited from our relationships, I would go on forever. I'd probably start blabbering about buffering effects, social support, and the physiology of stress. If I were to try to capture just the very important ways in which you have helped me, I would go on for at least as many pages as follow in this very lengthy dissertation. Instead, just know that you're loved, valued, and appreciated.

DEVELOPMENT AND VALIDATION

OF A DOG PERSONALITY QUESTIONNAIRE

Amanda Claire Jones, Ph.D.

The University of Texas at Austin, 2008

Supervisor: Samuel D. Gosling

Many groups, such as rescue and service-dog programs, are interested in assessing dogs' personalities. These groups often need to assess large numbers of dogs with limited resources (e.g., in terms of facilities, trained assessors, time, money). To meet these groups' requirements, an assessment tool that measures canine personality rapidly and is demonstrably reliable and valid is needed. The Dog Personality Questionnaire (DPQ) was developed to fill this gap. This dissertation describes a series of six studies designed to develop and evaluate the DPQ.

To ensure that the final instrument built on previous research and was based on a comprehensive item pool, 1,200 descriptions were culled from the dog-personality assessment literature, shelter assessments, and dog experts' input (e.g., researchers, trainers, veterinarians). Three expert judges narrowed this list to 360 items. In Study 1, these items were administered to 152 participants who gave feedback on the items' applicability and ease of use.

In Study 2, exploratory factor analysis was used to determine the number of factors underlying the 360-item questionnaire, based on 3,737 participants' ratings of their dogs. Convergent criteria favored five factors, labeled as Fearfulness, Aggression

towards People, Aggression towards Animals, Activity/Excitability, and Responsiveness to Training. Narrower facets within each factor were also identified. On the basis of item analyses, the questionnaire was shortened to 102 items.

In Study 3, the 102-item questionnaire was administered to 2,556 new participants and further exploratory and confirmatory factor analyses were used to examine the robustness of the five-factor solution. Items were then evaluated in terms of factor- and facet-loadings, content validity, internal consistency, and other criteria in order to shorten the questionnaire to a more manageable, 75-item form, and an even briefer 45-item form.

In Studies 4-6, the psychometric properties of the 75-item and 45-item DPQ were further evaluated. The DPQ was shown to have acceptable levels of inter-rater reliability (Study 4), test-retest reliability (Study 5), and predictive validity (Study 6). Discussion focuses on evaluating how well the DPQ meets the criteria that guided its development.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

It is now widely accepted that non-human animals can be characterized in terms of personality (Gosling & Vazire, 2002). One species in which personality has been examined extensively is the domestic dog (*Canis familiaris*). It is easy to understand why dogs and their personalities have garnered extensive research attention. Not only are they the most common pet in the United States (Humane Society of the United States, 2007), ubiquitous wherever humans live, but they are also used to realize a number of essential applied goals, such as guiding visually impaired people and searching for explosives. As a result, many groups of people are interested in assessing the dogs' personalities efficiently, reliably, and accurately. These groups include: (a) potential pet owners wanting to find a dog suitable for their particular circumstances (e.g., family dog vs. guard dog), (b) shelters seeking to identify suitable homes for dogs, (c) service-dog programs (e.g., guide dogs, hearing dogs), and (d) working-dog programs (e.g., patrol dogs, detection dogs). Many of these groups must evaluate dogs using very limited resources. They may have little time with each dog, few trained evaluators, minimal facilities, and other limitations. However, a well-validated, reliable, effective instrument for measuring dog personality that is easy and quick to use and widely applicable has yet to be developed. The goal of the research described here is to develop such an instrument and evaluate its psychometric properties, including aspects of reliability and validity.

Before a tool for assessing the broad construct of dog personality could be developed, the construct to be studied had to be explicitly defined (e.g., DeVellis, 2003). For the term "personality" to be used to characterize non-human animals, and dogs specifically, "personality" must be defined and how (or whether) it differs from "temperament" must be determined. "Personality" is often reserved for discussing adult humans, and "temperament" for discussing human infants and non-human animals. However, these uses are not consistently maintained, and the terms tend to be used interchangeably (McCrae et al., 2000). I use the term "personality" because the

distinction between the two terms is not maintained sufficiently, nor is a distinction between them generally useful for the current purposes.

Finding a definition of *personality* to suit all applications of the term is challenging. The phenomena studied by personality psychologists include temperament and character traits, dispositions, goals, personal projects, abilities, attitudes, physical and bodily states, moods, and life stories (John & Gosling, 2000). Only a very broad (and thus somewhat vague) definition could satisfy many personality psychologists simultaneously. For example, *personality* can be defined as those characteristics of individuals that describe and account for consistent patterns of feeling, thinking, and behaving (Pervin & John, 1997), a definition broad enough to capture most phenomena studied by personality psychologists. I adopt this broad definition, and my use of the term personality includes personality in all non-human animals as well as humans. The dog personality assessment tool developed in this dissertation will assess personality in terms of traits (also called factors or dimensions). I adopt Gosling's (1998) definition of traits as "aggregate summary trends in behavior". An individual's traits are also consistent and reflected in the individual's feeling, thinking, and behaving across time and situations.

In order to ground the current studies in the existing literature, in Chapter 2 I summarize the dog personality literature published prior to the onset of the current study and providing the basis for the current study. Chapter 2 includes a summary of my published review (Jones & Gosling, 2005) of the literature, supplemented with four studies published between the review's completion and the onset of this study (April, 2004-June, 2005). In Chapter 3, I define and discuss the criteria that guided my selection of an assessment method, and which I attempt to satisfy with the design of a new tool for assessing personality in dogs. In the next six chapters, I describe Studies 1 through 6, detailing the development of a questionnaire for assessing personality in individual dogs, from initial item generation to validation of the tool. Studies 1, 2, and 3 lay the groundwork for the latter three. Each study is outlined in greater detail below.

STUDY 1: PILOT TESTING THE INITIAL ITEM POOL

In Chapter 4, I describe Study 1, in which my goal was the development of an initial questionnaire for assessing dog personality; this study had two parts. First, I generated a pool of potential items for use in the questionnaire. The pool of items generated was intended to be as comprehensive as possible, representing as many aspects of dog behavior and personality as possible. In order to compile a very comprehensive list of descriptors, I drew together descriptors from multiple sources, including the dog personality and temperament research literature and tools used in applied settings (e.g., shelters); these sources were supplemented with items generated by dog experts. This process resulted in an initial list of 1,284 descriptors. These 1,284 descriptors served as the starting point for the process of sorting potential items based on content, eliminating items that did not fit my criteria (e.g., were applicable to very narrow contexts or only certain types of dogs, like guide dogs), and creating a list of 360 questionnaire items.

In the second part of the study, I administered the items to a small sample of participants in order to attain feedback that would then help me to identify and revise questionnaire items that were difficult for participants to understand or that described situations participants' dogs did not encounter. In this part of Study 1, the 360 items were piloted online with a sample of 152 dog owners who volunteered to fill out the questionnaire, and the questionnaire items were modified based on their feedback.

STUDY 2: FACTOR STRUCTURE IDENTIFICATION

Study 2 (described in Chapter 5) had two goals: to determine the number of factors underlying the behaviors and descriptors in the 360-item questionnaire, and to begin developing a more concise and coherent scale. I administered the questionnaire online to dog owners who volunteered to participate, then I used exploratory factor analysis (EFA; principle components analysis [PCA] with varimax rotation) of the responses from 3,737 participants to the lengthy questionnaire. Convergent criteria indicate that both the four- and five-factor solutions were robust. For reasons of interpretability, the five-factor solution was selected.

Once the five factors (or traits) had been confirmed, items loading on each trait were analyzed using EFA (PCA, oblique [promax] rotation) in order to determine the number of facets within each trait. Fifteen facets were identified. In addition to other criteria (e.g., item univocality), the results of the trait and facet analyses provide guidance for creating a new, briefer and more manageable 102-item form of the questionnaire to be administered in Study 3.

To assess the fit of the five-factor solution to the revised list of 102 items, I divided the participant set into two randomly selected halves, then repeated the EFA procedure on one half of the data (N = 1,868). Again, the five-factor solution was found to be robust and interpretable. I then confirmed, using structural equation modeling (SEM) to perform confirmatory factor analysis (CFA), that the five-factor model adequately fit the second half of the data (N = 1,869).

STUDY 3: FACTOR STRUCTURE CONFIRMATION

In Chapter 6, I describe Study 3, in which my primary goal is to determine how well the five-factor structure model found in Study 2 replicates in the new, shorter questionnaire and a new sample of participants. Replication of the factor solution in a new sample is key to establishing the solution's generalizability; if the solution does not generalize to the new sample and questionnaire, then it may be idiosyncratic to Study 2. I administered the 102-item questionnaire online to a new online sample of volunteer participants, who each rated a single dog (N = 2,556). I then used SEM to perform confirmatory factor analyses (CFA) to test the fit of the hypothesized five-factor model (derived in Study 2) to the newly collected data.

Next, I examined the replicability of the five-factor model in the data from Study 3. I used the same procedure as in Study 2. First, I divided the participant set into two randomly selected halves. Then I repeated the EFA procedure on one half of the data (N = 1,278). Again, the five-factor solution was found to be the most robust and interpretable. Finally, I verified, using SEM to perform CFA, that the five-factor model adequately fit the second half of the data (N = 1,278).

I also examined psychometric properties of the questionnaire. In addition to the CFA results, examination of the psychometric properties (e.g., content validity, construct validity, including discriminant and convergent validity; internal consistency) guided the creation of two final forms of the Dog Personality Questionnaire (DPQ). The questionnaire from Study 2 was shortened to a "long form" of five items per facet (or 75 items) and a "short form" with three items per facet (or 45 items).

STUDY 4: INTER-RATER RELIABILITY

The goal of Study 4, described in Chapter 7, was to address another type of reliability: inter-rater reliability. If a tool has high inter-rater reliability, then the ratings different observers ascribe to a target when using the tool are highly correlated. To assess how well the ratings of dogs' personalities made using the DPQ generalize across observers, 99 participant pairs in which both people were familiar with the same dog completed the online questionnaire rating that dog. I then examined how highly the pairs of participants' ratings correlate on each item, facet, and factor of both the long form and the short form of the questionnaire. The DPQ was found to have inter-rater reliability rates generally comparable to those found in human personality rating studies.

STUDY 5: TEST-RETEST RELIABILITY

The goal of Study 5, described in Chapter 8, was to address a third type of reliability: test-retest reliability. For a questionnaire, test-retest reliability, or reliability across time, addresses the consistency of a single observer's reports taken at different points in time. (This differs from a behavioral test, e.g., a Test Battery or Observational Test, in which test-retest reliability refers to the consistency of the dog's behavior at two or more different assessment times.) If a tool has high test-retest reliability, then the scores that are obtained when the test is administered at time 1 agree, or are highly correlated, with the scores obtained when the test is administered again (i.e., at time 2, time 3, and so on). To assess the DPQ's test-retest reliability, I administered the online questionnaire twice, with approximately four to six weeks between administrations, to 100 participants. I then examined how well the pairs of ratings correlate on each item,

facet, and factor of both forms of the questionnaire. The DPQ was found to have test-retest reliability rates generally comparable to those found in human personality rating studies.

STUDY 6: PREDICTIVE VALIDITY

The aspects of reliability assessed in Studies 3-5 are crucial prerequisites for predictive validity. Predictive validity is the extent to which scores on a given measure are related to some external, independent measure. In Chapter 9, I describe Study 6, the goal of which was to address how well participants' ratings of their dogs on the DPQ predict the dogs' behavior on independent measures. For purpose of this study, I devised a new Test Battery to assess behavioral traits thought to be related to items on the long form of the DPQ (the five dog personality factors and their facets). One hundred participant-dog pairs were recruited to take part in the study. Owners and kennel staff completed the long form of the DPQ on paper, and the dogs were assessed using the new Test Battery. Specific behavioral descriptions assessed on the Test Battery were predicted to be related to and unrelated to specific factors and facets measured on the DPQ; these relationships were assessed using convergent and discriminant correlations, respectively. The DPQ was found to have relatively high convergent validity, as compared with other dog personality assessments; discriminant validity results were mixed.

CHAPTER 2

Literature Review

INTRODUCTION

Early in the twentieth century, Nobel laureate Ivan Pavlov began a research program designed to identify the basic types of canine personality (e.g., Pavlov, 1906). Despite this auspicious start, the study of temperament and personality in animals did not evolve into a major area of research except, of course, in humans. Yet pet owners and practitioners working with dogs have long recognized that canine personality is important. It influences dogs' behavior and responses to their environments. Studies of dog personality have striven to fulfill many goals, from identifying a puppy test that will predict adult guide-dog behavior (e.g., Goddard & Beilharz, 1984a, 1984b, 1986), to examining the heritability of personality traits (e.g., Reuterwall & Ryman, 1973; Wilsson & Sundgren, 1998).

These studies are unified by a common interest in dog personality, but the researchers conducting these studies come from a wide variety of backgrounds, bringing with them assorted perspectives and publishing in a broad range of journals. As a consequence of their distinct disciplinary affiliations and research goals, these efforts at understanding personality in dogs have followed largely independent paths. The result is that it is hard to keep track of the various findings—the studies are scattered across journals in anthrozoology, psychology, biology, animal behavior, and veterinary medicine, among others.

Each of these discipline-bound studies is interesting and valuable in its own right, but it provides only a relatively narrow glimpse of dog personality. Taken together, the studies provide broader insight not only into dog personality, but also into the strengths and weaknesses of the methods used to assess dog personality. The goal of the current chapter is to describe and discuss the various methods used to assess dog personality, to summarize the major findings from the dog personality literature, to pinpoint major gaps in science's understanding of dog personality, and to use those gaps to inform

suggestions about the research challenges that lay ahead.

Specifically, this chapter starts by examining general trends in research on dog personality. What methods have been used, what breeds have been assessed, and what other trends can be identified? The next issue addressed is the specific domains, or traits, of personality that have been identified in dogs. Specifically, which traits have received the most cross-study support? Next, meta-analyses of past work on the reliability and validity of personality tests are used to evaluate the effectiveness of personality measures. Finally, the findings are drawn together to offer 18 broad conclusions about the field and identify the major questions that remain to be addressed.

The research reviewed in this chapter should be of interest both to practitioners and to researchers. Relevant practitioners include those interested in the practical task of identifying dogs whose personalities are well-suited to working as guide dogs, hearing dogs, or police dogs, and for various other jobs in which dogs assist people in their daily lives. The findings will also be relevant to dog shelters and rescue centers, which often rely on personality tests as a guide for placing dogs in suitable homes, and for individual pet owners interested in finding a pet suitable for their lifestyle (e.g., Coren, 1995, 1998; Hart & Hart, 1985, 1995; Tortora, 1983). With the recent moves in the United States to pass breed-specific legislation, intended to limit and control the ownership of specific breeds, this work will also be of interest to workers in animal welfare and social policy. Finally, the review will be useful to the growing body of research scientists interested in using animal models to examine basic issues in human psychology (Gosling, 2001) and animal behavior (Dugatkin, 2004).

LITERATURE REVIEW

To be certain that I included as many potentially relevant studies as possible, I searched the PsycINFO, Biosis, and Web of Science databases for articles in which either personality or temperament in dogs were examined. I included only those studies in the published empirical research literature. As such, my review did not include the methods that are frequently used and well-regarded in applied settings (e.g., Sue Sternberg's

Assess-a-Pet and Emily Weiss's the SAFER and SAFER-II tests) but for which data are not yet published.

Included in this chapter are only those dog personality or temperament studies published prior to the beginning of the development of the dog personality assessment tool that is the focus of the current research. That is, all articles examined in this review were published prior to June, 2005; articles published since are excluded.

Literature search procedures

The literature search used two basic procedures: Generating a large pool of potentially relevant articles, and selecting a smaller subset of articles for inclusion in the final review. These two procedures were used iteratively, such that one cycle generated a pool of potential articles and selected a subset of them for review, and this subset of articles provided the starting point for a subsequent cycle.

In the initial search cycle searched PsycINFO, Biosis, and Web of Science databases for all articles containing the keywords "dog" and "temperament," or "dog" and "personality." Searches for descriptors such as "aggressive" or "fearful" were not included because almost all behavior can be described as related to some domain; including these articles would have meant capturing a vast number of articles that did not focused personality constructs but merely included behaviors related to a personality domain. For example, the study of dogs' preference for humans by Topál et al. (1998) examined attachment behavior, including nervousness-related behaviors, but had no interest in individual differences in temperament per se. If an article did not even mention personality or temperament in the title, list of keywords, or abstract (i.e., the fields scanned in a keyword search), it was concluded that it was highly unlikely the research would be relevant to this review.

After eliminating duplicates, I examined the abstracts of the remaining reports to eliminate irrelevant articles. Articles varied in their relevance to research on dog temperament; some focused directly on temperament assessment but others clearly fell beyond the domain of this review. For example, one article examined the personalities of

people who strongly dislike dogs (Stubbs & Cook, 1999), and could therefore easily be classified as irrelevant. Although most articles could be unambiguously classified as clearly relevant or clearly irrelevant, there were a number of borderline reports that were distantly or obliquely related to temperament but did not fall neatly into the core set of clearly relevant papers. I retained these borderline articles for closer inspection.

This review cannot include every vaguely relevant reference so only the most important borderline studies were retained. Given the goals of my review, I selected those articles that were empirical, that were consistent with the definitions of temperament and personality described above, and that had a substantial focus on temperament or personality in dogs. Studies with only a cursory link to temperament were excluded. For example, I did not retain an article that described the working requirements for an Animal-Assisted Therapy dog (Hart, 2000); it explained the functional significance and role of the therapy dog, touching only briefly on the temperament requirements.

Inspection of the references cited in the selected articles revealed several studies that had not been identified in the initial search. Therefore, each time a new article was identified, I searched its references for other relevant articles. After repeating this process several times, my leads began to run dry and I was satisfied that I had captured the vast majority of relevant research. Nonetheless, given the great diversity of research, I wanted to make sure my own disciplinary perspective did not bias the review. Therefore, I asked colleagues in other fields and who study dog behavior to check the reference list and bring to my attention any studies I had missed. By the end of these search procedures, I had identified 55 articles, all but one of which are summarized in Table 2.1. This study (Campbell, 1972) was retained, because it is frequently referenced by and discussed in other studies, and because it seems to mark the beginning of a revival of interest in dog temperament.

Of course although I took care to identify all relevant articles, no selection procedure is flawless and I acknowledge that a few relevant studies will inevitably have slipped through my net. Nonetheless, I believe my review represents the most comprehensive summary to date of research on temperament and personality in dogs.

A General Survey of the Field

When reviewing a new field, the first major task is to step back and survey the general state of the field and identify the major trends. To this end, Table 2.1 summarizes the basic features of the studies included in my review. The first thing to note is that, with one exception, there is a great diversity of research. The one exception comes in terms of the constructs studied; as in Gosling's (2001) review of temperament in all non-human species, almost all the canine research has been on temperament traits, with almost no research on goals, motives, and other constructs.

In other respects, the studies are tremendously varied. They are drawn from a wide variety of disciplines, including animal behavior, biology, psychology, animal welfare, and veterinary medicine. The studies also have many different purposes, ranging from assessing temperament in specific breeds (e.g., Reuterwall & Ryman, 1973), to evaluating the domestic dog as a more general model of animal personality (e.g., Svartberg & Forkman, 2002). To help identify some specific patterns in this fragmented field, I propose several ways of summarizing the literature. These summaries are based on the methods of assessment, the breeds examined, the purpose of the studies, the age at which the dogs were tested, the breeding and rearing environment, and the sexual status of the animals.

Assessment methods

Table 2.1 is organized in terms of the four main methods by which dog temperament has been assessed: *Test Batteries*, *Ratings of Individual Dogs*, *Expert Ratings of Breed Prototypes,* and *Observational Tests*. A fifth category was composed of studies that combined more than one assessment method.

Test Batteries

As shown in the table, the most common method of assessment was the Test Battery, which appeared as the primary assessment method in 29% of the 55 studies reviewed. The core goal of studies using this method was to document dogs' reactions to specific stimuli. The tests were performed by presenting various, usually novel, stimuli

one at a time to a canine subject and recording its reaction(s). Thus, Test Batteries had two components: the tests themselves and the system for coding the dogs' reactions to the tests.

In theory, Test Batteries were the closest of the four methods to achieving objectivity, but in practice the levels of objectivity actually attained varied substantially. One of the more objective Test-Battery studies examined the relationship between Fearfulness and breed (Mahut, 1958). After presenting novel stimuli to target dogs, the researchers described the dogs' subsequent behaviors purely in terms of what was visually and auditorily observed over the next ten seconds.

Ratings of Individual Dogs

Ratings of Individual Dogs appeared in 18% of the studies reviewed. The goal of these studies was to gather information about individual dogs' behaviors and histories from an informant. One such data-gathering technique was to have a particular dog's owner state whether or not, or how often, his or her dog had engaged in a specified behavior (e.g., snapping at children). The owners who participated in such studies were usually preselected on the basis of group membership (e.g., owners of a specific dog breed). For example, Podberscek and Serpell (1996) asked English Cocker Spaniel (ECS) owners how likely, on a 5-point scale, their ECS was to act aggressively towards strange dogs, when reached for by a person, and in other situations. Although these methods are sometimes described as "subjective" approaches, Block (1961) long ago showed that the combined ratings of observers are largely independent of the idiosyncrasies of any one observer; therefore, when such ratings are aggregated, they are not appropriately characterized as "subjective."

Expert Ratings of Breed Prototypes

Expert Ratings of Breed Prototypes appeared in 16% of the studies reviewed. In these assessments, informants deemed by the researchers to be experts on dogs (e.g., American Kennel Club judges, veterinarians, dog trainers), described, ranked, or rated breeds of dogs as a whole rather than specific individual dogs. In these studies, the

experts could also make sex-specific judgments. Four of the nine reports included in this review are reanalyses of a single data set (Draper, 1995; Hart & Miller, 1985; Hart & Hart, 1985; Hart, 1995). These data were collected through 96 telephone interviews, conducted by three veterinary students (Hart et al., 1983; Hart & Miller, 1985). The students asked 48 obedience judges and 48 small-animal veterinarians to compare and rank a selection of seven breeds on 13 questions. When the data were combined, this resulted in the ranking of 56 total breeds on 13 behavioral traits, with 12 independent ratings of each breed on each item.

Observational Tests

Observational Tests were used in 13% of the studies. The overall goal of Observational Tests was to assess and describe relatively broad traits discernible in naturalistic environments, thus drawing broader conclusions about the dogs' temperaments and behavior patterns than is possible using Test Batteries. Like Test Batteries, Observational Tests had two major components: the test itself, and the system for scoring the dogs' performance on the test. Unlike Test Batteries, Observational Tests were usually conducted in carefully selected, but not controlled, environments and involved the fortuitous presentation of naturally occurring stimuli. For example, in one study, dogs were walked through a shopping center because it is an uncontrolled public area (Goddard & Beilharz, 1984b). Some Observational Tests also included the presentation of the kinds of experimental stimuli sometimes used in Test Batteries. The target dogs were usually assigned scores on various predetermined temperament traits based on overall observations; for example, in a series of studies, potential guide dogs were judged on cooperativeness based on all behaviors displayed during videotaped walks (Murphy, 1995, 1998).

Some of the studies reviewed (16%) did not fit neatly into any one of these categories because they used combinations of the assessment methods. An example of a study using combined methods was reported by Stephen and Ledger (2003). Dog owners filled out a questionnaire about their dog's behavior (i.e., Ratings of Individual Dogs) and

in a separate phase, unfamiliar testers put the dogs through a series of situations in a controlled environment and rated their behaviors (i.e., a Test Battery). In a final step, the researchers compared the scores derived from the two methods.

Breeds assessed

Another way to summarize the literature is in terms of the breeds assessed. Dogs come in an enormous variety of breeds, with as many as 150 breeds officially recognized by the American Kennel Club (AKC; http://www.akc.org/breeds/reg_stats.cfm, 2004) and many others not recognized by the AKC but described elsewhere (Morris, 2002; Wilcox & Walkowski, 1995). Given this variety, I examined whether the breeds assessed in these dog-temperament studies are representative of the breeds that exist, or whether there is a bias with some breeds particularly likely to garner research attention. To address this question, the breed composition of the studies is recorded in Table 2.1.

The Labrador Retriever, Golden Retriever, Beagle, and German Shepherd Dog (GSD) are, respectively, the first, second, third, and fourth most commonly registered breeds in the AKC (http://www.akc.org/breeds/reg_stats.cfm, 2004). As purebred pets and show dogs, they are extremely common. Overall, 90% of dogs examined were purebred. Two of these breeds—the Labrador Retriever and the GSD—were studied particularly frequently. Labradors and GSDs combined dominated the research literature, comprising 30% of the subjects in the studies reviewed. The GSD, which has been surpassed in popularity by the Beagle over the last few years according to the AKC registration records (http://www.akc.org/breeds/reg_stats.cfm, 2004), was the most frequently tested breed, comprising 24% of the dogs tested (9,253 dogs). Some studies examined huge numbers of these dogs. For example Reuterwall and Ryman's (1973) study involved 958 GSDs, tested at the Army Dog Training Center of Sollefteå, Sweden. The Labrador Retriever, the most commonly registered breed in the AKC, is the second most frequently tested breed, comprising 6% of the subjects. They too were occasionally present in large numbers in single studies. For example, 731 Labradors were in Goddard and Beilharz's (1982/83) study of animals with the Royal Guide Dogs for the Blind

Association of Australia.

As shown in Table 2.1, dog temperament assessment studies did not always rely on purebred dogs. Some of the dogs studied were the planned offspring of two purebreds of different breeds. In the studies reviewed, intentional crosses included 16 dogs evenly divided among all possible combinations of Labrador, GSD, Boxer, and Kelpie (Goddard & Beilharz, 1984a, 1986), and 145 Labrador/Golden Retriever crosses (140 in Serpell & Hsu, 2001; 5 in Murphy, 1995).

Also represented in the studies were less common purebred dogs (e.g., Bernese Mountain Dogs; Roll & Unshelm, 1997), and unintentional or unknown mixes of breeds. These studies are different from those not reporting breed in that they make clear that the dogs' involved are not just purebreds of unreported breed, but are actually mixed breeds. Only six studies reporting breed examined unintentional or unknown mixes, totaling 856 dogs (Goodloe & Borchelt, 1998; Hsu & Serpell, 2003; Rooney & Bradshaw, 2004; Seksel et al., 1999; Wahlgren & Lester, 2003; Weiss & Greenberg, 1997). Of these, 837 of them were in three studies using Ratings of Individual Dogs (Goodloe & Borchelt, 1998; Hsu & Serpell, 2003; Rooney & Bradshaw, 2004; Wahlgren & Lester, 2003). Of the remaining, ten were in a Test Battery which also included 50 purebred dogs (Seksel et al., 1999), and nine were in a study composed entirely of mixed breeds (Weiss & Greenberg, 1997).

Are some method-breed combinations more common than others? The breakdown of breeds by assessment method is clearly not random. The most salient patterns appear where huge numbers of dogs are assessed. For example, approximately one third (8,794 total dogs) of the dogs in Test Battery studies are GSDs (the most commonly assessed breed overall), tested for their potential as police and working dogs. More than 75% of all dogs in Observational Testing studies are Labrador Retrievers (the second most commonly assessed breed overall), tested for their potential as guide dogs (831 out of 1,089 dogs).

Purpose of study

Not surprisingly, given the diversity of fields doing research on dog temperament and personality, the studies reviewed varied widely in their goals. These goals included determining the suitability of a dog for guide-type work, selecting breeding stock for police-dog training centers, and assessing pet dogs' Fearfulness levels.

Ten of the studies reviewed focused on determining the suitability of a dog for guide-dog service work. For example, Goddard and Beilharz (1984a) devised a study to attempt to predict adult Fearfulness in potential guide dogs from tests conducted while they were still puppies.

Five studies aimed to determine suitability for police work and three others focused on suitability for related tasks, such as field work or tracking. For example, a Test Battery was developed for predicting adult police-dog effectiveness from the performance of approximately two-month-old puppies at the South African Police Service Dog Breeding Centre (Slabbert & Odendaal, 1999). This Test Battery included crossing obstacle courses, retrieving objects, novel and startling visual and auditory stimuli, and situations attempting to provoke aggressive behavior. High scores on the retrieval test at two months and the aggression test at nine months significantly predicted success as an adult police dog.

Three of the studies focused on determining the factors involved in aggressive behavior. For example, one study used Ratings of Individual Dogs to investigate whether red and golden ECSs display more aggressive behaviors than do other black and multi-colored ECSs (Podberscek & Serpell, 1996).

The goal of some puppy-temperament assessment methods was to help potential puppy buyers or adopters in selecting a suitable breed and a suitable individual puppy for themselves and their families. There are two types of assessment for this purpose: the breed profile created from Expert Ratings of Breed (e.g., Coren, 1995; Hart & Miller, 1985; Hart & Hart, 1985), and the puppy-behavior test, a type of Test Battery, to be performed by the puppy buyer (e.g., Campbell, 1972; examined in Beaudet et al., 1994).

A handful of other studies have scattered purposes, including developing

assessment tools for screening dogs for the presence or prevalence of behavior and temperament problems (Goodloe & Borchelt, 1998; Hsu & Serpell, 2003; Serpell & Hsu, 2001), evaluating previous tests (Beaudet et al., 1994; Weiss & Greenberg, 1995), evaluating the presence of personality traits in dogs (Draper, 1995; Gosling et al., 2003a; Royce 1955; Svartberg & Forkman, 2002), predicting post-adoption behavior problems in shelter dogs (Hennessy et al., 2001), and determining the relationship between physical build and temperament traits (Keeler, 1947; Lester, 1983).

Age at testing

As noted above, the goal of many studies has been to predict adult behavior from puppy temperament. This implies an age-related bias in the studies. To examine the extent of this bias, it is instructive to organize the studies in terms of the age at which the dogs were assessed. To facilitate this goal, in the text and tables I have converted the age information to a common metric of months. Of the studies reporting age at testing, over 20% of the assessments in this review were performed for the first time when the dogs were puppies between .23 months (i.e., 1 week) and six months of age. Ten were performed when the dog was between 10 and 24 months. Six of those that first assessed the puppies at six months old or younger also assessed the dogs on multiple subsequent occasions, with a final test at 12-24 months old; in these studies researchers tried to use scores from the puppy tests to predict behavior or aptitude when the dog was older (e.g., Wilsson & Sundgren, 1998).

The studies that tested dogs only once tended to test the dogs when they were older. Ten of these studies reported the age at which their first assessment took place as older than six months, and of these, eight were 12 months or older. Age is reported in only four of the Ratings of Individual Dogs studies (Hsu & Serpell, 2003; Podberscek & Serpell, 1996; Rooney & Bradshaw, 2004; Serpell & Hsu, 2001), and is not discussed in any of the Expert Ratings of Breeds.

Overall, there is a strong tendency towards testing puppies and young dogs. Tests of adult dogs were typically of dogs who were barely adults at just a few years old. A

single study did examine dogs with a mean age of 62.2 months (Hsu & Serpell, 2003), and two other studies report ages ranging up to 120 and 204 months (Mahut, 1958; Podberscek & Serpell, 1996, respectively). However, these studies have a minimal effect on the overall mean age, which is still less than 24 months. Thus, one striking pattern to emerge is the tendency of researchers to examine young dogs, usually no more than a couple years old.

Breeding and rearing environment

Our review reveals an interesting pattern in terms of the composition of breeding and rearing environments. More than one-third of the studies in my review focused on dogs bred and reared for particular programs. Many of these programs, such as the Swedish Dog Training Center (SDTC), Jackson Laboratories, the Australian Guide Dog Association, and the American Guide Dog Association, attempted to select dogs for breeding. The effects of this temperament-based selective breeding can be seen in various programs. For example, selective breeding based on puppy-test performance scores at the Guide Dogs for the Blind training center in San Rafael, California, (Scott & Bielfelt, 1976) lead to an improvement in puppy-test scores over successive generations; interestingly, this increase in puppy-test scores was not matched in the rates at which adult dogs became successful guide dogs, suggesting the puppy tests may not be an ideal criterion for selecting guide dogs, at least in this high-functioning group of subjects.

Many of the dogs in these studies are purebred dogs living as privately owned pets or show dogs. Others were bred to be guide dogs, police dogs, other working dogs, or as research subjects. Only a minority of the dogs studied were from the large populations of rescued and shelter dogs that hope to benefit from temperament research. A disproportionately large number of the dogs examined were dogs specially bred and specially trained for specific working programs. This is an important point to be borne in mind by people seeking to use the research on temperament to understand and predict the behavior of pet or shelter dogs.

Sexual status of subjects

As noted above, many of the dogs assessed were from programs seeking to breed dogs suitable for specific tasks, such as guide work or police work, and most of the privately-owned dogs were intact. Thus, most animals were not spayed or neutered and the effects of castration were addressed in only a few studies. The rare studies that assessed the effects of castration indicated that intact male dogs were the most likely to show aggressive behavior, and intact female dogs were the least likely (Podberscek & Serpell, 1996; Roll & Unshelm, 1997). Podberscek and Serpell's study also revealed that neutering an adult dog in reaction to his aggressive behavior does not reduce future aggression. Overall, however, researchers know little about the effects of spaying and neutering on dog temperament in general, and even less about how the animal's age at castration affects its later temperament. With the increasing prevalence of laws requiring spay and neuter surgeries before a pet dog can be adopted from a shelter or rescue and the prevalence of spayed and neutered dogs in people's daily lives, the effects of these surgeries on temperament is another area needing more research.

Summary of general survey

To provide some coherence to the enormously varied work on dog temperament, I organized the literature in terms of six frameworks. Organizing the studies in this way allowed me to make several observations about the state of the field. First, there is great diversity in most components of the research, including such features as the goals and the disciplinary bases of the studies. Second, the studies can be usefully categorized in terms of four assessment methods (*Test Batteries, Ratings of Individual Dogs, Expert Ratings of Breed Prototypes,* and *Observational Tests*). Third, most of the dogs studied (90%) were purebred, with Labrador Retrievers and GSDs composing 30% of the subjects. Only six studies reported examining unintentional or unknown mixes, totaling only 856 dogs (Goodloe & Borchelt, 1998; Hsu & Serpell, 2003; Rooney & Bradshaw, 2004; Seksel et al., 1999; Weiss & Greenberg, 1997; Wahlgren & Lester, 2003). Fourth, there is a systematic pattern in which certain breeds are associated with particular types of studies; approximately a third (8,794 total dogs) of the dogs in Test Battery studies were GSDs,

and more than 75% of all dogs in Observational Testing studies were Labrador Retrievers. Consequently, very few breeds other than Labrador Retrievers have been examined by Observational Testing. Fifth, there is a tendency in the research towards testing puppies and young dogs, with older adult dogs (over four years old) infrequently studied and elderly dogs almost entirely neglected by the research literature. Sixth, most of the studies in this review focused on dogs bred and reared for particular programs while tests selecting dogs as pets (e.g., from shelters or rescues) were rare. And last, most dogs involved in these studies were not spayed or neutered and the effects of castration were addressed in only a couple of studies.

Table 2.1. Summary of empirical research on dog temperament: Study design, breed, sex, age, and assessment purpose

		Breed Composition					Sex		Age at assessment (months)					Purpose of assessment					
Study	N	GSD	Lab	Pure	Mixed	Unk.	M(Neut)	F(Spay)	1st	2nd	3rd	4th	5th	Guide	Police	Work	Pet	Other	Population of Dogs
TEST BATTERIES																			
Cattell et al., 1973, Cattell & Korth, 1973	101	0	0	101	0	0	NR(NR)	NR(NR)	.23-3.72	12				0	0	0	0	101	research
Lindberg et al., 2004	1,159[a,b] (192-972)	0	0	1,159[c]	0	0	NR(NR)	NR(NR)	NR					0	0	0	0	1,159	privately owned
Mahut, 1958	230	11	0	230	0	0	96(NR)	134(NR)	7-120					0	0	0	0	230	202 privately owned show dogs, 8 research
Netto & Planta 1997	112	NR	yes	112	0	0	59(NR)	53(NR)	NR					0	0	0	112	0	privately owned
Reuterwall &, Ryman, 1973	958	958	0	958	0	0	NR[d](NR)	NR[d](NR)	18					yes	yes	yes	0	0	working dogs[e]
Royce, 1955	53	0	0	53	0	0	20(NR)	33(NR)	NR					0	0	0	0	53	research
Ruefenacht et al., 2002	3,497	3,497	0	3,497	0	0	1,679 (NR)	1,818 (NR)	21.5[f]					0	0	0	0	3,497[g]	privately owned
Seksel et al., 1999	60	NR	NR	50	10	0	32(NR)	28(NR)	1.38-3.91	+.46[h]	+.92[h]	+4 to 6[h]		0	0	0	60	0	privately owned
Slabbert & Odendaal, 1999	167	167	0	167	0	0	NR(0)	NR(0)	1.85	2.77	3.70	6	9	0	167	0	0	0	police work
Svartberg, 2002	2,655	2,219	0	2,655	0	0	1,381 (NR)	1,274 (NR)	12-18					0	0	2,655	0	0	privately owned
Svartberg & Forkman, 2002;	15,329[i]	NR	NR	15,329	0	0	7,878[j] (NR)	7,451[j] (NR)	19.72[f, k]					0	0	0	0	15,329	privately owned
Svartberg et al. 2005	40[l]	2	3	40	0	0	20(NR)	20(NR)	12-23.42 mean=14.9	+.99[h]	+1.17[h]			0	0	0	0	40	privately owned

Table 2.1. (*Continued*)

		Breed Composition					Sex		Age at assessment (months)					Purpose of assessment					
Study	N	GSD	Lab	Pure	Mixed	Unk.	M(Neut)	F(Spay)	1st	2nd	3rd	4th	5th	Guide	Police	Work	Pet	Other	Population of Dogs
van der Borg et al., 1991	81	NR	NR	NR	NR	81	NR(NR)	NR(NR)	NR					0	0	0	81	0	shelter dogs, adopted
Wilsson & Sundgren, 1998	630[l]	630	0	630	0	0	320(0)	310(0)	1.84	14.8-19.74				0	0	0	0	630	police work, guide work,
Wilsson & Sundgren, 1997	2,107	1,310	797	2,107	0	0	1,073(0)	1,034(0)	14.8-19.74					797	1,310	0	0	0	work/ service, breeding
Total	**27,179**	**8,794**	**800**	**27,008**	**10**	**81**	**12,558 (0)**	**12,155 (0)**	**16.21**[m] **19.31**[n]					**797**	**1,477**	**2,655**	**253**	**19,840**	
RATINGS OF INDIVIDUAL DOGS																			
Goodloe & Borchelt, 1998	2,018	NR	NR	1,412	588	18	916[o] (613)	1,084[o] (896)	NR					0	0	0	2,018	0	privately owned, show dogs
Gosling & Bonnenburg, 1998	1,022	NR	NR	NR	NR	NR	NR(NR)	NR(NR)	NR					0	0	0	0	1,022	privately owned
Hsu & Serpell, 2003	2,054[p]	48	94	1806	173	75	998 (NR[q])	1,047 (NR[q])	62.20[f]					0	0	0	2,054	0	privately owned, show dogs
Ledger, 2003	234	15	0	234	0	0	NR(NR)	NR(NR)	NR					0	0	0	234	0	privately owned
Podberscek & Serpell, 1996	1,109	0	0	1,109	0	0	545(94)	564(187)	3-204 mean=32.40					0	0	0	1,109	0	privately owned, show dogs
Rooney & Bradshaw, 2004	275[r]	2	79	245[r]	28	NR	238(53)[s]	37(29)[t]	13-138 mean=54					0	275[u]	275[u]	0	0	search dogs
Serpell, 1983	25	NR	NR	NR	NR	NR	NR(NR)	NR(NR)	NR					0	0	0	25	0	privately owned
Serpell & Hsu, 2001	1,067[v]	293	369	926	140	0	NR (NR[w])	NR (NR[w])	6	12	14-24			1,067	0	0	0	0	guide work

Table 2.1. (*Continued*)

		Breed Composition					Sex		Age at assessment (months)					Purpose of assessment					
Study	N	GSD	Lab	Pure	Mixed	Unk.	M(Neut)	F(Spay)	1st	2nd	3rd	4th	5th	Guide	Police	Work	Pet	Other	Population of Dogs
Stephen et al., 2001	14	NR	NR	NR	NR	NR	NR(NR)	NR(NR)	NR					0	0	0	14	0	privately owned
Wahlgren & Lester, 2003	264	<10[x]	37	216	48	0	119(NR)	145(NR)	NR					0	0	0	264	0	privately owned
Total	**8,082**	**368**	**579**	**5,948**	**977**	**93**	**2,816 (760)**	**2,877 (1,112)**	**38.65[m] 41.05[n]**					**1,067**	**275**	**275**	**5,718**	**1,022**	
EXPERT RATINGS OF BREED PROTOTYPES[y]																			
Bradshaw & Goodwin, 1998	49	1	1	49	0	0								0	0	0	49	0	
Coren, 1995	79	1	1	79	0	0								0	0	0	79	0	
Draper, 1995[z]	56	1	1	56	0	0								0	0	0	0	56	
Hart & Miller, 1985; Hart & Hart, 1985; Hart, 1995; Hart et al., 1983	56	1	1	56	0	0								0	0	0	56	0	
Keeler, 1947	5	0	0	5	0	0								0	0	0	0	5	
Lester, 1983	24	NR	NR	24	0	0								0	0	0	0	24	
Total	**213**	**3**	**3**	**213**	**0**	**0**								**0**	**0**	**0**	**184**	**85**	
OBSERVATIONAL TESTS																			
Goddard & Beilharz, 1984a, 1984b, 1985	102[1]	16	16	64	yes	0	51(51[2])	51(0)	2.77	4	6	12	12-18	102	0	0	0	0	guide dogs
Goddard & Beilharz, 1982/83	887[1]	0	731	NR	NR	76	436(227)	451(0)	12-18					887	0	0	0	0	guide dogs
Humphrey, 1934	NR	NR	0	NR	0	0	NR(NR)	NR(NR)	NR					0	0	yes	0	0	working dogs
James, 1951	11	0	0	11	0	0	5(0)	6(0)	NR					0	0	0	0	11	research

Table 2.1. (*Continued*)

		Breed Composition					Sex		Age at assessment (months)					Purpose of assessment					
Study	N	GSD	Lab	Pure	Mixed	Unk.	M(Neut)	F(Spay)	1st	2nd	3rd	4th	5th	Guide	Police	Work	Pet	Other	Population of Dogs
Murphy, 1998, 1995	89	0	84	84	5	0	38(NR[3])	51(NR[3])	12[3]					89	0	0	0	0	guide dogs
Total	**1,089**	**16**	**831**	**159**	**5**	**76**	**530(278)**	**559(0)**	**9.92[m] 13.60[n]**					**1,078**	**0**	**yes**	**0**	**11**	
STUDIES THAT USED A COMBINATION OF METHODS																			
Beaudet et al., 1994	39	15	0	0	0	0	15(0)	24(0)	1.61	3.68				0	0	0	0	39	privately owned
Goddard & Beilharz, 1986	102[1]	16	16	64	yes	0	51(51[2])	51(0)	.92[4]	1.15	1.38	1.62	1.85	102	0	0	0	0	guide dogs
Gosling et al., 2003a	78	NR	NR	NR	NR	NR	39(NR)	39(NR)	NR					0	0	0	0	78	privately owned
Hennessy et al., 2001	166	NR	NR	NR	NR	NR	70(NR)	96(NR)	NR[5]	+.46[h]	+6[h]			0	0	0	166	0	shelter dogs
Ledger et al., 1995	120	NR	NR	NR	NR	NR	NR(NR)	NR(NR)	NR					0	0	0	120	0	shelter dogs
Ledger & Baxter, 1996, 1997	56[6]	NR	NR	NR	NR	NR	NR(NR)	NR(NR)	NR					0	0	0	56	0	shelter dogs
Stephen & Ledger, 2003	40	NR	NR	NR	NR	NR	NR(NR)	NR(NR)	NR					0	0	0	40	0	shelter dogs
Svartberg, 2005	697	41	0	697	0	0	352(NR)	345(NR)	12.10-23.98[7] mean=16.49	NR[8]				0	0	697[9]	0	0	privately owned
Weiss & Greenberg, 1997	9	0	0	0	9	0	6(NR)	3(NR)	10-24					9	0	0	0	0	shelter dogs
Total	**1,307**	**72**	**16**	**761**	**9**	**0**	**533(51)**	**558(0)**	**9.00[m] 13.94[n]**					**111**	**0**	**697**	**382**	**117**	
Grand Total	**37,870[10]**	**9,253**	**2,229**	**34,089**	**1,001**	**250**	**16,437 (1,089)**	**16,149 (1,112)**	**18.82[m] 22.02[n]**					**3,053**	**1,752**	**3,627**	**6,537**	**21,075**	

Note. N indicates the number of subjects in each study. GSD indicates German Shepherd Dog. Lab indicates Labrador Retriever. Pure indicates dogs of specific, unmixed breeds including GSDs and Labs. Mixed indicates dogs known to be of mixed breeding. Unk. indicates dogs' breeds were unknown or not recorded, and guesses about mixed breeds were not made. M indicates male dogs, and Neut indicates neutered. F indicates female dogs, and Spay indicates the number spayed. Age at assessment has the sub-groupings of 1st, 2nd, 3rd, 4th, and 5th because dogs may be tested more than once, at different ages. Guide indicates that these dogs were assessed for possible use or utility as guide dogs. Police indicates that these dogs were assessed for possible use or utility as police dogs. Work indicates that these dogs were assessed for possible use or utility in other types of work (e.g., field work, search and rescue, tracking, protection work). Pet indicates that these puppies or dogs were assessed for selection as a pet, or that they already were pets at the time of assessment. The dogs in the category Other do not fit into any of the previous categories; they may be in studies seeking to learn more about personality itself. NR indicates that the authors did not report that particular piece of information, whereas "yes" indicates that the authors reported that there were in fact dogs of that type involved but did not report a number or percentage. When there is no entry, that calculation or report of the particular statistic is not appropriate or not applicable for the given study. One Test Battery in this review (Campbell, 1972) is excluded from this table because it includes descriptions of how to test dog temperament, not actual evaluations; other articles in this review (e.g., Keeler, 1947; Roll & Unshelm, 1997) are not included because, though they discuss dog temperament, they do not present assessments of temperament.

[a] The number of dogs assessed varied (from 192-972) by subtest in the Test Battery.
[b] A total of 1,159 dogs were tested, but some were removed from the sample due to issues of identity and pedigree.
[c] Some of these dogs were missing information on their pedigrees, but all were members of the Swedish Flatcoated Retrievers club.
[d] The number of male and female dogs in this study varies between analyses.
[e] The goal of Reuterwall and Ryman's (1973) article was to study the genetic components of behavior in German Shepherd Dogs; the test used was the Army Dog Training Center test which was designed to identify dogs suitable as working dogs and potentially to breed future generations of working dogs.
[f] These studies reported the average age of dogs assessed.
[g] Ruefenacht et al. (2002) gathered data through the Swedish German Shepherd Dog breeding club, which strives to evaluate whether dogs are physically and temperamentally sound enough for future breeding for many purposes (police work, guide work, protection work, etc.).
[h] The ages at each subsequent testing are reported in terms of number of days or months since the first testing.

[i] 5 Dachshunds and 5 Sight hounds were excluded from the analyses because their breed groups were under-represented in the sample

[j] The numbers of males and females in these studies are calculated using the percentages given by the authors, and are then rounded to the nearest whole number.

[k] All dogs were at least 12 months old when tested

[l] 81 dogs took part in the first test, 44 dogs took part in the first and second test, and 40 dogs took part in all three tests.

[m] The averages are calculated by adding together all the ages in one particular category (e.g. Test Batteries) and then dividing by the number of studies that reported age information; thus, studies that did not report an N do not skew (shrink) the average age. When the age in a given study is reported as a range (e.g. Cattell et al., 1973), the midpoint is used in calculating the overall average.

[n] Weighted average, weighted by number of dogs in each study.

[o] The authors note that the sex of 18 dogs is missing from the surveys they collected.

[p] The authors report 2,054 dogs total, but also report 998 males and 1,047 females, for a total of 2,045 dogs.

[q] Hsu & Serpell (2003) report that 59% of the dogs in their study are neutered (castrated) but do not report how many of these dogs were male and female.

[r] The authors surveyed 244 dog handlers and trainers, gaining information about 275 dogs; the breeds of 273 dogs were reported.

[s] 86.4% of 275 dogs were reported as male. 22.4% of male dogs were reported to have been neutered.

[t] 13.6% of 275 dogs were reported as female. 78.4% of female dogs were reported to have been spayed.

[u] Assessing differences in specialty search dogs working for the Ministry of Defense Police, HM Prison Service, HM Customs and Excise, the UK Army, the Royal Air Force, and civilian police forces.

[v] The authors report the total number of dogs in this study is reported inconsistently as both 1,067 and 1,097 without explaining the discrepancy. I am reporting an N of 1,067 because this is closer to the sum of the authors' report of dogs when broken down by breed (1,066).

[w] All but 10 dogs are intact.

[x] The authors specify that there are >10 GSDs included in the study but do not give a precise number; they have not been included in the totals.

[y] The numbers in this section represent the breeds evaluated; no actual dogs were involved in the studies.

[z] This study is a reanalysis of the data collected by Hart et al. (1983); the 56 dogs in that study are included only once in the totals.

[1] The authors report that not all original subjects were maintained throughout the study, but do not indicate how many subjects were maintained. Where applicable, the number of dogs per breed is thus also uncertain, because I do not know the breed of individuals who dropped out.
[2] All male dogs were castrated at approximately six months of age.
[3] All but 4 dogs were castrated; those 4 dogs were ex-show or ex-breeding dogs, were donated to the guide dog program, and were several years older than the other dogs assessed.
[4] Assessments were conducted weekly until the puppy reached 6 months of age, then another was performed at 12 months.
[5] Hennessy et al. (2001) include dogs of varying ages, divided into two groups: "puppies," who still have milk teeth, and "juveniles/adults", who have their adult teeth. Ages are not reported.
[6] 56 dogs were originally tested, but follow-up surveys assessed only 40 of the original 56.
[7] Test Battery
[8] The author reports that the time between the Test Battery (conducted when dogs were approximately 12-24 months of age) and the Rating of Individual Dogs ranged from 352 to 716 days, or 11.6-23.6 months, later. However, the author does not report the actual age of the dogs at this second assessment, the Rating of Individual Dogs.
[9] The purpose of this study could be seen as validating the Dog Mentality Assessment (DMA), a test previously used and designed for studying working dogs (e.g., Svartberg, 2002), or as validating the personality traits and structure uncovered through previous use of the DMA.
[10] The 102 dogs from the Goddard & Beilharz studies are counted towards the total number of dogs each time there is a separate listing for them, because different tests at different ages are analyzed.

REVIEW AND EVALUATION OF THE EMPIRICAL FINDINGS

Our review has identified enormous variability in the field in terms of the types of assessments, research purposes, and other attributes of the studies themselves. I next extend my review to the substantive findings of the studies. Specifically I investigated which traits have been studied and evaluate the evidence for the reliability and validity of the assessment methods developed so far.

What Traits Have Been Studied in Dogs?

To determine which traits have been identified in studies of non-human animals, Gosling and John (1999) reviewed the structural studies of temperament and personality in non-human species, ranging from chimpanzees to octopuses. They found evidence for several basic dimensions that recurred across species, with especially strong cross-species evidence for Anxiety/Nervousness, Sociability, and Aggression. What can be learned from the present, more focused review of the temperament traits that have been studied in dogs? In this section, I describe the findings of a systematic analysis of the traits and behaviors examined in past dog research.

Dog-temperament researchers have studied a broad array of traits ranging from gun shyness to snapping at children. Summarizing these findings is not a straightforward task because, as discussed above, the studies used different methods, different populations, and are grounded in different disciplines, resulting in a non-standardized vocabulary. On occasion, the same term was used to refer to different behaviors. For example, in one study "temperament" was defined as "character, sensitivity, discrimination, spirit and intellect" (Slabbert & Odendaal, 1999), in another study as "a combination of underlying traits" (Humphrey, 1934), and in yet another study as "physical flexibility and intensity of reaction to different environmental stimuli" (Ruefenacht et al., 2002). In addition, different terms have been used to refer to very similar behaviors. For example, in one study a dog that "goes up to any stranger on sight and makes friends" was scored as high on "confidence" (Humphrey, 1934, p. 133), but the similar behavior of exhibiting "friendly greetings to strangers" (with friendly tail-

wagging, for example), was scored by other researchers as high on "friendliness" (Goodloe & Borchelt, 1998), or "sociability" (Hennessy et al., 2001). In short, no standard lexicon of dog traits and behaviors exists, with the result that traits and behaviors have not been defined consistently across studies. The idiosyncratic terms used in the different studies impede attempts to make cross-study comparisons of what has been learned.

There is clearly a need to develop a common language with which to describe canine temperament. Despite an attempt by Goodloe and Borchelt (1998) to develop a standard lexicon of dog traits and behaviors, none have yet been widely adopted. Therefore, to allow me to summarize the findings across all articles, I used a systematic procedure in which expert judges categorized the varied constructs with a standardized set of terms. The procedure involved three major steps. The procedure was completed once to examine the traits assessed in the 51 studies included in the original review (Jones & Gosling, 2005); a second procedure, using the same basic methodology, was completed to add the four studies that have been added to the current literature review.

Step 1: Extracting behavioral descriptions

The first step was to gather the descriptions of the behaviors that had been studied but to avoid any biases introduced by the researchers' idiosyncratic choice of labels. In each study, I located the descriptions of the behaviors and wrote the descriptions on index cards with one index card for each behavior. The descriptions of behaviors provided in the articles varied in the detail of the descriptions and the degree to which the descriptions included trait-related terms. Terms indicative of the dogs' internal motivations or emotional states and terms directly referring to traits, such as "fearful," "timid," and "curious," were removed. This left behavioral descriptions that were less biased by the researchers' theories about which traits underlie the behaviors. For example, instead of "Social Attraction," a term used by Campbell (1972) and later by Beaudet et al. (1994), the card would be based purely on the behavioral descriptions provided by the researcher: "a puppy's tendency to move towards a human tester who has

placed the puppy in a corner of an observation area, moved to the opposite corner, crouched, and clapped his/her hands quietly."

It was sometimes impossible to separate descriptions of behaviors from labels describing a temperament trait. For example, although Mahut (1958) reports making detailed, objective notes about dogs' behaviors, all that is reported is the classification of these notes into categories such as "curiosity," and "wariness." I was unable to extract more detailed descriptions, so I used these non-descriptive classifications in the index-card task.

For the original review of 51 articles, this procedure resulted in a total of 623 different index cards. The index cards were assigned code numbers associated with the article from which they were drawn, but the key to the code was not known by the judges. This ensured that the judges in Steps 2 and 3 were blind to the identity of the researchers and study from which the descriptions were taken.

In updating this literature review, the same procedure was used to create cards describing new traits from the four additional articles. Traits that were exactly the same as traits sorted in the original task were not recreated. For example, the majority of traits studied in Svartberg (2005) had also been examined in Svartberg (2002) and Svartberg and Forkman (2002). Instead of creating new cards and redoing the procedure for these traits, the sorting results from the original task were used.

Step 2: Development of temperament categories

The first author and a research assistant/veterinary technician served as judges in a sorting task designed to identify the major temperament dimensions. Both judges had a moderate amount of professional work and research experience (at least five years each) with dogs. The cards were shuffled and the judges were instructed to sort them into groups based on the temperament traits likely to be underlying the behaviors described. For example, the cards displaying "Is 'spooked' by odd or unexpected things or objects" (from Serpell & Hsu, 2001) and "avoids or is fearful of unfamiliar children" (from Goodloe & Borchelt, 1998) were placed, by both judges, together in a single pile. The

judges were under no time pressure. Judges were allowed to place one behavior in more than one pile to indicate that the behavior is potentially related to more than one underlying temperament dimension. To do this, the judges copied the code number from the back of the relevant index card onto a new index card and placed cards in each pile; the same process was repeated if a description was deemed to fit into more than two categories.

Once all the cards had been grouped in this way, there were seven piles, with 92% agreement across the two judges. The two judges worked together to provide consensual labels for the seven piles. The final labels were Reactivity/Excitability-Stability, Fearfulness-Courage/Confidence, Aggression-Agreeableness, Sociability/Friendliness-Lack of Interest in Others, Openness-Non-openness (later renamed Responsiveness to Training), Dominance-Submission, and Activity Level.

Step 3: Classification of behaviors by a panel of experts

To ensure these categorizations were not attributable to the Step 2 judges' idiosyncratic experiences, I designed a second categorizing task undertaken by additional expert judges. The panel of judges were selected on the basis of their experience with dogs, the variation of situations in which they observed dogs, and the number of years they had worked with dogs. The complete panel was composed of a veterinarian, a public-shelter dog-temperament tester, three dog trainers with varying specialties, a professor studying animal social behavior, and a graduate student studying dog behavior. They had between seven and twenty years of experience working with dogs and at least three years of formal education in canine or animal behavior. Only the temperament tester specialized in researching or assessing temperament.

For the original sorting task, all of the expert judges were given identical sets of 623 index cards and sorting instructions. They were also given the list of the seven temperament dimensions derived in the previous step. To allow the judges to disagree with the classifications provided by the judges in Step 2, there were two additional categories labeled Other, and Not Temperament-Related. The expert judges were told to

take their time in separating the cards into groups corresponding to the nine categories. The judges were told that a behavioral description written on any one index card could be indicative of more than one temperament dimension or ambiguous as to the underlying dimension leading to the behavior. In such cases, the experts were instructed to copy the number from the back of the relevant index card onto a new card and place the trait in two temperament dimension piles; the same process could be used if a description fit into more than two temperament dimension categories. If the judges thought the behavioral descriptions did not fit into any of the seven temperament dimensions, the judges were instructed to place the card in the category "Not Temperament-Related" or "Other" and provide an explanation for why they had selected this category.

The results of this Expert Temperament Categorizing task were reassuringly consistent across the expert judges. Average pair-wise agreement across judges was 89%, with a maximum agreement between two judges of 95% and a minimum agreement of 80%. Points of disagreement among judges included what dimensions underlie the traits Barking, Problem Solving, and Fearfulness. Typically, the more detail present on the index cards, the more agreement among judges. For example, judges were in less agreement about how to categorize "Barking" than on how to categorize "Barks and sometimes growls when approached by a male stranger."

As noted above, the seven judges in Step 3 were at liberty to disagree with the categorizations developed by the two judges in Step 2. An inspection of the Step-3 judges' categorizations showed that they did indeed disagree with a distinction made in Step 2. In particular, the panel of seven dog experts saw less distinction between the Reactivity/Excitability dimension and the Fearfulness dimension than between the other dimensions, at least in the context of temperament-testing studies. Cards were quite frequently categorized as falling into both the Reactivity and Fearfulness categories. This overlap of dimensions is consistent with research in the human domain, where Reactivity and Fearfulness are components of the same Emotional Stability dimension. Further investigation of the dimensions of Reactivity and Fearfulness in dogs would need to be conducted to make it clear whether the two are indeed independent, or whether they

might fall under an even broader super-ordinate category.

The steps of the sorting task were repeated when the additional four studies were added to the review, and 34 trait cards were sorted using the Expert Temperament Categorizing task. However, a subset of only three of the original judges completed this task. Again, however, the sorting results were reassuringly consistent. Of the 34 traits that these five judges sorted, they disagreed on the sorting a total of five times, and no new issues were raised regarding the trait categories that were provided.

Potential limitations of sorting method

Although these multi-stage procedures were designed to reduce the impact of any single judge and are consistent with standard meta-analytic procedures (Lipsey & Wilson, 1996; Rosenthal, 1991), and very similar procedures have been utilized in various other meta-analyses related to personality in humans (e.g., Barrick & Mount, 1991; Bogg & Roberts, 2004; Heller et al., 2004) it is important to recognize the limitations of this method. One potential limitation is the possibility that the results are influenced by idiosyncratic experiences of the judges, such that a different group of judges might produce different results. Another potential limitation is that the labels generated in Step 2 could have biased the sorting task in Step 3; specifically the choice of labels in Step 2 could have influenced the views or limited the options of the judges in Step 3. To minimize these potential limitations, multiple safeguards were implemented. First, it was made very clear to the judges in Step 3 that the category labels with which they were provided were merely suggestions, so the Step-3 judges could choose not to use these labels if the labels were inadequate or inappropriate. Judges were also provided with "None" (or "Not Temperament-Related") and an "Other" category for cards that did not fit into the categories suggested in Step 2. After the judges had completed the sorting task, I asked them to describe each of the temperament dimension categories to make certain that the judges were using the labels similarly. Reassuringly, 59 of the 63 descriptions given by these Step-3 judges were almost exact matches to those Step-2 judges had used when they selected labels for the categories.

Of course, there is a danger that the safeguards would not be effective if the Step-3 judges felt they could neither use the "Not Temperament-Related" and "Other" categories nor generate their own categories. However, the results of the sorting task showed the judges were willing to use these two categories. An analysis of the frequencies with which the Step-3 judges used the various categories in the sorting task showed they used these two categories almost as frequently as they had used the other seven categories.

This frequency of use suggested both that the judges in Step 3 were comfortable using the categories, and that they agreed with the judges in Step 2 that some of the traits studied were simply not temperament traits (e.g., body sensitivity). In addition, as noted above, four of the judges questioned the Step-2 judges' separation of Reactivity and Fearfulness, suggesting the Step-3 judges were not constrained by the categories generated by the Step-2 judges. These four judges recommended the two categories be combined and relabeled as "Nervousness" or "Nerve Stability." This recommendation demonstrates that the judges took the provided labels as suggestions and not as final labels.

In addition, a number of safeguards were implemented against the danger of generating idiosyncratic categories. First, when selecting judges for Step 2, I strove to identify judges with different kinds of professional experience with dogs. Second, I had these two judges complete their sorting task entirely independently and, if discrepancies arose, discuss them until consensus was reached. This limited the impact of each judge on the results and safeguarded against the categories the judges generated being unique to this analysis. Third, when selecting the group of judges to participate in Step 3, I again strove to build a group with diverse professional experience with dogs. Fourth, this group of judges also completed the sorting task entirely independently from one another. In addition, both sets of judges were under no time constraints. Despite these safeguards, I acknowledge that it is still possible that these results might be unique to this group of judges and I caution the reader to interpret the findings with these caveats in mind.

Results from the sorting task

Table 2.2 summarizes the results of my analyses. For this summary I have combined the "Not Temperament Related" and "Other" categories. Thus, the eight column headings show the eight categories identified in the analyses. I relabeled the Openness category as Responsiveness to Training to avoid confusion with Openness as defined in the human-personality literature. As shown in the table, Reactivity, Fearfulness, Sociability, Responsiveness to Training, and Aggression have been examined more frequently than the other dimensions.

Traits related to the Reactivity dimension were studied quite frequently, in 43 of the studies in my review. High Reactivity was indexed by such behaviors as repeated approach/avoidance of novel objects, raised hackles, and increased activity in novel situations. Low Reactivity was characterized by such behaviors as a relative lack of change of behavior in new situations, and approaching novel stimuli without backing away. In the tests, Reactivity was assessed through such procedures as presenting a novel object or series of novel objects to a puppy and recording its subsequent behavior (Goddard & Beilharz, 1986). The labels and descriptors given to this dimension by researchers included "excitability" (Goddard & Beilharz, 1982/83; Hart et al., 1983; Hart & Miller, 1985; Hart, 1995; Hsu & Serpell, 2003; Ledger & Baxter, 1996, 1997; Murphy, 1995; Serpell, 1983; Stephen et al., 2001; Stephen & Ledger, 2003), "sound reaction" (Coren, 1995), and "heart reactivity" (Royce, 1955).

Fearfulness was studied even more frequently, addressed in 47 studies, and frequently overlapped with Reactivity. One possible reason for this is that dogs may exhibit similar or indistinguishable behaviors as a result of differing emotional states. A dog may exhibit signs of excitement, pacing or running around, approaching objects and then avoiding them, barking, and so on making it difficult to decipher behavioral reactions due to Fearfulness versus Reactivity (Hoffman, 1999). According to the sorting task, shaking and a tendency to avoid novel stimuli without approaching them are associated with high levels of Fearfulness. In the personality tests, Fearfulness was often assessed by recording reactions to novel stimuli or situations (Murphy, 1995, 1998). Low

levels of Fearfulness were sometimes labeled as “courage” (Wilsson & Sundgren, 1997, 1998; Reuterwall & Ryman, 1973), “confidence” (Goddard & Beilharz, 1985; Humphrey, 1934), and “self-confidence” (Ruefenacht et al., 2002). Some labels given to Fearfulness include “apprehension” (Cattell & Korth, 1973), “dog-directed fear or aggression” (Hsu & Serpell, 2003; Serpell & Hsu, 2001), and “timidity” (Hennessy et al., 2001; Ledger & Baxter, 1996, 1997; Royce, 1955; Stephen et al., 2001; Stephen & Ledger, 2003; Wahlgren & Lester, 2003).

Table 2.2. Which traits have been studied in dogs? A review of past research

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Beaudet et al. (1994)			Activity level	Following Social Attraction	Following	Restraint Dominance Elevation Dominance Social Dominance		Restraint Dominance Elevation Dominance Social Dominance
Bradshaw & Goodwin (1998)	Reactivity		Immaturity	Immaturity	Housetrainability		Aggressivity	
Campbell (1972)				Following Social Attraction	Following	Restraint Dominance Elevation Dominance Social Dominance		Restraint Dominance Elevation Dominance Social Dominance
Cattell et al. (1973)	Calmness Emotionality Excitation	Timidity Apprehension	Exuberance	Self-Sufficiency Aloofness			Aggressiveness	Competence
Cattell & Korth (1973)[a]	EII (Social Reactivity) EIII (Affective Arousal)	EI (Extraversion) EVIII (Apprehension) EIII (Affective Arousal)		AII (Cooperation) EIV (Independence)	AI (Un-named) EVI (Calmness)	EV (Timidity) EIV (Independence)		AVI (Breed Aptitude) EVII (Un-named)
Coren (1995)	Sound Reaction Stability Reaction to Novel Stimuli			Social Attraction (Approaching, Following)	Willingness to Work (Retrieval)	Social Dominance (Restraint, Forgiveness, Loss of control)		Touch Sensitivity Response to Food Incentive
Draper (1995)	Reactivity-Surgency			Reactivity-Surgency	Trainability-Openness	Aggression-Nonagreeableness [Dominance over Owner]	Aggression-Nonagreeableness [Dominance over Owner]	Investigation

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Goddard & Beilharz (1986)	Fearfulness	Fearfulness (Approach/Avoid)						
Goddard & Beilharz (1985)		Confidence Aggression-Dominance (Hackles, Biting)				Submissiveness Aggression-Dominance	Aggression-Dominance	
Goddard & Beilharz (1984a)	Fearfulness	Fearfulness						
Goddard & Beilharz (1984b)	PC3 (Fearful & Excitable)	PC1 (General Fearfulness)	PC2 (Activity on Walk) PC5 (Activity in Home, on Free Run) PC6 (Actiivty in Home)		PC4 (Recall) PC7 (Repetitions of Name During Recall)			PC7 (Repetitions of Name During Recall)
Goddard & Beilharz (1982/83)	Distraction Sensitivity Fearfulness & High Activity	Fearfulness Fearfulness & High Activity					Nervous Aggression Aggression	General Performance Sensitivity (Body, Sound)
Goodloe & Borchelt (1998)	Barking 1 Barking 2	Fear/Avoidance of Strangers		Fear/Avoidance of Strangers Friendliness Affiliation	Play 1 Compliance	Submission Compliance Mounting Other Dogs	Aggression to Family/Strangers/ Unfamiliar Dogs Biting	Barking 1 Separation Vocalization Play 2 Play 3 Destruction Digging/Burying Eating Sensitivity Male-related Behaviors Mounting Other Dogs Mounting Objects

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Gosling & Bonnenburg (1998)	Disorganized/ Organized Relaxed Temperamental Touchy Moody [Rude]	Withdrawn Fretful Nervous Anxious	Quiet	Withdrawn Cold/Warm Extraverted Unkind/Kind Shy Untalkative/ Talkative Verbal Bashful	Considerate Cooperative Prompt	Bold Bashful	Kind/Unkind	Artistic Careless Complex Uncreative/Creative Deep Inefficient/Efficient Harsh Imaginative Intelligent Unintelligent Unenvious/Jealous Philosophical Practical [Rude] Sloppy Unsympathetic/ Sympathetic
Gosling et al. (2003a)	Neuroticism			Extraversion	Openness		Agreeableness	
Hart et al. (1983); Hart & Miller (1985); Hart (1995)	Excitability Excessive Barking	Snapping at Children	General Activity	Snapping at Children Affection Demand	Obedience Training Playfulness Housebreaking Ease	[Dominance over Owner]	[Dominance over Owner] Snapping at Children Territorial Defense Aggressive to Dogs	Destructiveness Watchdog Barking
Hart & Hart (1985)	Reactivity	Reactivity		Reactivity (Affection Demand)	Trainability Playfulness	[Dominance over Owner]	[Dominance over Owner] Aggression	Destructiveness
Hennessy et al. (2001)		Flight Wariness Timidity	Locomotor Activity	Sociability Timidity				Solicitation

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Hsu & Serpell (2003)	Excitability	Stranger-directed Fear Dog-directed Fear or Aggression Non-social Fear		Attachment or Attention-seeking Behavior	Trainability		Stranger-directed Aggression Dog-directed Fear or Aggression Chasing Owner-directed Aggression	Separation-related Behavior Chasing Pain Sensitivity
Humphrey (1934)	Energy	Confidence	Energy	Confidence (Approaching to make friends)	Nose Ability Intelligence Willingness	Self-right	Sharpness Fighting (own kind)	Sensitivity (body, ear) Intelligence Willingness
James (1951)						Dominance over Other Pups Guarding Food Area		Which pups' company each prefers
Keeler (1947)		Nervous Courageous						Agile Tame
Ledger (2003)							Aggression	
Ledger & Baxter (1996, 1997)	Excitability	Timidity			Obedience		Aggression	Separation-related Problems
Lester (1983)	Lethargic [Emotional]	Timid	Lethargic Active	Friendly	Easy to Train Obedient		Aggressive	Curious [Emotional]
[c]Lindberg et al. (2004)	Excitement	Excitement		[Independence]	Willing to retrieve	[Independence]		
Mahut (1958)	[Fearfulness]	[Fearfulness]			[Fearfulness] (coming if called by mask-wearer) Interest in Stimuli			

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Murphy (1998, 1995)	Low Concentration Dog Distraction Excitability	Anxiety Suspicion Nervousness			Low Concentration Dog Distraction Low Willingness		Pure Aggression Nervous Aggression Dog Aggression	Low Body Sensitivity Immaturity
Netto & Planta (1997)		Aggression					Aggression	
Podberscek & Serpell (1996)							Aggression	
Reuterwall & Ryman (1973)	Adaptiveness to Different Situations & Environments Ability to Meet with Sudden, Strong Auditory Disturbances	Ability to Meet with Sudden, Strong Auditory Disturbances		Affability Disposition for Fighting in a Playful Manner			Disposition for Self-Defense Disposition for Handler Defense	Disposition for Fighting in a Playful Manner Disposition for Forgetting Unpleasant Incidents
Roll & Unshelm (1997)							Aggression	

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
ʿRooney & Bradshaw (2004)	[Motivation to Chase an Object] Reaction to Sudden Loud Noises Ease of Adaptation to Kennel Environment [Interest in Toys or Objects] Excitability Body Sensitivity [Ease of Adaptation to New Handler] Consistency of Behaviour from Day to Day	Boldness Reaction to Sudden Loud Noises Ease of Adaptation to Kennel Environment Fear of Specific Things [Ease of Adaptation to New Handler]	[Motivation to Chase an Object] Playfulness	Friendliness to People Independence Willingness to Bring and Object back to a Person [Ease of Adaptation to New Handler]	Obedience to Human Command Tendency to be Distracted when Searching Ability to Learn from Being Rewarded Willingness to Bring and Object back to a Person Motivation to Obtain Food Intelligence Incentive to Find an Object which is Out of Sight	Motivation to Retain Possession of an Object Independence	Level of Aggression towards Other Dogs Level of Aggression towards Humans	Agility Tendency to Hunt by Smell Alone Stamina [Interest in Toys or Objects] Acuity of Sense of Smell Health Size Travel Ability
Ruefenacht et al. (2002)	[Reaction to gunfire] Hardness Nerve Stability	[Reaction to gunfire] Self-confidence			Temperament		Sharpness Defense Drive Fighting Drive	Fighting Drive (tolerating stick beats)
Serpell (1983)	Excitability Reaction to Owner's Homecoming	Nervousness		Friendliness to Strangers Friendliness to Other Dogs Loyalty/Affection Sensitivity to Owner's Moods Expressiveness Attachment (1 person)	Obedience on Walks Obedience at Home Attentiveness		Territorial Barking Protectiveness Possessiveness	Attitude on Walks Attitude about Food Intelligence/ Aptitude Reaction to Separation Sense of Humor Attachment (1 person)

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Serpell & Hsu (2001)	[Chasing]	Dog-directed Fear/Aggression Non-social Fear Stranger-directed Fear/Aggression	Energy Level	Stranger-directed Fear/Aggression Attachment (1 person)	Trainability		Stranger-directed Fear/Aggression Owner-directed Aggression Dog-directed Fear/Aggression [Chasing]	Attachment (1 person)
Slabbert & Odendaal (1999)	Startle test Gunshot test	Startle test Gunshot test		[Retrieval test]	[Retrieval test] Obstacle test		Aggression test	[Retrieval test]
Stephen & Ledger (2003), Stephen et al. (2001)	Excitability	Timidity		Playfulness	Obedience		Aggression	
Svartberg (2002)	Boldness/Shyness	Boldness/Shyness		Boldness/Shyness				Boldness/Shyness
Svartberg & Forkman (2002), Saetre et al. (2004), [d]Svartberg et al. (2005)	Curiosity/ Fearlessness Chase-proneness	Curiosity/ Fearlessness					Aggressiveness Chase-proneness	Playfulness
Svartberg (2005)	Boldness/Shyness Curiosity/ Fearlessness Chase-proneness [c][Distance Playfulness]	Boldness/Shyness Curiosity/ Fearlessness	[c][Distance Playfulness]	Boldness/Shyness Sociability Playfulness			Aggressiveness Chase-proneness	Boldness/Shyness [c][Distance Playfulness] Playfulness
van der Borg et al. (1991)		Fear Fear-induced Aggression			Disobedience		Aggression	Disobedience (due to lack of training) Separation Anxiety Misc. Behavior

Table 2.2. (*Continued*)

	Traits							
Study	**Reactivity**	**Fearfulness**	**Activity**	**Sociability**	**Responsiveness to Training**	**Submissiveness**	**Aggression**	**None/Other**
Wahlgren & Lester (2003)		Factor IV (Timid, Emotional)		Factor I (High Sociability & Friendliness; Low Aggression & Bad-temperedness)	Factor II (Curious, Active, Independent) Factor III (Obedient, Clever, Protective)		Factor I (High Sociability & Friendliness; Low Aggression & Bad-temperedness)	
Weiss & Greenberg (1997)	Attention/ Distraction Excitement	Fear/Submission			Attention/ Distraction	Fear/Submission Dominance		
Wilsson & Sundgren (1997)	Nerve Stability Hardness Prey Drive	Courage	[Temperament]	Affability	Cooperativeness [Temperament]		Sharpness Defense Drive Prey Drive	
Wilsson & Sundgren (1998)	Nerve Stability Hardness Prey Drive	Courage	[Temperament] Energy Level	Affability	Cooperativeness [Temperament]		Sharpness Defense Drive Prey Drive	

Note. All dimension labels are those used by the authors. The study authors' definitions of temperament have been used, so I have not excluded items that would not normally be considered temperament constructs (i.e., specific behaviors). Those traits that fell into more than one category are underlined. I list in square brackets those traits that did not elicit 100% agreement among the expert judges in terms of category membership. I provide in standard brackets, where appropriate, more information about traits.

[a]The authors listed and described, but did not always label, the factors derived from their analyses.
[b]PC indicates Principle Component.
[c]These articles were not included in the original sorting task; the behavioral descriptors were sorted by a subset of the original judges and is not included in calculation of inter-judge reliability.
[d]This article contains a subset of traits that were included, with the same descriptions, in the author's previous work, and so the previous trait category assignments were used. The traits were not resorted.

Sociability was studied quite frequently, in 35 studies. The traits categorized under this dimension were also sometimes categorized under Responsiveness to Training, perhaps because interest in people is central to Sociability and to interest in training. Sociability was indexed by such behaviors as initiating friendly interactions with people and other dogs. In personality tests, Sociability was primarily assessed by setting up a meeting between the dog and an unfamiliar person (Humphrey, 1934) or dog (Goddard & Beilharz, 1986). Sociability was given a variety of different labels by researchers, including “extraversion” (Gosling et al., 2003a), “affection demand” (Hart, 1985), and “affability” (Reuterwall & Ryman, 1973).

Responsiveness to Training was studied in 36 of the articles reviewed, and was indexed by such behaviors as working with people, learning quickly in new situations, playfulness, and overall reaction to the environment. This dimension seems very closely related to the dog’s tendency to stay focused and engaged in a given activity, and thus may be more suitably termed “Distractibility” or “Focus.” The trait was assessed through such procedures as giving puppies puzzles to solve (Cattell & Korth, 1973) and “willingness” to work with a person (e.g., Goddard & Beilharz, 1982/83). Tests for this trait vary from specific to broad. For example, a very specific test was the retrieval test, said to be a test of how willing a puppy is to comply with a human in going to get an object and then returning with it (Slabbert & Odendaal, 1999). In contrast, a broad method of assessing Responsiveness to Training, labeled "temperament," examined the dogs’ reactions over a variety of situations, looking for whether the dog exhibited varying reactions and interest in its environment (Ruefenacht et al., 2002). Low Responsiveness to Training was a lack of cooperation or responsiveness to training, or a lack of interest in the situation, while high Responsiveness to Training was the reverse. Labels used to define Responsiveness to Training included “problem solving” (Cattell & Korth, 1973), “willingness to work” (Coren, 1995), and “cooperative” (Gosling & Bonnenburg, 1998).

Aggression was studied in 33 of the articles reviewed. It was indexed by behaviors such as biting, growling, and snapping at people or other dogs. These behaviors could also be caused by fear and may be performed in self-defense. In such cases, the

trait is also related to Fearfulness, but reflects a specific aggressive response to a fearful stimulus. The more dramatic testing procedures used to assess Aggression included such activities as having strangers approach and attack either the dog or the dog's handler (Reuterwall & Ryman, 1973; Ruefenacht et al., 2002). Aggressive behavior was also sometimes divided into subcategories, or types of Aggression, usually on the basis of the cause of the Aggression. For example, Aggression was divided into categories representing Aggression in the service of dominance (Goddard & Beilharz, 1985) and Aggression as a result of nervousness (Goddard & Beilharz, 1982/83). Other studies divided types of Aggression by targets, such as stranger directed fear/aggression, owner-directed aggression, dog-directed fear/aggression (Serpell & Hsu, 2001). Also, in studies looking for dogs that can work as police dogs a very specific subset of Aggression is tested; it was called "sharpness," and defined as the willingness to bite a human being (Humphrey, 1934; Ruefenacht et al., 2002; Wilsson & Sundgren, 1997, 1998).

There is some debate about whether the Dominance and its opposite, Submission, should be considered a trait or a social outcome (Gosling & John, 1999). Nonetheless, behaviors reflecting this dimension were present in 18 of the articles reviewed. Dominance was reflected in such behaviors as refusing to move out of a person's path, or "self-right" (Humphrey, 1934). Dominance and Submission with other dogs was assessed in James' (1951) study of the development of hierarchy in puppies, in which Dominance was judged by observing which dogs bullied others, and which guarded the food area and ate first, and Submission was judged by puppies getting out of a bully's way. Submission was also reflected by such behaviors as urination upon greeting people (Wilsson & Sundgren, 1998). However, clear behavioral definitions are not provided for all conceptualizations of dominance; I was unable to find clear and specific descriptions of the behaviors meant to characterize a type of Dominance called "dominance over owner" (e.g., Draper, 1995).

Activity was discussed in 16 studies. Activity has often been assessed by placing a puppy or dog in an empty arena with gridlines on the floor and seeing how many times the puppy or dog crosses the lines (see Wilsson & Sundgren's arena test, 1998). The

studies offered various labels including "activity" (Cattell & Korth, 1973; Goddard & Beilharz, 1984b; Reuterwall & Ryman, 1973), "general activity" (Hart & Miller, 1985), or "locomotor activity" (Hennessy et al., 2001). Activity level significantly drops between six and twelve months of age (Serpell & Hsu, 2001). Goddard and Beilharz (1984b) found that puppy general activity level is a weak predictor of adult activity level due to a decrease in activity as age increases. They also found that activity level "is of relatively little importance compared to traits which control activity in specific situations" (Goddard & Beilharz, 1984b, p. 275). However, Activity has been identified as a potentially useful, though weak, predictor of adult Dominance/Submission when paired with another test of puppy behavior (Beaudet et al., 1994). This is because Activity seems to moderate the predictive value of the other traits. Therefore, even if activity level does not directly predict adult outcomes, it may still be useful to assess activity as a potential moderator variable.

The categories of "Other" and "Not Temperament-Related" are represented in Table 2.2 as the final column, "None/Other." This category was used for variables examined by 26 different articles, and I have listed each individual variable. The two groups were condensed to one because experts did not identify an additional personality dimension, but rather classified the behavior as being associated with factors that are not based on personality. For example, "disobedience" (van der Borg et al., 1991) initially appeared as if it fit under Responsiveness to Training, however, in this example, the disobedience was due to a lack of training; the shelter dogs assessed in this study may not have ever been trained to know the commands they were asked to perform during testing so I cannot attribute their lack of obedience to their personalities. Sometimes the authors of an article labeled a behavior variable in a way that made it appear to be personality- or temperament-related, such as "dominance tests" (Beaudet et al., 1994), but the judges (who were blind to the label provided) agreed that the tests were not actually assessing Dominance/Submission. Other examples of variables in this category include "touch sensitivity" (Coren, 1995) and "hearing sensitivity" (Goddard & Beilharz, 1982/83).

The fact that the enormous number of terms in Table 2.2 can be classified into

seven categories of personality underscores the need for a standard language for describing personality traits and dimensions in dogs. When each author creates a new set of words with which to discuss the same personality traits, it not only makes comparisons across studies difficult, but is also a process of recreating the wheel. I propose that the seven categories derived from my review of the literature represent a sensible starting point for the development of such a lexicon of canine personality descriptors.

ARE ASSESSMENTS OF DOG PERSONALITY RELIABLE?

If personality tests are to be of any value, they must be shown to be both reliable and valid. Reliability is a prerequisite for validity, and so the evidence for reliability is presented first.

The first thing to conclude about reliability is that, with the few exceptions I will discuss in more detail, researchers have rarely reported reliability of any kind. Those studies that have examined reliability have done so from a variety of perspectives, using different statistical indices, assessing different types of reliability, and computing these reliabilities differently. I culled all measures of reliability from the studies in this review. Unfortunately, most studies that addressed reliability did so without references to numerical indices. For example, Lester (1983) described interjudge reliability as "acceptable" on all but three traits assessed, but did not specify the standard by which "acceptable" was evaluated. Slabbert and Odendaal (1999) discussed reliability in the context of attempts to improve reliability by using trainers (vs. dog owners) as raters but they did not provide any numerical indices of reliability. Weiss and Greenberg (1997) had raters train together, prior to personality testing, in order to ensure what they termed "high" inter-rater reliability but again, no numerical reliability standard was reported. None of these studies could be included in the quantitative review because they did not report reliability numerically.

In addition, I had to exclude from my analyses studies that reported incomplete, incomparable or unique measures of reliability. For example, although Murphy's (1995) study aimed to explicitly address the consistency with which guide dogs' personalities

could be rated by trained judges, the article did not provide a quantitative index of reliability and I could not compute reliability because the data set provided was incomplete. The data provided were a small subset of the whole sample. Although these data were described as representative, they consisted of only a handful of ratings so I considered them incomplete and did not compute a reliability coefficient.

I also had to exclude some studies that reported correlations between tests and retests in a way that did not address the tests' reliability. For example, Goddard and Beilharz (1986) reported some, but not all, correlations between scores at the various ages at which they assessed guide-dog puppies, making the point that the correlations increased as the dogs aged. The purpose of Goddard and Beilharz's studies (1986, 1984a, 1984b) was to evaluate this change and to determine how old a puppy must be for the puppy's Fearfulness level to predict its adult Fearfulness. Because these tests sought to index change in the subjects, not repeatability of the test, the scores were not appropriate for my analyses. In addition, reliability could not be computed in the studies that simply obtained frequency estimates of certain behaviors (e.g., Podberscek & Serpell, 1996) or obtained ratings from a single judge (e.g., Goodloe & Borchelt, 1998).

The remaining reliability coefficients that I was able to uncover or compute are reported in Tables 2.3 and 2.4. The tables are divided by type of analysis; Table 2.3 shows the inter-observer (or inter-rater) agreement and test-retest reliability correlations, and Table 2.4 shows internal consistency (associated with factor analyses) as indexed by Cronbach's alpha.

Table 2.3 is divided into two types of reliability: inter-observer agreement and test-retest reliability. The studies using inter-observer agreement used the traditional method of analysis in which each variable is analyzed across subjects (instead of computing reliability within subjects). The correlations suggest that interjudge agreement varies greatly across studies and traits. Although strong agreement is possible, it is by no means guaranteed; the sample-weighted mean agreement correlation was .60, but the agreement correlations ranged from .00 to .86. A study by Goodloe and Borchelt (1998) was excluded from the table because the data are not complete or precise enough to allow

me to integrate them into the quantitative analyses. However, their data also support the idea that dog personality can be assessed reliably. Ninety-six percent of their inter-observer correlations were above .60, and of those 55% were above .80; their lowest correlations were on four items, reported only as less than .50. Goodloe and Borchelt (1998) emphasized the point that dogs may interact with raters differently, which would diminish the apparent reliability, as the dogs may behave consistently with each person, but differently across people. Clearly, given that reliability is a fundamental standard of all measurement studies, future research is badly needed on this possibility and others. In general, whenever appropriate data are available, reliability indices should be reported, as is standard practice in research on human personality.

Three studies appear in the test-retest reliability category, listed in the lower section of Table 2.3, examining the correlation between scores when dogs were tested two or, in the case of Svartberg et al. (2005), three times. One of these studies, by Goddard and Beilharz (1986), reveals Activity level is reliable from test to test, but that this reliability decreases as puppies age. The other study, by Netto and Planta (1997), shows a strong mean correlation, but also included many insignificant correlations. Closer examination reveals that many of the Kappa coefficients reported are zero, indicating no reliability. However, this is partially an artifact of the testing situation because the subtests were not intended to elicit Aggression, so it makes little sense to assess the reliability with which they elicited aggression. Of the subtests in this study which were intended to elicit aggression, the lowest Kappa coefficient is -.03 for reaction to an artificial hand taking away food, and reaction to a stranger being mildly threatening when meeting the dog's handler. However, Netto and Planta's study should be commended for fully reporting their reliability data; when interpreted against an understanding of the testing situations, these are data are very valuable.

The third study (Svartberg et al., 2005), reports Cronbach's alpha values greater than .80 for four traits (Playfulness: .87, Chase-proneness: .84; Curiosity/Fearfulness: .80, Sociability: .89); a lower, but still relatively high alpha when compared with other values in Table 2.3, of .67 for Aggression; and an alpha of .71 associated with the broader trait

of Boldness. In addition to these analyses of consistency of ratings across three tests, Svartberg et al. (2005) examined the consistency of the dogs' scores relative to each other, or rank-order consistency. The rank-order consistency data are not presented in depth here because they are as readily comparable with other examinations of reliability, however, the researchers found evidence for consistency of rank order, with Spearman rank order correlation analyses of trait scores ranging from .57 (Sociability from test 1 to test 3, test 2 to test 3) to .90 (the broad, composite trait of Boldness from test 2 to test 3).

Table 2.3. How reliable are personality measures of dogs?: Inter-observer agreement and test-retest reliability

			Inter-observer agreement								
				Maximum		Minimum				
Study	**Assessment Method[a]**	**Mean cor.[b]**	**SE_z[c]**	**cor**	**Item Label**	**cor**	**Item Label**	**Sample Size**	**No. of Indicators**	**Retest Interval**
			Inter-observer agreement computed for each variable across subjects							
Gosling et al. (2003a)	Combination	.62	.12	.76	Extraversion	.55 .55	Agreeableness Openness	78	4[d]	
Goddard & Beilharz (1982/83)	Observational	.47	.41	.70	Nervousness	.00	Willingness	9	14	
Stephen et al. (2001) (also Stephen & Ledger, 2003)[f]	Combination	.75	.30	.86	Aggressiveness	.49	Excitability	13-14	75[g]	
Unweighted Mean		**.56**		**.78**		**.37**				
Sample-weighted Mean		**.60**		**.77**		**.50[e]**				
			Test-Retest Reliability							
Goddard & Beilharz (1986)	Combination	.39	.10	.52	Activity 7 and 9 weeks	.21	Activity 5 and 9 weeks	102	1	variable
Netto & Planta (1997)	Test Battery	.77	.17	.79	unfamiliar female dominant dog in area b	many non-significant (effect sizes not reported)		37	43	6 months
Svartberg et al. (2005) (consistency across 3 test points)	Test Battery	.83 .81[h]	.16	.87	Playfulness	.67	Aggressivity	40	33 on test 1; 32 on test 2, 3	test 1-test 2: 30 days test 2-test 3: 35 days
(rank-order consistency; 3 test points)		.72 .75[h]	.16	.57	Sociability test 2 to 3, 1 to 3	.90	Boldness test 2 to 3	40	33 on test 1 32 on test 2, 3	test 1-test 2: 30 days test 2-test 3: 35 days
Unweighted Mean		**.71 (.71[h])**		**.72**		**.68**				
Sample-weighted Mean		**.63 (.63[h])**		**.67**		**.55**				

Note.
[a] The categories used here refer to the types of assessment method identified earlier and summarized in Table 2.1.
[b] Mean correlations are computed using Fisher's r-to-z transformation.
[c] The standard errors reported are for the standardized scores and are computed by $SE_z = 1/\text{sqrt}[n-3]$.
[d] Gosling et al. (2003a) used scales, not individual items, as indicators.
[e] The weighted mean of the minimum correlation for inter-observer agreement are calculated using only one of the two scores from Gosling et al. (2003a).
[f] Stephen & Ledger (2001) used Spearman's rank test, and thus reported correlations as rho.
[g] This study included a 75-item questionnaire, which was analyzed to reveal five temperament dimensions.
[h] These means include the broad trait of Boldness, which is a composite of the 5 traits examined in the study.

Table 2.4 summarizes all the internal-consistency estimates reported in the studies reviewed. Internal-consistency measures estimate the degree to which items on a scale assess the same construct. In human personality research, they are often used following factor analyses to determine the internal coherence of the derived factors. Of the 18 studies in the current review to focus on factor analysis, only four reported internal consistency. Three of these studies (Hsu & Serpell, 2003; Serpell & Hsu, 2001; Svartberg, 2005) gathered data using various forms of the C-BARQ (see Hsu & Serpell, 2003), a questionnaire that uses a 5-point frequency or rating (Likert) scales. The fourth study that reported internal consistency (Seksel et al., 1999) used a 100-point scale. One additional study that did not focus on factor analysis also reported internal consistency (Gosling et al., 2003a) and is included in Table 2.4.

Internal consistency varied greatly across studies and factors, ranging from .42 for "Handling," to .93 for "Stranger-directed Aggression." Although high consistency is possible, it is by no means guaranteed. Nonetheless, the internal consistency measures had an unweighted mean of .76, and sample-weighted mean of .64, both well within the limits acceptable in most human personality research (John & Benet-Martinez, 2000).

Table 2.4. How reliable are personality measures of dogs?: Internal consistency

	Internal Consistency of Factors						
	MEAN	Maximum		Minimum			
Study	α	α	Item Label	α	Item Label	Sample Size	Total Number of Items in Study[a]
Gosling, Kwan, & John (2003)							
owner judgment (own dog)	.83	.89	Neuroticism	.77	Extraversion	78	43
peer judgments of dog	.82	.86	Neuroticism	.75	Openness	78	43
Hsu & Serpell (2003)	.81	.93	Stranger-directed Aggression	.67	Pain Sensitivity	2,054	132
Serpell & Hsu (2001)	.65	.84	Stranger-directed Fear/Aggression	.53	Attachment	1,067	38
Seksel et al. (1999)	.56	.73	Novel Stimuli	.42	Handling	60	21
Svartberg (2005)							
behavioral test	.78	.87	Distance-playfulness	.56	Aggressiveness	697	33
owner judgment (own dog)	.76	.85	Stranger-directed interest	.60	Pain Sensitivity	697	122
Unweighted Mean	**.76**	**.86**		**.62**			
Sample-Weighted Mean	**.64**	**.71**		**.55**			

Note. All Cronbach's alphas are as reported by the authors, not standardized.

[a] All of the studies except Gosling, Kwan, and John (2003) and Svartberg (2005) reported dropping items for various reasons.

Summary of reliability findings

As a whole, the review of reported reliability coefficients is both encouraging and disappointing. The findings are encouraging because they show it is possible to measure dog personality reliably using a variety of assessment methods. The findings are disappointing because they show that very few articles report reliability scores. Fortunately, there is an easy remedy—future researcher should compute and report the reliability of their measures.

ARE ASSESSMENTS OF DOG PERSONALITY VALID?

Once the reliability of a test has been established, the next step is to evaluate its construct validity. Validity is an index of how well an instrument is measuring what it is designed to measure. The construct-validation process involves determining how well a measure assesses a construct (e.g., Fearfulness) as that construct has been conceptualized. A full conceptualization of a construct involves specifying the things to which the construct should be related and also the things to which the construct should be unrelated (Cronbach & Meehl, 1955). These two components are known as convergent and discriminant validity. Convergent validity is supported when a measure correlates with other measures to which it should be related. Discriminant validity is supported when a measure is empirically unrelated to other measures that are theoretically unrelated (Campbell and Fiske, 1959). Thus, for example, the construct validity of a measure of Fearfulness would be supported by strong correlations with other measures of Fearfulness (i.e., convergent validity) and weak correlations with measures of theoretically unrelated traits, such as Sociability (i.e., discriminant validity; Devellis, 2003). To evaluate the validity of the tests in this review, I culled all potentially relevant validity data from the articles.

Obtaining and categorizing the validity coefficients

My goal was to summarize the validity findings for each the seven broad personality dimensions identified above (Reactivity, Fearfulness, Activity, Sociability,

Responsiveness to Training, Submissiveness, and Aggression). Given these meta-analytic goals, I had to exclude from my analyses validity indices that were unique or could not be compared with other indices. For example, although Serpell and Hsu (2001; Hsu & Serpell, 2003) addressed validity directly they report only the significance levels of the Mann-Whitney U-statistics, not effect sizes, so their results could not be combined with the eleven other studies reporting validity, all of which report effect sizes.

Most studies did not explicitly conceptualize their findings in terms of convergent and discriminant validity and even those that did assess convergent validity or discriminant validity typically did not employ these terms. Therefore, after identifying all the potentially relevant validity coefficients, I devised a procedure for dividing them into three categories: convergent correlations, discriminant correlations, and indeterminate correlations. In studies where clear predictions were made (e.g., Svartberg, 2005), I could easily classify the correlations. Specifically, where a trait was predicted to correlate with a behavior, the resulting correlation was considered as evidence for convergent validity, and where a trait was predicted not to correlate with a behavior, the resulting correlation was considered evidence for discriminant validity.

However, when studies examined relationships between assessment scores and future behavior or future assessments but did not make predictions about these relationships, I needed a systematic procedure for assigning the correlations to the convergent, discriminant, or indeterminate categories. Thus, for each of these studies, I extracted descriptions of (1) the predictor variables (the trait or behavior assessed and how it was assessed), and (2) the validity criteria (the future behavior or later assessment results). Two expert judges who were blind to the actual findings of these studies made judgments about whether the predictor-criterion pairs should theoretically be related or unrelated. The two judges first made their judgments independently, then compared their judgments and discussed points of disagreement until consensus was reached. Those correlations associated with predictor-criterion matches were assigned to the convergent-validity category and those correlations associated with predictor-criterion mismatches were assigned to the discriminant-validity category. For example, the predictor-criterion

pair in which adult dogs' wariness was a predictor of later problem behavior (Hennessy et al., 2001) was assigned to the convergent-validity category, and the predictor-criterion pair in which the number of objects a puppy investigated in a given period of time was a predictor of the adult dog's ability to cooperate (Wilsson & Sundgren, 1998) was assigned to the discriminant-validity category. Of course, it should be noted that despite my best efforts to be comprehensive and systematic, the validity coefficients I report are inevitably influenced by my procedures for selecting coefficients and these findings should be evaluated in this light.

Convergent validity

Table 2.5 summarizes the available evidence for convergent validity. The convergent validity coefficients are organized in terms of the seven personality dimensions plus two additional broader dimensions (Problem Behavior and Broader Evaluations of Temperament), which are listed in the first column of the table. The second column lists the relevant citation. The third column lists the trait evaluated, as it was labeled by the original authors. The fourth column lists the criterion against which the trait was evaluated. The fifth column briefly summarizes the procedures by which the criterion data were obtained. The sixth column provides the validity coefficient as Pearson correlations or Spearman's rho. The final column lists the sample size on which the validity coefficient was based.

Thus, the table shows, for example, that Ledger and Baxter (1996) examined the validity of Excitability ratings of 40 dogs made by their owners after adoption. The criterion by which the Excitability ratings were evaluated was behavior in response to an unfamiliar tester entering the dog's kennel. The sixth column shows that the owners' Excitability ratings correlated .64 with the dogs' behaviors when a stranger entered the kennel.

The summary statistics presented in Table 2.5 include both unweighted and a sample-weighted means. Both estimates are included because the sample sizes varied substantially across studies. For example, studies reporting convergent-validity data on

Reactivity had sample sizes ranging from nine (Weiss & Greenberg, 1997) to 630 (Wilsson & Sundgren, 1998). The study of nine dogs reported a Reactivity convergent validity estimate of .36, whereas the study of 630 dogs reported Reactivity convergent-validity estimates of .01 and .05. The mean validity coefficient for Reactivity is .28 if averaged across all studies, but .16 if weighted by the number of dogs in each study. Both estimates are potentially interesting, with the first estimate giving equal weight to each study and the second estimate giving equal weight to each individual dog tested.

Overall, the evidence for convergent validity is reasonably promising, with the various estimates averaging about .51 across the nine dimensions examined here. However, the findings do show some variability across the dimensions. The dimensions with the fewest studies will tend to provide the least stable estimates so it is not surprising that highest and lowest validity estimates are associated with the dimensions with very studies. In particular, the strongest convergent-validity coefficients (unweighted mean = .88, sample-weighted mean = .88) are associated with the Submissiveness dimension. However, with rather divergent evidence from only two studies, the confidence intervals around this mean are enormous, ranging from 0 to 1. Therefore, I do not feel confident providing a validity estimate for this dimension. Clearly, more research is needed before estimates can be made about the validity of Submissiveness assessments.

The lowest validity coefficients (unweighted mean = .15; sample-weighted mean = .21) are associated with the Activity dimension. However, only a few studies report convergent validity coefficients associated with this dimension, for a total of six coefficients, again suggesting the need for further research.

Table 2.5. Convergent validity: How well do dog personality tests predict future behavior or scores on other assessments?

		Criterion Measure			
Dimension					
Study	**Trait**	**Criterion behavior**	**Basis for Scoring**	**Validity Coefficient**	**Number of Subjects**
Reactivity					
Goddard & Beilharz (1986)[b]	Excitability score (rated by trainers)	Composite of scores on sit, activity tests	Repetitions dog needs to "sit" on command; number of movements	.22	102[c]
Ledger & Baxter (1996)[a]	Excitability (rated by owners after adoption)	Un-named	Unspecified behavioral response to an unfamiliar tester in kennel;	.64	40
Stephen & Ledger (2003)[a]	Behavior problems towards strangers (rated by owners after adoption)	Excitability towards tester	Behavior when tester greets/meets the dog Unspecified behavioral response to grooming	.32 .66	40 40
Svartberg (2005)	Chase-proneness (behavior test rated by observers)	"Chasing" score on questionnaire	"Chasing" factor on modified C-BARQ	.05	697
Svartberg (2005)	Curiosity/Fearlessness (behavior test rated by observers)	"Nonsocial fear" score on questionnaire	"Nonsocial fear" factor on modified C-BARQ	.26[j]	697
Svartberg (2005)	Distance-playfulness (behavior test rated by observers)	"Human-directed play interest" score on questionnaire "Stranger-directed fear" score on questionnaire "Stranger-directed interest" score on questionnaire	"Human-directed play interest" factor on modified C-BARQ "Stranger-directed fear" factor on modified C-BARQ "Stranger-directed interest" factor on modified C-BARQ	.29 .19 .16	697 697 697
Weiss & Greenberg (1997)[a]	Excitement (rated by 3 observers)	Excitement-related behaviors	Scoring method not specified, but behaviors included steady high level of jumping, pawing, barking, etc.	.36	9
Wilsson & Sundgren (1998)[d]	Prey Drive[e] (rated by trainers)	Fetching Retrieving	Time until puppy picks up tossed ball Willingness scored with set criteria	.01[f] .05	630 630
Unweighted Mean				**.28**	
Sample-weighted Mean				**.16**	

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Fearfulness					
Goddard & Beilharz (1986)[b]	General Fearfulness (rated by trainers)	Composite of fear on walk, Reactions to specific stimuli at different ages	Fear on walk – see Goddard & Beilharz (1984a) other tests include reaction to surfboard at 10 weeks, activity level during handling at 9 weeks, etc	.57	102[c]
Goddard & Beilharz (1984a)	General Nervousness (rated by trainers)	Fear on walk (3 months)	Ratings by trainers based on a combination of reactions to various stimuli, including clap noise, toy horse, gun shot, party whistle, rapid head movement, ear position, stranger entering house	.24	102[c]
	General Nervousness (rated by trainers)	Fear on walk (4 months)		.35	102[c]
	General Nervousness (rated by trainers)	Fear on walk (6 months)		.42	102[c]
	General Nervousness (rated by trainers)	Fear on walk (12 months)		.58	102[c]
	General Nervousness (rated by trainers)	Fear on walk (day 3 of final evaluation)		.59	102[c]
	General Nervousness (rated by trainers)	Fear on walk (day 4 of final evaluation)		.64	102[c]
Gosling et al. (2003a)	Extraversion	Extraversion-related behavior (rated by owner)	Observer rating based on a variety of field-test behaviors, during greetings, etc.	.32	78
Hennessy et al. (2001; puppies)[g, h]	Part of overall problem index (rated by owners after adoption)	Flight	Number of movements to escape; time in door well, jumps	.02/.34[j, k]	23/18[l]
Hennessy et al. (2001; adults)[g, i]	Part of overall problem index (rated by owners after adoption)	Flight	Number of attempts to escape, time spent in door well, jumps	.47/.74	10/7[l]
Hennessy et al. (2001; puppies)[g, h]	Part of overall problem index (rated by owners after adoption)	Timidity	Time spent in door well	.39[j, k]/.11	23/18[l]
Hennessy et al. (2001; adults)[g, i]	Part of overall problem index (rated by owners after adoption)	Timidity	Time spent in door well	.03/.37	10/7[l]
Hennessy et al. (2001; puppies)[g, h]	Part of overall problem index (rated by owners after adoption)	Wariness	Latency to contact toy car, horn	.43[j]/.31[k]	23/18[l]
Ledger & Baxter (1996)[a]	Timidity	Un-named (rated by owners after adoption)	Unspecified behavioral response to being walked on-leash;	.68	40
			Unspecified behavioral response to being approached by a person with a "titbit"	.79	40

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Fearfulness continued…					
Stephen & Ledger (2003)[a]	Fearfulness (rated by owners after adoption)	Tester observations through-out test		reported as "not correlated"	
Svartberg (2005)	Curiosity/Fearlessness (behavior test rated by observers)	"Nonsocial fear" score on questionnaire	"Nonsocial fear" factor on modified C-BARQ	.26[j]	697
Unweighted Mean				**.44**	
Sample-Weighted Mean				**.42**	
Activity					
Beaudet et al. (1994)	Activity Level (retested using same assessment at 1.61 and 3.68 months)	Locomotor activity	Number of line crossings in test chamber	.04[m]	39
Hennessy et al. (2001; adults)[g, i]	Part of overall problem index (rated by owners after adoption)	Locomotor activity	Number of movements to withdraw, approach, line crossing	.05[k]/.16[j]	10/7[l]
Svartberg (2005)	Distance-playfulness (behavior test rated by observers)	"Human-directed play interest" score on questionnaire	"Human-directed play interest" factor on modified C-BARQ	.29	697
		"Stranger-directed fear" score on questionnaire	"Stranger-directed fear" factor on modified C-BARQ	.19	697
		"Stranger-directed interest" score on questionnaire	"Stranger-directed interest" factor on modified C-BARQ	.16	697
Unweighted Mean				**.15**	
Sample-Weighted Mean				**.21**	
Sociability					
Gosling et al. (2003a)	Neuroticism (rated by owners)	Neuroticism-related behavior	Observer rating based on a variety of field-test behaviors, during greetings, etc.	.21	78
Hennessy et al. (2001; adults)[g, i]	Part of overall problem index (rated by owners after adoption)	Sociability	Time spent in proximity to a person, stationary, latency to contact person, time in door well, person contact, etc.	.21/.63	10/7[l]

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Hennessy et al. (2001; puppies)[g, h]	Part of overall problem index (rated by owners after adoption)	Timidity	Time spent in door well (includes avoiding people)	.39[j, k]/.11	23/18[l]
Hennessy et al. (2001; adults)[g, i]	Part of overall problem index (rated by owners after adoption)	Timidity	Time spent in door well (includes avoiding people)	.03/.37	10/7[l]
Stephen & Ledger (2003)[a]	Playfulness (rated by owners after adoption)	Play behaviors with tester	Play displays during tug-o-war with tester	.53	40
			grooming by tester	.44	40
			having lead put on by tester	.33	40
			being walked on lead by tester	.49	40
Svartberg (2005)	Playfulness (behavior test rated by observers)	"Human-directed play interest" score score on questionnaire	"Human-directed play interest" factor on modified C-BARQ	.36	697
Svartberg (2005)	Sociability (behavior test rated by observers)	"Stranger-directed fear" score on questionnaire	"Stranger-directed fear" factor on modified C-BARQ	.27[j]	697
Svartberg (2005)	Sociability (behavior test rated by observers)	"Stranger-directed interest" score on questionnaire	"Stranger-directed interest" factor on modified C-BARQ	.36	697
Wilsson & Sundgren (1998)[d]	Affability (rated by trainers)	Yelping	Time until puppy (alone) whines/yelps	.00[k, l, o]	630
Unweighted Mean				**.33**	
Sample-Weighted Mean				**.27**	

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Responsiveness to Training					
van der Borg, et al. (1991)	Disobedience (rated by owners after adoption)	Disobedience	Questionnaire to care-taker	.41[p]	81
			Testers' evaluation of disobedience	.27[p]	81
van der Borg, et al. (1991)	Pulling on leash (rated by owners after adoption)	Pulling on leash	Questionnaire to care-taker	.51[p]	81
			Testers' evaluation of pulling	.16[p]	81
Gosling et al. (2003a)	Openness	Openness-related behavior (rated by owner)	Observer rating based on a variety of field-test behaviors, during greetings, etc.	.23	78
Ledger & Baxter (1996)[a]	Obedience (rated by owners after adoption)	Un-named	Unspecified behavioral response to showing dog its leash, saying "walkies"	.72	40
Stephen & Ledger (2003)[a]	Obedience (rated by owners after adoption)		Tester observations through-out test	reported as "not correlated"	
Weiss & Greenberg (1997)[a]	Attention/Distraction (rated by 3 observers)	Attention/Distraction-related behaviors	Scoring method not specified, but behavior described: "dog's attention should be on the handler"	.00	9
Wilsson & Sundgren (1998)[d]	Ability to Cooperate (rated by trainers)	Contact	Reaction to, attempt to contact person	.17[f, k]	630
Unweighted Mean				**.33**	
Sample-Weighted Mean				**.25**	

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Submissiveness					
Weiss & Greenberg (1997)[a]	Dominance (rated by 3 observers)	Dominance-related behaviors	Scoring method not specified, but behaviors included front paws on handler, mounting, placing body above handler, growling while making eye contact	.13	9
Weiss & Greenberg (1997)[a]	Fear/Submission (rated by 3 observers)	Fear/Submission-related	Scoring method not specified, but behaviors behaviors included crouching, submissive urination, shoulder roll, prolonged startle/fear to noise, etc.	1.00[q]	9
Unweighted Mean				**.88**	
Sample-Weighted Mean				**.88**	
Aggression					
van der Borg, et al. (1991)	Aggression towards (rated by owners after adoption)	Aggression towards adults	Questionnaire to care-taker Testers' evaluation of aggression	.45[p] .26[p]	81 81
van der Borg, et al (1991)	Aggression towards dogs (rated by owners after adoption)	Aggression towards dogs	Questionnaire to care-taker Testers' evaluation of dog-related aggression	.55[p] .23[p]	81 81
Gosling et al. (2003a)	Agreeableness (rated by owner)	Agreeableness-related behavior	Observer rating based on a variety of field-test behaviors, during greetings, etc.	.33	78
Ledger & Baxter (1996)[a]	Aggression (rated by owners after adoption)	Un-named	Unspecified behavioral response to showing dog its leash, saying "walkies"; Unspecified behavioral response to playing tug-o-war	.82 .82	40 40
Netto & Planta (1997)	Bite History	Aggression, Tendency to bite (reported by owner)	Observed biting attempts and snapping during 43 subtests of test battery	.25[j]	112
Netto & Planta (1997)	Bite History	Aggression, Tendency to bite (reported by owner)	Observed biting attempts (without snapping) during 43 subtests of test battery	.31[r]	112
Svartberg (2005)	Aggressiveness (behavior test rated by observers)	"Stranger-directed aggression" score on questionnaire	"Stranger-directed aggression" factor on modified C-BARQ	.12	697

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Aggression continued…					
Wilsson & Sundgren (1998)[d]	Prey Drive[e] (rated by trainers)	Fetching	Time until puppy picks up tossed ball	.01[f, k]	630
		Retrieving	Willingness scored by set criteria	.05	630
Unweighted Mean				**.42**	
Sample-Weighted Mean				**.18**	
Unweighted Mean Across All 7 Temperament Dimensions				**.53**	
Sample-Weighted Mean Across All 7 Temperament Dimensions				**.23**	
Problem behaviors					
van der Borg, et al. (1991)	Car-related problems (rated by owners after adoption)	Car-related problems	Questionnaire to care-taker	.20[p]	81
			Testers' evaluation of car-related problems	.23[p]	81
van der Borg, et al. (1991)	Separation anxiety (rated by owners after adoption)	Separation anxiety	Questionnaire to care-taker	.66[p]	81
			Testers' evaluation of separation problems	.22[p]	81
Hennessy et al. (2001; puppies)[g, h]	Part of overall problem index (rated by owners after adoption)	Solicitation	Number of jumps against observation platform	.09/.18	23/18[l]
Hennessy et al. (2001; adults)[g, i]	Part of overall problem index (rated by owners after adoption)	Solicitation	Number of jumps against observation platform	.54/.72	10/7[l]
Ledger & Baxter (1996)[a]	Separation-Related Problems	Un-named (rated by owners after adoption)	Unspecified behavioral response to being approached by a person with a "titbit"	.82	40
Unweighted Mean				**.45**	
Sample-Weighted Mean				**.41**	

Table 2.5. (*Continued*)

Dimension / Study	Trait	Criterion Measure: Criterion behavior	Criterion Measure: Basis for Scoring	Validity Coefficient	Number of Subjects
Broad Evaluation of Temperament					
Beaudet et al. (1994)	Cumulative Social Tendency Score (Submissiveness/Dominance) (retested using same assessment at 1.61 and 3.68 months)	Social Attraction	Puppy's reaction during 30s of tester crouching, coaxing puppy to the tester	.29	39
		Following	Puppy's reaction when tester tries to coax puppy to walk by the tester		
		Restraint Dominance	Puppy's reaction when tester holds puppy on its back for 30s		
		Elevation Dominance	Puppy's reaction when tester holds puppy 15cm off the ground for 30s		
		Social Dominance	Puppy's reaction to being stroked from head to tail for 30s		
Beaudet et al. (1994)	Cumulative Social Tendency Score	Locomotor activity at 1.61 months (Submissiveness/Dominance) (tested at 1.61 months)	Number of line crossings in test chamber	.45	39
Beaudet et al. (1994)	Cumulative Social Tendency Score	Locomotor activity at 3.68 months (Submissiveness/Dominance) (tested at 3.68 months)	Number of line crossings in test chamber	.70	39
Weiss & Greenberg (1997)[a]	Completion of a set of tasks in final test	General Selection Test	Scored by tester on various tasks and subjective "feeling"	.18	9
Weiss & Greenberg (1997)[a]	Number of corrections needed to complete tasks in final test	General Selection Test	Scored by tester on various tasks and subjective "feeling"	.21	9
Unweighted Mean				**.39**	
Sample-Weighted Mean				**.46**	
Unweighted Mean Across All Dimensions, Including Problem Behaviors and Broad Evaluations of Temperament				**.51**	
Sample-Weighted Mean Across All Dimensions, Including Problem Behaviors and Broad Evaluations of Temperament				**.24**	

Note. Mean correlations are computed using Fisher's r-to-z transformation.

[a] These correlations are rho values from Spearman's rank analysis.
[b] Goddard and Beilharz (1986) report extensively on the correlations between components to these overall scores and the trait they were used to predict. I have not reported all of these coefficients individually because they are components to the overall scores and to do so would skew my overall correlations. Please see the original source for more details of these component correlation coefficients.
[c] Goddard & Beilharz (1984a, 1986) reported an original N of 102 before an unspecified number of subjects that dropped out.
[d] Wilsson & Sundgren (1998) examined all possible correlations but reported effect sizes for significant correlations only.
[e] In the sorting procedure, the behavior of prey drive was categorized into both aggression and reactivity and is thus listed twice here.
[f] The correlations between Yelping and Affability, Contact and Ability to Cooperate, and Fetching and Prey Drive were all reported as negative such that a shorter latency (less time) to Yelping correlates with higher Affability, a shorter latency to make contact correlates with a higher adult score on Ability to Cooperate, and a shorter latency to pick up a thrown ball correlates with a higher adult Prey Drive. These correlations have been rekeyed so that a higher correlation reflects greater validity.
[g] Hennessy et al. (2001) received so few reports of problem behaviors that it was deemed necessary to create a combined "behavior problems" score instead of attempting to examine prediction of individual types of behavior problems.
[h] These assessments were performed with puppies or juvenile dogs who still have their milk teeth.
[i] These assessments were performed with juvenile or adult dogs, or dogs who have their adult teeth.
[j] These correlations were all reported as negative such that, for example, a higher level of Locomotor Activity as a puppy correlated with fewer behavior problems as an adult. These correlations have been rekeyed so that a higher correlation reflects greater validity.
[k] These correlations are opposite what was predicted (e.g., a positive correlation was expected, but a negative was found).
[l] Owners were asked to rate their new pets 2 weeks after adoption, and then at 6 months after adoption. The Ns 2 weeks after adoption are larger than 6 months later for both puppies and juvenile/adult dogs.
[m] The correlation between number of movements at 1.61 and 3.68 months was reported as negative (but not significant).
[n] I have truncated these confidence intervals to reflect the range of possible convergent validity coefficients. Calculation of the intervals from the correlations provided yields confidence intervals ranging from less than zero, which is clearly impossible when addressing validity.
[o] Due to rounding, this correlation is reported as 0, but it is actually .001 and significant.
[p] I have calculated the validity coefficients for van der Borg et al. (1991) from the data the authors provided.

[q] Calculations involving the reported r = 1 are calculated using r = .99; when r = 1, Fisher's r-to-z yields a z of infinity, because a true correlation of r = 1 occurs with the probability of 0.
[r] I have calculated the validity coefficients for Netto & Planta (1997) from the data the authors provided. Netto & Planta (1997) also report 15.4% false positives, or that 15.4% of the dogs they predicted from their test to have a bite history do not/have never bitten before.

Strong, interpretable validity coefficients (unweighted mean = .44, sample-weighted mean = .42) are associated with the Fearfulness dimension. Fearfulness was examined in many studies and with many different predictors. This dimension has been shown to be relatively highly predictable, even from early puppyhood to later adulthood (e.g., Goddard & Beilharz, 1984b).

What criteria should be used to evaluate these validity coefficients? One potential benchmark is provided by equivalent research in the human literature. In human studies, trait-behavior correlations are typically in the order of .20-.30. For example, in one human study, correlations between self-reported personality and ratings made by observers after a 20-minute discussion task averaged .24 across the Big Five human personality dimensions (Paulhus & Bruce, 1992). Measured against this human standard, the dog validity coefficients seem satisfactory at the very least.

As shown in Table 2.5, the convergent validity coefficients varied substantially across the studies, with some studies obtaining much stronger validity estimates than others. What factors could be driving the cross-study differences in validity? One possibility is the age of the dogs. Indeed, indirect support for the idea that puppies are harder to test than older dogs is provided by the fact that the study with the lowest average validity coefficient (less than .05) involved puppies (Wilsson & Sundgren, 1998). More generally, there is a marked difference between the validity coefficients for tests of puppies (unweighted mean = .30; sample-weighted mean r = .14; Beaudet et al., 1994; Hennessy et al., 2001; Goddard & Beilharz, 1984a, 1986; Wilsson & Sundgren, 1998) versus adult dogs (unweighted mean = .61; sample-weighted mean r = .28).

Two of the studies provide more direct support for this idea. Hennessy et al. (2001) evaluated the validity of assessments administered in the same way both to puppies and to older dogs; the mean validity coefficients for the puppies (unweighted mean = .25; sample-weighted mean = .25) was much lower than that for the older dogs (unweighted mean = .41; sample-weighted mean = .37). This is consistent with another study, which identified a nearly linear relationship between age and test validity (Goddard & Beilharz, 1984a). Together these studies strongly suggest that tests of young

puppies are relatively poor predictors of their future behavior compared to tests of older dogs. These tests suggest that the inclusion of puppy studies in the meta-analysis biases the estimate of validity in older dogs. Indeed, if the results of the large study of 630 puppies (Wilsson & Sundgren, 1998) are removed from the meta-analysis, the overall sample-weighted validity estimate assessed across all seven personality dimensions jumps from .23 to .42.

Discriminant validity

Although discriminant validity has largely been neglected, a few articles (Hsu & Serpell, 2003; Serpell & Hsu, 2001; Svartberg, 2005) did examine and report this facet of construct validity. It should be noted that all three of these studies utilize a variant of Serpell and Hsu's questionnaire for assessing personality, and specifically behavioral problems: the C-BARQ. This leaves the question of discriminant validity in other assessment methods and tools relatively unexamined. However, all three of these studies found good evidence for the discriminant validity of the measures, with a few exceptions (e.g., an unpredicted association between Attachment and Stranger Fear/Stranger Aggression). However, even these exceptions are useful because they can serve as a launching point for future studies that investigate these unexpected links.

In addition to Serpell and Hsu's studies and Svartberg's (2005) study there were some other studies that mentioned discriminant validity but did not report the relevant correlations (Goddard & Beilharz, 1986; van der Borg, 1991) and there were some studies that reported the relevant correlations although they did not describe them in terms of discriminant validity (Hennessy et al., 2001; Wilsson & Sundgren, 1998). These latter studies were identified by the procedures described above ("Obtaining and categorizing the validity coefficients").

In Hennessy et al.'s (2001) study of personality in shelter animals, the validity-categorization procedures identified six potential discriminant correlations. For example, this study reported correlations between puppies' locomotor activity and the conceptually unrelated incidence of problem behavior measured two weeks (r = -.25) and six months

(r = -.30) after adoption. The absolute values of the discriminant correlations averaged .37 across the six estimates. Although none of these values were significant, these values were no lower than the convergent correlations from the same study (which averaged .36). This pattern of findings did not match the pattern of findings required to support discriminant validity, in which the convergent correlations should substantially exceed the discriminant correlations. Thus, there was no support for the discriminant validity of these assessments in this study.

Wilsson and Sundgren (1998) computed a very large number of validity correlations but reported the effect sizes only for those correlations that were statistically significant. However, because their sample size was very large, even very small coefficients reached significance. Indeed, the one statistically significant discriminant-validity estimate (between the number of objects puppies visited when placed in a room containing novel objects and adult defense drive) had a very small effect size (.024). Of course, the numerous discriminant-validity correlations that did not reach statistical significance can also reasonably be taken as evidence for the discriminant validity of the corresponding measures (because these measures also exhibited convergent validity). Unfortunately, however, as in the Hennessy et al (2001) study, the convergent correlations in the Wilsson and Sundgren study did not substantially exceed the discriminant correlations.

Summary of validity findings

Taken as a whole, the evidence broadly supports the convergent validity of personality assessments in dogs, especially adult dogs, but there was only mixed evidence for discriminant validity. However, these conclusions are based on a rather small proportion of the literature as most studies did not address validity issues. Given the centrality of validity in any assessment context, further examination of validity should remain a top priority for dog-personality researchers. In particular, research is needed to establish the parameters (e.g., dog age, testing context) that could affect validity; such findings will be essential for future work in both research and applied contexts.

In addition to furnishing numerical estimates of validity, a couple notable trends became apparent in my review of the validity literature. First, an unusually large proportion (around 90%) of the validity studies were based on Ratings of Individual Dogs; this should be contrasted with the fact that Ratings of Individual Dogs are relatively rare (18%) compared with the other methods of assessment. Second, although studies of shelter dogs constitute a small proportion of the studies in my review (13%), they were assessed in half the studies of validity. It would seem that researchers working in shelter contexts are particularly concerned with measurement issues; indeed, five of the seven shelter-dog studies reported the validity of their personality tests, and, of the two that did not, one focused on the reliability of personality testing.

We conclude by noting a trend that pervades temperament and personality research on other species (Gosling et al., 2003b). I highlight it here because although it is typically missed or ignored, it has substantial implications for validity. Research on the reliability of the measures of the criterion behaviors (against which the ratings are tested) is almost nonexistent; the reliability of behavioral codings such as the number of movements to escape (see Table 2.5) is often assumed but is rarely tested. I suspect that researchers assume that the reliability of behavioral codings will be high because such codings seem objective. That is, behavioral codings like the number of movements a puppy makes in a given time period (Hennessy et al., 2001) appear more objective than do ratings of personality, but research on humans has shown substantial variability in the reliability of such behavioral codings (Gosling et al., 1998). Therefore, it is essential that future validity research should assess and report the reliability of the criterion measures against which the validity of other means of assessment are to be estimated. Without this information, it is impossible to know whether low validity correlations are low due to genuinely low validity or due to the low reliability of the criterion measures.

SUMMARY AND CONCLUSIONS

By bringing together the disparate research on personality in dogs, my review allowed me to summarize what is known about canine personality and to identify some

trends and gaps in the field. Below I summarize my conclusions and, where appropriate, highlight some directions for future research.

(1) An extensive literature search identified 55 empirical publications on dog personality or temperament. The articles, published between 1934 and 2005, varied greatly in their assessment methods, research goals, and the attributes of their subjects (in terms of breed, age, breeding and rearing environment, and sexual status). In addition, the studies also varied in their methodological rigor, with some studies being little more than a few informal observations of a handful of dogs and others being large-scale systematic multi-phase assessments.

(2) I found that dog-personality assessment methods can be usefully grouped into four categories, which I have called Test Batteries, Ratings of Individual Dogs, Expert Ratings of Breed Prototypes, and Observational Tests. A fifth category represents studies that combined more than one assessment method. The most common assessment method was the Test Battery, which was, in theory, the closest of the four methods to achieving objectivity. In practice, however, the levels of objectivity attained differed considerably. Future research should focus on direct comparisons of the methods in terms of reliability, validity, and efficiency in different research contexts.

(3) The current review showed that dog-personality studies varied in their research goals (e.g., examining behavioral tendencies specific to breeds, Hart & Miller, 1985; Mahut, 1958; Svartberg & Forkman, 2002; predicting adult police-dog performance from puppy behavior, Slabbert & Odendaal, 1999). The vast majority of dogs tested were in working contexts (e.g., as guide or police dogs), with a relatively small number of pet or shelter dogs being studied. Given the high demand for personality testing in shelters and to assess whether dogs are fit to be adopted, greater research attention should be directed towards pet and shelter dogs. And until studies have been done to establish the generalizability of findings from working dogs to pet dogs, generalizations from one population to another should be made with caution.

(4) In the studies in my review reporting breed, at least 90% of the dogs were purebred. The Labrador Retriever and the GSD were the most frequently represented

breeds, combining to compose 30% of the subjects. The GSD was by far the most frequently tested breed, composing 24% of the dogs tested (9,253 dogs). A small minority of dogs were the planned offspring of two purebreds of different breeds, and there were also very few dogs of unintentional or unknown breed mixtures. Although this makes sense insofar as the Labrador and the GSD are two of the most frequently registered breeds in the AKC, little work has been done to examine the generalizability of these findings to different breeds. One of the few studies to compare personality across breed examined large populations of both of these commonly assessed breeds, the GSD and the Labrador Retriever, and found substantial differences in personality (Wilsson & Sundgren, 1997). Another study to examine differences among groups of breeds (e.g., Terriers, Scent hounds, Sheepdogs, etc.) again found significant differences among the breeds, indicating that some breed groups display unique patterns of personality (Svartberg & Forkman, 2002). Unfortunately, a substantial number of studies failed to report breed information. By neglecting to examine breed as a potentially important influence on personality, the value of such studies is diminished. Future research should concern itself with gaining a fuller representation of dog breeds and with providing breed information, further elucidating breed- and breed group-specific personality patterns.

(5) I also found that some method-breed combinations are more common than others. About one third of the dogs in Test Battery studies are GSDs being tested for their potential as police and working dogs. Eighty percent of all dogs in studies using Observational Tests are Labrador Retrievers, tested for their potential as guide dogs. Future research should examine the effectiveness of these two test methods, particularly Observational Tests, for other breeds and other purposes, because their ability to generalizability beyond such specific contexts cannot be assumed.

(6) There is also an age-related bias in the studies. Most studies examine dogs who were young or still in puppyhood when tested, and only few studies looked at dogs over the age of four years. In addition, age effects were rarely examined in studies using Ratings of Individual Dogs and Expert Ratings of Breeds. Consequently, I know little about how aging may shape personality in dogs. Future research should focus on this

question, and examine the developmental trajectory of personality in dogs. In particular, future research should identify the point at which personality stabilizes, such that adult traits can be predicted from puppy behavior.

(7) Eighteen of the studies in the review examined dogs bred for particular programs. Some of these studies used scores on personality tests as the basis for selective breeding. After several generations, such selective-breeding programs may shape responses to personality tests. Indeed, in one study, selective breeding lead to an increase in puppy-test scores over successive generations, but the rates at which adult dogs became successful guide dogs did not match this increase (Scott & Bielfelt, 1976).

(8) Although most pet and shelter dogs are spayed or neutered, the vast majority of dogs assessed were intact. The rare studies that did examine the effects of castration indicated that intact male dogs were the most likely to show aggressive behavior, and intact female dogs were the least likely (Podberscek & Serpell, 1996; Roll & Unshelm, 1997). However, Podberscek and Serpell's study also revealed that neutering a dog in reaction to his aggressive behavior does not reduce future aggression. Obviously, given that aggressive behavior is a concern in many programs and to many private dog owners, additional systematic research is needed in this area.

(9) A systematic multi-step procedure for summarizing the traits that have been examined in previous canine research identified seven broad personality dimensions: Reactivity, Fearfulness, Activity, Sociability, Responsiveness to Training, Submissiveness, and Aggression. The sorting procedures revealed very little standardization in the terms used to describe dog personality. Different studies often used the same terms to refer to different behaviors and different terms were often used to refer to very similar behaviors. There is clearly a need to develop a common language with which to describe personality traits in dogs (Goodloe & Borchelt, 1998). I propose that the seven categories derived from my review of the literature represents a sensible starting point for developing such a standard lexicon of canine-personality descriptors. However, substantial work to determine what factors, or trains, construct canine personality is needed.

(10) The most frequently examined personality dimension was Fearfulness, with traits related to this dimension appearing in 47 studies. Traits in the Fearfulness dimension were frequently also categorized in the Reactivity dimension, suggesting some conceptual and empirical overlap between these two dimensions. Further research on the traits of Reactivity and Fearfulness in dogs is needed to determine whether the two can be usefully distinguished or are better considered as two facets of an even broader superordinate category.

(11) Sociability was also studied frequently, in 35 studies. The traits categorized under this dimensions were sometimes also categorized under the Responsiveness to Training dimension. I suggest this overlap may be driven by the fact that an interest in people is central to both Sociability and interest in training. Future research should examine the extent to which Sociability determines Responsiveness to Training, and how best to isolate Responsiveness to Training as a separate dimension.

(12) Numerous studies included traits related to Activity. My review showed that level of Activity changes dramatically with age. However, there was also some evidence that Activity can moderate the expression of other traits. Future research should directly examine this important possibility.

(13) The studies that reported reliability were encouraging, showing that it is possible to assess dog personality reliably. However, these findings must be tempered by the fact that these conclusions are based on a lamentably small number of studies. I was shocked to discover that very few studies even report the reliability of the measures they used. Clearly, given the importance of reliability in all assessment contexts, future research should examine and report reliability.

(14) Taken as a whole, the evidence broadly supports the convergent validity of personality assessments in dogs. However, this conclusion is based on a rather small proportion of the literature as most studies do not address validity issues. Given the centrality of validity in any assessment context, further examination of validity should remain a top priority for dog-personality researchers. In particular, research is needed to establish the parameters that affect validity; such findings will be essential for future

work in research and applied contexts.

(15) Although the overall convergent validity findings were generally encouraging, my review suggests that tests of young puppies are not valid predictors of their future behavior. Given that puppy tests are widely used but their validity is rarely examined, this finding has huge implications for work in applied and research contexts. Future research is urgently needed to examine this possibility directly.

(16) My review showed that unusually large proportions of the validity studies were based on Ratings of Individual Dogs and used shelter dogs. It seems that researchers working in shelter contexts are particularly concerned with measurement issues. However, such basic issues should be of concern to all dog-personality researchers.

(17) Although rating methods (e.g., of "Fearfulness") were well represented in the studies examining reliability and validity, studies examining the reliability and validity of behavioral codings (e.g., number of time the dog scratches) are almost nonexistent. The reliability and validity of codings is often assumed but rarely tested. However, research on humans has shown substantial variability in the reliability and validity of such codings (Gosling et al., 1998). Therefore, future research should also assess and report the reliability and validity of behavioral codings. This is important in the context of validity because behavioral codings are often used as the criterion against which ratings are evaluated; but if the criterion behavioral codings are not measured reliably the ratings would appear to have low validity irrespective of their true validity.

(18) Past validity research has focused on convergent validity and generally neglected discriminant validity. Overall, the reported discriminant validity results were mixed. If the construct validity of dog personality measures is to be established, it is important that future research examine both types of validity.

Over the past 70 years great strides have been made in understanding of personality and temperament in dogs. Review of the published empirical research over this period generally supports the viability of assessing canine personality. The literature reviewed in this chapter provides a roadmap specifying the major empirical questions that need to be addressed in the next generation of studies and indicates areas of weakness

that need to be addressed, including those addressed above (e.g., lack of reporting psychometric properties like reliability). However, other weaknesses arise from the diverse nature of the studies that compose the field of dog personality research and assessment. For example, one weakness arises from the fact that the field consists of a variety of studies and tools that are each designed to address specific problems or issues. The studies assess specific types of dogs (e.g., potential guide dogs, pet dogs, potentially aggressive dogs) with specific types of behavior in mind (e.g., that suitable for a guide dog, potential behavior problems in the home, aggressive behavior). As a result, the field lacks a tool suited to assessing canine personality generally, regardless of the situation in which that dog may live. Having such a tool could afford a number of advantages, from the theoretical, broad understanding of dog personality and its structure, to the more applied comparison of the personalities of dogs who serve different roles (e.g., guide dogs, explosives detection dogs, pet dogs).

CHAPTER 3

Criteria for Selecting an Assessment Method

INTRODUCTION

To determine which of the assessment methods described in Chapter 2 to adopt, I evaluated each against a set of ten criteria. These criteria were developed to help ensure my research goals could be met by the selected assessment method. For some assessment methods, meeting all of my criteria in principle is a challenge (e.g., making a Test Battery comprehensive requires compromises likely to make the Test Battery lengthier and more difficult to administer). Other test methods are able to meet criteria in principle that they fail to meet—or only rarely meet—in practice. For example, Observational Tests could be designed to be widely applicable, but, in practice, Observational Tests are typically designed for use in constrained contexts with select populations (e.g., potential guide dogs, Murphy, 1995, 1998). Furthermore, it is sometimes difficult to evaluate how well a test method has, in practice, met my criteria due to how rarely crucial information (e.g., reliability statistics) is reported. Table 3.1 summarizes how the four methods stand up to the ten criteria both in principle and in practice.

It should be noted that my goals and criteria differ from those driving the development of many previous dog personality assessments. For example, some dog personality assessments have been developed for the very specific purpose of predicting adult guide-dog behavior (e.g., Goddard & Beilharz, 1984a, 1984b, 1986). Others have focused specifically on a single personality dimension (e.g., aggression, Netto & Planta, 1997; fearfulness, Mahut, 1958). Although these tests would not meet all of my criteria, they may have been suitable for the purposes for which they were developed.

CRITERIA

What are the qualities that would maximize the usefulness of a dog personality assessment tool? Ideally, a tool designed to assess personality and behavior in dogs meets the following ten criteria.

Reliability

The first criterion is that a dog personality assessment tool should be reliable. Indeed, in order for a dog personality test to measure personality, which has patterned and predictable effects on behavior (Pervin & John, 1997), the assessment must yield consistent (or reliable) results (Hsu & Serpell, 2003). A tool that is reliable yields results that generalize across items (or subtests) within the tool, observers, time, situations, and so on. Reliability across items, also called internal consistency, is attained when items within a measure that are purported to measure the same construct are shown to yield consistent results (e.g., two items intended to measure aggression yield similar scores). Reliability across observers, also called inter-observer or inter-rater reliability, is attained when a tool yields consistent results from different observers who use the same tool to assess the same target (e.g., a specific dog). Reliability across time, or test-retest reliability, is attained when a tool yields results that area consistent in repeated testing in which the same person uses the same tool to assess the same dog. Reliability across situations is attained when an assessment tool yields consistent results when it is used to assess the same target in different contexts (e.g., at home and a park).

Validity

Second, an assessment tool must be shown to be valid in order for its results to be meaningful. Without the prerequisite of high reliability, a personality assessment tool cannot be valid. There are various types of validity, including content, construct, and predictive validity. Content validity is the extent to which the set of items or subtests within an assessment tool represents all facets of the construct being measured. Content validity is often examined by having expert judges review the items or subtests a researcher plans to include in the assessment tool.

Construct validity is the extent to which the items or subtests within an assessment tool measure the broad construct (e.g., personality trait) they were intended to measure. This type of validity is often examined by using factor analysis to evaluate whether unique, unrelated factors underlie groups of items or subtests. For a tool to have construct validity, each item or subtest must be strongly related to its underlying factor,

but only weakly related to other factors in the assessment tool. For a tool to have construct validity, items or subtests must be strongly correlated to theoretically related items or subtests on another measure (convergent validity), but only weakly related to theoretically unrelated items or subtests on the other measure (discriminant validity).

Predictive validity is extent to which the results of an assessment tool are consistent with results from another, independent measure of the same constructs (e.g., Fearfulness-related items on a questionnaire might be compared with Fearfulness-related subtests in a Test Battery). A complete evaluation of a tool's predictive validity involves relationships not only with things to which each construct is expected to relate (convergent validity), but also those to which each construct is expected to be unrelated (discriminant validity).

Sensitivity

Third, a dog personality assessment must show a high level of measurement sensitivity. That is, the tool should differentiate not only among dogs whose[1] personalities and behaviors are very different from one another, but also among dogs whose personalities or behaviors are similar but not identical (i.e., differ relatively minimally, but still meaningfully). For example, an assessment should be able to distinguish between dogs who are pervasively aggressive and dogs who are aggressive in only a few, specific situations.

[1] In referring to dogs, I have chosen to use personal pronouns such as "he" and "she" instead of the neutral pronoun "it," the personal pronoun "who" in place of the more conventional demonstrative pronouns "that," and so forth. Writers (e.g., Dunayer, 2001) and researchers (e.g., Gilquin & Jacobs, 2006) have argued that language is often used to as a subtle means of denying the individuality and unique selves of nonhuman animals. In their reviews of linguistic corpora, Gilquin and Jacobs (2006) found "who" was used to refer to nonhuman animals in particular when there was a sense of psychological closeness (e.g., with pets) and when a feature shared with humans was being discussed. On these bases, the use of impersonal language to refer to non-human animals is at odds with the widely accepted notion that nonhuman animals can be characterized in terms of individual differences and personality.

Minimal order effects

Fourth, effect of each item or subtest on subsequent items or subtests should be addressed. Ideally, the items on a test should be independent of one another so that a dog would earn the same score on any given item regardless of whether it was presented early or late and regardless of the items immediately before and after it. That is, development of the tool would control for order effects, and the instrument would be shown to be free of order effects of have only minimal order effects.

Availability of psychometric properties to the public

Fifth, information about the tool's psychometric properties, including aspects of reliability and validity described above, should be readily available to the public so that potential users of the tool can evaluate its effectiveness and suitability for their purposes.

Comprehensiveness and detail

Sixth, a dog personality assessment tool should be comprehensive and detailed in terms of the behaviors and characteristics it includes. The tool must assess all pertinent personality traits and their facets, collecting enough information about each so as to be generally useful and informative. For example, for the trait of Fearfulness, which is often very situation-dependent, the results of the assessment tool should be able to specify in some degree of detail the type(s) of situations in which a dog who was assessed with the tool is likely to exhibit fearful behavior. In order to collect information about the situations that provoke fear in different dogs, the test must assess the fearfulness in a wide variety of situations. These situations must vary by factors that are likely to influence whether or not a dog behaves fearfully, including individuals with whom the dog interacts (e.g., veterinarians, unfamiliar adults of both sexes, unfamiliar children, people in uniform, other animals), how others treat the dog (e.g., threaten, reach for, ignore), and the physical environment (e.g., familiar, unfamiliar, home, business). Comprehensiveness and detail are closely tied to content validity, but differ from content validity in that a tool can collect additional detail above and beyond what is necessary for content validity to be attained. Making a tool more comprehensive can be difficult to

balance with other criteria (e.g., manageability, ease of use, efficiency) because making the tool more comprehensive and detailed necessarily includes and increase in the number of situations, observations, or ratings.

Wide applicability

Seventh, the tool should be widely applicable, or useable across a range of dogs and contexts. It should be able to assess the personalities of specific types of working dogs, like military patrol dogs, and also general pet dogs. The tool should also be suitable for assessing dog personality in a wide variety of contexts, including kennels, shelters, and homes.

Ease of use

Eighth, a dog personality measurement tool should be easy for people to use -- in terms of administration, scoring, and application of the results -- without extensive expertise or training. For a tool to be generally useful, it should not depend on experts (e.g., veterinarians, behaviorists) and should be relatively transparent. Preferably, use of the tool would not require special training (e.g., training courses, extensive reading, extensive practice administering the test). The language of the instrument should be readily comprehended by the layperson and extremely complex tests or calculations should be excluded.

Efficiency

Ninth, the tool must be efficient, or require a minimum of time to administer to collect a maximum amount of information about the dogs. Many situations, such as those in which a limited number of people have little time to devote to assessing a very large number of dogs (e.g., shelters, working dog programs), require a quick and efficient tool.

Manageability

Finally, the test must not require extensive resources to administer. Testers should be able to assess a dog with a minimum of monetary resources, physical space, time, and

test administrators. That is, the assessment should not require exacting environments that demand large amounts of time and effort to prepare.

Summary of Criteria for Selecting an Assessment Method

In summary, an ideal instrument should yield the same results regardless of when it is used, who uses it, and the situation in which it is used – that is, it should be reliable. Items (or subtests) in the instrument that purport to measure the same feature of a dog's personality should yield corresponding results, indicating a high level internal consistency. These items (or subtests) should also yield the same results regardless of the order in which they are presented. The tool should also have high validity, measuring all aspects of pertinent constructs, showing that items purported to measure a given construct are (statistically) related and those purported to measure different constructs are (statistically) unrelated, and predicting outcome of theoretically related, but independent, measures. Furthermore, the tool should also be highly sensitive, able to distinguish between dogs who differ relatively minimally. Information about how well the instrument meets these and other criteria should be available to the public and other potential users. In order to be as universally useful as possible, the instrument should also be comprehensive, detailed, and widely applicable to a variety of dogs and contexts. In order to be as accessible as possible to a wide variety of people and situations, the tool should be easily used without extensive training or specialized backgrounds, its length should be suitable to a variety of situations, and its logistics should be manageable so that it does not require resources beyond the scope of the average situation in which many dogs' personalities need to be tested.

Assessment Methods

As reviewed in Chapter 2, previous assessments of dog personality fall into four types: Test Batteries, Observational Tests, Expert Ratings of Breed Prototypes, and Ratings of Individual Dogs. Each of these assessment methods has the potential to, in principle, fulfill a variety of the criteria outlined above. However, the criteria each method could meet in principle and the criteria each method does meet in practice vary.

Test Batteries

As described in Chapter 2, Test Batteries aim to document dogs' reactions to specific stimuli or situations. Typically, stimuli are presented to a dog and the dog's reactions are recorded or scored. Thus, Test Batteries have two primary components: the behavioral tests (or subtests) and the systems for assigning scores to dogs' reactions to stimuli.

In principle, Test Batteries can be designed to meet eight of the ten criteria: reliability, validity, sensitivity, availability of psychometric properties, comprehensiveness and detail, wide applicability, efficiency, and manageability (see the second column of Table 3.1). However, meeting many of the criteria (e.g., reliability across situations, content validity, comprehensiveness and detail, efficiency) may require compromising other criteria. Meeting the criteria of ease of use and minimal order effects may be impossible for Test Batteries. Some of these challenges may be most prevalent for Test Batteries, but most are specific to both types of behavioral tests (Test Batteries and Observational Tests), and others are general challenges associated with a criterion.

It may be possible, but seems challenging and unlikely, that Test Batteries could meet the criterion of ease of use. They tend to require training to administer, although that is not always the case and depends on the scoring system and other factors.

Minimizing order effects is a particularly challenging hurdle for both types of behavioral tests (Test Batteries and Observational Tests) because a dog's experience in one situation may impact his or her reaction to the next situation. For example, if a Test Battery exposes dogs to the abrupt opening of an umbrella, then to an unfamiliar person, the response elicited by the umbrella may shape the dog's reaction to the unfamiliar person. If the dog is scared by the umbrella, he or she may be physiologically aroused (e.g., have elevated cortisol, a racing heartbeat), and the unfamiliar person may elicit more fear than if the order of the two subtests were reversed.

It is difficult for a behavioral assessment to balance the criterion of content validity (and comprehensiveness and detail) with ease of use, efficiency, and manageability. To meet the criterion of content validity, containing items or subtests

pertinent to all aspects of canine personality, a Test Battery is likely to require a large number of subtests in which different, specific stimuli are presented and the dog is placed in different situations. Administering each subtest, then scoring each subtest, would likely result in a Test Battery that was comprehensive, but too lengthy and burdensome to meet the criteria of ease of use, efficiency, and manageability.

In practice, it is unclear whether Test Batteries meet many of the criteria that they could, in principle, meet. Indeed, Test Batteries might meet only one criterion: sensitivity. It is impossible to determine how well Test Batteries generally meet the criteria of reliability, validity, minimal order effects, and efficiency, because statistics and other information pertinent to these criteria are rarely, if ever, reported. However, as reviewed in Chapter 2 and summarized in Table 2.1, some Test Batteries are associated with high reliability statistics.

Existing Test Batteries also fail to meet a number of the criteria expounded in the previous section. The psychometric properties of Test Batteries are, to my knowledge, very rarely reported and not made available to the public. Test Batteries simply do not, in practice, meet the criteria of comprehensiveness and detail, wide applicability, ease of use, and manageability. In practice, Test Batteries likely fail to meet the criterion of comprehensiveness and detail and the criterion of wide applicability because they are designed to suit narrow contexts and meet specific goals (e.g., accessing aggression level in potentially aggressive breeds; Netto & Planta, 1997). The research goals for which many Test Batteries have been developed do not require measurement of all aspects of dog personality, but rather of a subset thought important to a specific task or situation (e.g., guide dog work, being a domestic pet). Tests with more focused goals do not need to assess every aspect of dogs' personalities, but, as a result, do not meet the criterion of comprehensiveness and detail and are unlikely to meet the criterion of content validity. Furthermore, Test Batteries have limited applicability because they are designed to meet very specific goals. For example, Slabbert and Odendaal's (1999) study using a Test Battery of puppies to predict adult police dog efficacy meets their study's goals well, but

factors such as the behaviors examined (e.g., reaction to gunfire) limit how widely applicable the tool is.

Test Batteries do not meet the criterion of ease of use, because most are designed for administration and scoring by testers with a lot of training and expertise. Even those Test Batteries more readily available for public use tend to require training. For example, there are Test Batteries available for public use in dog shelters, such as Rebecca Ledger's DTA-IV and Emily Weiss's SAFER-II, but the designers of each test recommend that potential users do not use the test prior to undergoing training in the specific administration and scoring. A potential user cannot simply pick up the test and apply it.

Observational Tests

Observational Tests aim to assess and describe traits discernible in more naturalistic environments. They seek to draw broader pictures of dogs' personalities and behavioral patterns than is possible using Test Batteries. Like Test Batteries, Observational Tests can also be considered as having two primary components: the tests themselves and the systems for scoring the dogs' behavior during the test. Unlike Test Batteries, however, Observational Tests are usually conducted in carefully selected, but not controlled, environments and involve the fortuitous presentation of naturally occurring stimuli.

Observational Tests are similar to Test Batteries in that they both depend on the measurement of behavior that dogs display during the test. Thus, although Observational Tests and Test Batteries differ in how they elicit behavior, they both depend on provoking dogs to respond to their environments behaviorally and scoring the resulting behavior. As a consequence, Observational Tests fail to meet many of the same criteria and face many of the same challenges as Test Batteries.

In principle, Observational Tests can be designed to meet the same eight criteria that Test Batteries can meet: reliability, validity, sensitivity, availability of psychometric properties, comprehensiveness and detail, wide applicability, efficiency, and manageability (see the fourth column of Table 3.1). Again, meeting many of these criteria

(e.g., reliability across situations, efficiency) may require a decrease in how well the tool meets other criteria. Meeting the remaining two criteria, ease of use and minimal order effects, may be possible only in narrow circumstances.

In principle, an Observational Test is unlikely to be easy to use because administration and scoring of the test generally depends on experts observing dogs as they go through and respond to naturalistic environments. It may be possible to make such a test easy to use, if the scoring system used is readily accessible to the general public. For example, a scoring system on which observers assigned trait-based ratings to each dog at the end of the Observational Test might be easy to use.

Order effects are difficult to minimize in Observational Tests for the same reason they are difficult to minimize in Test Batteries: A dog's experiences are likely to affect his or her responses to subsequent situations. For example, if a dog is put through an Observational Test in which the dog is walked through a crowded shopping mall, the dog might encounter small child who is yelling and running around. The dog could be excited by this movement, then a more excitable in response to later stimuli (e.g., drinking fountains, statues, people) that would not have elicited excitable behavior if they had been encountered before the child.

It is difficult for an Observational Test to simultaneously satisfy the criterion of content validity (and comprehensiveness and detail) and other criteria, such as ease of use, efficiency, and, in particular, manageability. For an Observational Test to have increased content validity, the dog must be exposed to additional stimuli, and most likely be taken to additional physical locations (e.g., home, busy shopping mall, park). This increase in stimuli and situations necessarily decrease how manageable the test is, because it will require more time, money, and other resources to administer than a single-location test would require.

In practice, it is unclear whether Observational Tests meet many of the criteria that they could, in principle, meet. Observational Tests might, like Test Batteries, meet only one criterion: sensitivity. It is impossible to determine how well Observational Tests meet the criteria of reliability, validity, minimal order effects, efficiency, and

manageability because the pertinent statistics and other information are rarely, if ever, reported. In addition, Observational Tests typically fail to meet the criterion of making their psychometric properties available to the public because they simply do not report many of the necessary statistics, nor do they make the reported statistics accessible or interpretable to the public.

For Observational Tests, manageability is a particular issue, because Observational Tests generally depend on studying dogs' responses to stimuli in relatively naturalistic environments. For example, Goddard and Beilharz's (1984b) Observational Test in which dogs were observed during walks through shopping malls may be, for many potential users, manageable. For other potential users (e.g., shelters that would have to transport hundreds of dogs to the test site), however, such tests are not manageable. Unfortunately, an increase in manageability means a decrease in the amount of information that is gathered, because it necessitates a reduction in the variety of test situations included in the assessment.

Published Observational Tests generally fail to meet the criteria of comprehensiveness and detail, wide applicability, and ease of use. In practice, Observational Tests may fail to meet these criteria because the tests have been, like Test Batteries, designed to meet very specific goals or to be used on a specific type of dog. For example, Murphy (1995, 1998) used Observational Tests to assess potential guide dogs for their suitability as guide dogs. Observational Tests like those described by Murphy (1995, 1998) are not required to measure all aspects of dog personality; they only need to measure those pertinent to the task at hand (e.g., guide dog work). They do not need to be widely applicable, because they are designed exclusively for assessing a specific type of dog in a specific context. It might be argued, however, that the people administering such assessments (who have limited time) would benefit were the tests easy to use, because the test administrators would then require minimal training to use the test.

Expert Ratings of Breed Prototypes

In Expert Ratings of Breed Prototypes, researchers select people whom they consider experts on dogs (e.g., American Kennel Club judges, veterinarians, dog trainers) and ask those people to describe, rank, or rate dog breeds rather that specific individual dogs. Expert Ratings of Breed Prototypes face limitations in terms of the criteria they can meet because they do not assess the personalities of individual dogs.

In principle, Expert Ratings of Breed Prototypes can meet the criteria of reliability, some aspects of validity (content and construct validity), minimal order effects, availability of psychometric properties to the public, comprehensiveness and detail, efficiency, and manageability. That is, Expert Ratings of Breed Prototypes cannot fully meet the criteria of predictive validity, sensitivity, ease of use, or wide applicability. Expert Ratings of Breed Prototypes, however, always require a certain level of expertise (e.g., extensive familiarity with a breed, veterinary or behavioral training) and cannot, in principle or in practice, be used without expertise, making them fail to meet the criterion of ease of use.

There are also, in principle, problems associated with Expert Ratings of Breed Prototypes. First, it should be noted that Expert Ratings of Breed Prototypes could achieve high reliability across observers if ratings are largely based on stereotypes, but these stereotypes may or may not accurately predict real dogs' behavior. The degree to which Expert Ratings of Breed Prototypes have predictive validity is limited by the degree to which a dog's breed determines his or her behavior. Expert Ratings of Breed Prototypes do not assess the personalities of individual dogs and, therefore, cannot meet the criterion of sensitivity defined as the ability to differentiate between dogs whose personalities differ from each other relatively minimally; this type of assessment can, instead, distinguish between dog breeds associated with personalities that differ minimally from each other (which no other assessment method can address). The applicability of Expert Ratings of Breed Prototypes is limited by the applicability and the predictive validity the prototype itself. Because of the potentially large limitation of

applicability, serious consideration must be given to the usefulness of Expert Ratings of Breed Prototypes.

In practice, Expert Ratings of Breed Prototypes meet only a subset of the criteria they could meet. They meet some, but not all, reliability criteria (reliability across observers and situations), are comprehensive and detailed, and are manageable. It is impossible to determine how well this assessment method meets the criteria of reliability across items, reliability across time, content validity, construct validity, predictive validity, and minimal order effects because very few studies have used and examined Expert Ratings of Breed Prototypes (see Table 2.1), and the studies that have used Expert Ratings of Breed Prototypes have rarely, if ever, reported the statistics and other information necessary for evaluating how well the assessments meet these criteria. Like Test Batteries and Observational Tests, the psychometric properties of Expert Ratings of Breed Prototypes are not available to the public. Those studies that discuss efficiency of data collection using Expert Ratings of Breed Prototypes (e.g., Hart & Miller, 1985) indicated that the assessment method is efficient. However, in practice, as well as in principle, the usefulness of Expert Ratings of Breed Prototypes is tempered by a lack of information about the assessment method's predictive validity.

Ratings of Individual Dogs

Ratings of Individual Dogs gather information about specific dogs' behaviors and histories from an informant (e.g., the dog's owner). The informant states whether or not, or how often, his or her dog has engaged in specified behaviors (e.g., snapping at children, barking when someone knocks on the door) or provides a rating of the dog on a list of traits (e.g., Fearfulness, Aggression).

In principle, Ratings of Individual Dogs can meet all ten of the criteria summarized above and listed in Table 3.1. However, it should be noted that, like all the other methods of assessment, Ratings of Individual Dogs will become lengthier, thus requiring greater amounts of time and attention to complete, as they become more comprehensive and detailed. Thus, the criteria of efficiency and manageability must be

carefully balanced with a questionnaire's content validity, comprehensiveness and the level of detail is collects. One benefit associated with Ratings of Individual dogs is that they are the only method of assessment which is, in principle, generally usable without training or expertise, and which can be designed to be easily used by the layperson.

In practice, Ratings of Individual Dogs are known to meet the criteria of sensitivity, ease of use, efficiency, and manageability. It is challenging to estimate the degree to which Ratings of Individual Dogs generally meet the criteria of reliability, validity, and minimal order effects because the pertinent statistics are rarely reported. However, some dog personality assessment studies have examined various aspects of reliability in Ratings of Individual Dogs (e.g., Serpell & Hsu, 2001; Gosling et al., 2003a). For example, Gosling et al. (2003a) found rating of dogs were consistent across items within a scale, across observers, and across tests and time (ratings to a behavioral test several weeks later).

Ratings of Individual Dogs generally fail to meet the criteria of availability of psychometric properties to the public, comprehensiveness and detail, and of wide applicability. They fail to meet the first of these criteria because report of reliability and validity statistic is rare. Most Ratings of Individual Dogs fail to meet the criterion of comprehensiveness and detail because they are very brief questionnaires designed for narrow, specific purposes. Notable exceptions include Goodloe and Borchelt's (1998) and Serpell's C-BARQ, which are lengthy and thus relatively comprehensive, but both of which focus explicitly on behavioral problem assessment. As a consequence of such measures' focus on specific goals, Ratings of Individual Dogs also have limited applicability. Unlike all previously discussed assessment methods, Ratings of Individual Dogs are generally reported to be very easily used by the layperson (e.g., Goodloe & Borchelt, 1998; Hsu & Serpell, 2003; Podberscek & Serpell, 1996).

Table 3.1. Dog personality assessment methods' fulfillment of 10 criteria in principle and in practice

	Test Batteries		Observational Tests		Expert Ratings of Breed Prototypes		Ratings of Individual Dogs	
Criterion	**Principle**	**Practice**	**Principle**	**Practice**	**Principle**	**Practice**	**Principle**	**Practice**
Reliability…								
internal consistency (across items or subtests)	Yes	Rarely reported	Yes	Rarely reported	Yes	Rarely reported	Yes	Rarely reported
test-retest reliability (across time)	Yes	Rarely reported	Yes	Not reported	Yes	Not reported	Yes	Rarely reported
inter-observer reliability (across observers)	Yes	Not reported	Yes	Not reported	Yes	Yes	Yes	Rarely reported
across situations	Yes	Rarely reported	Yes, but less manageable	Rarely reported	Yes	Yes	Yes	Rarely reported
Validity…								
content validity	Yes	Not reported	Yes	Not reported	Yes	Rarely reported	Yes	Rarely reported
construct validity (convergent & discriminant)	Yes	Rarely reported	Yes	Rarely reported	Yes	Rarely reported	Yes	Rarely reported
predictive validity (convergent & discriminant)	Yes	Rarely reported	Yes	Rarely reported	Limited amount breed determines behavior	Not reported	Yes	Rarely reported
Sensitivity	Yes	Yes	Yes	Yes	Limited	Limited	Yes	Yes
Minimal order effects	Maybe, but difficult	Not reported	Maybe, but difficult	Not reported	Yes	Not reported	Yes	Not reported
Availability of psychometric properties to the public	Yes	Rarely reported	Yes	Rarely reported	Yes	Rarely reported	Yes	Rarely reported

Table 3.1. (*Continued*)

	Test Batteries		Observational Tests		Expert Ratings of Breed Prototypes		Ratings of Individual Dogs	
Criterion	**Principle**	**Practice**	**Principle**	**Practice**	**Principle**	**Practice**	**Principle**	**Practice**
Comprehensiveness and detail	Yes, but longer and less efficient	No	Yes, but longer and less efficient	No	Yes	Yes	Yes	No (designed for specific situation)
Wide applicability	Yes	No	Yes	No	Limited	Limited	Yes	No (designed for specific situation)
Ease of use (e.g., without training, expertise)	Training usually needed	No	Depends on construct assessed, coding system and other factors	No	No	No	Yes	Yes
Efficiency	Yes, but less comprehensive	Rarely reported	Yes, but less comprehensive	Rarely reported	Yes	Yes (few studies; rarely reported)	Yes	Yes
Manageability (logistics)	Yes, but compromise other criteria	No	Yes, but compromise other criteria (e.g., efficiency)	Rarely reported	Yes	Yes	Yes	Yes

Evaluation and Selection of an Assessment Method

Only Ratings of Individual Dogs stand out as having the potential to meet the ten criteria described above with minimal sacrifice and compromise. Test Batteries and Observational Tests both fall short because they require large amounts of time in order to be comprehensive and collect detailed data, they typically require expertise and training to administer, and it may be impossible to limit order effects that arise during testing. Expert Ratings of Breed Prototypes have two apparent, unavoidable limitations. First, they require a dog expert to assign ratings. Second, they do not allow for measurement of individual dogs' behavior and personality, but rather aggregate across all dogs of a given breed with whom the expert rater has experience.

However, it is apparent from the review of existing dog personality assessment tools (see Chapter 2) that rating systems used less frequently than Test Batteries. Why might researchers prefer other assessment methods? One reason is that, historically, ratings have been criticized as too subjective and described as inappropriate tools for scientific use. There is also a relatively long history of research contradicting this, showing that ratings, rather than hindering measurement, actually allow researchers to utilize human perceptions and experiences to record and interpret very large amounts of information effectively (see Vazire, Gosling, Dickey, & Schapiro, 2007). Buirski et al. (1978) argued trait terms (e.g., aggressive, timid) are no more subjective and no less useful than most terms used in psychology or ethology (p. 127). Furthermore, aggregate observations composed of ratings by several independent observers meet the standards required of any measurement instrument; they are reliable and largely independent of idiosyncrasies of individual observers (e.g., Block, 1961; Epstein, 1983).

Despite the historical objection, some researchers have sought to take advantage of the benefits afforded by Ratings of Individual Dogs. In particular, Serpell, Hsu and colleagues (e.g., Serpell & Hsu, 2001; Hsu & Serpell, 2003) and Goodloe and Borchelt (Goodloe & Borchelt, 1998) have developed the broadest and most comprehensive of the existing Ratings for Individual Dogs; both were developed for use in assessing behavioral

problems. Serpell, Hsu, and their colleagues have performed numerous studies to develop a questionnaire for assessing consistent patterns in pet and assistance dogs' behavior (arguably personality), the Canine Behavioral Assessment and Research Questionnaire (C-BARQ; e.g., Serpell & Hsu, 2001; Hsu & Serpell, 2003). The C-BARQ is available to the public and the research community online at www.vet.upenn.edu/cbarq, where users can also compare their dogs' scores to those of other dogs. In the development of this questionnaire, the authors paid explicit attention to and reported statistics relevant to the questionnaires' psychometric properties, including reliability and validity. However, the C-BARQ is a very extensive and lengthy questionnaire, with items separated into 13 categories by the factor they measure, making the questionnaire too time-consuming for many applications. Also, as mentioned above, the C-BARQ's design is motivated by an interest in evaluating and screening dogs for behavioral problems (www.vet.penn.edu/cbarq). Similarly, Goodloe and Borchelt (1998) designed a very extensive questionnaire to assess dog behavior, which also focused on behavioral problems. While the assessment of dogs for behavioral problems is clearly a worthwhile purpose and one that makes the questionnaire very useful, it limits the general applicability of the tool.

The existing Ratings of Individual Dogs have only limited application and breadth, and thus meet only some of my criteria. In addition, many studies using other assessment methods have yielded findings about dog personality that have never explored through trait ratings. So, although previous research does not satisfy the field's need for a widely applicable, comprehensive questionnaire about dog personality, it does suggest that Ratings of Individual Dogs is the best method to meet the ten criteria I have specified, and it does provide an adequate starting point – in terms of personality constructs and item content -- for the development of the tool. The goal of this study, therefore, is to build a questionnaire that draws from past research and is thus grounded in research and theory, that is comparable with other systems of assessment, that is useful across a broad array of applications (from pets to working dogs), that can be completed in a reasonable amount of time to aid in efficiently gathering data, that is amenable to a long

and short version, and that is understandable by laypersons so researchers may draw on their everyday experiences and observations of their dogs.

INSIGHT FROM HUMAN PERSONALITY RESEARCH

Human personality psychologists face the same type of decisions regarding which method of assessment to use, and, within each method, how to best develop an assessment tool or system. Some methods for studying personality in human animals and non-human animals overlap. The majority of methods that can be used to assess personality in non-human animals can be used with humans. Of the four methods of assessment discussed here, Test Batteries, Observational Tests, and Ratings of Individuals are readily used for assessing most individuals, whether human or non-human. However, because humans are not as readily or acceptably divided into groups as dogs are divided into breeds, Expert Ratings of Breed Prototypes are unlikely to easily translate to use in humans.

When developing tools for rating individuals, or a taxonomy of personality-descriptive terms to be used as items in a questionnaire, human personality psychologists face many of the same challenges and issues that are inherent to the process of developing a dog personality questionnaire. The overlap of challenges, regardless of the species assessed, suggests that that these issues are not tied to the assessment of one species or the other, but rather to the assessment method itself.

Twenty years ago John, Angleitner and Ostendorf (1988) examined the process of building complete and comprehensive collections of human personality attributes, as needed for a personality questionnaire. They described a number of basic issues, or challenges. At a very basic level, before beginning to develop an assessment, researchers must specify what they will measure, or define personality. Inherent to this first step is setting limits of what will be studied, and possibly creating limitations. A common approach is to focus on a limited domain (e.g., stable traits), but this is often still too broad and researchers further limit the scope of their tools to, for example, extraversion, or interpersonal traits, or risk-taking behaviors.

Furthermore, researchers must determine from where they will compile a comprehensive list of instantiations or descriptions relative to the domain they study. Some researchers have collected personality descriptions from subjects' responses to questionnaires. Other researchers have relied on clinical expertise (e.g., Block, 1861). Still others have used the lexical approach, assuming that natural language can serve as a source for personality attributes because people will have encoded in language ways of describing the most important or salient differences between people (e.g., Klages, 1926/1932; Allport, 1937; Cattell, 1943; Goldberg, 1982).

Researchers have recently suggested a construct-validity based approach to scale construction (Simms & Watson, 2007). Because field of human personality assessment is arguably more developed than that of animal personality assessment, but the two share very similar challenges, human personality assessment can advise animal personality assessment methodologies. To that end, much of the construct-validity based approach over-viewed here has been adopted in developing the dog personality questionnaire.

The construct-validity based approach to scale construction divides scale construction into three phases: the substantive validity phase, the structural validity phase, and the external validity phase. The substantive validity phase is built upon reviewing existing literature to determine if a new assessment tool is needed, then defining exactly what is to be measured by the new tool, compiling and writing items, then examining the items through pilot testing and/or expert review. The structural validity phase targets the goal of determining the structure of the items (e.g., their grouping into personality factors), which is often done through data collection and factor analysis. This phase may suggest changes to the questionnaire item list. The final stage, the external validity phase, is the phase in which the researcher determines whether the assessment tool predicts results of an independent assessment (a criterion), and whether items on the scale that should be related (i.e., items purported to measure extraversion) are statistically related while those that should not be related (i.e., items purported to measure openness and items purported to measure neuroticism) are statistically shown to be unrelated. This phase, too, may suggest changes to the questionnaire item list.

CHAPTER 4

Study 1: Pilot Testing the Initial Item Pool

INTRODUCTION

The goal of Study 1 was to build the initial form of the DPQ. The initial list of questionnaire items that served as the launching point for the DPQ's development was generated, then the item list was revised and administered to participants. Participants' feedback was used to guide further modification of the DPQ items.

I employed a multi-phase procedure for generating a pool of descriptors and then selecting items from the pool. Essentially, the phases involved generating an item pool; categorizing the items; removing items that were clearly redundant, too broad, or too narrow in focus; then creating a questionnaire to collect empirical data from volunteer participants.

PART 1: GENERATING AN ITEM POOL

My first goal was to generate a comprehensive pool of descriptors from which questionnaire items could be drawn. To ground the questionnaire in current research and practice and to maximize comprehensiveness, descriptors were chosen from both research and applied settings; in particular the item pool was based on descriptors used in (a) the dog personality and temperament research literature, (b) human personality questionnaires (e.g., the Big Five Inventory [BFI]; John, 1990), (c) instruments used in applied settings (e.g., shelter intake forms, shelter dog personality tests), and (d) supplemental items generated by dog experts.

In Chapter 2, I identified numerous studies in their review of the dog temperament and personality literature. Their review identified articles that would serve as one source of descriptors. Specifically, I collated the articles reviewed in Chapter 2 as well as the questionnaires (e.g., the C-BARQ, the Dog-Big Five Inventory or D-BFI) used in the research reported by these articles. In addition, I supplemented this pool with items that were not included in the review because they are still under development by other

researchers. From each of these articles and questionnaires, I extracted behavior and trait descriptions. To be as comprehensive as possible, I extracted all behavior and trait descriptions, regardless of whether they were the focus of the article. For example, in one factor analytic study (Serpell & Hsu, 2001), I extracted all the traits analyzed, and also the researchers' descriptions of common reasons for potential guide dogs' failing to become working guide dogs.

I also gathered a separate representative set of animal shelter intake forms and frequently-used shelter dog personality tests. The selection of these tests was based on how distinctive the instruments were from each other, how frequently they were used, and how well they, collectively, represented the variety of instruments used in shelters. From each intake form and personality test, I again extracted behavior and trait descriptions.

After extracting behavior and trait descriptions from all my sources, I eliminated exact repetitions (e.g., "dominance over owner" from Draper, 1995; Bradshaw & Goodwin, 1998; Hart & Miller, 1985). Once repetitions were removed, the behavioral and trait descriptions from the research and applied contexts yielded a pool of 1,284 descriptions.

Of course although I took care to gather all relevant behavioral and trait descriptions, my questionnaire item list's content is largely dependent on the behaviors and traits examined in past research and/or identified as important in applied settings, and therefore may suffer associated limitations.

Initial groupings of items into categories

Given such a large pool of potential items for the questionnaire, a challenge was finding and identifying items with similar or matching content but slightly varied wording (e.g., "In house/apartment, follows owner/family member from room to room," Goodloe & Borchelt, 1998; "How often does the dog follow the owner around the house?," Ledger, personal communication). To facilitate the search for redundant items, the items were grouped into broad, content-based categories based on types of behavior

(e.g., aggression, friendliness). Through this process, the 1,284 items were grouped into 51 categories, such as “Mounting behavior,” “Chasing,” “Aggression,” and “Fearfulness” (for a complete list, see Appendix A). The categories ranged in size from one item (e.g., “Alertness”) to 240 items (”Aggression”). Seven of the categories (e.g., Aggression, Fearfulness, Separation-Related Behaviors) were still too broad to determine the redundancy so they were divided into narrower subcategories, usually based on significant features of the items (e.g., involve children), stimuli described in the items (e.g., cars, traffic), or environments (e.g., vet’s office). For example, the Separation-Related Behaviors category was subdivided into seven categories: Vocalizations, Destructiveness, Food-Related, Restlessness, Elimination Behavior, and General, with a “Miscellaneous” subcategory for items that did not fit other subcategories. It should be noted that the categories were provisional and designed to help in reducing redundancy; they did not determine the later phases of the questionnaire development. Judges were free to consider issues of cross-category redundancy, such as overlap in items in the Hearing Sensitivity category and the Fearfulness category (e.g., “Dog acts anxious or fearful in response to sudden or loud noises;” “Is dog fearful or startled when a car horn sounds?”).

Removal of redundant items

One expert judge reviewed all 1,248 items category by category. The judge had three goals: (1) to identify and remove redundant behavioral and trait descriptions by consolidating descriptions that were clearly similar, (2) to remove descriptions that were too specific to personality-testing contexts, and (3) to maintain the breadth of the original 1,284-item list. This expert made consolidation and removal decisions only when the decision did not require much judgment and could therefore be performed by a single judge. For example, “(Would like to) chase cats” (Serpell & Hsu, 2001), “Dog chases cats if given the chance” (Hsu & Serpell, 2003), and “Chases cats (given the chance)” (C-BARQ) were consolidated without consulting other judges. However, “Does your dog chase strange cats” (Stephen, personal communication) was retained as a separate item

for further evaluation with other judges because of the introduction of the issue of strange, familiar, and unspecified cats. The descriptions “fear of a toy car traveling in circles with lights flashing” (Goddard & Beilharz, 1984a) and “reaction to a stranger that hides and assaults the dog when passing the hiding place” (Svartberg, 2002) were removed because they were seen as specific to personality-testing contexts. When there was doubt whether to retain an item, the item was retained. This process reduced the pool of items to 750 items, grouped into 51 categories and their subcategories.

Review of the item pool

Three expert judges reviewed the 750 items category-by-category, examining each item individually and relative to the others in its category. The process of expert review to assess item quality has been described as a critical step in the process of item pool development. In particular, Worthington and Whittaker (2006) recommend experts assess the content validity (e.g., the extent to which the set of items represents all facets of the concept being measured) of the individual items and the item pool during this phase of questionnaire development. Therefore, the aim of this phase was to reduce the number of behavioral and trait descriptions to a more manageable set, while retaining all meaningful, unique, and potentially important descriptions for the empirical phases. Careful attention was paid to be sure that items with both positive valence (e.g., “Dog is friendly”) and negative valence (e.g., “Dog ignores commands”) were included and were roughly counterbalanced to overcome possible acquiescence response bias. Both traits (e.g., “Dog is clever”) and behaviorally descriptive items (e.g., “Dog learns readily,” “Dog is slow to learn new tricks or tasks”) were included for as many categories of items as possible. This process resulted in a list of 358 potential items, covering all 51 categories and their subcategories.

The same three expert judges then reviewed the list to determine whether any potentially important behaviors or traits had been excluded. A trait or behavior was deemed potentially important if dogs could be differentiated on the trait or behavior dimension, it had practical importance (e.g., related to learning or training), or it had

theoretical importance (e.g., captured something that might be an additional facet of a trait). Two items were added: "Dog appears to remember an object when it is out of sight (e.g., in your pocket, behind your back)," and "Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden." A list of 360 questionnaire items resulted.

The questionnaire item list needed to be screened for items that were difficult to understand/interpret, ambiguous, too technical, or otherwise problematic. I therefore needed to administer the questions, with ample opportunity for feedback, to a population representative of those who would be taking the questionnaire and possibly using it and any future versions (i.e., a variety of lay dog owners with differing levels of experience and areas of expertise).

PART 2: ADMINISTRATION AND FEEDBACK-DRIVEN REVISION OF THE DPQ

My second goal in Study 1 was to attain and use feedback from a population of dog owners. Because major goals in the general design of the DPQ include meeting criteria such as ease of use and wide applicability, it was of vital importance to get feedback from a variety of dog-owner participants approximating those who might ultimately use the DPQ. To help facilitate reaching a wide variety of dog owners, and to take advantage of other benefits afforded by web-based data collection (Gosling et al., 2004), participants completed the DPQ and gave feedback online. Their feedback guided revision of the DPQ.

Method

Participants

The first version of the questionnaire was lengthy, with 360 items and background information, so completing it was an arduous task. To respond to each of the items on the questionnaire, participants were required (1) to be highly familiar with a particular dog's behavior, and dog behavior in general, and (2) to be highly engaged in completing the task. Thus, common sources of participants (e.g., undergraduate students) would not be appropriate because they would not have sufficient knowledge of canine behavior and

would not be sufficiently engaged to respond to all the items carefully. The expertise criterion suggested I use experienced dog owners. The engagement criterion suggested I use people who freely volunteered to participate in the research with a full understanding that an arduous task lay ahead.

Volunteers signed up through a webpage. At the site potential volunteers were informed about the reasons for the study and what participation would entail. If an individual wanted to volunteer, he or she entered an email address that was sent to a database. An email message was sent to all volunteers thanking them for their interest in the study and providing additional information about the research. I used four means of publicizing my call for research volunteers. First, I posted a link to information about the study on the website of the Animal Personality Institute, www.animalpersonality.org, a University of Texas-based research group founded for the study of non-human animal personality. Second, I posted messages to online discussion and e-mail groups focused on topics like dog training, dog health, and rare-breed rescue. Third, I posted flyers at Austin-area dog training and boarding facilities. Last, I e-mailed rescue groups focused on rare breeds. In all of these recruitment efforts, I gave permission for people to pass on information about the study. As a result, messages about the study circulated through various other online dog discussion groups and training groups, in addition to various organizations' flyers and newsletters. All potential volunteers contacted me directly to sign up to participate in the study and were notified when questionnaires became available.

When initially volunteering, 70% of volunteers reported how they had learned about the study. Of the 70% who reported this information, 14% learned about the study from reading the Animal Personality Institute website. Sixty-seven percent learned about the study from an online discussion group, including discussion groups focused on rare dog breeds and dog rescue groups. Fifteen percent learned about it from a friend or relative, breeder, or trainer with whom they work, and the remaining four percent learned about the study from a flyer.

A random subsample of 225 participants was selected from a larger list of people who had volunteered to fill out the DPQ. At the end of May, 2005, these participants were contacted via e-mail, asked to fill out the DPQ, and directed to the online questionnaire form. By the deadline two weeks later, 152 volunteers had responded. The data from these 152 participants were consolidated and analyzed.

Materials and Procedures

Utilizing web-based methods

The format of the DPQ – items followed by Likert-scale rating systems – lends itself to both a traditional paper-and-pencil form and an online form. Web-based studies and online self-selected Internet samples are becoming more and more popular as the Internet becomes more widespread and accessible. This method of data collection is associated with various concerns, but also with numerous benefits. Most concerns associated with web-based data collection have been refuted, including concerns about impact of self-selection (Walsh et al., 1992); the correspondence between surveys conducted online with those conducted through more traditional methods (e.g., McCabe et al., 2006; Gosling et al., 2004); whether Internet samples are composed of maladjusted, isolated, and/or depressed individuals (e.g., Gosling et al., 2004); and the integrity of the data (McGraw et al., 2000).

Web-based data collection also affords many advantages over traditional methods of data collection. For example, online data collection tends to be very efficient, because many participants can be reached and can complete the study at once without requiring an experimenter to administer the study. Online data collection also removes the necessity of entering data and, with it, the risk of data entry mistakes. In addition, using the Internet enables researchers to reach people outside of the typical subject pool (e.g., Gosling et al., 2004), including people who are disabled, geographically distant, elderly, or in a specific and rare population.

The DPQ was placed online to take advantage of these general advantages, but also because online data collection afforded benefits specific to administration of the

DPQ and similar questionnaires. First, the goal of creating a questionnaire that is widely applicable requires piloting and evaluating the questionnaire in a diverse population. Respondents (dog owners) should vary in terms of age, sex, occupation, type and extent of experience with dogs, geographic region, etc. The target dogs should also compose a diverse sample; they should vary in breed, sex, age, size, castration status, health, training, geographic location, etc. Second, development of this questionnaire requires a very large sample, and a specialized group of people who are both knowledgeable about dogs and their dog in particular, and interested enough in dog behavior to complete a lengthy questionnaire. Third, an algorithm could be used to randomly select the target dog for a questionnaire when a participant had more than one dog, thus avoiding bias from the participants. Finally, the questionnaire items could be presented in a uniquely random order for each participant.

To ensure data quality, individuals will be asked as a part of the DPQ whether they have completed the form before, repeat responders will be removed from the sample, and recruitment with conducted the goal of collecting data from a diverse population of people and dogs.

Administration of the DPQ

The online questionnaire consisted of four parts. Part 1 asked participants to provide the basic background information that allowed me to assess the diversity of my sample. These questions included where the person lives, the capacity in which the person interacts with dogs, how many dogs the person has had in his or her lifetime, and how many dogs the person currently owns. After the person indicated how many dogs he or she had, the questionnaire asked for the names of up to five of those dogs, then randomly selected one dog as the target dog for the questionnaire.

Part 2 of the questionnaire asked background questions specific to the target dog, allowing me to assess the diversity of the dogs in the study. For example, these questions included the dog's breed, sex, age, weight, whether the dog was castrated and, if so, at what age. Further questions included whether the dog performed any specific jobs (e.g.,

guide dog), whether the dog had any illnesses or disabilities (e.g., blind, deaf, arthritic), and whether the dog had ever bitten a person or a dog, or been bitten by another dog.

Part 3 of the questionnaire included the 360 questionnaire items, plus one accidental repetition of the item "Dog is anxious," for a total of 361 questionnaire items. Fifty questions appeared on each page. The answer options for each question appeared below that question. The options were a seven-point Likert scale ranging from "strongly disagree" to "strongly agree," plus "not applicable." Participants were instructed to indicate an item was "not applicable" if their dogs had never been in the situation described (e.g., had never been around children). In addition, there was an option for participants to tick a box indicating when an item was "difficult to interpret". The 361 questionnaire items were in random order for each participant such that no two participants' questions appeared in the same order. I intentionally grouped some questions, however, so that they would always appear together. Items were grouped to enable participants to compare and contrast the questions. For example, "Dog exhibits submissive behaviors (e.g., rolls over, avoids eye contact, yawns, licks lips) when greeting **familiar** people" and "Dog exhibits submissive behaviors (e.g., rolls over, avoids eye contact, yawns, licks lips) when greeting **unfamiliar** people" always appeared consecutively, in this order.

After participants completed Part 3 of the questionnaire, I thanked them for their participation, gave them the opportunity to sign up to receive updates about the study, and gave them the opportunity either to be finished with the questionnaire or to fill out further background information and/or give feedback on the questionnaire. If participants volunteered to fill out more background information about their dogs, they were forwarded to Part 4 of the questionnaire. Part 4 asked participants about their dogs' diets, amount of time their dogs spend alone or with people, the type of veterinary care they use (e.g., traditional, holistic), and gave participants the opportunity to give any feedback they felt was important.

Results

The main goals in Study 1 were to assess whether participants had trouble understanding any of the items and to incorporate other participant feedback to improve the questionnaire. I therefore analyzed participants' responses to the "difficult to interpret," "not applicable," and free-response items. I did not analyze their responses to the 360 Likert scales.

Difficult to interpret

Fifty-one participants never used the "difficult to interpret" option, indicating that 101 participants used the option at least once. Sixty-four participants used the "difficult to interpret" response 1-5 times, 21 used it 6-10 times, 11 used it 11-15 times, 3 used it 16-20 times, and 2 used it more than 20 times. One subject used the "difficult to interpret" response more than 36 times; only this last participant was removed from the dataset. The following analyses are based on 151 participants' feedback.

Not applicable

If a participant's dog had no experience like that described in an item, then participant was able to indicate that the item was not applicable to the dog by clicking "not applicable" in lieu of rating the dog on that item. For example, if the dog had never come into contact with livestock, then the participant could indicate the item "Likes to chase livestock" did not apply. Inclusion of this response option made certain that I was not forcing participants to choose a response. I was also able to make certain participants made use of this response. A total of only 16 of the 151 participants never used the "not applicable" response; the remaining 135 participants used the response at least once. Fifty-nine participants used the "not applicable" response 1-5 times, 25 used it 6-10 times, 24 used it 11-15 times, 12 used it 16-20 times, seven used it 21-25 times, and eight used it more than 25 times.

Free-response suggestions

The free-response portion of the questionnaire gave participants the opportunity to give feedback, share comments, provide further background information, and even tell

stories about their dogs. The vast majority (more than 90%) of participants completed Part 4 of the DPQ and provided free responses, and more than 75% indicated that they enjoyed filling out the questionnaire. Many participants also indicated they believed they had gained a greater understanding of their dogs' behavior from the extensive and detailed consideration required in responding to the questionnaire items. Few participants (four) complained about the length of the questionnaire, and of those two indicated that they felt the length was necessary.

Participants also gave specific feedback pertaining to questionnaire items. A number of participants said that, for specific types of behavior (assertiveness, submission), they had more trouble understanding items that did not include examples. Participants also indicated that they were unsure how to rate their dogs on items assessing common types of behavior (e.g., friendliness, aggression) if they thought their dog fit a description (e.g., friendly to unfamiliar people) but the dog did not display any or all of the example behaviors (e.g., tail wagging). Many participants also indicated that descriptors in items (e.g., "fearful" behavior, "aggressive" behavior) should be emphasized typographically to make then stand out. Some participants also suggested additional items that could be added to the questionnaire, including, "Dog is attuned to owner's routine (e.g., predicts homecoming)," "Dog appears aware of owner's emotions (e.g., comes when owner is sad, crying)," and "Dog feeds off owner's emotions (e.g., nervous if owner is nervous)".

Sample demographics

Demographic information collected from the 151 participants retained for analyses in Study 1 is presented in Table 4.1. Column 1 indicates the type of information reported in the corresponding rows. Study 1 included both men and women, though the majority of participants (89.4%) were women. The mean age of participants was 44 years, and the standard deviation of their ages was 12 years. The majority (84.8%) of participants were from the United States (US). Although 43 participants indicated they currently lived in Texas, a total of 31 of the 50 states were represented by at least one

participant. All participants indicated that they were native English speakers. Just over 14% (22) of the 151 participants reported having dog-related careers, which would likely entail some degree of expertise and increased familiarity with dogs and dog behavior. The mean number of dogs each participant had owned over his or her lifetime, previous to his or her current dog and including childhood family pets, was 4.4, but was variable (s.d. = 3.4). The mean number of dogs currently owned and living in-residence with the participants was 2.2 (s.d. = 1.2). On the basis of demographic information, participants in Study 1 were comparable to those in Study 2 (run later); Study 1 participants were seen as a representative sub-sample.

Table 4.1. Demographic information about participants in Study 1

		Study 1
Sex (dog owner)	Male	16 (10.6%)
	Female	135 (89.4%)
Age (dog owner)	Mean	44
	standard deviation	12
	Age not reported	2 (1.3%)
Country of residence	U.S.	128 (84.8%)
	Canada	16 (10.6%)
	Australia	2 (1.3%)
	U.K.	2 (1.3%)
	Other (Native language is English)	2 (1.3%)
	Country of residence not reported	1 (.7%)
Dog-related careers	Breeder and/or Exhibitor	9 (5.9%)
	Trainer and/or Behaviorist	8 (5.2%)
	Dog rescue worker/volunteer	3 (2%)
	Veterinarian	1 (.7%)
	Groomer	1 (.7%)
	Kennel/Daycare owner or employee	0
	Assistance dog partner	0
		Mean (s.d.)
Number of dogs owned	Past (mean; s.d.)	4.4 (3.4)
	Number not reported	0
	Currently (mean; s.d.)	2.2 (1.2)
	Number not reported	0
Total number of participants		**151**

It was also important to confirm that a diverse sample of dogs was assessed in Study 1. I examined the composition of the sample in terms of various characteristics, including age, sex, castration, and breed, among others. In Study 1, the dogs' average age was 5.1 (s.d. = 3.5) with two dogs' ages unreported and one reported as unknown. Other demographic information about the dogs is presented in Table 4.2. The characteristic examined is listed in bold in the first column (e.g., castration status), subcategories of the characteristic (e.g., male, female) are listed in the second and, as necessary, third columns (e.g., spayed, intact, not reported). The number of dogs (out of 151) identified as fitting each category is listed in the column titled "Number of dogs". The sample of 151 dogs was relatively diverse. Approximately half (51.7%) of the dogs were male. The majority (119, or 78.8%) are castrated, but intact animals are also represented. Both purebred and mixed-breed dogs are included in the sample, with 41 different breeds included in the purebred portion of the sample. Labrador Retrievers are strongly represented; 20 are including, making up approximately 13% of the sample. This large number of Labrador Retrievers is representative of both the population in general and of dog personality research (see Chapter 2). I also examined whether the dogs had any disabilities or health issues, whether the dogs had bitten a person, what sort of role they played in their owners' lives (e.g., pet, guide dog), and what dog sports they were in. The collective demographic information presented in Table 4.2 suggests the 151-dog sample included a wide variety of dogs who differed along many dimensions. It is likely that these dogs also differ in terms of their personalities. For example, some of the dogs serve as guide dogs or certified Animal Assistance Therapy dogs and are thus likely to be friendly towards people and other animals and show stable behavior patterns, but other dogs are guard dogs or compete in dog sports that require a willingness to bite people (e.g., Schutzhund).

Table 4.2. Demographic information about target dogs in Study 1

			Number of dogs
Sex	Male		78 (51.7%)
	Female		73 (48.3%)
	Unreported		0
Castration Status	Males	Neutered	61 (40.4%)
		Intact	17 (11.3%)
		Unreported	0
	Females	Spayed	58 (38.4%)
		Intact	15 (9.9%)
		Unreported	0
Purebred	Unknown		7 (4.6%)
	No (mixed)		50 (33.1%)
	Yes		94 (62.9%)
	Unreported		0
Bitten a Person	Unknown		4 (2.6%)
	No		130 (86.1%)
	Yes		15 (9.9%)
	Unreported		2 (1.3%)
Disability and Health issues	Unilaterally deaf		0
	Bilaterally deaf		1 (.7%)
	Blind in one eye		1 (.7%)
	Blind in both eyes		1 (.7%)
	Arthritic		15 (9.9%)
	Hip dysplasia		7 (4.6%)
	Elbow dysplasia		2 (1.3%)
	Other disabilities		34 (22.5%)
Job or Role	**Pet/Companion**		144 (95.3%)
	Assistance dog		3 (2.0%)
	Guide dog		0
	Hearing ear dog		1 (.7%)
	Medical assistance dog		2 (1.3%)
	Search and rescue		1 (.7%)
	Guard dog		12 (7.9%)
	Animal Assisted Therapy		21 (13.9%)
	Dam or sire for breeding		8 (5.3%)

Table 4.2 (*Continued*)

		Number of dogs
Sport	Obedience	21 (13.9%)
	Sledding	0
	Carting	6 (3.9%)
	Frisbee	5 (3.3%)
	Earth dog	4 (2.6%)
	Show/Conformation	12 (7.9%)
	Schutzhund	2 (1.3%)
	Agility	29 (19.2%)
	Herding (competitive)	4 (2.6%)
	Flyball	2 (1.3%)
	Hunting	8 (5.3%)
Total number of dogs		**151**

Note. The sum of the number of dogs who participate in sports will not equal the total number of dogs in the study; some dogs are not in sports, and others are in multiple. Similarly, participants may consider a dog to be both a pet/companion and to have another job or role (e.g., assistance dog, search and rescue dog).

Item Re-examination and Revision

Difficult to interpret

Because one of my goals was to design a questionnaire that participants find easy to understand, I examined the items participants indicated were difficult to interpret. To be assured that I was not using feedback from participants who over-used the "difficult to interpret" option, I looked to see if there were any participants who indicated a particularly high number of items as difficult to interpret. I decided to remove any participants who used this option on more than 10%, or 36, of the 360 items. I then examined items that a large percentage of participants indicated were difficult to interpret. I set a high threshold for items to be included in the questionnaire without undergoing this scrutiny, requiring all items marked as difficult to interpret by 5% or more of participants to be re-examined. Only if 95% (or more) of participants did not indicate an item was difficult to interpret did I presume the item was readily understood. At this threshold, 25 of the 360 items had to be re-examined. Six researchers independently examined the difficult-to-interpret items, generated possible re-wordings

of each, and then consensually reworded the items. When rewording items, the objective was to use more precise and simpler language, and also to provide example behaviors. For example, 39 participants indicated that "Dog is reacting to coming into physical contact with objects" was difficult to interpret. I therefore reworded the item to "Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a doorframe)".

Of the 25 items included for revision, seven (28%) were from a pool of items modeled after the human Conscientiousness factor, drawn from the BFI (e.g., John, 1990). Of all of the items in the original 1,284-item DPQ item pool, these seven were the only ones modeled after items used to assess Conscientiousness in human personality research, or, indeed, posited to assess Conscientiousness. The items proved quite difficult to reword, in part because providing example behaviors to illustrate them was challenging. For example, "Dog does things efficiently" was reworded to "Dog achieves tasks (e.g., fetches objects) quickly and easily." However, the finding that it was difficult for participants to interpret traditional Conscientiousness-related items with respect to their dogs is consistent with Gosling and John's (1999) suggestion that Conscientiousness does not appear as a separate personality dimension in species other than humans, chimpanzees, and possibly other closely related apes.

Free-response suggestions

A similar method was used to examine free-response feedback given by the participants. First, I compiled all of the free-response feedback, separate from all other data provided by each participant. Next, four researchers examined each free response independently. These researchers aimed to extract information on how the participant felt the questionnaire could be improved, how the participant felt about the questionnaire, whether anything was missing from the questionnaire, and so on. After completing this step independently, the four researchers collaborated to generate a list of the suggestions and feedback generated from these free responses. Two dog personality researchers reviewed the list and decided how to modify the questionnaire items. For example, many

participants suggested adding items pertaining to how aware dogs are of their owners and how much their behavior is affected by their owners. To resolve the issue of whether this topic was adequately covered in the questionnaire, I tried to make certain that questions about how socially attentive dogs were to their owners (e.g., "When in the home, dog follows owner/family member from room to room," "Dog is easily influenced by owner without being given direct commands or cues," "Dog is attentive to actions and words of its owner") were included and that each was clear (i.e., not marked as "difficult to interpret," or carefully revised if they had been).

INTERMEDIATE QUESTIONNAIRE #1

At the end of Study 1, the questionnaire retained 360 items, 25 of which had been revised for clarity. In addition, instructions that dogs need not display any or all example behaviors to be rated high on an item were emphasized. Items that had been marked as "not applicable" and the "not applicable" response option were retained for further examination in the next study.

SUMMARY OF FINDINGS

In Study 1, 151 participants rated their dogs on a 360-item questionnaire and gave feedback on how easily interpreted and generally applicable the 360 items were. Because the goal of this study was to make certain that items were readily understood by lay dog owners, items that were marked as difficult to understand by 5% or more of the participants were re-examined. At this threshold, 25 items were reworded to be more precise and to have simpler phrasing.

CHAPTER 5

Study 2: Factor Structure Identification

INTRODUCTION

The goals of Study 2 were to determine the number of factors underlying the behaviors and characteristics assessed in the DPQ, to identify the facets of each trait, to begin to reduce the number of items in the questionnaire from 360 to a more practical length and to assess the stability of the factor solution.

To determine the factors in the original questionnaire, I used EFA (PCA, varimax rotation) on data collected from a new, larger sample. I then selected the number of factors based on the convergence of four criteria. I then sought to assess the repeatability of the factor solution and to evaluation the adequacy of the solution to the Study 2 data by performing an EFA on one half of the data with 353 items, then fitting the derived model to the second half of the data. The model was fit using SEM to perform CFA and was evaluated with a number of fit indices.

Once the factor model was confirmed, I sought to determine the number of facets that composed each factor. I used EFA (PCA, promax rotation) on the whole dataset, separately analyzing groups of items that loaded on each factor. I then selected the number of facets in each factor based on three criteria.

Finally, to achieve the goal to have a shorter and more manageable questionnaire, I evaluated each item in terms of its contribution to the content validity of the scale and the questionnaire, its loadings on factor and facet scales, and its contribution to the internal consistency the scale onto which it loaded. A total of 102 items were retained for further evaluation in Study 3.

To assess the replicability of the factor solution within Study 2 with the newly trimmed item list, I divided the sample in half for a new EFA and then CFA. With the first half of the data, I performed an EFA and selected the number of factors based on the

three criteria used in the previous EFA of the same data. I performed a CFA using SEM to fit the five-factor model from the first half of the data to the second half of the data.

METHOD

Design

My intentions to use EFA in Study 2 dictated requirements of the sample composition and size. Specifically, a large and diverse sample was needed. There are two major risks to having a homogeneous sample or few participants: (1) the sample may not be representative of the intended population, and (2) chance can substantially influence correlations among items when the ratio of participants to items is low, creating unstable patterns of covariation and leading to factor structures that do not generalize to new samples (DeVellis, 2003; Worthington & Whittaker, 2006).

Gorsuch (1997) addressed sample composition, claiming that it is not necessary for the sample to closely represent the intended population, if individuals who would score high and low on the scale are well represented. If participants share certain characteristics (e.g., sex, age, race), even very large samples may not control for the systematic variance produced by these characteristics. Thus, in the current study it was important to recruit a heterogeneous sample, and also to verify that the sample was diverse before conducting EFA.

There are many recommendations for determining sample size requirements based on study design, most of which err on the side of recommending larger sample sizes. Large sample sizes are needed in questionnaire development research so that variance that is introduced by specific participants might be cancelled out (i.e., by random effects that tend to occur in large samples; e.g., Tabachnick & Fidell, 2001). Recommendations have been based on sample size alone (e.g., an excellent sample comprises 1,000 participants; Comrey, 1973), on the ratio of participant number to items (e.g., ratios of 5-10 participants per item are adequate, Gorsuch, 1983; ratios of fewer than 3 participants per item are inadequate, Reise et al., 2000), and item communalities and ratios of items to factors (e.g., Guadagnoli & Velicer, 1988; MacCallum et al., 1999).

In Study 2, I elected to be conservative by using a large sample size, seeking a ratio of at least 10 participants per item for the first EFA (performed on the whole sample), or a total of at least 3,600 participants in the whole sample.

Participants

A new group of participants, drawn from the same population as those who took part in Study 1 (but not including any participants who had taken part in Study 1), filled out the questionnaire online for Study 2. These volunteer participants responded to the online questionnaire between August, 2005, and January, 2006. A total of 4,105 cases were compiled; initial analyses (described below) refined the dataset to 3,737 participants.

Materials and Procedures

The procedure for Study 2 was identical to that for Study 1, and using web-based data collection afforded the same advantages. The questionnaire for Study 2 had the same four parts as for Study 1, modified as per the feedback from participants from Study 1. As in Study 1, the questionnaire in Study 2 included the accidental repetition of the item "Dog is anxious" for a total of 361 items.

RESULTS

It was necessary to refine the data sample and examine it for errors prior to analysis. First, I removed duplicated cases in which participants had filled out the questionnaire more than once (either about the same dog or about different dogs). In the case that a single participant filled out the questionnaire about two different dogs, I retained the questionnaire information that participant had completed first. This process reduced the dataset from 4,105 to 3,830 participants.

Second, I examined the frequency of responses for each questionnaire item to check for possible errors in saving the data (i.e., items for which there was only one response saved from all participants). No anomalies were found in the dataset, though

responses to each item were not regularly distributed (i.e., responses to some items were positively or negatively skewed).

Third, I examined the number of participants who had left items blank. I removed 79 of the 3,830 participants because they had failed to respond to 10 or more (approximately 3%) of the items; they had neither rated their dog on these items, nor indicated the item was "not applicable". The resulting dataset contained 3,751 unique cases.

Fourth, I examined participants' use of the "not applicable" response to items. I examined use of this response by both items and by participants. I removed 14 of the 3,751 participants from the dataset because they indicated that more than 20% of the items were not applicable. Because of the frequency of this response to items about aggression and other socially undesirable behaviors, I think it is possible some of these participants misunderstood my directions and used the "not applicable" response in place of "strongly disagree" when they believed an item did not describe their dogs. Removal of these datasets resulted in retaining the data from 3,737 participants.

Sample demographics

Demographic information for the 3,737 participants in Study 2 is presented in Table 5.1. As in Study 1, both male and female dog owners were represented, but men made up the minority (15.9%, or 595 men). The average reported age of participants was 43 years, with a standard deviation of 12 years; 29 participants did not report their ages. The majority of participants were from the U.S., though Canada, Australia, and the United Kingdom (U.K.) were also represented. Every U.S. state was represented in the sample, though there were only two participants from North Dakota and three participants from Delaware. Likely due to the increased attention the questionnaire received after a related story in *The Des Moines Register*, a large number of participants (663) were from Iowa. Large numbers of participants also came from New York (435) and from Texas (371). In addition, 51 participants reported that they were native English speakers but currently living outside those four countries. Slightly less than 10% (N =

370) of the 3,737 participants reported having dog-related careers (e.g., 138 were dog trainers or behaviorists, 38 were veterinarians). The mean number of dogs each participant had owned in his or her lifetime, previous to the current dog or dogs but including childhood family dogs, was 5.3, though the number was highly variable (s.d. = 3.5). On average, the participants each currently owned 2.3 dogs (s.d. = 1.3).

Table 5.1. Demographic information about participants in Study 2

		Study 2
Sex (dog owner)	Male	595 (15.9%)
	Female	3,142 (84.1%)
Age (dog owner)	Mean	43
	standard deviation	12
	Age not reported	29 (.8%)
Country of residence	U.S.	3,447 (92.2%)
	Canada	150 (4%)
	Australia	39 (1%)
	U.K.	46 (1.2%)
	Other (Native language is English)	51 (1.4%)
	Country of residence not reported	4 (.1%)
Dog-related careers	Breeder and/or Exhibitor	139 (3.7%)
	Trainer and/or Behaviorist	138 (3.7%)
	Dog rescue worker/volunteer	27 (.7%)
	Veterinarian	38 (1%)
	Groomer	10 (.3%)
	Kennel/Daycare owner or employee	14 (.4%)
	Assistance dog partner	4 (.1%)
		Mean (s.d.)
Number of dogs owned	Past (mean; s.d.)	5.3 (3.5)
	Number not reported	0
	Currently	2.3 (1.3)
	Number not reported	0
Total number of participants		**3,737**

As in Study 1, it was also important to confirm that a diverse sample of target dogs was assessed in Study 2. I examined the composition of the sample in terms of the same characteristics that I examined in Study 1. In Study 2, the dogs' average age was slightly older: 5.9 (s.d. = 3.6) with all but four dogs' ages reported (those dogs' ages were reported as "unknown"). Other demographic information about the dogs is presented in Table 5.2. The number of dogs (out of the 3,737 retained in Study 2) identified as fitting each category is listed in the column titled "Number of dogs". The sample of 3,737 dogs appears to be relatively diverse. Approximately half (50.8%) of the dogs are male. The majority (82.0%, or 3,049 of the 3,717 whose castration status was reported) were castrated, but intact animals were also represented (N = 656). Purebred dogs (N = 2,703), mixed-breed dogs (N = 883), and dogs whose breeding was unknown (N = 94) were included in the sample, with 182 breeds represented in the purebred portion of the sample. As in Study 1, Labrador Retrievers are the breed represented in the largest number; 195 Labradors are included, composing approximately 5% of the sample.

I also looked at whether the dogs had any disabilities or health issues, whether the dogs had bitten a person, what sort of role they played in their owners' lives (e.g., pet, guide dog), and what dog sports they were in. No cumulative total is presented for the number of dogs who participate in sports, because many of these dogs participate in a more than one sport. A large number of dogs serve as Animal Assisted Therapy dogs (N = 506), or compete in agility (N = 624), formal obedience (N = 524), and/or in American Kennel Club conformation (N = 520), reflecting these activities' and competitions' current popularity, especially among dog enthusiasts. The demographic information about the dogs in Study 2 suggested that they compose a diverse sample, satisfactory for EFA.

Table 5.2. Demographic information about target dogs in Study 2

			Number of dogs
Sex	Male		1,897 (50.8%)
	Female		1,828 (48.9%)
	Unreported		12 (.3%)
Castration Status	Males	Neutered	1,469 (39.3%)
		Intact	421 (11.3%)
		Unreported	7 (.2%)
	Females	Spayed	1,580 (42.3%)
		Intact	235 (6.3%)
		Unreported	13 (.3%)
Purebred	Unknown		94 (2.5%)
	No (mixed)		883 (23.6%)
	Yes		2,703 (72.3%)
	Unreported		57 (1.5%)
Bitten a Person	Unknown		123 (3.3%)
	No		3,152 (84.3%)
	Yes		447 (12.0%)
	Unreported		15 (.4%)
Disability and Health issues	Unilaterally deaf		6 (.2%)
	Bilaterally deaf		51 (1.4%)
	Blind in one eye		21 (.6%)
	Blind in both eyes		24 (.6%)
	Arthritic		306 (8.2%)
	Hip dysplasia		201 (5.4%)
	Elbow dysplasia		56 (1.5%)
	Other disabilities		728 (19.5%)
Job or Role	**Pet/Companion**		3,648 (97.6%)
	Assistance dog		66 (1.8%)
	Guide dog		19 (.5%)
	Hearing ear dog		19 (.5%)
	Medical assistance dog		28 (.7%)
	Search and rescue		31 (.8%)
	Guard dog		413 (11.1%)
	Animal Assisted Therapy		506 (13.5%)
	Dam or sire for breeding		301 (8.1%)

Table 5.2 (*Continued*)

		Number of dogs
Sport	Obedience	524 (14.0%)
	Sledding	16 (.4%)
	Carting	56 (1.5%)
	Frisbee	127 (3.4%)
	Earth dog	55 (1.5%)
	Show/Conformation	520 (13.9%)
	Schutzhund	36 (1.0%)
	Agility	624 (16.7%)
	Herding (competitive)	210 (5.6%)
	Flyball	120 (3.2%)
	Hunting	195 (5.2%)
Total number of dogs		**3,737**

Note. The sum of the number of dogs who participate in sports will not equal the total number of dogs in the study; some dogs are not in sports, and others are in multiple. Similarly, participants may consider a dog to be both a pet/companion and to have another job or role (e.g., assistance dog, search and rescue dog).

Item refinement

I examined the items in terms of the percentage of the retained 3,737 participants who indicated each item was "not applicable". Because I have the goal of designing a questionnaire that is broadly applicable with items that are descriptive for the majority of dogs, I opted to remove items that were indicated as not applicable by 20% or more of the 3,737 participants. This resulted in the removal of eight items, leaving 352 items (or 353 items if the two appearances of "Dog is anxious" are counted as two items). Items that were removed are listed in Table 5.3.

Table 5.3 Items removed from the 360-item pool due to narrow applicability

Item text	Participants who marked item as "not applicable"
Dog likes to chase livestock.	1,736 (46.5%)
Dog exhibits less aggression towards objects or situations after repeated exposure to them.	1,060 (28.4%)
Dog is quick to calm down after showing aggression.	1,029 (27.5%)
Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when an unfamiliar person threatens a family member.	947 (25.3%)
Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards large animals (e.g., horses, cattle).	916 (24.5%)
Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards large animals (e.g., horses, cattle).	865 (23.2%)
Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when cornered by an unfamiliar child.	798 (21.4%)
When off leash and away from home, dog barks at bicycles, children running, or joggers.	778 (20.8%)

ANALYSES

Part 1: Exploratory Factor Analysis

In the first phase of Study 2, my goal was to assess and identify the major dimensions underlying the personality ratings through EFA (PCA, varimax rotation) on the participants' ratings of their dogs. My analyses were exploratory so I made no specific predictions concerning the number of factors that would emerge. However, I had no reason to believe that one primary (or principal) factor would account for the majority of the variance in the data, so varimax rotation was the most appropriate choice (Hetzel, 1996). The primary reason for using EFA is that it allows items to be related to any of the factors underlying participants' responses. It also helps in the process of item reduction because it enables the researcher to identify items that do not measure an intended factor or that measure more than one factor at once (i.e., items that are not univocal). These traits may be poor indicators of the construct being measured and can be eliminated from the questionnaire in the process of shortening it.

As noted above, as a part of my selection procedures I included participants who had left questionnaire items blank. However, when items that were left blank and items

that were marked as "not applicable" were both treated as missing data, most participants had at least one missing data point. Indeed, only 342 cases were entirely complete. The mean number of missing data points per participant when items left blank and items marked as "not applicable" were combined was 8.6 (2.3%), with 72.5% of participants having 10 or fewer missing data points. In the cases where data were missing, I used mean substitution.

Before proceeding with the EFA, I needed to assess the sampling adequacy to make certain that the data were amenable to EFA. Having an adequate sample size addresses one aspect of this issue, but another is the magnitude of correlations between variables, as displayed in the correlation matrix. Bartlett's (1950) test of sphericity is a very common test used for this purpose; it estimates the probability that the correlations in the matrix are 0. Unfortunately, Bartlett's test of sphericity is largely dependent on sample size such that it is likely to be significant for large samples even when correlations in the matrix are small (Tabachnick & Fidell, 2001). To avoid this problem, Tabachnick and Fidell (2001) recommend using Bartlett's test of sphericity only if the ratio of participants to items is less than 5:1. In Study 2, the ratio of participants to items is greater than 10:1, so Bartlett's test of sphericity is likely to be inflated and inappropriate. Indeed Bartlett's test was significant (approximate $\chi^2 = 522368$, $df = 62128$, $p < .001$). Instead, I relied on the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. The KMO indicates the extent to which a correlation matrix contains factors vs. chance correlations by looking at the relationship between partial correlations and the sum of squared correlations. If two variables share a common underlying factor with other variables, the two variables' partial correlation with be small. In order for a sample to be adequate for factor analysis, the KMO value (which ranges from 0 to 1) must be high. If factor analysis is conducted on a correlation matrix with a high KMO, the factors extracted will account for a large amount of the variance in among the variables. But what does "high" mean in this context? Conservative cut-offs suggest at least .6 (e.g., Tabachnick & Fidell, 2001). The KMO associated with the correlation matrix for the data in Study 2 was .95, indicating that the matrix was good for EFA.

There are many rules for determining the number of underlying factors in a dataset. Many of these rules lead to severe over-extraction (i.e., extracting too many factors), especially in datasets with large numbers of items. This tendency for over-extraction is partially due to the fact that many of the extraction rules were developed for analyses based on much smaller numbers of items, before computers facilitated analyses with large samples and large numbers of items (Zwick & Velicer, 1982). To illustrate the severity of the over-extraction problem, consider what happens when the popular Kaiser eigenvalue rule is applied in the present data. The Kaiser rule, in which components with eigenvalues greater than one are retained (Kaiser, 1960; Velicer & Jackson, 1990), suggests retaining 72 factors. Extracting 72 factors results in 12 components consisting of just one item each and numerous very small factors that are highly related to one another. For example, one three-item factor includes "Dog's behavior is predictable," "Dog is highly predictable," and "Dog behaves consistently across different situations" but excludes the closely related item "Dog's behavior varies from situation to situation (e.g., dog is quiet when others are quiet but more excited when invited to play)," which forms its own single-item factor.

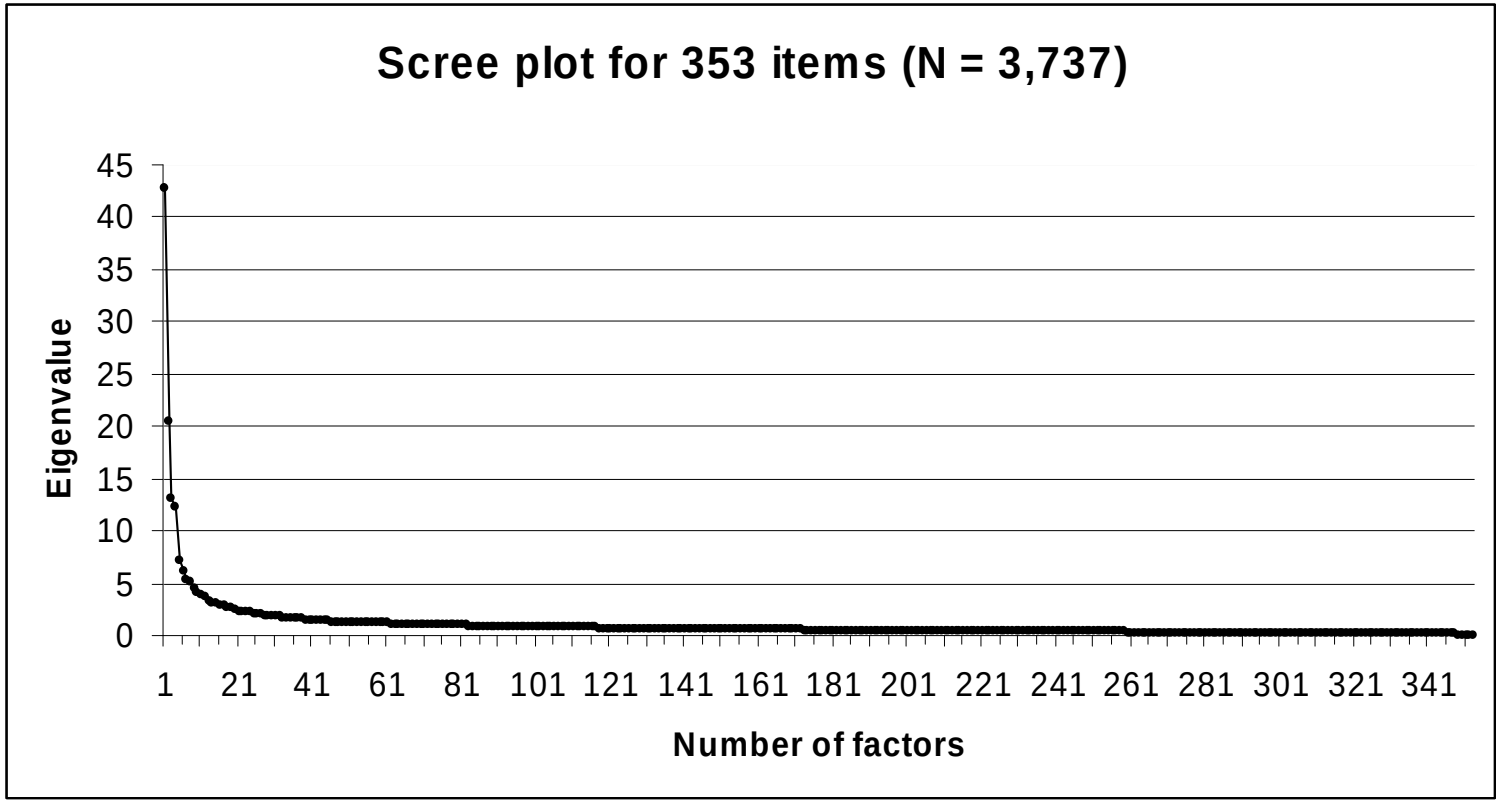
Similarly, far too many factors would be retained using parallel analyses of Monte Carlo simulations (Horn, 1965), another common method for determining the number of factors to retain (e.g., Lance et al., 2006). Parallel analyses of Monte Carlo simulations provide a comparison standard in terms of the eigenvalues that would be expected were the data purely random with no underlying structure. In using this method, only those factors that have eigenvalues greater than their randomly derived counterparts are retained. In the present data, the parallel analyses of one-hundred independent Monte Carlo simulations suggested retaining 42 factors. Clearly, such rules are not appropriate for meeting my goal of identifying the major dimensions underlying dog personality.

I therefore sought convergence across a number of other methods better suited to the large number of participants and items in my sample: A graphical scree test (Cattell, 1966), factor replicability across items and samples, a top-down method in which correlations between orthogonal factor scores from different factor solutions are viewed

as path coefficients in a hierarchical structure (Goldberg, 2006), and interpretability of the solutions (see Zwick & Velicer, 1986; Worthington & Whittaker, 2006). Each method and associated conclusions are elaborated below.

The scree test is based on a graphic plot of the eigenvalues of the factors; only those factors above a noticeable break or "elbow" in the line joining the eigenvalues, after which the values tend to level off horizontally, are retained. Interpretation of the scree plot suggested retaining four or five factors (see Figure 5.1).

Figure 5.1. Scree plot of eigenvalues for 353 items (N = 3,737)



A valid factor solution should reflect an underlying factor structure that is not contingent on the particular items tapping each factor. Thus, as long as the item pool is sufficiently broad to capture the full range of relevant behaviors, the same factor structure should be obtained regardless of the specific items used. Therefore, to test the replicability of the four and five-factor solutions I divided the item pool into non-overlapping item sets, repeated the factor analysis, and examined the convergence across solutions. Specifically, I broke the pool of questionnaire items into three randomly selected subsets of 117, 117, and 118 items, and into four randomly selected subsets of 88 items each. (One occurrence of the item "Dog is anxious" was randomly selected to be included in the questionnaire item pool for these analyses, yielding a total of 352 items divided into the three and four subsets.) I performed PCA on each subset and saved each participant's factor scores resulting in 63 factor scores per participant (i.e., 28 scores for the four-factor solutions [seven four-factor solutions] and 35 scores for the five-factor solution [seven five-factor solutions]). To determine the degree to which the major factors replicated across the different item subsets, I examined the correlations among individuals' factor scores. If the factors replicated, then a pattern of strong convergent correlations (between the factors that were replicated across solutions) and weak discriminant correlations will emerge. As shown in Table 5.4, there was strong evidence that both the four and the five-factor solutions are robust: Across seven replications of the four-factor solution, the mean convergent correlation was .909 (based on 28 individual correlations), much stronger than the mean discriminant correlation of .095 (based on 84 individual correlations).

Across seven replications of the five-factor solution, the mean convergent correlation was .849 (based on 35 individual correlations), much stronger than the mean discriminant correlation of .127 (based on 140 individual correlations). It should be noted, however, that these convergent and discriminant correlations are inflated because each item in the data subsets also appears in the whole dataset.

Table 5.4. Convergent and discriminant correlations for the 4- and the 5-factor solutions

			Data divided in thirds			Data divided in quarters			
			1	**2**	**3**	**1**	**2**	**3**	**4**
Whole dataset (all items)	**Factor**	**1**	.951	.957	.948	.946	.921	.908	.958
		2	.947	.960	.934	.953	.915	.924	.946
		3	.923	.758	.884	.838	.672	.925	.877
		4	.934	.675	.89	.856	.667	.913	.897
	mean convergent corr.		**.940**	**.889**	**.918**	**.911**	**.832**	**.918**	**.927**
	mean discriminant corr. [a]		**.067**	**.123**	**.083**	**.090**	**.175**	**.050**	**.077**
	Factor	**1**	.924	.930	.943	.953	.907	.901	.957
		2	.839	.761	.811	.896	.75	.829	.885
		3	.933	.738	.756	.792	.798	.827	.930
		4	.893	.727	.741	.765	.867	.717	.919
		5	.548	.824	.869	.847	.564	.692	.774
	mean convergent corr.		**.865**	**.814**	**.843**	**.869**	**.803**	**.807**	**.908**
	mean discriminant corr. [b]		**.123**	**.171**	**.121**	**.090**	**.174**	**.152**	**.054**

Note. All values are absolute values. Fisher's r-to-z transformation was used to convert all original correlations to z scores before averaging; means reported here have been converted back to r.

[a] Each of these mean discriminant correlations is the average of 12 correlations (not shown).

[b] Each of these mean discriminant correlations is the average of 20 correlations (not shown).

By this point, the two methods pointed to between four and five major dimensions underlying the personality ratings. But was it four or five? How could this ambiguity be resolved? The third and fourth methods for examining the factor structure illuminated the reason for the lack of clarity. Specifically, Goldberg's hierarchical tree analysis, supported by the interpretability of the solutions, showed that there were four basic factors underlying the data, but one of those factors was a super-factor, itself comprising two major, related but separable dimensions.

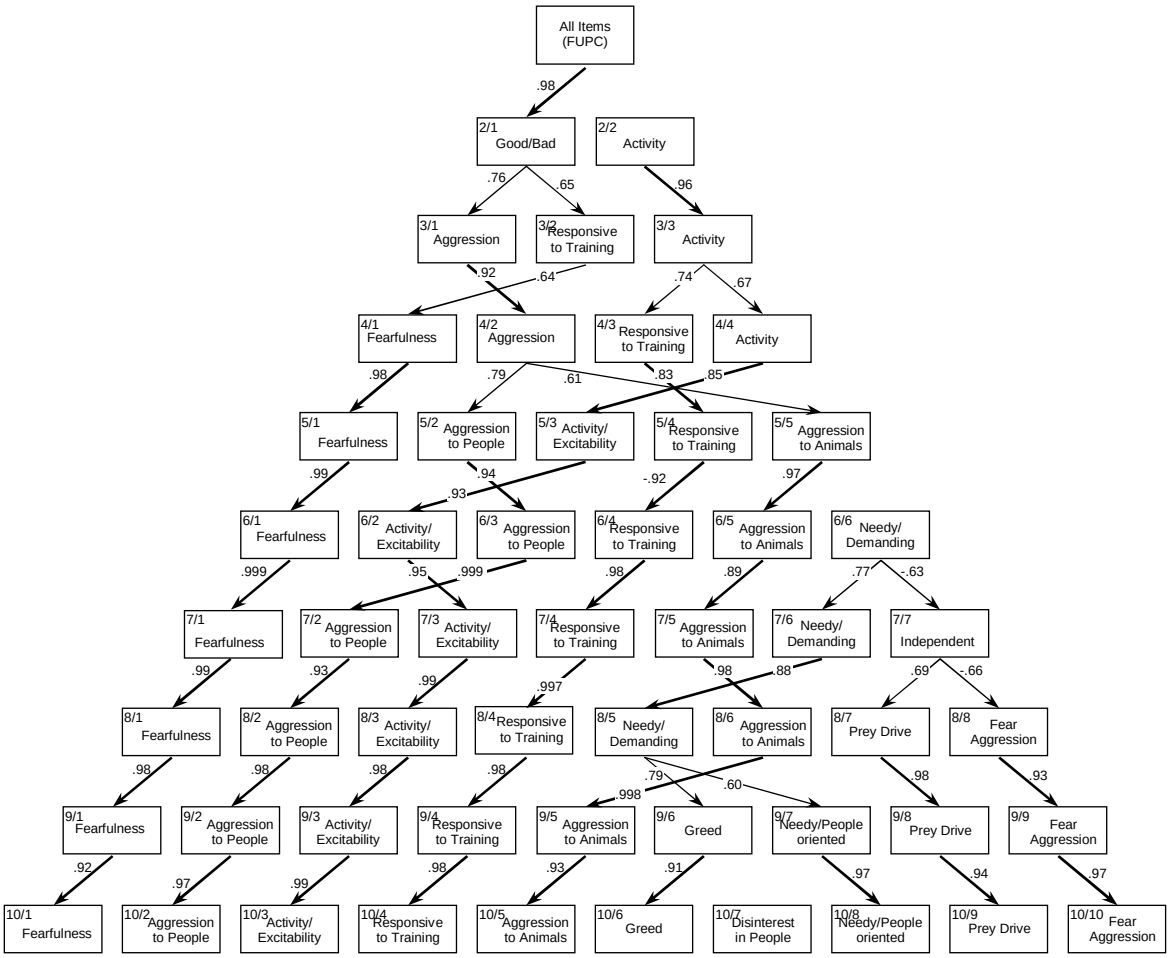
The principle behind Goldberg's hierarchical tree analysis is based on what happens when too few or too many factors are extracted. When too few factors have been extracted, factors that are separate are forced together statistically; as a result, when

additional factors are extracted, major new factors appear that show substantial differences from their parent factors. When too many factors have been extracted, factors that are not separate are forced apart statistically; as a result, when additional factors are extracted, minor new factors appear but the major factors remain essentially unchanged from their parent factors. Thus, Goldberg's technique utilizes a hierarchical tree structure for examining the degree to which the factors in an N-solution (i.e., the parent solution, or a row in the hierarchical tree) statistically resemble the factors in an (N+1)-solution (i.e., the child solution, of the next row down in a hierarchical tree). The ideal solution is found where the major factors stop breaking apart into major factors when a larger number of factors are extracted.

The hierarchical tree is presented in Figure 5.2. The factors in each row are presented in an order corresponding to the amount of variance for which each factor accounts, with descending variance associated with factors as from left to right. For example, in the three-factor solution, factor 3/1 accounts for the most variance, then 3/2, then 3/3. Correlations between factor are presented along the arrows connecting those factors. To ease interpretability, only correlations of .60 or greater were included. As a result, some factors are not shown to be correlated with factors from the previous (higher) level of the tree. To further facilitate interpretability, each factor was assigned a descriptive label, which is presented in the box representing that factor. These labels were generated by having two personality experts examine the factor structures for every solution (i.e., the 2 factor solution, the 3 factor solution) and then examine the relationships between the factor structures as presented in the hierarchical tree. At least one of the dimensions (labeled as Fearfulness) seems similar to a human dimension (Neuroticism); in this case, the label "Fearfulness" was selected to avoid issues of anthropomorphism and because it seemed more clearly descriptive of the items that compose the factor. However, as in most factor-analytic research, Fearfulness and the other labels employed are only broadly descriptive and inevitability fail to capture some of the facets that make up each factor (John, 1990).

The hierarchical tree suggested a five-factor solution (see Figure 5.2). The factors present when fewer (e.g., 3, 4) factors were extracted were not consistent through the various lower or child levels of the tree. For example, the four-factor solution was problematic because the second factor ([4/2] Aggression) subdivided into "Aggression to people" (correlated .79) and "Aggression to dogs" (correlated .61) when five factors were extracted. In contrast, all the factors from the five-factor solution are retained, essentially unchanged, even when six or more factors are extracted. The factors in the five-factor model are correlated with the first five factors in the six-factor solution at .92 (absolute value) or higher. The hierarchical tree thus suggests five factors: Fearfulness (5/1), Aggression towards People (5/2), Activity/Excitability (5/3), Responsiveness to Training (5/4), and Aggression towards Animals (5/5). These five factors are presented in order of descending variance accounted for.

Figure 5.2. Hierarchical structure of factor score correlations in 353-item Study 2 data set



Note. For clarity, only correlations of .60 and greater are included in the hierarchical tree. Correlations of .80 or greater are indicated by boldface arrows.

Finally, I examined the interpretability of the solutions. Conceptual interpretability is a definitive criterion for retaining factors, despite being subjective (e.g., Worthington & Whittaker, 2006). In the end, a factor can only be retained if it is interpretable, regardless of what the other criteria indicate. I extracted two- through ten-factor solutions from each of the seven data subsets and inspected the specific item loadings to make sure the factors were consistent with factors identified in the literature (see Chapter 2) and with three experts' understanding of dog behavior; these inspections were conducted independently of the labeling of factors in hierarchical tree. Assessment of the factor solutions' interpretability confirmed the findings of the other extraction methods (i.e., those using the whole dataset), pointing to four or five factors. In the three-factor solution, items that do not intuitively belong together loaded on the same factor (e.g., "Dog learns readily" and "Dog is timid"). Also, in the three-factor solution there are many items that are not univocal; these items (e.g., "Dog is anxious," "Dog reacts appropriately to various situations") load strongly on more than one factor, suggesting that more factors needed to be extracted. In solutions with six or more factors, some of the factors were difficult to interpret and appeared composed of unrelated items (e.g., "Dog likes to chase squirrels, birds, and other small animals" and "Dog tends to be independent" from the nine-factor solution).

The interpretability exercise was useful in illuminating the reason why both the four- and five-factor solutions were supported across methods and was consistent with the findings of the hierarchical tree shown in Figure 5.2. Specifically, these analyses showed that the structure consists of four basic factors: Fearfulness, Aggression, Responsiveness to training, and Activity/Energy. However, the Aggression factor was composed of two major factors: Aggression towards Animals and Aggression towards People. The four-factor solution makes sense because it combines these two components of aggression, which are conceptually and empirically related to each other. However, the five-factor solution also makes sense because the two components are also readily separated. The five-factor solution also divides the items that load onto the factors Responsiveness to Training and Activity/Excitability such that the two factors each

consist of items that are more clearly related. In the four-factor solution, the corresponding factors are labeled “Responsiveness to Training” and “Activity/Energy.” The difference between Activity/Energy and Activity/Excitability is that the latter, from the five-factor solution, includes items describing a dog being very active in response to a stimulus (e.g., a person’s arriving at the dog’s home, a doorbell’s ringing) instead of simple high energy (e.g., the opposite of lethargy).

Given my goal to create a practical assessment instrument, I decided to proceed with the five-factor solution because the added fidelity could be of substantial practical and scientific significance for the many cases in which the two forms of aggression need to be separated.

I extracted five factors using PCA with varimax rotation. The resulting rotated solution is presented in Appendix A and serves as the basis for characterizing the five factors.

The first of the five factors comprises a total of 99 items (100, if “Dog is anxious” is counted twice) with absolute factor loadings ranging from .141 to .721. Items that loaded strongly on this factor include “Dog is fearful” (.721), “Dog is timid” (.712), “Dog is nervous” (.697), “Dog is shy” (.662), “Dog is confident” (-.653), “Dog remains calm in stressful situations” (-.562), and “Dog adapts easily to new situations and environments” (-.538). Six of the first nine items are short trait descriptions (e.g., “Dog is shy”). Collectively, the items that load strongly on this factor appear to be driven by fear and/or emotional instability, and thus I labeled the factor for one end of that dimension, “Fearfulness”.

The second factor is composed of 66 items, with absolute factor loadings ranging from .229 to .716. The twelve strongest items that load on this factor load in the same direction (positively) and are all explicitly associated with Aggression towards People. These items include “Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar women” (.716), “Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar men” (.697), and “Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when directly approached by an unfamiliar person”

(.693). The strongest negatively loading items are “Dog is friendly” (-.588), “Dog is friendly towards unfamiliar men” (-.540), and “Dog is not aggressive” (-.490). Taken together these items appear to measure an underlying dimension of aggression, and the first twelve indicate that it is aggression specifically targeting people. In addition, the combination of aggression and friendliness towards people on a single factor suggests that these are opposites (i.e., that aggression’s opposite may not be lack of aggression, but actual overt friendliness). The inclusion of general items such as “Dog is not aggressive” and “Dog is friendly” on this factor may indicate that when people provide a general rating of whether a dog is aggressive, they are weighting the dog’s behavior with people more than the dog’s behavior with other animals. The factor is labeled “Aggression towards People”.

The third factor is composed of 86 items with absolute factor loadings ranging from .121 to .678. Items associated with the strongest absolute loadings include “Dog is energetic” (.678), “Dog is active during play with people” (.654), “Dog is active” (.650), and “Dog is very excitable during play with toys” (.633). The first 28 items on this factor loaded in one direction (positively), and only eleven load negatively. Some of the items that load negatively are “Dog is lethargic” (-.446), “Dog gets bored in play quickly” (-.379), and “Dog tends to be calm” (-.362). Other items that load on this factor describe specific behaviors (e.g., “Dog chases after thrown objects [e.g., sticks, balls, or toys]”). Together, the items on this factor seem to be driven by an underlying characteristic of activity, energy, playfulness, and excitability. I labeled the factor “Activity/Excitability”.

The fourth factor is composed of 55 items with absolute factor loadings ranging from .127 to .631. Items associated with the strongest absolute loadings on this factor include “Dog is responsive to training, readily trained” (-.631), “Dog is easy to train” (-.616), “Dog learns readily” (-.607), “Dog is willing and able to react to signals and cues from the handler” (-.596), and “Dog is slow to respond to corrections” (.589). The majority of items that load onto this factor are about learning, responding to cues from the handler, and obedience, so I labeled the factor “Responsiveness to Training”.

The fifth and final factor comprises 46 items with absolute factor loadings ranging from .088 to .651. The two strongest items to load on the factor have loadings of equal magnitude but opposite valence: “Dog is friendly towards unfamiliar dogs” (-.651) and “While on leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar large dogs” (.651). A separate item that differed only in the size of the unfamiliar dog (“small” vs. “large”) had a similar loading of .644. A majority of the items that load strongly on this factor describe dogs’ being aggressive or pushy in their interactions with other dogs, while some also describe dogs’ displaying aggressive behavior towards other animals (e.g., “Dog kills other animals [e.g., squirrels, rabbits, birds]”). Seven of the items loading on this factor describe dogs’ being friendly or playful with other dogs. Taken together, these items indicate that the factor is driven by underlying tendencies to be aggressive or friendly towards dogs and other animals.

Part 2: Split-sample EFA and CFA with 353 items

For a factor solution to be generalizable, it is vital that it is stable. In the current section of Study 2, and again in Part 5 of study 2, my goal was to determine whether the five-factor model selected in the first phase of Study 2 replicates. In Part 2, the question is whether the five-factor model is found again when half the data are analyzed by EFA, and then how well that five-factor model fits the second half of the data. That is, how well does the five-factor model replicate with the full selection of items, within the Study 2 data?

I divided the participant sample in half (first half, N = 1,868; second half, N = 1,869), then performed an EFA on half of the participants’ ratings of their dogs on the 353 items. I used SEM to perform CFA on the second half of the participants’ ratings of their dogs. That is, I used SEM to fit the factor solution derived from the EFA of the first half of the data to fit that factor solution to the second half of the data.

EFA

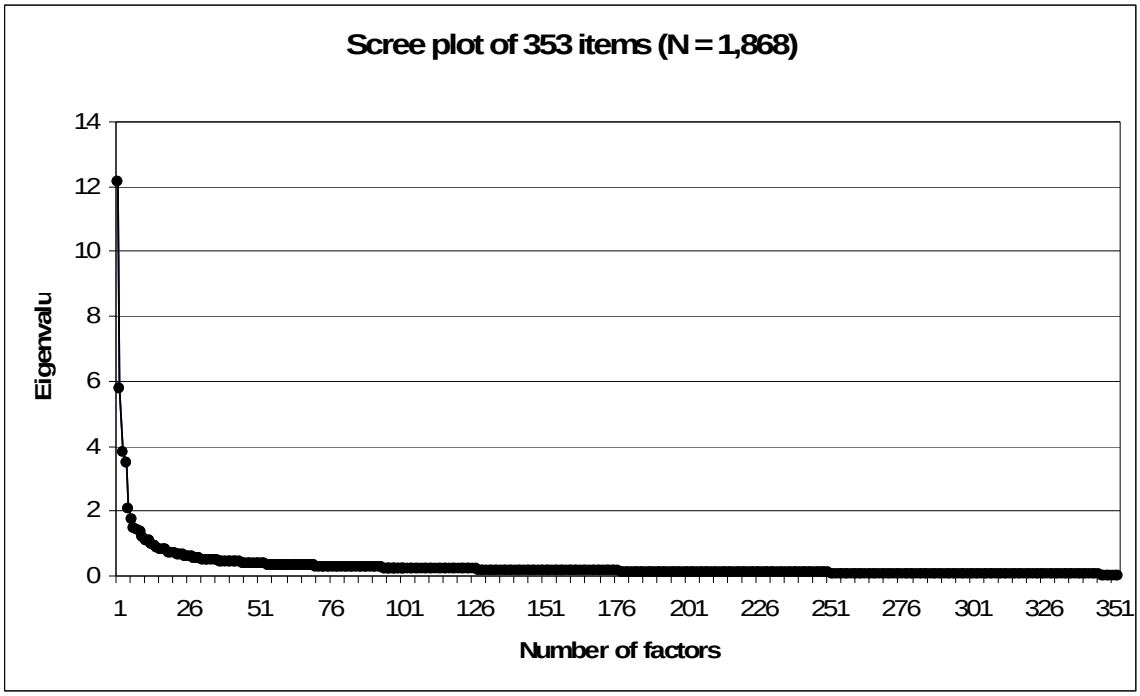
The EFA was performed using the same procedures as in Part 1 of Study 2; I used PCA with varimax rotation, and I dealt with missing data points through mean substitution.

As in Part 1 of Study 2, before proceeding with the EFA, I needed to assess the sampling adequacy to make certain that the data were amenable to EFA. This procedure needed to be repeated because the sample in Part 2 of Study 2 utilized half the participant sample. The ratio of participants to items is slightly greater than 5 : 1, which is relatively low for factor analysis. However, because the ratio of participants to items in Part 2 of Study 2 is slightly greater than 5:1, I relied again on the KMO measure of sampling adequacy. Conservative cut-offs suggest that a KMO of at least .6 is adequate (e.g., Tabachnick & Fidell, 2001). The KMO associated with the correlation matrix resulting from half the data set for Study 2 (N = 1,868) is .934, indicating that the matrix was good for EFA.

I used the convergence of three criteria (described in Part 1 of Study 2) to determine the number of factors in the solution: the graphical scree test (Cattell, 1966), the top-down method in which correlations between orthogonal factors scores from different factor solutions are viewed as path coefficients in a hierarchical structure (Goldberg, 2006), and interpretability of the solutions (see Zwick & Velicer, 1986; Worthington & Whittaker, 2006). Results from each method and associated conclusions are elaborated below.

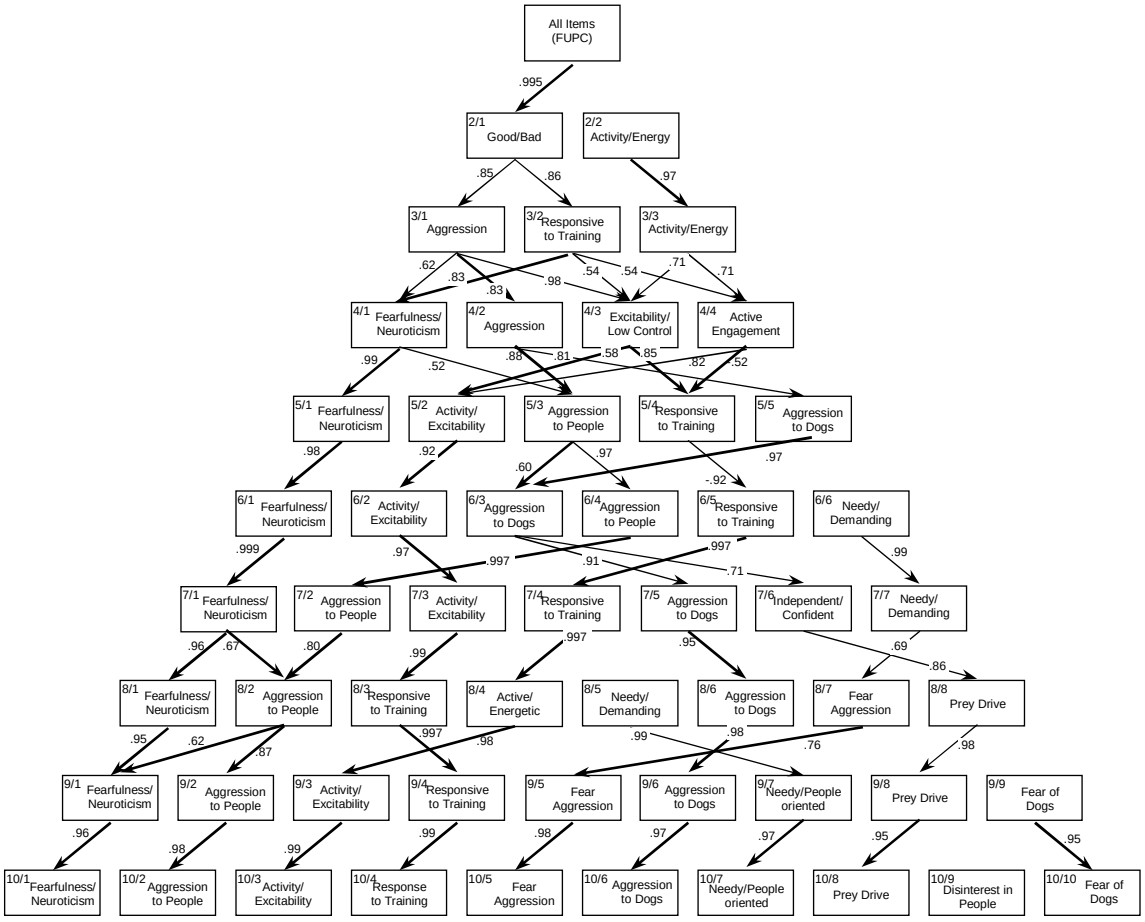
Interpretation of the scree plot suggested retaining four or five factors; there is a break in the values after four, and an elbow or bend appears after five.

Figure 5.3. Scree plot of eigenvalues for 353 items in half of Study 2 sample



The hierarchical tree is presented in Figure 5.4. To ease interpretability, only correlations of .60 or greater were included. As a result, some factors are not shown to be correlated with factors from the previous (higher) level of the tree. To further facilitate interpretability, each factor was assigned a descriptive label, which is presented in the box representing that factor. These labels were generated by having two personality experts examine the factor structures for every solution (i.e., the 2 factor solution, the 3 factor solution, etc.) and then examine the relationships between the factor structures as presented in the hierarchical tree.

Figure 5.4. Hierarchical structure of factor score correlations in 353-item Study 2 data set



Note. For clarity, only correlations of .60 and greater are included in the hierarchical tree. Correlations of .80 or greater are indicated by boldface arrows.

Although there are subtle differences, the hierarchical tree presented in Figure 5.4 is relatively close replication of the tree derived in Part 1 of Study 2. Ultimately, the hierarchical tree thus suggests five factors: Fearfulness (5/1), Activity/Excitability (5/2), Aggression towards People (5/3), Responsiveness to Training (5/4), and Aggression towards Animals (5/5).

Finally, I examined the interpretability of the solutions. I followed the same procedures as outlined in Part 1 of Study 2, examining factor interpretability independent of the hierarchical tree structure. The interpretability exercise closely duplicated that of examining interpretability in Part 1 of Study 2 and resulted in the same five-factors derived. These factors were ultimately labeled as they appear in the hierarchical tree: Fearfulness, Activity/Excitability, Aggression towards People, Responsiveness to Training, and Aggression towards Animals.

Across the three criteria, the supported factor solution is the same structure as was derived in analysis of the whole sample. Next, the solution derived in this analysis was used in a CFA performed on the second half of the data set, briefly described in the next section.

CFA

The most current approach to conducting CFA is to use SEM. SEM is a powerful confirmatory technique because it provides the researcher with a lot of control over the constraints placed on the items and factors when analyzing the hypothesized model. Another benefit to using SEM is that it allows researchers to examine competing models and assess the extent to which one model fits the data better than an alternative model (Worthington & Whittaker, 2006). I used AMOS to conduct SEM, which enabled me to compare the fit of various models.

Using SEM to perform CFA requires a large ratio of participants to parameters (or items). The statistical theory underlying SEM implies large samples are needed to provide parameter estimates (Bentler, 1995). If the samples are too small, there are two major risks: (1) the SEM may not be able to be computed because the parameters cannot

be estimated (Bentler, 1995), and (2) idiosyncrasies in the sample can substantially influence correlations among items, creating unstable patterns of covariation and leading to unique factor structures (DeVellis, 2003). Various guidelines have been provided for determining the required ratio of participants to parameters, based on the number of participants, on items per factor, or on ratio of participants per item (Worthington & Whittaker, 2006). It is unclear which guideline is best, because which is best depends on the characteristics of the model being fit (e.g., items per factor, normality of the data, strength of the relationship between the items and latent factors). Bentler and Chou (1987) stated that researchers can use as few as five participants per parameter if the data are normally distributed, no data are missing, and there are no outliers. Stevens (1996) indicated that 15 participants per item is a good guideline. The conventional rule of thumb recommends a range from five to 10 participants per parameter (e.g., Worthington & Whittaker, 2006). In the current CFA, there are only 5.3 participants per parameter (N = 1,869, parameters = 353).

However, number of participants per parameter is only one factor is determining whether a sample is adequate for factor analysis. As in the EFA described above, I used KMO to assess the adequacy of the sample for use in factor analysis. The KMO associated with the correlation matrix for the derived from the second half of the Study 2 data, those to be used in the current CFA, is .932, greater than the cut-off of .6 for a matrix to be acceptable for factor analysis (e.g., Tabachnick & Fidell, 2001).

Prior to conducting the CFA using SEM, a researcher must indicate (1) how many factors are present, (2) which items are related to each factor, and (3) whether the factors are correlated or uncorrelated. These issues were largely revealed through the EFAs performed in previous phases of Study 2 (Part 1 and the previous phase of Part 2 of Study 2). For the current analyses, the fit of the five-factor model structure derived in Part 2 of Study 2 was assessed; that is, which items were expected to load onto which factor were determined in the EFA on the first half of the data in Study 2. The items predicted to load on each factor were specified prior to CFA. In the EFA, I forced varimax rotation, which focuses on “cleaning up” the factors so that they have high correlations with one set of

items and little or no correlation with other items (Stevens, 1996). Thus I will include the assumption that the factors are uncorrelated in one CFA model. However, the challenge in deciding between the four- and five-factor models was due to items that load onto the five-factor model's Aggression towards People and Aggression towards Animals being related to one another and grouping as a single factor in the four-factor model. Thus, a model in which Aggression towards People and Aggression towards Animals were related was also fit. Furthermore, evidence suggests that aggression, particularly towards people, is often fear-related, hence the terms fear-aggression and defensive aggression. This suggests a relationship between the Fearfulness and Aggression towards People factors, which was also added to the model in an additional analysis of fit.

After performing the CFA, the non-trivial task of evaluating the model fits remained. The task is complicated because there is no single fit statistic with consistently low Type I and Type II error rates across all conditions of model complexity, sample size, item distribution, and so on (Recklitis et al., 2006). Most investigators rely on multiple indices to evaluate model fit. Commonly used goodness-of-fit indices include chi-square, the root mean square residual (RMR), the Goodness of Fit Index (GFI), the Bentler-Bonett non-normed fit index (NNFI; also called the Tucker-Lewis index, or TLI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA) with a corresponding 90% confidence interval, among others (e.g., Recklitis et al., 2006; Worthington & Whittaker, 2006; Kline, 2005). Each goodness-of-fit index has common guidelines for values indicating an acceptable model fit.

Chi-square is probably the most commonly used test statistic for assessing the overall fit of a model in SEM (e.g., Recklitis et al., 2006). The chi-square statistic tests the null hypothesis of a perfect model fit in which the residual covariance is zero. For a model to be accepted as a good fit, the chi-square test statistic should not be significant. Unfortunately, the chi-square test statistic is very sensitive to sample size (Jöreskog & Sörbom, 1989; Bentler & Bonett, 1980; Recklitis et al., 2006) and to even small amounts of unexplained variance (Bentler, 1990; Bollen & Long, 1992); as sample size increases, so does the likelihood of rejecting an adequately fitting model. Consequently, with a

sample size that is large enough for conducting SEM for CFA, most models will have a significant chi-square value (e.g., Recklitis et al., 2006). In fact, Carmines and Melver (1981) have suggested that, for this type of analysis, a chi-square two to three times the value of the associated degrees of freedom is acceptable. Because the chi-square is so sensitive to sample size, I did not use it to evaluate model fit, but I reported it because it is still commonly used and reported.

I depended more heavily on the other test statistics – the RMR, GFI, NNFI, CFI, and especially RMSEA – to determine whether the five-factor model is a good fit and whether the factors are better fit if correlated or uncorrelated. The better a model fits the data, the smaller the RMR value will be. The GFI, which indicates the amount of variance jointly accounted for by the model, ranges from 0 to 1 with values greater than .85 generally considered to indicate an acceptable fit (see Jolliffe & Farrington, 2006). The NNFI and CFI also range from 0 to 1, with higher values being better fits and values greater than .9 generally being accepted (e.g., Recklitis et al., 2006). However, these two incremental fit indices are used in particular to determine how much better a model fits compared to an alternate model (e.g., Bollen, 1989). The NNFI and CFI values represent the amount of covariance explained by the model and are considered appropriate for larger samples (Bentler, 1990; Floyd & Widaman, 1995; MacCallum, 1990). RMSEA, often called a badness-of-fit statistic, represents the covariance that is not explained by the model. Smaller RMSEA values indicate a better model fit; values below .08 are considered adequate and values less than .05 are considered good (Browne & Cudeck, 1992; Recklitis et al., 2006). The RMSEA is considered an appropriate test for use with large samples because it is less affected by sample size than the chi-square test statistic (Browne & Cudeck, 1992; Floyd & Widaman, 1995).

Results

The fit indices associated with the three different models evaluated through CFA, along with standard cut-offs for each index, are presented in Table 5.5. Examination of the fit indices indicated that all three represented a good fit and can be accepted

according to RMSEA. Other indices (RMR, GFI, CFI, NNFI) show that including a correlation between the Aggression towards Animals and the Aggression towards People factors in the model is associated with improved model fit. Including a correlation between Aggression towards Animals and the Aggression towards People factors and between the Fearfulness and Aggression towards People factors in the model may be associated with a slightly more improved fit.

Table 5.5 Confirmatory factor analysis of second half of Study 2 data (353 items)

Five-factor models					
Fit Index	**Model fit estimate**				**Standard cut-off**
	factors uncorrelated	**F corr AP**	**AP corr AA**	**F corr AP; AP corr AA**	
Chi-squared	303,166.544, df = 61,424, p < .001	303,164.841, df = 61,423, p < .001	303,151.525, df = 61,423, p<.001	303,144.136, df = 61,422, p<.001	n.s.
RMR	.305	.305	.303	.303	smaller indicates better fit
GFI	.654	.675	.695	.710	>.85
CFI	.554	.587	.577	.631	>.9; larger indicates better fit
NNFI	.453	.498	.510	.552	>.9; larger indicates better fit
RMSEA	.046 (lo = .046, hi = .046)	.046 (lo = .046, hi = .046)	.046 (lo = .046, hi = .046)	.045 (lo = .045, hi = .046)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr AP = .434	AP corr AA = .501	F corr AP = .421 AP corr AA = 489	

Note. F = Fearfulness, AP = Aggression towards People; AA = Aggression towards Animals. The abbreviation "n/a" indicates that statistics are not applicable to the model in that column, and "n.s." indicates "not significant".

Part 3: Determination of trait facets

In the third part of Study 2, I aimed to determine the facets of each of the five factors by applying EFA (PCA with oblique [promax] rotation so that the factors on each factor would be allowed to correlate) to the items comprising each factor. Items that loaded comparably on multiple factors were included in the analysis only for the factor on which they most strongly loaded, but were noted to have multiple comparable loadings. This criterion was used for selecting items to analyze on each factor because

items that were not univocal were poor candidates for the final questionnaire. Therefore, for example, the item “Dog likes to chase bicycles, joggers, or skateboarders” was included only in the analysis of Aggression towards People, but was noted to load comparably on Factor 2 (.259; Aggression towards People), Factor 3 (.224; Activity/Excitability) and Factor 5 (.241; Aggression towards Animals). Notes of comparable cross-loadings were later used to in assessing whether the removal of items with multiple loadings comparable in size resulted in a reduction of content validity.

Two expert judges used the following criteria to determine the number of facets to retain in each factor’s EFA: A graphical scree test (Cattell, 1966), a hierarchical tree structure (Goldberg, 2006), and interpretability of the solutions (see Zwick & Velicer, 1986; Worthington & Whittaker, 2006). These criteria are described in detail in Part 1 of the current study. The fourth criterion used in the EFA presented in Part 1 of this study, replicability of the solution, was not used here because there were not enough items in each factor to allow for subdividing the item sets and then repeating the analysis.

A major goal in conducting these analyses was to determine the facets of each factor so that the facets could guide the selection of items to retain in the next step. Therefore, when the scree plot (Cattell, 1966) and the hierarchical tree (Goldberg, 2006) criteria did not clearly converge on the number of facets in a factor, the experts very carefully considered the interpretability of the facets and based their decisions heavily on interpretability. In doing so, they considered whether the items in each facet composed a cohesive group, and whether the items composed a meaningful and useful facet (i.e., one that people assessing canine personality might want or need to measure). Below, I briefly describe each facet solution, including pattern matrix factor coefficients where useful to illustrate the relationship between the item and facet; the pattern matrix loadings are best understood as regression coefficients of the items on the facets. It should be noted that not all of the items included in each facet can be retained in the final questionnaire, and therefore not all of the items listed below appear in the final questionnaire.

The trait Fearfulness was determined to have four facets. Evaluation of the scree plot (Cattell, 1966) indicated three facets, but evaluation of the hierarchical tree

(Goldberg, 2006) indicated four factors. In the four-facet solution, four conceptually unique facets (Fear of People, Non-social Fear, Fear/Submission towards Dogs, and Fear of Handling) emerged, whereas in the three-facet solution Non-social Fear and Fear of Handling were combined. Because all four facets appeared to be conceptually distinct and of potential use to dog personality assessors, the four-facet solution was preferable. The first facet, Fear of People, is characterized by items such as "Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar women" and "Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar men". The second facet, Non-social Fear, is characterized by a dogs' being able or unable to adapt to new situations and remain calm in new and/or stressful situations. Items in the second facet of Fearfulness include "Dog is able to adapt to various types of situations" and "Dog is fearful when walking near loud, heavy traffic". The six of the central (strongest loading) items on the third facet, Fear/Submission towards Dogs, describe fearful behavior in situations that vary by whether the target dog is on- or off-leash, whether the other dog is small or large, and whether the other dog is familiar or unfamiliar. Other items in this facet describe dogs who are submissive towards other dogs, and dogs who are fearful when other dogs greet or threaten them. Fear of Handling, the fourth facet, is characterized by items describing fearful behavior during grooming (e.g., when toe nails are trimmed), sensitivity to mild and moderate pain, and examination by a veterinarian.

Based on consensus among evaluation of the scree plot, hierarchical tree, and interpretability, the trait Aggression towards People (Factor 2) was determined to have two facets. The two facets describe Aggression towards People in general (e.g., "Dog exhibits aggressive behavior [e.g., bares teeth, growls, lunges] towards unfamiliar men,") and in reaction to specific situations (e.g., when stolen items are taken from the dog, the dog is scolded). The two facets of Aggression towards People have been labeled descriptively as General and Reactive Aggression towards People, but might also be considered offensive and defensive or situational aggression.

The evaluation of the hierarchical tree and of the interpretability of facets indicated that Factor 3, Activity/Excitability, was most cleanly divided into four facets.

The scree plot, however, indicated only two or three facets. Again, interpretability and content validity implications were considered the most important criteria, and so four facets were identified. The primary difference between the three- and four-facet solutions is the separation of items about playfulness. In the three facet solution, they are divided between the first and second facet, but in the four-facet solution playfulness-related items form a distinct facet. The first facet is Excitability and is characterized by items such as "Dog is boisterous," "Dog is very excitable in play with other dogs," and "Dog becomes wildly excited when owner returns home." The second facet is Playfulness, which is comprised of items such as "Dog chases after thrown objects (e.g., sticks, balls, or toys)" and "Dog enjoys play fighting or playing roughly." The third facet is Active Engagement and is characterized by items describing dogs' physical activity level and their interest in their environments. For example, items comprising this facet include "Dog is lethargic" and "Dog is very alert, watchful (e.g., monitors yard for squirrels, monitors noises around the home." The fourth and final facet of Activity/Excitability is Companionability. This facet includes items pertinent dogs' interest in spending time with people (e.g., "Dog seeks companionship with people," "Dog follows people around") and how affectionate the dogs are (e.g., "Dog seeks affection from family members," "Dog often expresses affection"). The inclusion of Companionability in the DPQ is important because it is a quality that many people desire in a dog and thus one that people want to assess.

For the Responsiveness to Training factor, the scree plot indicated two or three facets, the hierarchical tree suggested three or four, and the interpretability of the facets indicated two. The second facet of the two-facet solution is readily understood as one facet, but can also be divided into two facets. However, when it is divided into two facets – which could be described as Unruliness and Destructiveness – the latter facet, Destructiveness, is extremely narrow in focus. This argues against the three-facet solution, because the third facet is focused one very specific type of unruly behavior. In the four-facet solution, the one of the three facets is divided such that is it again difficult to interpret. For sake of interpretability, but to have meaningful and useful facets, the two-facet solution was selected. The first facet, Trainability, includes how easily the dog

is trained and how readily the dog learns. The second facet, Unruliness, includes items describing a number of behaviors characteristic related to whether dogs are under their owners' control (e.g., stealing food, destroying household items, coming when called).

For the Aggression towards Animals factor, the scree plot showed multiple elbows indicating there may be two, three, four, seven, or more facets. Evaluation of the hierarchical tree suggested four facets, and the interpretability of facets indicated three or four facets. In the four-facet solution, Dominance over Other Dogs and Food Guarding (or Food Motivation) are two separate facets. However, Food Guarding was conceptually strongly related to other guarding behaviors loading on the Dominance over Other Dogs facet. Therefore, the three facets were selected: Aggression towards Dogs, Prey Drive, and Dominance over Other Dogs. The first facet, Aggression towards Dogs, was composed of items that describe both friendly and aggressive behavior towards other dogs (e.g., "Dog is friendly towards unfamiliar dogs,"; "While on leash, dog exhibits aggressive behavior [e.g., bares teeth, growls, lunges] towards unfamiliar small dogs"). The second facet, Prey Drive, includes items describing chasing, aggressive, and/or predatory behavior towards smaller animals, such as "Dog likes to chase cats", and "Dog kills other animals (e.g., squirrels, rabbits, birds)". The third facet, Dominance over Other Dogs, is composed of items such as "Dog willingly shares toys with other dogs," "Dog exhibits assertive behaviors (e.g., stands erect, ears forward, direct stare, tail up) when greeting dogs," and "Dog guards good or treats from other dogs". The items in this facet reinforce that status-related behavior and aggression may be related to, but separable from, other types of aggression in dogs.

Taken together, these analyses indicate that the DPQ's five factors subdivide into a total of 15 facets. Fearfulness comprises four facets; Aggression towards People two, Activity/Excitability four, Responsiveness to Training two, and Aggression towards Animals three.

Part 4: Item reduction

In the fourth part of Study 2, my goal was to refine the questionnaire by identifying the most suitable items to retain for use in a more manageable and useful form of the instrument. I selected items to make certain that all of the 15 facets were relatively equally represented. I aimed for about eight items per facet, but often found entire facets were fully represented by fewer.

I based the elimination and retention of items primarily on four criteria: (1) content validity, (2) strength of item loadings, (3) cross-loading magnitude and difference from highest loadings, and (4) internal consistency. Alternatively, the second and third criteria can, together, be thought of as univocality of the items with respect to the five factors (or traits) derived in Part 1 of Study 2. In addition, I attempted to retain items with both positive and negative valence on each facet and factor if possible.

In selecting items, I first examined each facet to identify the best items to fully represent the content of the facet. Item quality was evaluated in terms of content validity, loading on the facet, and cross-loading across facets. I also examined each item's univocality with respect to the five broad personality traits. If multiple items appeared to measure the same thing or to be very closely related to one another, as the aggression towards other dogs items appeared to, I considered at those items as a group and carefully examined their correlations with each other. The goal was to retain items that assessed as disparate aspects of each facet as possible (though this is associated with a reduced level of internal consistency). I then combined items (e.g., "While on leash, dog exhibits aggressive behavior [e.g., bares teeth, growls, lunges] towards unfamiliar small dogs" with "While off leash, dog exhibits aggressive behavior [e.g., bares teeth, growls, lunges] towards unfamiliar small dogs") when the items were very closely conceptually and statistically related, because the items were empirically shown to tap the same behavior.

Additional consideration was given to each item's length and syntactic complexity, and to including both positively and negatively loading items. I prioritized choosing items that were shorter and worded more simply when their content and loadings were comparable with longer, more complex items. I aimed to select items that

loaded both positively (e.g., “Dog is curious” on Activity/Excitability) and negatively (e.g., “Dog is lethargic” on Activity/Excitability). Furthermore, if it was evident (i.e., through correlations of > .4 between items) that people responded similarly to items that lacked behavioral descriptions (e.g., “Dog is aggressive,” “Dog has a tendency to attack [or attempt to attack] people”) and items that included lists of example behaviors (e.g., “Dog exhibits aggressive behavior [e.g., bares teeth, growls, lunges] towards unfamiliar men”), retained items were simplified and example behaviors were removed. In cases in which there was little or no evidence that participants understood the items without examples (e.g., assertive), behavior examples were retained.

Content validity (facets)

As discussed in Chapter 6, content validity is the extent to which the set of items represents all facets of the construct (e.g., the personality trait) being measured. In order for the short forms of the questionnaire to have high content validity, the items that compose the questionnaires had to represent all facets within each of the five factors. Items were chosen to represent the four facets of Fearfulness: Fear of People, Non-social Fear, Fear of Dogs, and Fear of Handling; the two facets of Aggression towards People: General and Defensive Aggression towards People; the four facets of Activity/Excitability: Excitability, Playfulness, Engagement and Companionability; the two facets of Responsiveness to Training: Trainability, and Unruliness; and the three facets of Aggression towards Animals: Aggression towards Dogs, Prey Drive and Dominance over Other Dogs.

Item factor loadings

It is common to use a guideline for a lower limit on item factor loadings to determine whether to retain or delete items (Tabachnick & Fidell, 2001). Common cut-offs are absolute values of .3 and .4 (Floyd & Widaman, 1995), both of which are represented in Appendix B as horizontal lines between items. Tabachnick and Fidell (2001) describe the cut-off for factor loadings as a matter of preference, and Worthington and Whittaker (2006) recommend setting the lower limit as high as possible without

compromising scale length or factor structure. Because my initial item pool is so large and I seek to reduce it by at least two-thirds, I aimed to retain only those items with trait loadings greater than or equal to .4 (i.e., loadings on Fearfulness, Aggression towards People, and so on). However, value was considered adjustable as dictated by the other criteria (i.e., items with much lower primary trait loadings were retained as needed to preserve content validity).

Item cross-loadings

Other common guidelines address issues of cross-loading in an attempt to retain the most univocal items, or those items that load strongly on only one factor. Cross-loadings are also an issue because larger, more frequent cross-loadings contribute to factor intercorrelations (Worthington & Whittaker, 2006). One guideline for reducing cross-loadings is to set an upper limit for that absolute magnitude of cross-loadings. For example, if this limit is set to .25, then all items that have a second-highest loading of .25 will be eliminated. A second guideline for reducing cross-loadings is to set a minimum difference that must be present between the highest and second-highest factor loadings for items to be retained. For example, if the minimum difference is set to .15, then all items that have a difference of less than .15 between their highest and second-highest loadings will be eliminated. Worthington and Whittaker (2006) recommended setting these cut-offs to be as strict as possible without compromising scale length or factor structure. I aimed to retain only items with a difference of .2 between their highest and second-highest loadings on the broad personality traits (e.g., Fearfulness, Aggression towards Animals), but this rule was also flexible so that content validity could be maximized.

Internal consistency

The final criterion I used to retain or delete items was internal consistency, as measured by Cronbach's coefficient alpha. Internal consistency is a measure of the extent to which items composing a scale measure the underlying variable. Generally, Cronbach's alpha will increase as the correlations between items that measure a factor

increase, and it tends to increase as the scales increase in length. Therefore longer scales that consist of highly correlated items are typically more reliable (and their scales or factors are associated with higher values of Cronbach's alpha), but the drawbacks associated with having a long scale (e.g., time, participant fatigue; Worthington & Whittaker, 2006) are such that attempting to balance length and internal consistency is typically preferable to simply retaining a large number of items in order to have high internal consistency. When choosing between items that were similar in content and loadings, I examined how removing each item would affect the questionnaire's internal consistency, then removed the item that would have had the most deleterious effect on internal consistency if retained.

Resultant item pool

Item reduction following the procedures described above resulted in a pool of 102 items. Of the 102 items retained for Study 3, 53 were entirely unchanged. Multiple items were changed such that "men" and "women" were collapsed to "people" (e.g., "Dog behaves fearfully towards unfamiliar people), or so that different types of dogs (e.g., large, small, familiar, unfamiliar) were collapsed to simply "dogs". These items were collapsed based on the items' being strongly correlated such that, for example, dogs who feared men tended to fear women. Collapsing these very similar items helps to avoid inflating the reliability of the questionnaire; although some dogs may respond differently to men than to women, the data collected in Study 2 indicated that difference was rare and so the questions addressing men and women separately were so closely related as to be, functionally, repetitions of the same question.

The items, organized by facet and factor, are presented in Appendix C. In the second column, the loading of each item on its factor is indicated. The far right column in Appendix C lists arbitrarily assigned numbers used to identify the item or items that provided the basis for the item listed in that row. These numbers correspond to the items from Intermediate Questionnaire #1 (from Study 1), also used in Study 2, and correspond to the numbered items that appear in Appendix B.

In Appendix C, the items that are unchanged are indicated by a lack of parentheses around the item number listed in the far right column. Item numbers in parentheses indicate that the item text listed in that row (the item derived in Study 2 and for use in Study 3) is intended to address the content of all of the listed items, is a composite of those items, or is a modified version of the item. In cases in which one item number is outside parentheses but it is presented next to a set of parentheses, the items in parentheses are conceptually related and correlated to the item outside parentheses; the item outside parentheses was selected to be retained and address those items content.

Part 5: Split-sample EFA and CFA with 102 items

My goal in Part 5 of Study 2 was to determine whether the five-factor model selected in the first phase of Study 2 replicates using the Study 2 data pertinent to the 102 items (or their proxies) retained for Study 3. Modification of the questionnaire items was necessary to improve overall questionnaire quality, item length, and item interpretability, but reduction of the number of items and modification of the items' wording made it impossible to see analyses in Study 3 as a true replication of the factor solution in Study 2.

The items selected for inclusion in Part 5 of Study 2 included the 53 items that are exactly those retained for Study 3, and 49 items that served as proxies for the items that were revised for Study 3. For 11 proxy items, it was straightforward to select the item from the Study 2 item pool that would best represent the modified item of the Study 3 item pool. These Study 2 items differed from the modified Study 3 items only in that the Study 2 item included a behavioral description that was removed for Study 3. For example, Study 2 included the item, "Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when near crowds of people," which was used as a proxy for its modified form, "Dog behaves fearfully when near crowds of people" for Study 3.

Some Study 2 items were found to be so closely related to each other that including more than one of them would have been redundant. Instead, a modified or combined form of the items was created for Study 3, or only one of the items was

retained. In these instances, selecting a proxy item was more challenging. Proxy items were selected based on (1) which items wording and content was most similar to Study 3 item, and (2) which item was most strongly correlated with the related items and thus seemed most central or representative. For example, Study 2 included the items, "Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when examined by the vet," and "Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) during visits to the veterinary office." In Study 2, these two items were correlated at r = .788. Because the items were similar in content and highly correlated, they were collapsed into one item, and because the behavioral description was deemed unnecessary, the behavioral description was also removed. So the item for Study 3 became "Dog behaves fearfully during visits to the veterinarian," and the less specific item from Study 2 ("Dog exhibits fearful behavior [e.g., flees, tucks tail, trembles] during visits to the veterinary office"), with wording more similar to the new Study 3 item, was used as a proxy.

Once the proxy items were systematically selected, I randomly divided the participant sample in half again (first half, N = 1,868; second half, N = 1,869), then performed an EFA on half of the participants' ratings of their dogs on those 102 proxy items only. I used SEM to perform CFA on the second half of the participants' ratings of their dogs. That is, I used SEM to fit the factor solution derived from the EFA of the first half of the data to fit that factor solution to the second half of the data. This method allowed me to examine how well the five-factor model derived using the whole sample and the 353 items fit the subset of the sample using the 102 proxy items.

EFA

The EFA was performed using the same procedures as in Parts 1 and 2 of Study 2; I used PCA with varimax rotation, and I dealt with missing data points through mean substitution.

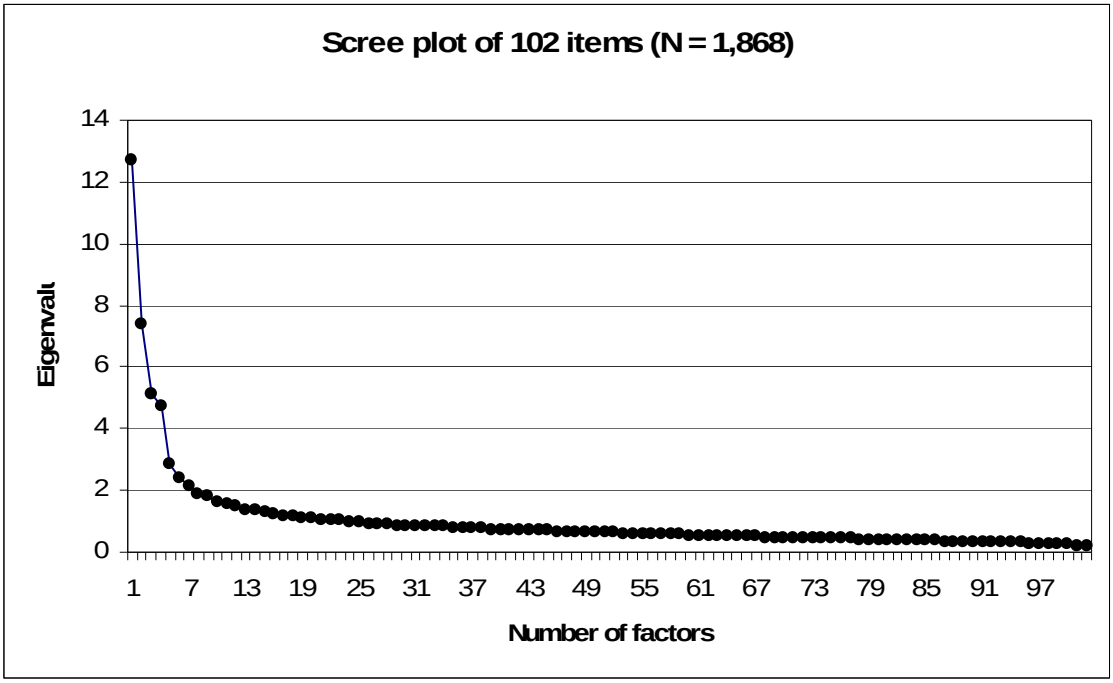
As in Parts 1 and 2 of Study 2, before proceeding with the EFA, I needed to assess the sampling adequacy to make certain that the data were amenable to EFA. This procedure needed to be repeated because the sample in Part 2 of Study 2 utilized half the

participant sample and a subset of the items from Part 1 of Study 2. The ratio of participants to items is large – great than 18 : 1. However, having an adequate sample size is only one aspect of the issue of sampling adequacy issue; another is the magnitude of correlations between variables, as displayed in the correlation matrix. Because the ratio of participants to items in Part 5 of Study 2 is greater than 5:1, I relied again on the KMO measure of sampling adequacy. Conservative cut-offs suggest that a KMO of at least .6 is adequate (e.g., Tabachnick & Fidell, 2001). The KMO associated with the correlation matrix resulting from half the data set for Study 2 (N = 1,868) is .909, indicating that the matrix was good for EFA.

As in Part 2 of Study 2, I used the convergence of three criteria (described in Part 1 of Study 2) to determine the number of factors in the solution: the graphical scree test (Cattell, 1966), the top-down method in which correlations between orthogonal factors scores from different factor solutions are viewed as path coefficients in a hierarchical structure (Goldberg, 2006), and interpretability of the solutions (see Zwick & Velicer, 1986; Worthington & Whittaker, 2006). Results from each method and associated conclusions are elaborated below.

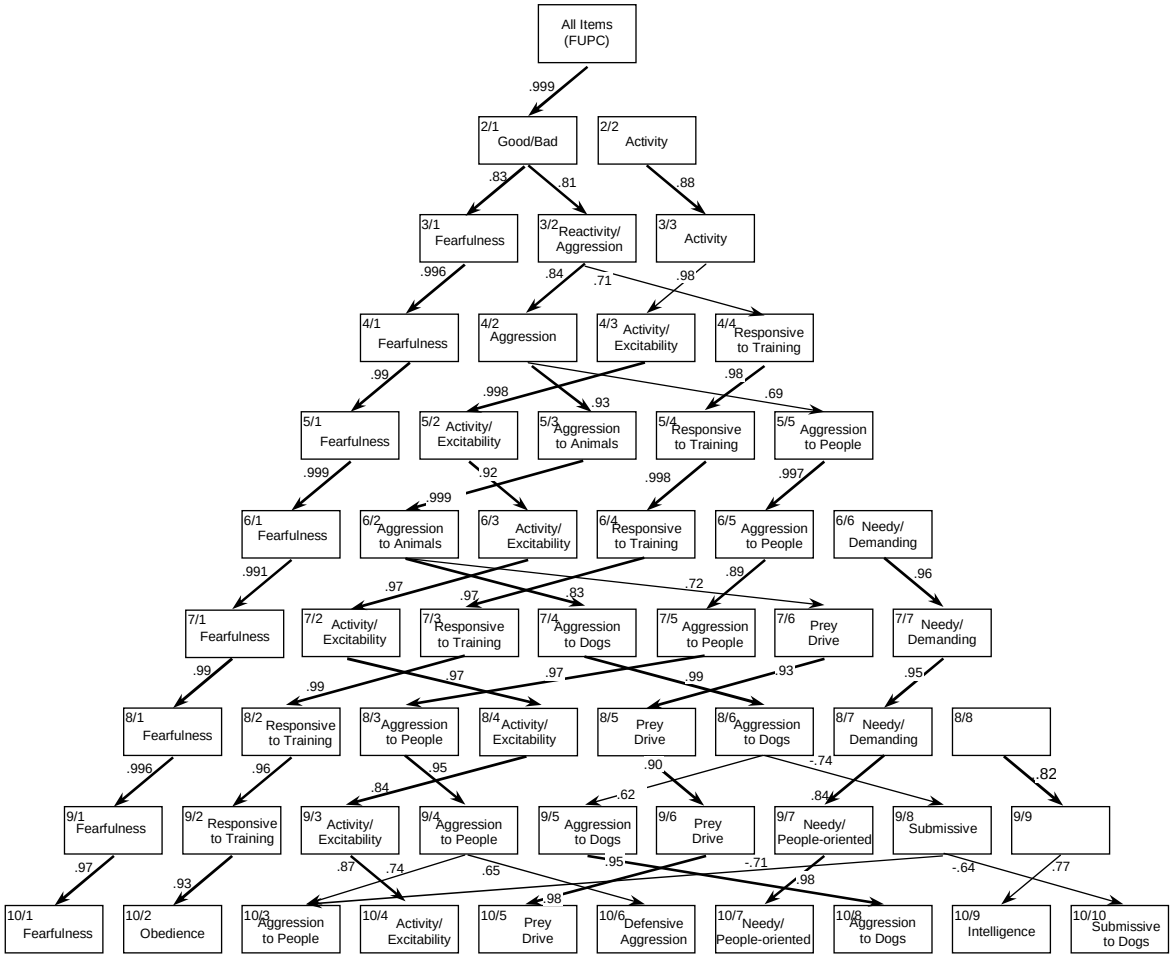
Interpretation of the scree plot suggested retaining four or five factors; there is a break in the values after four, and an elbow or bend appears after five.

Figure 5.5. Scree plot of eigenvalues for 102 items in Study 2



The hierarchical tree is presented in Figure 5.6. To ease interpretability, only correlations of .60 or greater were included. As a result, some factors are not shown to be correlated with factors from the previous (higher) level of the tree. To further facilitate interpretability, each factor was assigned a descriptive label, which is presented in the box representing that factor. These labels were generated by having two personality experts examine the factor structures for every solution (i.e., the 2 factor solution, the 3 factor solution, etc.) and then examine the relationships between the factor structures as presented in the hierarchical tree.

Figure 5.6. Hierarchical structure of factor score correlations in 102-item Study 2 data set



Note. For clarity, only correlations of .60 and greater are included in the hierarchical tree. Correlations of .80 or greater are indicated by boldface arrows.
Unlabeled boxes indicate difficult-to-name factors composed of 4 or fewer items.

The hierarchical tree is somewhat ambiguous, but is suggestive of a four- or five-factor solution (see Figure 5.6). Overall, the hierarchical tree is a replication of the tree derived in Part 1 of Study 2, using the whole sample and all 353 items. Subtle differences between the two exist. In the current hierarchical tree, Fearfulness appears as a strong and consistent factor very early on, earlier than it appeared in the hierarchical tree presented in Part 1. The order in which the factors are presented in each level of the hierarchy also differs, which indicates that the factors account for a different proportion of the variance in Part 1 and 2 analyses; in each level of the tree, the factors are presented in order of diminishing variance accounted for such that 5/1 accounts for the most variance in the five-factor solution, followed by 5/2, then 5/3 and so on.

The factors present when fewer (e.g., 3) factors were extracted were not consistent through the various levels of the tree. Specifically, the second factor of the three-factor solution ([3/2] Reactivity/Aggression) is broken up into two sizable, interpretable, and logically separable factors in the four-factor solution, (4/2) Aggression (correlated .84) and (4/4) Responsive to Training (correlated .71). The relationship between the four-factor level of the hierarchical tree and the five-factor level of the hierarchical tree is a close replication of the hierarchical tree in Part 1 of Study 2. The four-factor solution was problematic because the second factor ([4/2] Aggression) subdivided into (5/5) Aggression to people (correlated .69) and (5/3) Aggression to dogs (correlated .93) when five factors were extracted. In contrast, all the factors from the five-factor solution are retained, essentially unchanged, when six factors are extracted. When more factors are extracted, the picture becomes less clear, but close examination indicates that extracting more factors forces a solution in which very small groupings of items that do not create cohesive factors are extracted (e.g., [8/8], [9/9]), and very specific groups of items (e.g., those addressing submission towards dogs, [10/10]) are extracted and treated as factors. The hierarchical tree thus suggests five factors: Fearfulness (5/1), Activity/Excitability (5/2), Aggression towards People (5/3), Responsiveness to Training (5/4), and Aggression towards Animals (5/5).

Finally, I examined the interpretability of the solutions. I followed the same procedures as outlined in Part 1 of Study 2, examining factor interpretability independent of the hierarchical tree structure. The interpretability exercise closely duplicated that of examining interpretability in Part 1 of Study 2 and resulted in the same five-factors derived. These factors were ultimately labeled as they appear in the hierarchical tree: Fearfulness, Activity/Excitability, Aggression towards People, Responsiveness to Training, and Aggression towards Animals.

Across the three criteria, the supported factor solution is thus the same general structure as was derived in analysis of the whole sample. Next, the solution derived in this analysis was used in a CFA performed on the second half of the data set, outlined immediately below.

CFA

As in Part 2 of Study 2, I used AMOS to conduct SEM, because SEM is the most current approach to performing CFA. However, as discussed in Part 2, using SEM to perform CFA also requires a large ratio of participants to parameters (or items). In the current CFA, there were about 5.3 participants per parameter (N = 1,869, parameters = 353), so it was particularly important to examine the sample to determine whether it was adequate for the analyses. I used KMO to assess the adequacy of the sample for use in factor analysis. The KMO associated with the correlation matrix for the derived from the second half of the Study 2 data, those to be used in the current CFA, is .908, greater than the cut-off of .6 for a matrix to be acceptable for factor analysis (e.g., Tabachnick & Fidell, 2001).

The procedure for conducting the CFA in Part 5 of Study 2 was the same as that in Part 2, though the models that were fit and the number of item included differed. Prior to conducting the CFA using SEM, I specified each model, and I fit models with Fearfulness and Aggression towards People and with Aggression towards People and Aggression towards Animals included.

The model fits were also evaluated using the same fit indices as in Part 2 of Study 2, again depending most heavily on the RMSEA. See Part 2 for descriptions of each of the fit indices.

Results

The fit indices associated with the three different models evaluated through CFA, along with standard cut-offs for each index, are presented in Table 5.6. Examination of the fit indices indicated that all three represented a good fit and can be accepted according to RMSEA. Other indices (RMR, GFI, CFI, NNFI) show that including a correlation between the Aggression towards Animals and the Aggression towards People factors in the model is associated with improved model fit. Including a correlation between Aggression towards Animals and the Aggression towards People factors and between the Fearfulness and Aggression towards People factors in the model is associated with a still greater fit.

Table 5.6. Confirmatory factor analysis of second half of Study 2 data (102 items)

Five-factor models					
Fit Index	**Model fit estimate**				**Standard cut-off**
	factors uncorrelated	**F corr AP**	**AP corr AA**	**F corr AP; AP corr AA**	
Chi-squared	24,304.453, df = 5,049, p < .001	23,674.231, df = 5,048, p < .001	23,433.214, df = 5,048, p < .001	23,105.262, df = 5,047, p < .001	n.s.
RMR	.236	.224	.212	.201	smaller indicates better fit
GFI	.784	.789	.791	.823	>.85
CFI	.657	.676	.687	.840	>.9; larger indicates better fit
NNFI	.564	.597	.603	.678	>.9; larger indicates better fit
RMSEA	.043 (lo = .043, hi = .044)	.042 (lo = .042, hi = .043)	.042 (lo = .041, hi = .042)	.039 (lo = .038, hi = .039)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr AP = .587	AP corr AA = .534	F corr AP = .503 AP corr AA = .480	

Note. F = Fearfulness, AP = Aggression towards People; AA = Aggression towards Animals. The abbreviation “n/a” indicates that statistics are not applicable to the model in that column, and “n.s.” indicates “not significant”.

INTERMEDIATE QUESTIONNAIRE #2

Intermediate Questionnaire #2 contained a total of 102 items, which were unequally distributed. Each of the fifteen facets contained an average of between six and seven items. The facets Fear/Submission towards dogs, Companionability, and Aggression towards other dogs each consist of only five items. The facet Fear of people contains the most items (10). Because the factors do not consist of equal numbers of facets, ranging from two to four facets per factor, the factors are not uniform in the number of items they contain. Responsiveness to Training consists of the fewest items; it has only 14 items, divided equally between two facets. Fearfulness consists of the most items; it has 29 items, which are unequally distributed across four the four facets of Fearfulness.

SUMMARY OF FINDINGS

In Study 2, results of EFA (PCA, varimax rotation) of the responses from 3,737 participants to the 360-item questionnaire were evaluated using four criteria that, taken together, indicated a five-factor structure was underlying the questionnaire structure. The factors were discussed in order of the amount of variance they account for in the data, and were labeled as Fearfulness, Aggression towards People, Activity/Excitability, Responsiveness to Training, and Aggression towards Animals. The stability of the five-factor model was confirmed using two independent split-halves procedures in which I divided the participant set into randomly selected halves, repeated the EFA procedure on one half of the data (N = 1,868; with 353 items in Part 2, or 102 items in Part 5), then performed CFA to fit the model to the second half of the data (N = 1,869).

Each factor was further examined through factor analysis to determine the factors' facet structures. A total of fifteen facets were identified. Items loadings on factors and facets, along with other criteria (e.g., content validity, item valence) guided the shortening of the questionnaire to a more manageable instrument that strove to simultaneously maximize manageability, ease of use, and the psychometric property criteria (e.g., internal consistency, content validity).

CHAPTER 6

Study 3: Factor Structure Confirmation

INTRODUCTION

The primary goal of Study 3 is to confirm that the five-factor solution replicates in a new dataset collected using the new 102-item questionnaire. Other goals are to confirm that the five-factor solution replicated in this sample is stable, to evaluate the internal consistency, and to examine the construct validity-related discriminant and convergent validity of the questionnaire items. The ultimate goal is to derive two final, brief-but-comprehensive forms of the questionnaire.

METHOD

Design

As discussed in Part 2 of Study 2, SEM is a powerful and current approach to performing CFA, but using SEM to perform CFA requires a large number of participants per parameter to provide parameter estimates (Bentler, 1995). Based on Stevens' (1996) conservative recommendation of 15 participants per item, I aimed to administer the new questionnaire form (derived in Study 2 and comprised of 102 items) to at least 1,530 participants.

Participants

To attempt to recruit a diverse sample, I advertised the dog personality questionnaire in a variety of venues, including the Animal Personality Institute website (www.animalpersonality.org); online bulletin boards that target a variety of dog-related topics (e.g., Schutzhund, Animal Assisted Therapy, deaf dogs, dog training, general dog discussion); Austin-area parks, veterinary clinics, and training facilities; and similar venues in other cities (e.g., San Francisco, California; Detroit, Michigan).

In compensation for their participation, participants received feedback about their dogs' personalities with respect to the five personality factors (and their facets) assessed

in the DPQ and were entered into a drawing for a gift certificate. Participants responded to the online questionnaire between in April and May, 2007. A total of 3,282 cases were compiled; initial analyses (described below) refined the dataset to 2,556 participants.

Materials and Procedures

The procedure for Study 3 was very similar to that for Studies 1 and 2 in that it utilized web-based data collection and randomly ordered items for each participant. However, in Study 3 participants were allowed to select the dog about whom they fill out the questionnaire so that they could get feedback on the personality scores of the dog (or dogs) of their choosing. Participants were prohibited from submitting the questionnaire if any items are left blank, and the "not applicable" item response was eliminated. To avoid having the sample affected by repeat responders, at the beginning of the questionnaire I asked participants whether they have previously filled out the questionnaire.

The online questionnaire was composed of four parts. The first two parts were like Parts 1 and 2 of the questionnaire used in Studies 1 and 2. In Part 1, participants were asked to provide basic demographic information about themselves (e.g., age, sex, geographic location). These data allowed me to assess the diversity of the sample, and also to further screen for repeat responders. In Part 2, participants were asked the name of the dog about whom they were filling out the DPQ and were then asked to provide basic demographic information about that dog. These data allowed me to assess the demographic diversity of the dogs who were rated.

Part 3 of the questionnaire included the revised set of 102 questionnaire items. The item presentation and response options were the same as in Studies 1 and 2. The questionnaire items were presented in random order for each participant such that no two participants' questions were likely to appear in the same order.

After participants submitted their ratings of their dogs, they were presented with a brief summary of how their dogs were scored on the questionnaire. After recoding the reverse keyed items, scale scores were computed from the average score on items that comprise each factor (or facet). Participants were also informed that these results are

preliminary and that the questionnaire is under development. After they received these results, participants were given the opportunity to write free responses about how well the results suited their dog.

RESULTS

I examined and refined the participant sample collected in Study 3 in much the same way as outlined in Study 2, though missing data and "not applicable" response use analyses were made obsolete in Study 3. First, I removed repeat responders (i.e., the cases in which a participant completed a questionnaire more than once, either about the same dog or different dogs). This reduces the dataset from 3,282 cases to 2,630 unique participants. Based on examination of the owners' and dogs' names, demographic data and contact information, I retained only participants' first ratings of a dog, and only one participants' rating of each dog. If a person rated a dog more than once, only the first case was kept. If a person rated multiple dogs, only the first was kept. If two people rated the same dog, only the first was kept. The goal was to create a one-to-one ratio of raters to dogs in the sample.

Second, I examined the frequency of responses for each item to check for obvious errors (i.e., items for which there was only one response saved from all participants). No anomalies were found in the dataset. However, as in Studies 1 and 2, responses to each item were not regularly distributed.

Third, I removed cases in which participants or their dogs were not suitable for the study. These included participants who indicated that they were under 18 years of age (N = 49) or that their native language was not English (N = 15). Whether wolves and wolf hybrids can be suitably assessed using a personality assessment tool developed for assessing dog personality is an empirical question outside the constraints of this study, so participants who indicated their dogs were known wolf hybrids (2 participants) were deleted. I also removed nine participants who withheld responses (e.g., responded to the question of location by saying "none of your business" or "withheld").

The resulting dataset contained 2,556 unique cases.

Sample demographics

Demographic information for the 2,556 participants retained for analysis in Study 3 is presented in Table 6.1. As in Studies 1 and 2, both sexes of dog owners were represented, but there were far more women (N = 2,295, or 89.8%) than men (N = 261, or 10.2%). The average age of participants was 43 years, with a standard deviation of 12.6 years; this is very similar to the age of participants in Study 2 (mean = 43 years, s.d. = 12 years). Approximately 9% of participants were currently residing outside the U.S., including 156 participants in Canada, 32 in Australia, 27 in the United Kingdom, and 21 in other countries. All 50 states and Washington, D.C., were represented in the sample, though there was only one participant from Washington, D.C., and the states were not equally represented. There were large numbers of participants from Texas (N = 298) and California (N = 238), and very few from Hawaii (N = 2), North Dakota (N = 6), and Delaware (N = 6). The vast majority of participants (N = 2,404, or 94.1%) reported their race as Caucasian/White, though the sample also included participants who identified themselves as African American or Black (or African), Hispanic, East Asian, Native American, Asian Indian, and Multi-racial (or Other).

Just under 9% (or N = 227) of the 2,556 participants reported having at least one dog-related job or career. For example, 70 were trainers or behaviorists, 14 were groomers, and three had jobs in which they handled working dogs (e.g., as police officers). The mean number of dogs each participant had owned in his or her lifetime, previous to the current dog or dogs but including childhood family dogs, was 5.6 (s.d. = 6.9), though the number was skewed towards having had more dogs. On average, the participants each currently owned 2.4 dogs (s.d. = 1.3).

Table 6.1. Demographic information about participants in Study 3

		Dog owners
Sex (dog owner)	Male	261 (10.2%)
	Female	2,295 (89.8%)
Age (dog owner)	Mean	43
	standard deviation	12.6
	Range	18-84
Country of residence	U.S.	2,320 (90.8%)
	Canada	156 (6.1%)
	Australia	32 (1.3%)
	U.K.	27 (1.1%)
	Other (Native language is English)	21 (.8%)
Race	Caucasian	2,404 (94.1%)
	African American (or Black)	15 (.6%)
	Hispanic	36 (1.4%)
	East Asian	23 (.9%)
	Native American	17 (.7%)
	Asian Indian	2 (.1%)
	Multi-racial or Other	59 (2.3%)
Dog-related careers	Breeder and/or Exhibitor	75 (2.9%)
	Trainer and/or Behaviorist	70 (2.7%)
	Dog rescue worker/volunteer	24 (0.9%)
	Veterinarian	11 (0.4%)
	Groomer	14 (0.5%)
	Kennel/Daycare owner or employee	26 (1.0%)
	Assistance dog partner	4 (0.2%)
	Working dog handler	3 (0.1%)
		Mean (s.d.)
Number of dogs owned	Past (mean; s.d.)	5.6 (6.9)
	Currently	2.4 (1.3)
Total number of participants		**2,556**

As in Studies 1 and 2, it was also important to confirm that a diverse sample of target dogs was assessed in Study 3. I examined the composition of the sample in terms of the same characteristics I examined in Studies 1 and 2. In Study 3, the dogs' average age was 5.0 years (s.d. = 3.4) with all but one dogs' age reported; that dog's age was reported as "unknown". Unlike in the Study 2 sample, in this sample dogs aged less than one year were retained; 152 dogs of less than one year of age were included. The two oldest dogs included in the sample were 18 years old.

Other demographic information is presented in Table 6.2. The number of dogs identified as fitting each category is listed in the column titled "Number of dogs". The sample of 2,556 dogs appears to be relatively diverse. As in Study 2, approximately half (51.8%) of the dogs are male. Also, the majority of the dogs (79.8%, or 2,039 of the 2,556 dogs) were castrated, but intact dogs were also included (20.2%, or 517 of the 2,556 dogs).

Purebred dogs (N = 1,841), mixed-breed dogs (N = 632), and dogs whose breeding was unknown (N = 20) were included in the sample, with 153 unique breeds included in the purebred portion of the sample. GSDs are the breed represented in the largest number in Study 3; 156 GSDs are included, composing approximately 6% of the sample. There were also a large number of Siberian Huskies (N = 121), Boxers (N = 120), and Labrador Retrievers (N = 117). Together, these four breeds composed 20% of the sample.

I also examined whether the dogs had any reported disabilities or health issues, whether the dogs were reported to have bitten a person, what sort of role the dogs play in their owners lives (e.g., pet, guide dog), and what dog sports they were involved in. As in the previous studies, no total is presented for the number of dogs who participate in sports, because many of these dogs participate in more than one sport. As in Study 2, a large number of dogs serve as Animal Assisted Therapy dogs (N = 256), or compete in agility (N = 508), formal obedience (N = 345), and/or in American Kennel Club conformation (N = 249). The demographic information about the dogs in Study 3 indicated that they do compose a diverse sample, while also reflecting the current

popularity of some breeds of dogs (e.g., GSDs, Labradors Retrievers) and some dog sports (e.g., agility).

Table 6.2. Demographic information about target dogs in Study 3.

			Number of dogs
Sex	Male		1,324 (51.8%)
	Female		1,232 (48.2%)
Castration Status	Males	Neutered	985 (74.4%)
		Intact	339 (25.6%)
	Females	Spayed	1,054 (85.6%)
		Intact	178 (14.4%)
Purebred	Unknown		83 (3.2%)
	No (mixed)		632 (24.7%)
	Yes		1,841 (72%)
Bitten a Person	Unknown		20 (.8%)
	No		2,270 (88.8%)
	Yes		262 (10.3%)
Disability and Health issues	Unilaterally deaf		0 (0.%)
	Bilaterally deaf		2 (<.1%)
	Blind in one eye		1 (<.1%)
	Blind in both eyes		3 (.1%)
	Arthritic		175 (6.8%)
	Hip dysplasia		105 (4.1%)
	Elbow dysplasia		14 (5.4%)
	Other disabilities or health problems		463 (18.1)
Job or Role	**Pet/Companion**		1,944 (76.1%)
	Assistance dog		2 (<.1%)
	Guide dog		1 (<.1%)
	Hearing ear dog		0
	Medical assistance dog		1 (<.1%)
	Search and rescue		16 (.6%)
	Guard dog		113 (4.4%)
	Animal Assisted Therapy		256 (10%)
	Dam or sire for breeding		126 (4.9%)

Table 6.2. (*Continued*)

		Number of dogs
Sport	Obedience	345 (13.5%)
	Sledding	0
	Carting	52 (2%)
	Frisbee	104 (4.1%)
	Earth dog	53 (2.1%)
	Show/Conformation	249 (9.7%)
	Schutzhund	45 (1.8%)
	Agility	508 (19.9%)
	Herding (competitive)	126 (4.9%)
	Flyball	89 (3.5%)
	Hunting	124 (4.9%)
Total number of dogs		**2, 556**

Note. The sum of the number of dogs who participate in sports will not equal the total number of dogs in the study; some dogs are not in sports, and others are in multiple. Similarly, participants may consider a dog to be both a pet/companion and to have another job or role (e.g., assistance dog, search and rescue dog).

PART 1: CONFIRMATORY FACTOR ANALYSIS

Many of the details and much of theory behind performing CFA using SEM were described in the previous chapter's discussion of Part 2 of Study 2. Details are reviewed here if particularly useful in considering the CFA performed in Study 3.

CFA is commonly used during scale development, following EFA, to help support the generalizability of the scale and its structure to new samples (Worthington & Whittaker, 2006). CFA is also informative in determining a questionnaire's construct validity, or the extent to which the items within the questionnaire measure the broad construct (e.g., personality trait) that they were intended to measure (e.g., Worthington & Whittaker, 2006). CFA requires a researcher to have and specify a theoretically meaningful factor structure (typically derived through EFA, though sometimes hypothesized). Items are generally constrained to load on only one factor in CFA, so the method is appropriate for evaluating the extent to which a specified model is replicated in the data, but not for exploring whether a particular item measures no factors, one factor,

or so on. The researcher assesses how well the specified model fits the new data by assessing various goodness-of-fit indices (discussed below).

Prior to analyzing the data, a researcher must indicate (1) how many factors are present in the instrument, (2) which items are related to each factor, and (3) whether the factors are correlated or uncorrelated. These issues are typically revealed (or specified) through EFA. In the current study, the primary goal of the analyses was to confirm the five-factor structure underlying the questionnaire data by conducting a CFA using a new sample of participants and the revised questionnaire. However, I also compared the fit of the five-factor structure model to the fit of the four-factor structure model; in the four-factor model, Aggression towards People and Aggression towards Animals are collapsed into a single factor (Aggression).

Items related to each factor were determined in Study 2 and the items predicted to load on each factor were specified prior to CFA. In the original EFA of the full dataset in Study 2, I forced varimax rotation, which focuses on "cleaning up" the factors so that they have high correlations with one set of items and minimal correlation with other items (Stevens, 1996). Thus I included the assumption that the factors are uncorrelated in one model. However, as discussed in Study 2, the challenge in deciding between the four- and five-factor models was due to a relationship between the items that load on the five-factor model's Aggression towards People and Aggression towards Animals. Thus, an alternative model in which Aggression towards People and Aggression towards Animals were correlated was also fit. In addition, a model in which Fearfulness and Aggression towards People were related was also fit, because these two factors have been suggested to be related and to lead to fear-based aggression. And, finally, a model including both correlations was also fit.

Before proceeding with the CFA planned in Study 3, I assessed the adequacy of the sample for use in factor analysis. The KMO associated with the correlation matrix for the data in Study 3 is .926, greater than the cut-off of .6 for the data to be acceptable for factor analysis (e.g., Tabachnick & Fidell, 2001).

As in Study 2, I used AMOS to conduct SEM and took advantage of the opportunity to compare the goodness-of-fit indices of various models. Again, I relied on and compared multiple indices to evaluate model fit. Some fit indices are better suited to the current situation than others. (For a more detailed discussion of the goodness-of-fit indices, see Part 2 of Study 2.)

Chi-square is a very popular fit index, but it is very sensitive to sample size and prone to rejecting adequately fitting models if the sample size is large. I did not use chi-square to evaluate fit, but I reported it because it is still commonly used and reported.

I depended more heavily on the other test statistics – the RMR, GFI, NNFI, CFI, and RMSEA – to determine which five-factor model is the best fit and whether the best five-factor model is a better fit than the best four-factor model. The most weight was given to the RMSEA, because it is generally favored for large samples. The RMSEA represents the covariance that is not explained by a model (e.g., Recklitis et al., 2006). The NNFI and CFI, incremental indices of fit suitable for comparing the fits of different models, are also considered useful for larger samples. Each fit index examined and its standard cut-off is listed in Table 6.3 below. (See the previous chapter for further discussion of these indices and their cut-offs.)

Table 6.3 presents the standard cut-offs for model fit along side the fit indices associated with six tested models: the four-factor model with no factors correlated, the four-factor model with two factors correlated (Aggression with Fearfulness), the five-factor model with no factors correlated, the five-factor model with two factors correlated (Fearfulness and Aggression towards People), the five-factor model with a different two factors correlated (Aggression towards People with Aggression towards Animals), and the five-factor model with both pairs of factors correlated (Aggression towards People with Aggression towards Animals, Aggression towards People with Fearfulness). These six different models were fit because, along with previous hypotheses, the modification indices associated with the simplest four- and five-factor models suggested the additional factor correlations would substantially improve the fit of the models. It should be noted that the modification indices concur with previous analyses and hypotheses to indicate

that correlations between the other factors in each model would not improve model fit; that is, no other correlations between factors improve the models' fits substantially.

According to the RMSEA indices, all five models are adequate fits; all five have RMSEA values less than .08. The RMSEA values associated with the four-factor models suggest the four-factor models may be slightly better fits than the five-factor models. However, examination of the other fit indices (RMR, GFI, CFI, NNFI) indicates that the five-factor models are better fits than the four-factor model.

As shown in Table 6.3, the best model fits do appear to be those in which the Aggression towards People and Fearfulness factors and the Aggression towards People and Aggression towards Animals factors are, respectively, correlated. This is true in comparing four-factor models to each other and in comparing five-factor models to each other.

Based on the CFA and theoretical bases, the five-factor model with both sets of correlations was preferred. Appendix D presents a table of the loadings associated with this model, separated by factor, with items listed in numerical order down the left column and regression weights, or loadings, in the right column.

Table 6.3. Confirmatory factor analyses: Comparative model fits for Study 3 full sample

Four-factor models					
Fit Index	**Model fit estimate**				**Standard cut-off**
	factors uncorrelated	**F corr A**	**AP corr AA**	**F corr A; AP corr AA**	
Chi-squared	56,548.080, df = 5,053, p < .001	56,064.216, df = 5,052, p < .001	n/a	n/a	n.s.
RMR	.340	.315			smaller indicates better fit
GFI	.603	.604			>.85
CFI	.483	.488			>.9; larger indicates better fit
NNFI	.460	.465			>.9; larger indicates better fit
RMSEA	.063 (lo = .063, hi = .064)	.063 (lo = .062, hi = .063)			<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr A = .476			

Table 6.3. (*Continued*)

Five-factor models					
Fit Index	**Model fit estimate**				**Standard cut-off**
	factors uncorrelated	**F corr AP**	**AP corr AA**	**F corr AP; AP corr AA**	
Chi-squared	53,180.079, df = 5,054, p < .001	52,326.145, df = 5,053 p < .001	52,445.705, df = 5,053, p < .001	51,690.174, df = 5,052, p < .001	n.s.
RMR	.354	.326	.337	.309	smaller indicates better fit
GFI	.623	.625	.628	.629	>.85
CFI	.517	.525	.524	.532	>.9; larger indicates better fit
NNFI	.492	.501	.499	.507	>.9; larger indicates better fit
RMSEA	.061 (lo = .061, hi = .062)	.061 (lo = .060, hi = ..061)	.061 (lo = .060, hi = .061)	.060 (lo = .060, hi = .061)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr AP = .596	AP corr AA = .558	F corr AP = .528 AP corr AA = .475	

Note. F = Fearfulness, A = Aggression, AP = Aggression toward People, AA = Aggression toward Animals. The abbreviation "n/a" indicates that statistics are not applicable to the model in that column, and "n.s." indicates "not significant".

PART 2: SPLIT-SAMPLE EFA AND CFA

For a factor solution to be generalizable, it is vital that it is stable. In Part 2 of Study 3, my goal was to gain information about whether the five-factor model of the revised 102 items is stable. To do this, I divided the participant sample in half, then performed an EFA on one half of the participants' ratings of their dogs. I used SEM to perform CFA on the second half of the participants' ratings of their dogs. That is, I used SEM to fit the factor solution derived from the EFA of the first half of the data to fit that factor solution to the second half of the data.

EFA

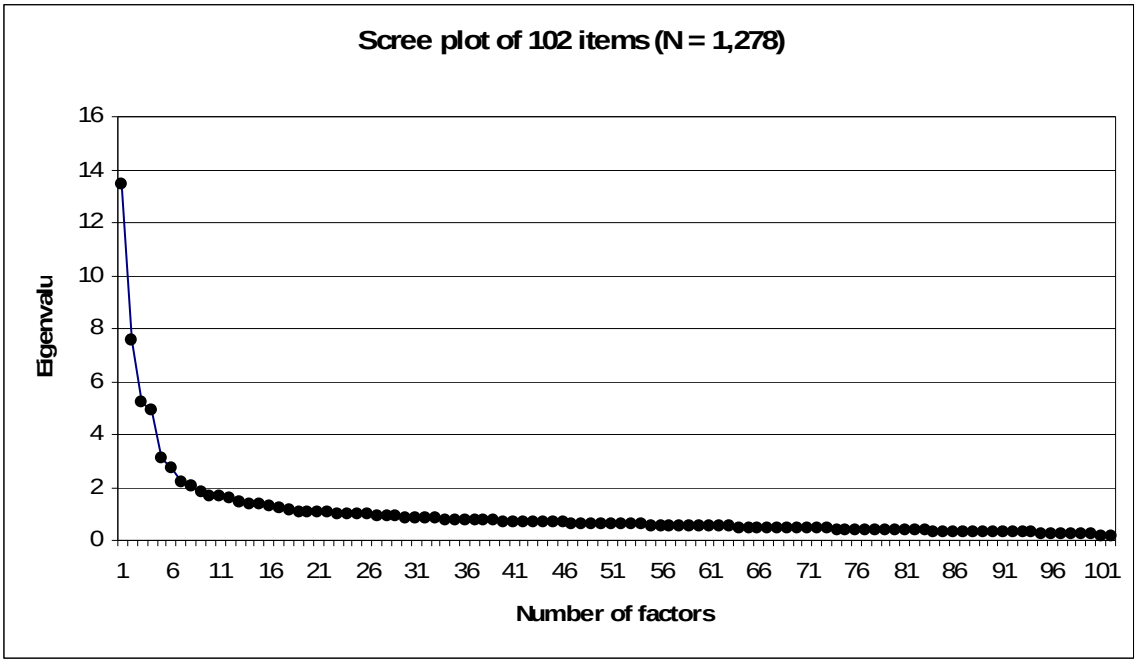
The EFA was performed using the same basic procedures as in Parts 1 and 2 of Study 2; I used PCA with varimax rotation. In Study 3, however, participants were prohibited from leaving items blank, so there were no missing data points.

Once again, before proceeding with the EFA, I needed to assess the sampling adequacy to make certain that the data were amenable to EFA. This procedure needed to be repeated because the sample in Part 2 of Study 3 was only half of the sample for which sampling adequacy was assessed in Part 1 of Study 3. In Part 2 of Study 3, the ratio of participants to items is relatively high – it is slightly greater than 12.5:1. However, having an adequate sample size does not ensure a sample is appropriate for factor analysis. A second issue is the magnitude of correlations between variables, as displayed in the correlation matrix, is also important. Because the Study 3 data were split in half for the analyses in this phase, the correlation matrices for Part 2 of Study 3 include different participants and are likely slightly different from the matrix associated with the whole sample in Part 1 of Study 3. Because the ratio of participants to items in Part 2 of Study 3 remains greater than 5:1, I relied on the KMO measure of sampling adequacy. Conservative cut-offs suggest that a KMO of at least .6 is adequate (e.g., Tabachnick & Fidell, 2001). The KMO associated with the correlation matrix resulting from the first half the data set for Study 3 (N = 1,278) is .910, indicating that the matrix was adequate for the analyses.

As in Part 2 of Study 3, I used the convergence of three criteria to determine the number of factors in the solution: graphical scree test (Cattell, 1966), the top-down method in which correlations between orthogonal factors scores from different factor solutions are viewed as path coefficients in a hierarchical structure (Goldberg, 2006), and interpretability of the solutions (see Zwick & Velicer, 1986; Worthington & Whittaker, 2006). Results from each method and suggested factor models for Part 2 of Study 3 are elaborated below.

Cattell's (1966) now classic scree test utilizes a graphical interpretation of the scree plot of the eigenvalues associated with the factors. The number of factors that this method indicates should be retained is the number of eigenvalues appearing before a visible break or elbow in the plot, after which the values tend to level off horizontally. Interpretation of the scree plot suggested retaining four or five factors; there is a break in the values after four, and an elbow or bend appears after five.

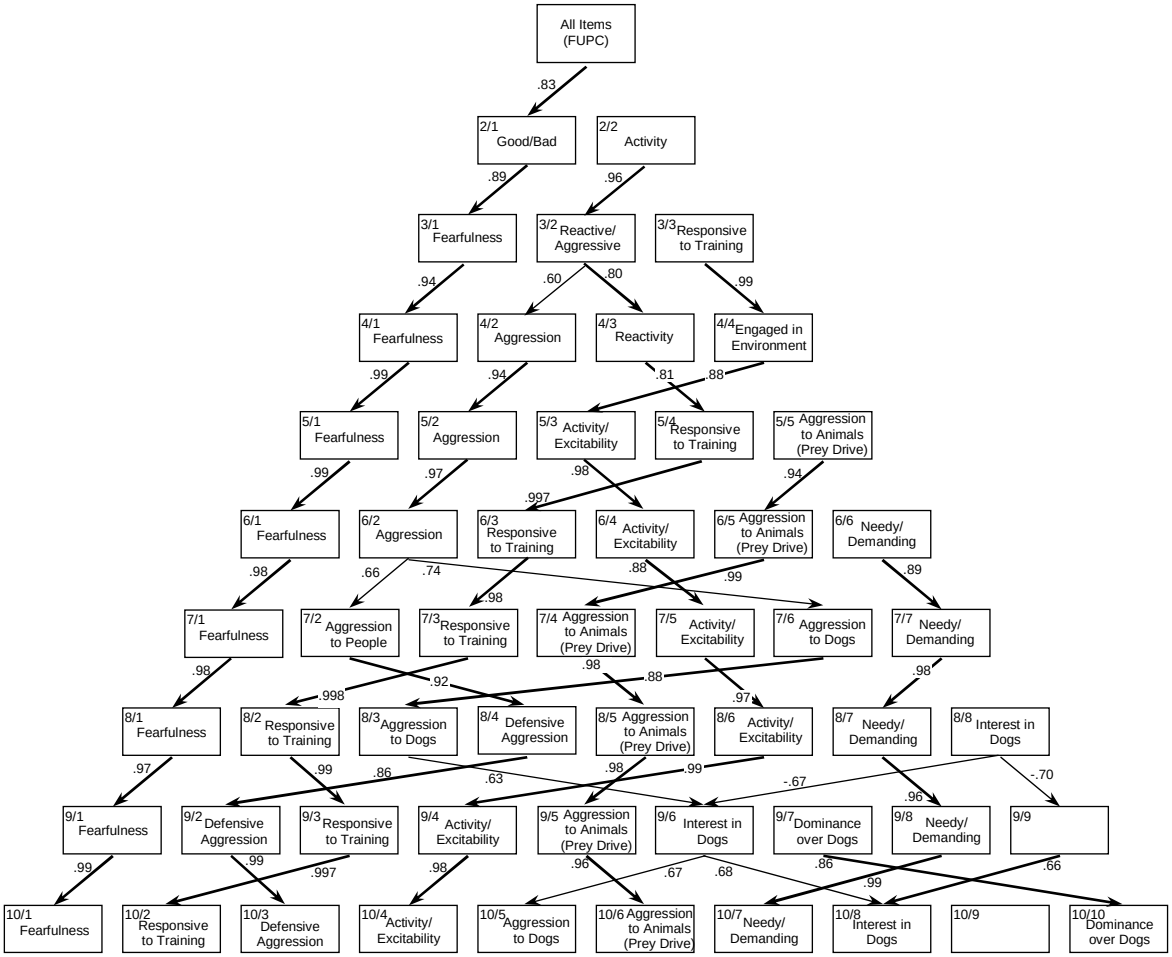
Figure 6.1. Scree plot of eigenvalues for 102 items in Study 3



Goldberg's (2006) hierarchical tree analysis, the top-down method in which correlations between orthogonal factors scores from different factor solutions are viewed as path coefficients in a hierarchical structure, is the next criterion I examined. In the hierarchical tree, correlations between orthogonal factors scores from different factor solutions are viewed as path coefficients in a graphic hierarchical structure. Using this criterion, the best factor solution is found where the major factors stop breaking apart when additional factors are extracted. (More details of this methodology can be found under Part 1 of Study 2.)

The hierarchical tree is presented in Figure 6.2. To ease interpretability, only correlations of .60 or greater were included. As a result, some factors are not shown to be correlated with factors from the previous (higher, or parent) level of the tree. To further facilitate interpretability, each factor was assigned a descriptive label, which is presented in the box representing that factor. These labels were generated by having two personality experts examine the factor loadings for every solution (i.e., the two-factor solution, the

Figure 6.2. Hierarchical structure of factor score correlations in 102-item Study 3 data set



Note. For clarity, only correlations of .60 and greater are included in the hierarchical tree. Correlations of .80 or greater are indicated by boldface arrows.
Unlabeled boxes indicate difficult-to-name factors composed of 4 or fewer items.

three-factor solution, etc.) and then examine the relationships between the factor structures as presented in the hierarchical tree.

Based purely on the correlations between factor scores at different levels of the hierarchical tree, the four-factor model appears to be supported. The factors that arise in the four-factor model correlate strongly (≥ .81) with the first four factors in the five-factor solution (5/1, 5/2, 5/3, 5/4), which in turn correlate very strongly (≥ .97) with the first four factors in the six-factor solution (6/1, 6/2, 6/3, 6/4). That is, according to the correlations alone, these four factors appear to be consistent through the lower (or child) levels of the tree. However, the factors that arise in the five-factor solution also correlate strongly (≥ .94) with the first five factors of the six-factor solution (6/1, 6/2, 6/3, 6/4, 6/5), suggesting that the five-factor solution might also be viable. So, the hierarchical tree structure can be interpreted as supporting the four-factor solution, or possibly the five-factor solution, if only the correlation of the factor scores (and not the interpretability of the factors) are considered.

When the cohesiveness and interpretability of the factor solutions presented in each level of the hierarchical tree is also considered, as is required in order to assign labels to each factor in the tree, the four-factor solution appears less viable. This is because the distribution of items in the four-factor solution's third and fourth factor (4/3, 4/4) makes those factors difficult to interpret. The third factor appears to relate to dogs' reactivity, activity, and obedience, with items such as "Dog is very excitable around other dogs" (.576), "Dog ignores commands" (.561), and "When off leash, dog comes immediately when called" (-559), "Dog is excitable when visitors arrive" (.501), "Dog is boisterous" (.464) loading strongly on the factor. Multiple of these items also load strongly on the fourth factor in the solution (e.g., "Dog ignores commands" (-.409), "When off leash, dog comes immediately when called" (.398), "Dog is boisterous" (.357)). Items that load on the fourth factor seem to have to do with how engaged and interested the dog is in the environment. For example, items loading strongly on the factor (4/4) include "Dog enjoys playing with toys" (.563), "Dog is attentive to owner's actions and words" (.479), and "Dog is very alert" (.506). Again, there is cross-loading

such that some items on the fourth factor also load strongly onto the third factor (e.g., “Dog is attentive to owner’s actions and words” loads -.449 on factor (4/3)).

In the five-factor solution, the factors become more cohesive and interpretable: the items group together in ways that are easy to interpret, and fewer items cross-load strongly. Factor (5/1) is clearly related to Fearlessness. Factor (5/2) is related to Aggression towards People and Dogs. Factor (5/3) consists of items related to activity level, playfulness, and excitability, such as “Dog enjoys playing with toys” (.632), “Dog seeks constant activity” (.530), and “Dog is lethargic” (-523). Factor (5/4) consists of items related more exclusively to training, obedience, and how readily a dog is trained, including “Dog ignores commands” (.678), “Dog is attentive to owner’s actions and words” (-.648), and “Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs)” (-.526). The final factor derived in the five-factor solution relates to dogs’ tendency towards chasing and aggressing towards animals other than dogs or people, and might be most easily interpreted as prey drive. Items loading on this factor include “Dog is very excitable when around squirrels, birds, and other small animals” (.728) and “Dog likes to chase cats” (.696). Conceptually, this final factor (5/5) may be thought of as a facet of Aggression but might also relate to how reactive, excitable, and generally active a dog is. Indeed, factor (5/5) Aggression towards Animals is most correlated with factor (4/3) Reactivity (.470), and is secondarily correlated with (4/2) Aggression (.265).

Across the three criteria, the supported factor solutions are the four- and the five-factor solutions. Based on the ultimate criteria of interpretability, the five-factor solution is favored. However, it is noteworthy that the five-factor solution here differs from the five-factor solutions in Study 2. In Parts 1 and 4 of Study 2 Aggression is divided into Aggression towards Dogs and Aggression towards People. In the EFA performed in Part 2 of Study 3, Aggression is divided into Aggression (towards People and Dogs) and Prey Drive, which is also shown to be strongly related to dogs’ tendencies towards reactivity or excitability and activity. This might suggest that the four-factor solution is more favorable, however, the difficulty of interpreting the third and fourth factor in the four-

factor solution prohibit its selection. Thus, the factor solution selected for the CFA in Study 3 is composed of five factors: Fearfulness, Aggression, Activity/Excitability, Responsiveness to Training, and Prey Drive.

CFA

Many of the details and much of theory behind performing CFA using SEM were described in the previous chapter's discussion of Part 2 of Study 2. Details are reviewed here as useful in considering the CFA performed in Part 3 of Study 3.

In the current study, the primary goal of the analyses is to confirm the five-factor structure underlying the questionnaire data in Study 3 is stable by conducting a CFA using the second half of the Study 3 participant sample and their responses to the 102 items included in Study 3.

As stated above, using SEM to perform CFA requires a large number of participants per parameter. Exactly how many participants per parameter is a "large" number is unclear, because it depends on the characteristics of the model being fit (e.g., items per factor, normality of the data, strength of the relationship between the items and latent factors). However, a very conservative cut-off is indicated that 15 participants per item is a good guideline (Stevens, 1996), and more conventional rules of thumb recommend a range from five to 10 participants per parameter (e.g., Worthington & Whittaker, 2006). In the current CFA, as in the EFA in this part of Study 3, there are 12.5 participants per parameter (N = 1,278, parameters = 102).

As in the EFA described above, I assessed the assessed the adequacy of the sample (using KMO) for use in factor analysis. The KMO associated with the correlation matrix for the data in Study 3 is .914, indicating that the matrix was adequate (e.g., Tabachnick & Fidell, 2001).

Prior to conducting the CFA using SEM, a researcher must indicate (1) how many factors are present, (2) which items are related to each factor, and (3) whether the factors are correlated or uncorrelated. In this analysis, the factors labeled as Fearfulness, Aggression, Activity/Excitability, Responsiveness to Training, and Prey Drive, derived in

the immediately previous EFA on the first half of the Study 3 data, were examined thorugh CFA. The items loading onto each of these factors were determined through the EFA in Part 2 of Study 3. As discussed in previous sections, there are reasons to suspect correlations between Aggression towards People and Fearfulness, and between Aggression and Prey Drive, so models including these correlations were also fit.

Results

According to RMSEA, all four version of the five-factor model are adequate fits to the data. However, the comparative and other fit indices (RMR, GFI, CFI, NNFI) all improve when correlations between the specified factors are added. It should be noted that modification indices do not suggest significant gains in model fit would be incurred if more factor correlations were added to the model.

Table 6.4. Confirmatory factor analysis of second half of Study 3 data

Five-factor models					
Fit Index	**Model fit estimate**				**Standard cut-off**
	factors uncorrelated	**F corr A**	**A corr Prey Drive**	**F corr A; A corr Prey Drive**	
Chi-squared	27,707.146, df = 5,049, p < .001	27,431.718, df = 5,048, p < .001	27,379.350, df = 5,048, p < .001	27,304.206, df = 5,047, p < .001	n.s.
RMR	.326	.298	.292	.285	smaller indicates better fit
GFI	.618	.621	.622	.623	>.85
CFI	.548	.554	.554	.556	>.9; larger indicates better fit
NNFI	.499	.504	.506	.507	>.9; larger indicates better fit
RMSEA	.059 (lo = .058, hi = .060)	.059 (lo = .058, hi = .060)	.059 (lo = .058, hi = .059)	.059 (lo = .058, hi = .059)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr A = .578	A corr Prey Drive = .564	F corr A = .533, A corr Prey Drive = .528	

Note. F = Fearfulness, A = Aggression. The abbreviation “n/a” indicates that statistics are not applicable to the model in that column, and “n.s.” indicates “not significant”.

PART 3: CONFIRMATION OF THE FIVE-FACTOR, 15-FACET MODEL

The ultimate goal of Study 3 was to devise shorter, more usable forms of the DPQ. To do this, a final revision of the item pool was required. In this final revision, the items were to be evaluated based on multiple criteria, including their loadings on both facets and factors. In order to examine item loadings on each facet, a full hierarchical model in which the 15 facets derived in Study 2, in addition to the five factors, were fit was required.

Table 6.5. Confirmatory factor analyses of full hierarchical model (factors and facets)

Fit Index	**Model fit estimate**				**Standard cut-off**
Five-factor models					
	Factors uncorrelated	**F corr AP**	**AP corr AA**	**F corr AP; AP corr AA**	
Chi-squared	40,655.301, df = 5,032, p < .001	39,826.512, df = 5,031, p < .001	39,964.957, df = 5,031, p < .001	39,208.514, df = 5,030, p < .001	n.s.
RMR	.323	.292	.306	.274	smaller indicates better fit
GFI	.715	.716	.721	.722	>.85
CFI	.642	.615	.649	.657	>.9; larger indicates better fit
NNFI	.612	.620	.619	.626	>.9; larger indicates better fit
RMSEA	.053 (lo = .052, hi = .053)	.052 (lo = .052, hi = .053)	.052 (lo = .052, hi = .053)	.052 (lo = .051, hi = .052)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr AP = .616	AP corr AA = .612	F corr AP = .556, AP corr AA = .536	

Note. F = Fearfulness, AP = Aggression towards People, AA = Aggression towards Animals. The abbreviation "n/a" indicates that statistics are not applicable to the model in that column, and "n.s." indicates "not significant".

As indicated by the fit indices displayed in Table 6.5, these models incorporating the 15 facets are improvements over the previous models according to RMR, GFI, CFI, NNFI, and RMSEA. The five-factor, 15-facet model with correlations between Fearfulness and Aggression towards People, and between Aggression towards People and Aggression towards Animals, was deemed best and used for further analyses pertaining to item selection.

PART 4: BUILDING THE DPQ LONG AND SHORT FORMS

Some of the items in the 102-item questionnaire performed better than others. In building the long (75-item) and short (45-item) forms of the questionnaire, I aimed to retain the best items, striking a balance among the sometimes conflicting criteria of strength of loadings on factors and facets, content validity, utility, discriminant and convergent validity, and internal consistency. I retained five items for each of the 15 facets for the long form. For the short form, I retained a subset of the items on the long form such that the short form contains three items per facet.

Factor and facet loadings

The first criterion I examined was the magnitude of items' loadings on their predicted (and primary) factors and facets. Items that loaded weakly (absolute value ≤ .500) on their intended factor and/or facet were seen as candidates for deletion from the questionnaire, but were not deleted prior to consideration of the other criteria (see below). The loadings resulting from the best-fitting models were used for this purpose. That is, the regression loadings from CFA of the five-factor model with Fearfulness correlated with Aggression towards People, and with Aggression towards People correlated with Aggression towards Animals, were used for the factor loadings. The regression loadings from CFA of the five-factor, 15-facet model with the same factor correlations were used for the facet loadings. The items with absolute loadings less than or equal to .500 on their primary facet or factor were identified. There were 16 items with absolute loadings less than or equal to .500 on their respective factors, and 14 with absolute loadings less than or equal to .500 on their respective facets. Eight of these items (number 2, 21, 29, 40, 43, 78, 95, and 101 from the Study 3 item pool) overlapped, showing absolute loadings of less than or equal to .500 on both their factor and facet. Thus, a list of 22 items suggested for deletion resulted from analysis of item loadings.

In removing items, attention was also paid to attempting to retain items with both positive valence (positive loadings) and negative valence (negative loadings). That is, I tried to retain items pertaining to aggression (+), lack of aggression (-), and friendliness (-) on Factor 2: Aggression towards People.

Content validity

Another important issue to consider in removing items from the questionnaire to create the long and short forms was that high content validity is maintained. That is, the items that were retained needed to represent diverse aspects of each facet of dog personality. By making certain that each facet was well represented, each trait was to be well-represented. The short (45-item) form of the questionnaire is an abbreviated version of the long questionnaire and thus less comprehensive, but the content represented by the items included is still very important.

Item utility

Third, I tried to retain items with high utility, or those that address specific behaviors that might be of great interest. For example, people who might use the DPQ would likely be more interested in learning how aggressively a dog behaves towards unfamiliar people in general than in the specific context of a person approaching the dog's house or yard. In the long form, items asking about both behaviors were retained as part of the General Aggression facet of the Aggression towards People factor. Based on item utility, the item pertinent to the specific situation of when people approach the house marking was a good candidate for deletion in creating the short form.

Discriminant and convergent validity

Another important issue to consider when shortening the questionnaire was the importance of the discriminant and convergent validity of the items. Discriminant validity is indicated by low and nonsignificant correlations with measures that assess different, unrelated constructs. It was expected, for example, that the correlation between an item assessing Fearfulness and an item assessing Activity/Excitability would be low because these two items assess different aspects of personality. However, given the relationship between Fearfulness and Aggression towards People and the relationship between Aggression towards People and Aggression towards Animals, items on each of those factors were predicted to show higher correlations with each other than with items on the other factors (e.g., Responsiveness to Training). Correlations at the item-to-item level

were predicted reflect the relationships that are evident at the factor-to-factor level. However, these items were still predicted to correlate more strongly with other items within the same factor than with other items on a related factor.

I correlated each of the items separately to examine whether items that load on different factors (e.g., Aggression towards Animals, Activity/Excitability) show weaker correlations with each other than items that load on the same factor (i.e., only Aggression towards Animals) or the same facet of a factor. Convergent validity is indicated by large and significant correlations between theoretically related constructs. For example, a relatively large correlation was expected between two items purported to assess aggression towards unfamiliar people, and between two items purported to assess activity level during play. Convergent validity correlations, or correlations between items that load on a single factor, are presented in Appendix E. Items showing poor convergent validity correlations were also likely candidates for deletion when building the long and short forms of the DPQ. For example, the item "When alone or about to be left alone, dog shakes, shivers, or trembles" (Item 29 in Study 3) was considered a good candidate for deletion, and ultimately was deleted, because of low convergent validity correlations ($\leq$ .282). However, in some cases items with poor convergent validity correlations (e.g., Item 21 in Study 3, from the Aggression towards Animals factor, "Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.") were retained to keep five items per factor or because they were important to maintaining comprehensive content validity.

Correlations that were expected to be low, or discriminant correlations between items, are presented in Appendix F. Items showing relatively poor discriminant validity were candidates for deletion when building the long and short forms of the DPQ, but these items were not always removed. For example, the Fearfulness item "Dog behaves fearfully towards unfamiliar people" (Item 1 in Study 3) shows correlations ranging from .029 to .640 with items that load on the Aggression towards People factor, which could be seen as consistent with the correlation between the two factors. However, due to content validity concerns, this item was ultimately retained.

Furthermore, items that showed both poor discriminant and poor convergent validity (e.g., Item 45 in Study 3, "Dog becomes very excited when owner returns home," were considered particularly good candidates for deletion when the DPQ was shortened).

Internal consistency

It is also important that each scale (or the items measuring each factor) maintain as much internal consistency, or reliability across items within a scale (measured with Cronbach's alpha), as possible while maximizing compliance with the other criteria. This aspect of reliability is a measure of the extent to which items that load on a single factor or scale measure the same underlying construct (Cronbach, 1951). If an item was suggested for deletion based on its loadings, its relevance to content validity, its divergent validity, or its convergent validity, before the item was removed I examined what impact the item's removal would have on internal consistency. Because Cronbach's alpha tends to increase as the number of items in a scale (or factor) increases, factors in the 102-item questionnaire were expected to have higher internal consistency than the long (75-item) form, and factors on the long form of the questionnaire are predicted to have higher internal consistency than corresponding factors in the short (45-item). Ultimately, as shown in Table 6.6, this was the case for most of the factors. However, the internal consistency associated with Fearfulness increased with the removal of items from the long form to create the short form. Presumably, the removal of eight items to create the short form from the long form decreased the amount of conceptual variability, and probably content validity, of the Fearfulness items when taken together.

Table 6.6. Scale reliabilities

Trait	DPQ Form		
	Study 3 questionnaire (102 items)	**Long form** (75 items)	**Short form** (45 items)
Fearfulness	.911 (29 items)	.882 (20 items)	.838 (12 items)
Aggression towards People	.852 (15 items)	.814 (10 items)	.742 (6 items)
Activity/Excitability	.818 (27 items)	.772 (20 items)	.728 (12 items)
Responsiveness to Training	.822 (14 items)	.797 (10 items)	.771 (12 items)
Aggression towards Animals	.836 (17 items)	.823 (15 items)	.748 (9 items)

The internal consistency of the resulting long and short forms of the questionnaire is comparable or slightly higher than those found in other studies of dog personality (see Chapter 2). Given the amount of variability generally covered in each factor, each of which have at least two facets, these alpha levels are even more impressive. As John and Benet-Martínez (2000), explain, a factor showing "high" Cronbach's alphas of .90 or above are less impressive if the items on that factor address only a narrow range of characteristics (e.g., aggression towards unfamiliar men) than if the items address a broad array of related characteristics (e.g., aggression towards unfamiliar people, aggression towards familiar people, aggression when threatened, aggression without apparent reason).

THE DOG PERSONALITY QUESTIONNAIRE FINAL FORMS

A listing of the items that were retained or removed for each facet and factor is presented in Appendix G. The final long form and short form of the questionnaire are presented in Appendices H and I. Each questionnaire contains items representing the 15 facets and five factors. The long form of the questionnaire contains five items per facet, for a total of 20 Fearfulness items, 10 Aggression towards People items, 20 Activity/Excitability items, 10 Responsiveness to Training items, and 15 Aggression towards Animals items. The short form contains three items per facet, for a total of 12 Fearfulness items, six Aggression towards People items, 12 Activity/Excitability items, 12 Responsiveness to Training items, and nine Aggression towards Animals items.

Once all of the items were selected based on the criteria described above, I used SEM to perform CFAs assessing the fit of the five-factor model to the new DPQ forms. The Results are presented in Tables 6.7 and 6.8, below. In keeping with the previous analyses, I fit models with all or only one pair of the predicted factor pairs correlated. Again, the models in which correlations between Fearfulness and Aggression towards People and between Aggression towards People and Aggression towards Animals are specified showed the best fits according to all fit indices. The model with more indicators – the 75-item Long Form – was associated with a better fit.

Table 6.7. 75-item DPQ confirmatory factor analyses

Five-factor model					
Fit Index	**Model fit estimate**				**Standard cut-off**
	Factors uncorrelated	**F corr AP**	**AP corr AA**	**F corr AP; AP corr AA**	
Chi-squared	35,088.121, df = 2,705, p < .001	34,218.193, df = 2,704, p < .001	34,413.082, df = 2,704, p < .001	33,667.269, df = 2,703, p < .001	n.s.
RMR	.388	.356	.366	.334	smaller indicates better fit
GFI	.661	.663	.667	.669	>.85
CFI	.544	.556	.554	.564	>.9; larger indicates better fit
NNFI	.525	.536	.534	.544	>.9; larger indicates better fit
RMSEA	.068 (lo = .068, hi = .069)	.068 (lo = .067, hi = .068)	.068 (lo = .067, hi = .068)	.067 (lo = .066, hi = .068)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr AP = .621	AP corr AA = .547	F corr AP = .550, AP corr AA = .455	

Note. F = Fearfulness, AP = Aggression towards People, and AA = Aggression towards Animals. The abbreviation “n/a” indicates that statistics are not applicable to the model in that column, and “n.s.” indicates “not significant”.

Table 6.8. 45-item DPQ confirmatory factor analyses

Five-factor model					
Fit Index	**Model fit estimate**				**Standard cut-off**
	Factors uncorrelated	**F corr A**	**AP corr AA**	**F corr A; AP corr AA**	
Chi-squared	17,617.700, df = 948, p < .001	16,712.292, df = 947, p < .001	17,031.717, df = 947, p < .001	16,336.029, df = 946, p < .001	n.s.
RMR	.429	.385	.408	.366	smaller indicates better fit
GFI	.724	.731	.735	.738	>.85
CFI	.561	.585	.576	.595	>.9; larger indicates better fit
NNFI	.548	.571	.563	.581	>.9; larger indicates better fit
RMSEA	.083 (lo = .082, Hi = .084)	.081 (lo = .080, h = .082)	.082 (lo = .080, hi = .083)	.080 (lo = .079, hi = .081)	<.08 = adequate fit <.05 = good fit
Factor Correlations	n/a	F corr AP = .674	AP corr AA = .542	F corr AP = .593, AP corr AA = .413	

Note. F = Fearfulness, AP = Aggression towards People, and AA = Aggression towards Animals. The abbreviation "n/a" indicates that statistics are not applicable to the model in that column, and "n.s." indicates "not significant".

As discussed in the above section on internal consistency, the long and short forms of the questionnaire maintained relatively high internal consistency. Other psychometric properties of the long and short forms of the DPQ, including inter-rater reliability, test-retest reliability, and predictive validity, are presented in the following chapters.

SUMMARY OF FINDINGS

In Study 3, I used SEM to perform CFA using data collected from 2,556 new participants' ratings of their dogs on the 102-item questionnaire to confirm that the best-fitting model was the five-factor model including correlations between factors 1 and 2 (Fearfulness and Aggression towards People) and factors 2 and 5 (Aggression towards People and Aggression towards Animals). As in Study 2, the stability of the five-factor model was also confirmed using a split-halves procedure in which I divided the participant set into two randomly selected halves, repeated the EFA procedure on one

half of the data (N = 1,278), then repeated the CFA procedure on the second half of the data (N = 1,278).

So that researchers and practitioners wishing to assess dog personality using the DPQ have a choice of tools to meet their needs, a more comprehensive long form and a briefer short form of the DPQ were developed. The long form consists of 75 items, and the short form has 45. Items for each form were selected to maximize the forms' psychometric properties (e.g., content validity, construct validity, including discriminant and convergent validity; internal consistency), and so that each facet has an equal number of items. The final forms are presented in Appendices H and I.

CHAPTER 7

Study 4: Inter-Rater Reliability

INTRODUCTION

As mentioned above, an assessment tool must demonstrate reliability across raters. If an assessment tool does not have high inter-rater reliability, then the ratings probably reflect idiosyncrasies of a particular observer, not attributes of the target whom the questionnaire purports to assess. The goal of Study 4 was to evaluate the inter-rater reliability of the long and short forms of the DPQ. To achieve this goal, I administered the long form, and therein the short form, of the questionnaire online to a group of 99 pairs of participants, with both individuals of each pair rating the same target dog. I assessed the short form of the questionnaire by extracting responses to items included on the short form from the data collected with the long form.

METHODS

Participants

Participants for Study 4 were all newly recruited participants. Pairs of participants who were familiar with the same target dog were recruited for Study 4. Pairs of participants signed up to participate by responding to online postings on the Animal Personality Institute website (www.animalpersonality.org) and dog-related Internet discussion boards or newsletters. E-mail groups were also given permission to recirculate postings through-out their groups. These online ads described the study's goal as evaluating how much people tended to agree when independently rating dogs on the DPQ, and emphasized that agreement is affected by various aspects of the questionnaire itself in addition to other factors. It was emphasized that participants would need to complete the questionnaire without talking to each other before or while they completed the questionnaire to ensure independence of the ratings. Participants completed an online version of the questionnaire between June and August of 2007. Participants received no compensation for their participation, but were given feedback about their dogs'

personalities with respect to the five personality factors (and their facets) assessed in the DPQ. A total of 213 participants completed the questionnaire during this time period; initial analyses (described below) refined the sample to 99 pairs of participants rated the same dog.

Materials and Procedures

In Study 4, participants completed the long form of the questionnaire (derived in Study 3) and were asked to provide the same demographic information about themselves and their dogs that participants provided in the previous studies. The questionnaire was presented online, following an introduction page that described the basic goals of the study and again emphasized the importance of the participants completing the questionnaire independently. Participants were also told that, at the end of the questionnaire, they would each receive summaries of how they had rated the target dog, so they could compare their ratings if they wished.

As in Study 3, participants were allowed to choose which of their dogs they rated, if they had more than one dog. The second participant in each pair was instructed to rate the same dog the first participant had rated. Participants were also asked whether the dog they were rating had previously been or would be rated by another participant, and, if so, were asked to provide that participant's first name. Unlike in the previous studies, the order in which questionnaire items were presented was not uniquely randomized for each participant in Study 4; in order to evaluate the reliability of the instrument as it would be administered in the paper-and-pencil format, items were presented in the same order as on the paper form of the DPQ's long form (see Appendix H). Participants were not allowed to leave answers blank.

As in previous studies, I refined the sample and then examined the demographic data in Study 4 to assure that the participants and the target dogs compose diverse samples.

RESULTS

First, I removed cases in which only a single participant had filled out the questionnaire about a given dog and cases in which single participants had filled out the questionnaire multiple times about a given dog. I identified these cases by matching cases in terms of the participants' names, dogs' names, and other demographic information (e.g., location, dog breed). This process reduced the dataset from 213 participants to 198 participants, all of whom were paired. In each of the 99 pairs that were retained, the participants had listed another participant in the dataset as the other individual who would rate or had rated a given dog. That is, if Ellen rated a male Labrador Retriever named Gunner and said that Rob would also rate Gunner, then there was another case in which Gunner was rated by Rob, and Rob said that Ellen had rated Gunner.

Second, as in previous studies, I examined the frequency of responses for each questionnaire item to check for obvious errors in saving the data. No anomalies were found in the dataset, though, as in previous studies, responses to each item were not regularly distributed (i.e., responses to some items were skewed).

Third, I examined each case for anomalies and problems of data omission (e.g., participants who had withheld their location). No problems were found, and so all cases were retained.

Sample demographics

Demographic information for the 198 participants who compose the 99 pairs of participants in Study 4 is presented in Table 7.1. Aside from the listings specific to the pairing of sexes (e.g., "Pairing of sexes," "Average difference in age between raters"), all entries in Table 7.1 treat participants separately (i.e., not as pairs).

As in previous studies, both sexes of dog owners were represented in Study 4. Again, men composed the minority (37.9%, or 75 men), though they composed a larger percentage of the sample in this study than in previous studies. In this sample, participants were paired, and so I also examined the composition of the pairs of participants. Only six of the 99 pairs were composed of two men, 30 of the pairs consisted of two women, and 63 pairs consisted of one man and one woman. (No

information regarding the relationships between raters was collected.) The average age of participants was 47 years, with a standard deviation of 13 years. The average difference in age between raters was 6.2 years.

The majority of participants were currently residing in the U.S., with only two participants residing in Canada and two who indicated they were living in an unlisted country (Finland). The sample included participants from 37 states; individuals of every pair indicated that they currently resided in the same state, and most (N = 94 pairs) in the same zip code. The majority of participants identified themselves as Caucasian/White (N = 171). Other participants identified themselves as African American/Black (N = 15), Hispanic (N = 6), East Asian (N = 3), and Multiracial or Other (N = 3).

Of the 198 participants, 27 (or 13.6%) indicated having a dog-related job or career. The majority of these participants (N = 15) reported working as dog trainers or behaviorists.

Because many of the pairs of participants may share ownership of their current dogs and past dogs, statistics about their dog ownership are likely to be highly correlated. However, these numbers still reflect the individuals' experience with different dogs and dog ownership, and so they are reported in Table 7.1 in terms of individual dog owners (not pairs of dog owners). The mean number of dogs each participant had owned in his or her lifetime, not including current dogs but including childhood family dogs, was 5.9 (s.d. = 3.8). The mean number of dogs currently owned by participants in Study 4 was 1.9 (s.d. = .9).

Table 7.1. Demographic information about participants in Study 4

		Dog owners
Sex (dog owner)	Male	75 (37.9%)
	Female	123 (62.1%)
Pairing of sexes	Both raters are male	6 pairs (6.1%)
	Both raters are female	30 pairs (30.3%)
	One male, one female	63 pairs (63.6%)
Age (dog owner)	Mean	47
	standard deviation	13
	Age not reported	0
	Average difference in age between raters	6.2
Race	Caucasian/White	171 (86.4%)
	African-American/Black	15 (7.6%)
	Hispanic	6 (3.0%)
	East Asian	3 (1.5%)
	Native American	0
	Asian Indian	0
	Multiracial or Other	3 (1.5%)
Country of residence	U.S.	196 (98.0%)
	Canada	2 (1.0%)
	Australia	0
	U.K.	0
	Other (Native language is English)	2 (1.0%)
	Country of residence not reported	0
Dog-related careers	Breeder and/or Exhibitor	3 (1.5%)
	Trainer and/or Behaviorist	15 (7.6%)
	Dog rescue worker/volunteer	3 (1.5%)
	Veterinarian	0
	Groomer	0
	Kennel/Daycare owner or employee	6 (6.0%)
	Assistance dog partner	0
		Mean (s.d.)
Number of dogs owned	Past (mean; s.d.)	5.9 (3.8)
	Number not reported	0
	Currently	1.9 (.9)
	Number not reported	0
Total number of participants		**198 (99 pairs)**

As in the previous studies, it was also important to confirm that a demographically diverse sample of dogs was rated in Study 4. Because each participant in the 99 pairs rated a single dog, there are 99 dogs in this. In Study 4, all dogs' ages were reported as known, and the dogs' average age was 4.8 years (s.d. = 2.9). The dogs ranged in age from one to 12 years.

Other demographic information about the dogs is presented in Table 7.2. The number of dogs (out of the 99 in Study 4) in each category is listed in the far right column, which is entitled "Number of dogs". Although a smaller sample than the samples in previous studies, the sample of 99 dogs appeared to be diverse. Slightly more than half the dogs (N = 55, or 55.6%) were male. Ninety of the dogs (90.9%) were castrated.

Purebred dogs made up more than half of the sample (N = 59, or 59.6%), with 29 breeds represented. The breeds most represented included the Australian Shepherd (N = 5), Labrador Retriever (N = 5), German Shepherd Dog (N = 4), and Border Collie (N = 4).

Information about whether the dogs were reported to have ever bitten a person, had any reported health issues, what role they played in their owners' lives (e.g., pet, guide dog), and whether they were involved in any sports is also included in Table 7.2.

Four dogs were reported to have bitten someone, one dog had an unknown bite history, and 94 dogs were reported to have never bitten a person (intentionally, outside play, and outside work or sport requiring biting; e.g., Schutzhund training or competition). Twenty-three dogs were indicated to have health problems or disabilities, including hip dysplasia (N = 7), arthritis (N = 4), elbow dysplasia (N = 1), and other disabilities (N = 11). Other disabilities listed included, for example, a heart murmur, incontinence as a result of a spay surgery, skin allergies, chronic ear infections, and cherry eye.

The majority of the dogs in this sample (N = 95, or 95.9%) were considered pets or companions by their owners. However, there were also three working assistance dogs. One assistance dog was a guide dog, rated by his blind partner and his blind partner's seeing spouse, and the other two are medical assistance dogs (e.g., seizure alert dogs). As

in the previous studies, the popularity of agility (N = 17) and American Kennel Club show/conformation (N = 6) was also represented in this sample.

Table 7.2. Demographic information about target dogs in Study 4

			Number of dogs
Sex	Male		55 (55.6%)
	Female		45 (45.4%)
Age	Mean (s.d.)		4.8 (3.1)
Castration Status	Males	Neutered	49 (49.5%)
		Intact	6 (6.1%)
	Females	Spayed	44 (44.4%)
		Intact	3 (3.0%)
Purebred	Unknown		3 (3.0%)
	No (mixed)		37 (37.4%)
	Yes		59 (59.6%)
Bitten a Person	Unknown		1 (1.0%)
	No		94 (95.0%)
	Yes		4 (4.0%)
Disability and Health issues	Unilaterally deaf		0
	Bilaterally deaf		0
	Blind in one eye		0
	Blind in both eyes		0
	Arthritic		4 (4.0%)
	Hip dysplasia		7 (7.1%)
	Elbow dysplasia		1 (1.0%)
	Other disabilities		11 (11.1%)
Job or Role	**Pet/Companion**		95 (95.9%)
	Assistance dog		3 (3.0%)
	Guide dog		1 (1.0%)
	Hearing ear dog		0
	Medical assistance dog		2 (2.0%)
	Search and rescue		0
	Guard dog		1 (1.0%)
	Animal Assisted Therapy		6 (6.1%)
	Dam or sire for breeding		0

Table 7.2. (*Continued*)

		Number of dogs
Sport	Obedience	0
	Sledding	0
	Carting	0
	Frisbee	0
	Earth dog	3 (3.0%)
	Show/Conformation	6 (6.1%)
	Schutzhund	1 (1.0%)
	Agility	17 (17.2%)
	Herding (competitive)	1 (1.0%)
	Flyball	0
	Hunting	1 (1.0%)
Total number of dogs		**99**

Note. The sum of the number of dogs who participate in sports will not equal the total number of dogs in the study; some dogs are not in sports, and others are in multiple. Similarly, participants may consider a dog to be both a pet/companion and to have another job or role (e.g., assistance dog, search and rescue dog).

Dogs with unknown histories (e.g., bite history) are typically dogs who were adopted from shelters or rescues.

Analyses

Inter-rater reliability

The goal of the analyses in Study 4 is to measure the strength of the relationship between target dogs' paired ratings on the DPQ's individual items and on the facets and factors of both the long and short form.

A very common method of assessing the strength of relationships between variables is the Pearson correlation. The Pearson correlation can be thought of as assessing the relationship between data organized in two columns in a meaningful way. These columns might include a first and second test (as in test-retest reliability), or factors such as height and income, husbands and wives, or any other meaningful pair of variables. In the current study, the columns would be Rater 1 and Rater 2. However, because the raters were not consistent types of pairs (e.g., husband/wife), because they

filled out the questionnaire independently, and so on, there is no meaningful way to identify one rater as Rater 1 and the other as Rater 2.

Because there is no obvious way to determine which rater to place in which column or group, the Pearson correlation is not an appropriate test for assessing inter-rater reliability in Study 4. A meaningful Pearson correlation cannot be calculated because the correlation would differ if, for example, the correlation were calculated, then recalculated after half of the raters were swapped from one column to the other. A correlation coefficient that is not affected by the ordering of raters is required, such as an intraclass correlation coefficient, or ICC (Griffin & Gonzales, 1995; Shrout & Fleiss, 1979). Along with the ICC, Cronbach's alpha provides additional information about the reliability of the ratings' means (e.g., John & Soto, 2007).

The ICC is an Analysis of Variance-based (ANOVA-based) correlation. It measures the relative homogeneity within groups in ratio to total variance; the covariance of the ratings is divided by a form of the total variance to obtain this sort of correlation coefficient.

ICCs assess inter-rater reliability by comparing the variability of different ratings of the same subject to the total variation across all ratings and all subjects. There are multiple types of ICCs, representing different study models and data characteristics. Similar to ANOVA, in the case of inter-rater reliability ICC varies depending on whether the raters are thought to be a sample of a larger population or the entire population of interest, whether the targets they rate are a sample of a larger population, or the entire population, and whether reliability is measured based on individual ratings or mean ratings of all judges (Shrout & Fleiss, 1979; McGraw & Wong, 1996).

In the current study, each dog is rated by two people, and each person, or rater, rates only one dog. Because the raters are different for every dog, it is impossible to distinguish variability due to rater, so a one-way random-effects model is most appropriate; for this model, it is only possible to evaluate consistency (not absolute agreement). It is, unfortunately, also impossible to extract information about the exact source of the inter-rater reliability. Are the reliability levels that occur due to independent

experiences with the dog, discussions between the two raters, shared past experiences (e.g., previous dogs owned), or other factors? How much is due to actual variation in the dog? It is assumed that various factors will impact the ratings that each member of each pair gives, and that a substantial portion of the variability is due to the individual differences in the target dogs themselves.

Cronbach's alpha can be seen, simply, as a coefficient of reliability or generalizability. It is affected by number of items or ratings, and by the homogeneity or average correlation among the ratings. In the context of ICC as a measure of inter-rater reliability, Cronbach's alpha is an estimate of how strongly the score obtained from the actual raters would correlate with the score that would have been or could be obtained from another sample of raters. That is, the ICC alpha is an estimate of the reproducibility of the raters' mean rating.

Because the ICC and the associated alpha are both estimates of the reliability of the ratings, it is expected that they will yield parallel results. That is, items, facets, or factors associated with high ICCs will also be associated with high alphas.

The item-level ICC *r*s indicate the strength of the relationship between paired participants' ratings of the target dogs on each item. The item-level alphas indicate the reliability of the ratings, as described above, on each item. Item-level ICC *r*s and alphas are presented in Table 7.3. All 75 items included in the long form of the DPQ are presented in Table 7.3. The items are numbered as they appear on the long form. The items retained for the 45-item short form are indicated by bolded item numbers. Items that are reverse keyed are indicated with an asterisk. The items are separated into labeled groups based on their loading onto each factor, and facets within each factor. For both the short form and the long form of the questionnaire, the mean correlations (calculated using Fisher's *r*-to-*z* transformation) and mean alphas across the items in each facet, across all items in each factor, and across all items in each DPQ form are presented.

Inter-rater reliability statistics were evaluated based on how they compared to those found in previous research on dog personality assessment (see Chapter 2). In previous research, the average reported inter-rater reliability statistics were .60 (weighted

by sample size) and .56 (unweighted). As shown at the bottom of Table 7.3, the average inter-rater reliability for the long form was .555, and for the short form was .547. These are slightly lower than the average reported in the previous dog personality literature.

The average minimums were .50 (weighted) and .37 (unweighted), and the average maximum inter-observer reliability correlations in previous dog personality studies was .77 (weighted by N per sample) or .78 (unweighted). The item-level inter-rater reliabilities associated with DPQ items ranged from .240 ("Dog behaves aggressively when restrained or handled (e.g., groomed)," from the Situational Aggression facet on the Aggression towards People factor) to .839, ("Dog enjoys playing with toys," from the Playfulness facet of Activity/Excitability).

In addition, I followed the suggestion of Gosling and Vazire (2002) by comparing inter-rater reliability of humans' ratings of dogs to those of humans' ratings of other humans. Gosling and Vazire (2002) cite a large amount of evidence supporting the inter-rater reliability of human personality ratings, with inter-rater agreement correlations of approximately .50 being typical. This value was used as a guideline for assessing the inter-rater reliability of the DPQ, which, with a few noticeable exceptions (e.g., items 16, 17, 44), compared favorably to the typical .50 observed in human personality rating studies. Indeed, fifty of the 75 correlations were greater than .500.

The average Cronbach's alpha, across all items in the 75-item from, was .701. Across all items in the 45-item form, the average Cronbach's alpha was .695. The lowest alpha was .393, associated with "("Dog behaves aggressively when restrained or handled (e.g., groomed)," from the Situational Aggression facet on the Aggression towards People factor, which also had the lowest item-level ICC (.240). The highest alpha was .916, associated with the item "Dog enjoys playing with toys," from the Playfulness facet of Activity/Excitability, which also had the highest ICC (.839).

Table 7.3. Item-level inter-rater reliability

Factor / Facet		Reliability: Single Measure ICC	Reliability: Alpha
Factor 1 – Fearfulness			
Facet 1 – Fear of People			
1*	Dog is relaxed when greeting people.	.684	.818
12	Dog is shy.	.638	.781
30	Dog behaves fearfully when near crowds of people.	.621	.767
47	Dog behaves fearfully towards unfamiliar people.	.540	.705
54	Dog behaves fearfully towards children.	.513	.723
Short Form Mean (Fear of People facet)		**.624**	**.768**
Long Form Mean (Fear of People facet)		**.603**	**.759**
Facet 2 – Nonsocial Fear			
6	Dog is anxious.	.360	.559
19*	Dog is confident.	.506	.672
24	Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a door frame).	.656	.805
38*	Dog adapts easily to new situations and environments.	.683	.820
58*	Dog is quick to recover after being startled or frightened.	.395	.570
Short Form Mean (Nonsocial Fear facet)		**.530**	**.684**
Long Form Mean (Nonsocial Fear facet)		**.533**	**.685**
Facet 3 – Fear of Dogs			
9*	Dog is bold.	.579	.731
21	Dog avoids other dogs.	.498	.669
36	Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.	.597	.755
66	Dog behaves fearfully when threatened by other dogs (e.g., growled or lunged at, cornered).	.459	.631
70	Dog behaves fearfully towards other dogs.	.548	.705
Short Form Mean (Fear of Dogs facet)		**.549**	**.710**
Long Form Mean (Fear of Dogs facet)		**.538**	**.698**
Facet 4 – Fear of Handling			
28	Dog behaves fearfully during visits to the veterinarian.	.600	.748
32	Dog is easily upset when corrected, scolded, or punished.	.587	.738
42	Dog is sensitive (and reactive) to pain.	.465	.662
61	Dog exhibits fearful behaviors when restrained.	.375	.542
74	Dog behaves fearfully when groomed (e.g., nails trimmed, brushed, bathed, ears cleaned).	.566	.721
Short Form Mean (Fear of Handling facet)		**.520**	**.670**
Long Form Mean (Fear of Handling facet)		**.524**	**.682**
Short Form Mean (Fearfulness factor)		**.557**	**.706**
Long Form Mean (Fearfulness factor)		**.550**	**.708**

Table 7.3. (*Continued*)

Factor / Facet		Reliability: Single Measure ICC	Reliability: Alpha
Factor 2 – Aggression towards People			
Facet 1 – General Aggression			
13	Dog behaves aggressively towards unfamiliar people.	.580	.733
23	Dog behaves aggressively when a person (e.g., visitor, delivery person) approaches the house or yard.	.418	.596
33*	Dog is friendly towards unfamiliar people.	.637	.776
68	Dog shows aggression when nervous or fearful.	.482	.647
73	Dog behaves aggressively towards children.	.801	.899
Short Form Mean (General Aggression facet)		**.570**	**.719**
Long Form Mean (General Aggression facet)		**.602**	**.730**
Facet 2 – Situational Aggression			
2	Dog behaves aggressively if disturbed or moved when resting.	.517	.709
17	Dog behaves aggressively when restrained or handled (e.g., groomed).	.240	.393
43	Dog behaves aggressively in response to perceived threats from people (e.g., being cornered, having collar reached for).	.324	.565
51	Dog behaves aggressively during visits to the veterinarian.	.658	.798
62	Dog aggressively guards coveted items (e.g., stolen item, treats, food bowl).	.466	.648
Short Form Mean (Situational Aggression facet)		**.496**	**.670**
Long Form Mean (Situational Aggression facet)		**.454**	**.623**
Short Form Mean (Aggression towards People factor)		**.532**	**.695**
Long Form Mean (Aggression towards People factor)		**.534**	**.676**
Factor 3 – Activity/Excitability			
Facet 1 – Excitability			
27	Dog is boisterous.	.363	.528
53	Dog seeks constant activity.	.691	.821
55	Dog is very excitable when visitors arrive.	.630	.772
69*	Dog tends to be calm.	.401	.568
72	Dog is very excitable around other dogs.	.583	.737
Short Form Mean (Excitability facet)		**.502**	**.639**
Long Form Mean (Excitability facet)		**.546**	**.685**
Facet 2 – Playfulness			
3*	Dog is aloof or indifferent towards other dogs.	.359	.523
16*	Dog gets bored in play quickly.	.250	.397
31	Dog enjoys playing with toys.	.839	.916
46	Dog is interested in playing tug-o-war with people or dogs.	.737	.852
59	Dog retrieves objects (e.g., balls, toys, sticks).	.812	.903
Short Form Mean (Playfulness facet)		**.701**	**.739**
Long Form Mean (Playfulness facet)		**.656**	**.720**

Table 7.3. (*Continued*)

Factor / Facet		Reliability: Single Measure ICC	Reliability: Alpha
Factor 3 – Activity/Excitability (*continued*)			
Facet 3 – Active Engagement			
10*	Dog is lethargic	.420	.588
14	Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden.	.719	.843
25	Dog works at tasks (e.g., getting treats out of a Kong, shredding toys) until entirely finished.	.522	.683
40	Dog is curious.	.576	.729
48	Dog is very alert.	.619	.763
Short Form Mean (Active Engagement facet)		**.509**	**.667**
Long Form Mean (Active Engagement facet)		**.580**	**.721**
Facet 4 – Companionability			
7	Dog loves to be praised.	.790	.884
35	Dog seeks companionship from people.	.559	.715
44*	Dog is aloof.	.265	.437
63	Dog is affectionate.	.521	.693
67	Dog follows people around.	.498	.673
Short Form Mean (Companionability facet)		**.457**	**.615**
Long Form Mean (Companionability facet)		**.551**	**.680**
Short Form Mean (Activity factor)		**.550**	**.665**
Long Form Mean (Activity factor)		**.585**	**.701**
Factor 4 – Responsiveness to Training			
Facet 1 – Trainability			
37	Dog is attentive to owner's actions and words.	.667	.802
45*	Dog is slow to learn new tricks or tasks.	.386	.567
50*	Dog is slow to respond to corrections.	.563	.718
64*	Dog ignores commands.	.452	.619
71	Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).	.464	.676
Short Form Mean (Trainability facet)		**.495**	**.671**
Long Form Mean (Trainability facet)		**.514**	**.676**

Table 7.3. (*Continued*)

Factor / Facet		Reliability: Single Measure ICC	Reliability: Alpha
Factor 4 – Responsiveness to Training (*continued*)			
Facet 2 – Controllability			
4*	Dog is destructive.	.704	.829
11	When off leash, dog comes immediately when called.	.596	.746
18*	Dog is quick to sneak out through open doors, gates.	.514	.682
29*	When walking on leash, dog tends to pull ahead.	.317	.510
56	Dog leaves food or objects alone when told to do so.	.479	.644
Short Form Mean (Controllability facet)		**.532**	**.691**
Long Form Mean (Controllability facet)		**.534**	**.682**
Short Form Mean (Responsiveness to Training factor)		**.513**	**.681**
Long Form Mean (Responsiveness to Training factor)		**.524**	**.679**
Factor 5 – Aggression towards Animals			
Facet 1 – Aggression towards Dogs			
5	Dog behaves aggressively toward dogs.	.578	.767
8	Dog responds aggressively when threatened by another dog (e.g., growled or lunged at, cornered).	.512	.709
34*	Dog is playful with other dogs.	.675	.804
57	Dog has a tendency to attack (or attempt to attack) other dogs.	.490	.685
60*	Dog is friendly towards other dogs.	.597	.763
Short Form Mean (Aggression towards Dogs facet)		**.619**	**.778**
Long Form Mean (Aggression towards Dogs facet)		**.574**	**.746**
Facet 2 – Prey Drive			
15	Dog likes to chase squirrels, birds, or other small animals.	.329	.519
22	Dog catches and kills other animals (e.g., squirrels, rabbits).	.616	.798
26	Dog is very excitable around cats.	.607	.763
39	Dog likes to chase bicycles, joggers, and skateboarders.	.535	.743
65	Dog behaves aggressively towards cats.	.516	.737
Short Form Mean (Prey Drive facet)		**.467**	**.666**
Long Form Mean (Prey Drive facet)		**.527**	**.712**

Table 7.3. (*Continued*)

Factor / Facet		Reliability: Single Measure ICC	Reliability: Alpha
Factor 5 – Aggression towards Animals (*continued*)			
Facet 3 – Dominance over other Dogs			
20	Dog is dominant over other dogs.	.505	.695
41	Dog guards food or treats from other dogs.	.519	.693
49*	Dog willingly shares toys with other dogs.	.699	.826
52	Dog jumps up on (e.g., in play) and/or mounts other dogs (outside appropriate mating).	.568	.722
75	Dog is assertive or pushy with other dogs (e.g., if in a home with other dogs, when greeting).	.527	.701
Short Form Mean (Dominance over Other Dogs facet)		**.584**	**.741**
Long Form Mean (Dominance over Other Dogs facet)		**.569**	**.727**
Short Form Mean (Aggression towards Animals factor)		**.559**	**.728**
Long Form Mean (Aggression towards Animals factor)		**.557**	**.728**
Short Form MEAN (all items on short form)		**.547**	**.695**
Long Form MEAN (all items)		**.555**	**.701**

Note. Items marked with an asterisk are reverse coded items. Items listed in boldface are on both the long and short form of the DPQ. Mean ICC *r*s were computed using Fisher's r-to-z transformation.

The facet scores are the means of the scores assigned to items that compose each facet, after the reverse keyed items have been rekeyed. Facet scores were calculated separately for the long form and the short form of the DPQ, because the facets on the short form consist of subsets of the items that compose the facets on the long form. Facet-level ICCs indicate the strength of the relationship between the pairs of facet scores calculated for each dog, based on averaging the item-level scores that load on each facet. Facet-level ICCs and alphas are presented in Table 7.4. The facet-level ICCs and alphas are organized by the factors with which they are associated and are presented underneath the bolded factor labels and factor-related statistics.

The factor scores are the mean score raters assigned to items composing each factor, after the reverse keyed items have been rekeyed. Factor scores were calculated separately for the long form and the short form of the DPQ, because, as with the facets,

the factors on the short form consist of subsets of items on the corresponding factors on the long form. Factor-level ICCs indicate the strength of the relationship between the paired factor-level ratings, which were calculated based on averaging the item-level scores that load on each factor. Factor-level ICCs and alphas are presented in rows and bolded in Table 7.4.

Table 7.4. Facet- and factor-level inter-rater reliability

Factor	**Inter-Rater Reliability**			
Facet	**Long form**		**Short form**	
	ICC	**Alpha**	**ICC**	**Alpha**
Factor 1 – Fearfulness	**.753**	**.861**	**.777**	**.875**
Facet 1 – Fear of People	.738	.854	.767	.868
Facet 2 – Nonsocial Fear	.732	.845	.675	.812
Facet 3 – Fear of Dogs	.503	.667	.492	.656
Facet 4 – Fear of Handling	.603	.755	.640	.778
Factor 2 – Aggression towards People	**.659**	**.806**	**.697**	**.832**
Facet 1 – General Aggression	.683	.810	.642	.780
Facet 2 – Situational Aggression	.468	.707	.597	.797
Factor 3 – Activity/Excitability	**.786**	**.879**	**.745**	**.854**
Facet 1 – Excitability	.725	.840	.669	.802
Facet 2 – Playfulness	.872	.933	.774	.880
Facet 3 – Active Engagement	.704	.825	.558	.714
Facet 4 – Companionability	.550	.707	.459	.632
Factor 4 – Responsiveness to Training	**.701**	**.823**	**.666**	**.797**
Facet 1 – Trainability	.689	.815	.627	.776
Facet 2 – Controllability	.679	.807	.587	.741
Factor 5 – Aggression towards Animals	**.685**	**.826**	**.734**	**.852**
Facet 1 – Aggression towards Dogs	.678	.834	.749	.869
Facet 2 – Prey Drive	.590	.747	.539	.698
Facet 3 – Dominance over Other Dogs	.669	.800	.628	.778
MEAN across Factors	**.720**	**.839**	**.726**	**.842**

I compared inter-rater reliability levels across the different facets and factors to address the question of whether differences in reliability exist between the facets or between the factors, which could be indicative of some facets or factors being more sensitive to observer-related effects than others. As displayed in Tabel 7.4, facet ICCs were generally high, but ranged from .468 (Situational Aggression) to .872 (Playfulness)

on the 75-item long form, and from .459 (Companionability) to .774 (Playfulness) on the 45-item short form. Factor ICCs ranged from .659 (Aggression towards People) to .786 (Activity/Excitability) on the 75-item long form, and from .666 (Responsivness to Training) to .777 (Fearfulness) on the 45-item short form. All of these, however, are relatively high. It should be noted that it is expected that ICCs across facets and factors are higher than those at the item level because facets and factors benefit from aggregation of ratings, creating a composite analogous to measuring behavior at multiple instances instead of a single instant in time.

Though still high, the facets Fear of dogs, Situational Aggression, and Companionability are associated with the lowest inter-rater reliability. ICCs associated with factor-level scores show a narrower range, but are high relative to ICCs in previous research (see Chapter 2).

Facet-level alphas ranged from .667 (Fear of Dogs) to .933 (Playfulness) on the long form and from .632 (Companionability) to .880 (Playfulness) on the short form. Factor-level alphas ranged from .806 (Aggression towards People) to .879 (Activity/Excitability) on the long from and from .797 (Responsiveness to Training) to .875 (Fearfulness) on the short form. As with the facet and factor ICCs, the alphas at the facet and factor level are higher than those at the item level because facets and factors benefit from aggregation of ratings; aggregated scores will typically be associated with higher inter-correlations than item-level scores, and higher inter-correlations lead to higher alphas.

SUMMARY OF FINDINGS

Inter-rater reliability levels associated with the items, facets, and factors on the DPQ long and short forms were shown to be acceptable. Most DPQ inter-rater reliability levels were comparable to or stronger than inter-rater reliability levels found in previous dog personality research and typical human personality research. The DPQ long form's item-level inter-rater reliability ranged from .240 to .839 with a mean of .555, facet-level inter-rater reliability ranged from .468 to .872 with a mean of .672, and factor-level inter-

rater reliability ranged from .659 to .786 with a mean of .720. The DPQ short form's item-level inter-rater reliability ranged from .265 to .839 with a mean of .547, facet-level inter-rater reliability ranged from .459 to .767 with a mean of .638, and factor-level inter-rater reliability ranged from .666 to .726 with a mean of .726.

CHAPTER 8

Study 5: Test-Retest Reliability

INTRODUCTION

Reliability across time is crucial to the value of an assessment, particularly one intended to measure a relatively temporally stable construct (e.g., personality). Reliability across time, also called test-retest reliability, is a necessary prerequisite to predictive validity (addressed in Study 6). The goal of Study 5 was to evaluate the temporal stability of the DPQ by estimating the degree to which responses vary within individuals across time, thus reflecting error due to uncontrolled ancillary variables (e.g., test administration). To achieve this goal, I administered the questionnaire twice to 100 participants who rated the same dog on the questionnaire both times.

METHODS

Participants

The 110 participants who took part in this study were newly recruited and reported having not participated in any previous studies using the DPQ. Participants signed up to take part in Study 5 in response to online advertisements posted on the Animal Personality Website (www.animalpersonality.org) or on dog-related Internet discussion boards, news groups, or newsletters. E-mail groups were also given permission to circulate postings through-out their groups.

Participants completed the questionnaire once in June or July of 2007, then were e-mailed reminders to complete the questionnaire a second time in August of 2007. All participants completed the questionnaire online twice between June and September of 2007. Again, participants were given no compensation for their participation, but were given feedback about their ratings of their dogs' personalities with respect to the five personality factors assessed in the DPQ. (As in previous studies, participants were informed that the feedback was preliminary and the reliability and validity of the DPQ

were under assessment.) A total of 100 participants completed the questionnaire twice and were included in the test-retest reliability analyses.

Materials and Procedures

Participants in Study 5 completed the long form of the questionnaire. Participants were also asked to provide the same demographic information about themselves and their dogs as participants in previous studies. Participants were not allowed to leave items blank and, during the first completion of the questionnaire, were required to provide an e-mail address to which a reminder could be sent four to six weeks later.

As in previous studies, I examined the demographic data in Study 5 to assure that the participants and the target dogs composed a diverse sample.

RESULTS

Study 5 participants were limited to those 100 who completed the questionnaire twice in the time allowed (between June and September of 2007). Cases were matched by owners' first names, dogs' names, locations, breeds, and owners' indication of whether they had completed the questionnaire previously.

As in previous studies, I looked at the frequency of responses for each questionnaire item to check for errors (e.g, from saving the data). As with the previous studies, no anomalies were found in the dataset, but responses to some of the items were not regularly distributed (i.e., responses to some items were skewed).

I also examined the data to determine whether the 100 participants retained had provided appropriate responses to background and free-response items. No anomalies or problems (e.g., withheld information) were found, so all 100 participants were retained.

Sample demographics

Demographic information for the 100 participants who took part in Study 5 is presented in Table 8.1.

Both sexes of dog owners were represented in Study 5. Again, men composed a small minority (9.0%, or 9 men), comparable to the percentage of men in Study 1

(10.6%) and Study 3 (10.2%). The average age of participants was 41.6 years, with a standard deviation of 10.3 years.

Ninety-one of participants were currently residing in the U.S., with only three participants residing in Canada, one residing in Australia, and five who indicated they living in an unlisted country (Italy (1), China (2), Germany (2)). The sample included participants from 22 states, with the largest percentage of participants form California (21%). The majority of participants identified themselves as Caucasian/White (N = 92). Other participants identified themselves as African American/Black (N = 3), East Asian (N = 2), and Multiracial or Other (N = 3).

Of the 100 participants, 25 (or 25%) indicated having a dog-related job or career, and one worked with an assistance dog partner (guide dog). The majority of these participants reported working as dog trainers and/or behaviorists (N = 9) or as veterinarians (N = 8).

The mean number of dogs each participant had owned in his or her lifetime, not including current dogs but including childhood family dogs, was 5.7 (s.d. = 4.5). The mean number of dogs currently owned by participants in Study 5 was 2.2 (s.d. = 1.2).

Table 8.1. Demographic information about participants in Study 5

		Study 5
Sex (dog owner)	Male	9 (9.0%)
	Female	91 (91.0%)
Age (dog owner)	Mean	41.6
	standard deviation	10.3
	Age not reported	0
Country of residence	U.S.	91 (91.0%)
	Canada	3 (3.0%)
	Australia	1 (1.0%)
	U.K.	0
	Other (Native language is English)	5 (5.0%)
	Country of residence not reported	0
Race	Caucasian/White	92 (92.0%)
	African-American/Black	3 (3.0%)
	Hispanic	0
	East Asian	2 (2.0%)
	Native American	0
	Asian Indian	0
	Multiracial or Other	3 (3.0%)
Dog-related careers	Breeder and/or Exhibitor	2 (2.0%)
	Trainer and/or Behaviorist	9 (9.0%)
	Dog rescue worker/volunteer	1 (1.0%)
	Veterinarian	8 (8.0%)
	Groomer	1 (1.0%)
	Kennel/Daycare owner or employee	4 (4.0%)
	Assistance dog partner	1 (1.0%)
		Mean (s.d.)
Number of dogs owned	Past (mean; s.d.)	5.7 (4.5)
	Number not reported	0
	Currently (mean; s.d.)	2.2 (1.2)
	Number not reported	0
Total number of participants		**100**

The 100 dogs in this sample were also examined for demographic diversity. First was age. In Study 5, all dogs' ages were reported as known, and the dogs' average age

was 5.9 years, but age was highly variable with a standard deviation of 4.1 years. The dogs ranged in age from one to 14 years.

Other demographic information about the dogs is presented in Table 7.2. The number of dogs (out of the 100 in Study 5) in each category is listed in the far right column, which is entitled “Number of dogs”. Although a smaller sample than in most of the previous studies, the sample of 100 dogs appeared to be diverse, including, for example, dogs of both sexes and many breeds. Slightly more than half the dogs (N = 56, or 56%) were male. Eighty-nine of the dogs (89%) were castrated. Purebred dogs composed exactly half the sample (N = 50, or 50%), with 24 breeds represented. The breeds most represented included the GSD (N = 6), Border Collie (N = 5), French Bulldog (N = 4), American Pit Bull Terrier (N = 4), and Labrador Retriever (N = 2).

Information about whether the dogs were reported to have ever bitten a person, had any reported health issues, what role they played in their owners’ lives (e.g., pet, guide dog), and whether they were involved in any sports is also included in Table 7.2.

Twelve dogs were reported to have bitten someone, two dogs had unknown bite histories, and the remaining 86 were reported to have never bitten a person (intentionally, outside play, and outside work or sport requiring biting; e.g., Schutzhund training or competition). Twenty-nine dogs were indicated to have health problems or disabilities, with nine dogs having more than one health problem or disability. Common health issues included hip dysplasia (N = 10) and arthritis (N = 9). Owners who indicated their dogs had other disabilities listed their dogs as having, for example, a Vitamin B-12 deficiency, low thyroid activity, heart worms, a broken leg in the process of healing, a tumor, incontinence, and allergies.

The majority of the dogs in this sample (N = 84, or 84%) were considered pets or companions by their owners. However, there was also one working assistance dog (a guide dog), and there were seven guard dogs, two animal assisted therapy dogs, and two dogs used in breeding. As in the previous studies, the popularity of agility was also represented in this sample (N = 12).

Table 8.2. Demographic information about target dogs in Study 5

			Number of dogs
Sex	Male		56 (56.0%)
	Female		44 (44.0%)
Castration Status	Males	Neutered	48 (48.0%)
		Intact	8 (8.0%)
	Females	Spayed	41 (41.0%)
		Intact	3 (3.0%)
Purebred	Unknown		7 (7.0%)
	No (mixed)		43 (43.0%)
	Yes		50 (50.0%)
Bitten a Person	Unknown		2 (2.0%)
	No		86 (86.0%)
	Yes		12 (12.0)
Disability and Health issues	Unilaterally deaf		0
	Bilaterally deaf		0
	Blind in one eye		0
	Blind in both eyes		0
	Arthritic		9 (9.0%)
	Hip dysplasia		10 (10.0%)
	Elbow dysplasia		0
	Other disabilities		16 (16.0%)
Job or Role	**Pet/Companion**		84 (84.0%)
	Assistance dog		1 (1.0%)
	Guide dog		0
	Hearing ear dog		0
	Medical assistance dog		1 (1.0%)
	Search and rescue		0
	Guard dog		7 (7.0%)
	Animal Assisted Therapy		2 (2.0%)
	Dam or sire for breeding		2 (2.0%)

Table 8.2. (*Continued*)

		Number of dogs
Sport	Obedience	5 (5.0%)
	Sledding	3 (3.0%)
	Carting	0
	Frisbee	2 (2.0%)
	Earth dog	1 (1.0%)
	Show/Conformation	5 (5.0%)
	Schutzhund	0
	Agility	12 (12.0%)
	Herding (competitive)	6 (6.0%)
	Flyball	4 (4.0%)
	Hunting	0
Total number of dogs		**100**

Note. The sum of the number of dogs who participate in sports will not equal the total number of dogs in the study; some dogs are not in sports, and others are in multiple. Similarly, participants may consider a dog to be both a pet/companion and to have another job or role (e.g., assistance dog, search and rescue dog).

Dogs with unknown histories (e.g., bite history) are typically dogs who were adopted from shelters or rescues.

Analyses

Test-retest reliability

I used Pearson's *r* correlation to measure the strength of the relationship between participants' ratings of their target dogs at Time 1 and Time 2. In Study 4, Pearson's *r* was inappropriate because there was no meaningful way to group the ratings into two categories. In Study 5, the ratings are logically and meaningfully grouped into Time 1 (the first rating) and Time 2 (the second rating). Like inter-rater reliability, test-retest reliability statistics were evaluated based on how they compare with those found in previous research on dog personality assessment (see Chapter 2).

I examined test-retest reliability at the item, facet, and the factor level. Item-level Pearson's *r* correlations for test-retest reliability are presented in Table 8.3, alongside the ICC associated with each item's inter-rater reliability.

Test-retest reliability statistics were evaluated based on how they compared to those found in previous research on dog personality assessment (see Chapter 2). In previous research, the averages across reported test-retest reliability statistics were .63 (weighted by sample size) and .71 (unweighted). As shown at the bottom of Table 8.3, the average item-level test-retest reliability for the short form of the DPQ was .753 and for the long form was .750. These are slightly higher than the average reported in the previous dog personality literature.

As shown in Chapter 2, the average minimum test-retest reliability correlations reported in previous dog personality studies were .55 (weighted by N per sample) and .68 (unweighted), and the average maximum test-retest reliability correlations were .67 (weighted) and .72 (unweighted). The test-retest reliabilities associated with DPQ items ranged from .325 ("Dog behaves aggressively during visits to the veterinarian" from the Situational Aggression facet of the Aggression towards People factor) to .923 ("Dog is affectionate," from the Companionability facet of the Activity/Excitability factor). Overall, it should be noted that, while the range of test-retest reliability correlations is large, only three items have test-retest reliability correlations less than .500. Indeed, 48 are greater than .700, and 20 are greater than .800.

Table 8.3. Item-level inter-rater and test-retest reliability

Factor / **Facet**		**Reliability**	
		Inter-Rater	**Test-Retest**
Factor 1 – Fearfulness			
Facet 1 – Fear of People			
1*	Dog is relaxed when greeting people.	.692	.737
12	Dog is shy.	.641	.822
30	Dog behaves fearfully when near crowds of people.	.623	.811
47	Dog behaves fearfully towards unfamiliar people.	.550	.734
54	Dog behaves fearfully towards children.	.585	.842
Short Form Mean (Fear of People facet)		**.631**	**.768**
Long Form Mean (Fear of People facet)		**.621**	**.793**

Table 8.3. (*Continued*)

Factor / Facet		Reliability: Inter-Rater	Reliability: Test-Retest
Factor 1 – Fearfulness (*continued*)			
Facet 2 – Nonsocial Fear			
6	Dog is anxious.	.388	.863
19*	Dog is confident.	.508	.867
24	Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a door frame).	.673	.674
38*	Dog adapts easily to new situations and environments.	.703	.778
58*	Dog is quick to recover after being startled or frightened.	.404	.801
Short Form Mean (Nonsocial Fear facet)		**.547**	**.840**
Long Form Mean (Nonsocial Fear facet)		**.550**	**.807**
Facet 3 – Fear of Dogs			
9*	Dog is bold.	.578	.820
21	Dog avoids other dogs.	.505	.748
36	Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.	.609	.702
66	Dog behaves fearfully when threatened by other dogs (e.g., growled or lunged at, cornered).	.465	.592
70	Dog behaves fearfully towards other dogs.	.546	.817
Short Form Mean (Fear of Dogs facet)		**.555**	**.760**
Long Form Mean (Fear of Dogs facet)		**.543**	**.746**
Facet 4 – Fear of Handling			
28	Dog behaves fearfully during visits to the veterinarian.	.606	.786
32	Dog is easily upset when corrected, scolded, or punished.	.586	.602
42	Dog is sensitive (and reactive) to pain.	.500	.666
61	Dog exhibits fearful behaviors when restrained.	.372	.684
74	Dog behaves fearfully when groomed (e.g., nails trimmed, brushed, bathed, ears cleaned).	.572	.691
Short Form Mean (Fear of Handling facet)		**.524**	**.724**
Long Form Mean (Fear of Handling facet)		**.532**	**.691**
Short Form Mean (Fearfulness factor)		**.566**	**.777**
Long Form Mean (Fearfulness factor)		**.562**	**.763**

Table 8.3. (*Continued*)

Factor / Facet		Reliability: Inter-Rater	Reliability: Test-Retest
Factor 2 – Aggression towards People			
Facet 1 – General Aggression			
13	Dog behaves aggressively towards unfamiliar people.	.578	.780
23	Dog behaves aggressively when a person (e.g., visitor, delivery person) approaches the house or yard.	.424	.834
33*	Dog is friendly towards unfamiliar people.	.642	.887
68	Dog shows aggression when nervous or fearful.	.483	.810
73	Dog behaves aggressively towards children.	.860	.772
Short Form Mean (General Aggression facet)		**.571**	**.832**
Long Form Mean (General Aggression facet)		**.628**	**.821**
Facet 2 – Situational Aggression			
2	Dog behaves aggressively if disturbed or moved when resting.	.552	.642
17	Dog behaves aggressively when restrained or handled (e.g., groomed).	.246	.645
43	Dog behaves aggressively in response to perceived threats from people (e.g., being cornered, having collar reached for).	.417	.686
51	Dog behaves aggressively during visits to the veterinarian.	.670	.325
62	Dog aggressively guards coveted items (e.g., stolen item, treats, food bowl).	.482	.697
Short Form Mean (Situational Aggression facet)		**.532**	**.591**
Long Form Mean (Situational Aggression facet)		**.486**	**.613**
Short Form Mean (Aggression towards People factor)		**.552**	**.734**
Long Form Mean (Aggression towards People factor)		**.561**	**.734**
Factor 3 – Activity/Excitability			
Facet 1 – Excitability			
27	Dog is boisterous.	.360	.537
53	Dog seeks constant activity.	.711	.796
55	Dog is very excitable when visitors arrive.	.631	.758
69*	Dog tends to be calm.	.397	.858
72	Dog is very excitable around other dogs.	.584	.652
Short Form Mean (Excitability facet)		**.510**	**.758**
Long Form Mean (Excitability facet)		**.551**	**.739**
Facet 2 – Playfulness			
3*	Dog is aloof or indifferent towards other dogs.	.356	.673
16*	Dog gets bored in play quickly.	.252	.782
31	Dog enjoys playing with toys.	.856	.686
46	Dog is interested in playing tug-o-war with people or dogs.	.743	.882
59	Dog retrieves objects (e.g., balls, toys, sticks).	.828	.771
Short Form Mean (Playfulness facet)		**.719**	**.750**
Long Form Mean (Playfulness facet)		**.670**	**.771**

Table 8.3. (*Continued*)

Factor / Facet		Reliability: Inter-Rater	Reliability: Test-Retest
Factor 3 – Activity/Excitability (*continued*)			
Facet 3 – Active Engagement			
10*	Dog is lethargic	.420	.706
14	Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden.	.728	.401
25	Dog works at tasks (e.g., getting treats out of a Kong, shredding toys) until entirely finished.	.525	.690
40	Dog is curious.	.579	.605
48	Dog is very alert.	.618	.720
Short Form Mean (Active Engagement facet)		**.511**	**.669**
Long Form Mean (Active Engagement facet)		**.583**	**.636**
Facet 4 – Companionability			
7	Dog loves to be praised.	.796	.353
35	Dog seeks companionship from people.	.559	.664
44*	Dog is aloof.	.280	.578
63	Dog is affectionate.	.534	.923
67	Dog follows people around.	.529	.770
Short Form Mean (Companionability facet)		**.466**	**.771**
Long Form Mean (Companionability facet)		**.564**	**.712**
Short Form Mean (Activity factor)		**.561**	**.739**
Long Form Mean (Activity factor)		**.594**	**.718**
Factor 4 – Responsiveness to Training			
Facet 1 – Trainability			
37	Dog is attentive to owner's actions and words.	.669	.793
45*	Dog is slow to learn new tricks or tasks.	.400	.774
50*	Dog is slow to respond to corrections.	.560	.609
64*	Dog ignores commands.	.458	.738
71	Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).	.569	.799
Short Form Mean (Trainability facet)		**.531**	**.724**
Long Form Mean (Trainability facet)		**.538**	**.749**
Facet 2 – Controllability			
4*	Dog is destructive.	.725	.832
11	When off leash, dog comes immediately when called.	.595	.793
18*	Dog is quick to sneak out through open doors, gates.	.518	.643
29*	When walking on leash, dog tends to pull ahead.	.344	.885
56	Dog leaves food or objects alone when told to do so.	.475	.781
Short Form Mean (Controllability facet)		**.531**	**.746**
Long Form Mean (Controllability facet)		**.544**	**.799**
Short Form Mean (Responsiveness to Training factor)		**.531**	**.735**
Long Form Mean (Responsiveness to Training factor)		**.541**	**.776**

Table 8.3. (*Continued*)

Factor / Facet		Reliability: Inter-Rater	Reliability: Test-Retest
Factor 5 – Aggression towards Animals			
Facet 1 – Aggression towards Dogs			
5	Dog behaves aggressively toward dogs.	.620	.764
8	Dog responds aggressively when threatened by another dog (e.g., growled or lunged at, cornered).	.551	.746
34*	Dog is playful with other dogs.	.686	.872
57	Dog has a tendency to attack (or attempt to attack) other dogs.	.527	.734
60*	Dog is friendly towards other dogs.	.626	.850
Short Form Mean (Aggression towards Dogs facet)		**.645**	**.834**
Long Form Mean (Aggression towards Dogs facet)		**.605**	**.801**
Facet 2 – Prey Drive			
15	Dog likes to chase squirrels, birds, or other small animals.	.400	.733
22	Dog catches and kills other animals (e.g., squirrels, rabbits).	.663	.838
26	Dog is very excitable around cats.	.617	.810
39	Dog likes to chase bicycles, joggers, and skateboarders.	.591	.709
65	Dog behaves aggressively towards cats.	.584	.736
Short Form Mean (Prey Drive facet)		**.525**	**.726**
Long Form Mean (Prey Drive facet)		**.577**	**.770**
Facet 3 – Dominance over Other Dogs			
20	Dog is dominant over other dogs.	.534	.781
41	Dog guards food or treats from other dogs.	.531	.663
49*	Dog willingly shares toys with other dogs.	.707	.682
52	Dog jumps up on (e.g., in play) and/or mounts other dogs (outside appropriate mating).	.565	.680
75	Dog is assertive or pushy with other dogs (e.g., if in a home with other dogs, when greeting).	.540	.640
Short Form Mean (Dominance over Other Dogs facet)		**.600**	**.706**
Long Form Mean (Dominance over Other Dogs facet)		**.580**	**.693**
Short Form Mean (Aggression towards Animals factor)		**.592**	**.762**
Long Form Mean (Aggression towards Animals factor)		**.588**	**.758**
Short Form MEAN (all items on short form)		**.564**	**.753**
Long Form MEAN (all items)		**.570**	**.750**

Note. Items marked with an asterisk are reverse coded items. Items listed in boldface are on both the long and short form of the DPQ. Mean correlations were computed using Fisher's r-to-z transformation.

As in the inter-rater reliability calculations, the facet scores are the means of the scores assigned to items that compose each facet, after the reverse keyed items have been rekeyed. Facet scores were calculated separately for the long form and the short form of the DPQ, because the facets on the short form consist of subsets of the items that compose the facets on the long form. Facet-level test-retest reliability correlations indicate the strength of the relationship between first and second assessment for each dog, based on averaging the item-level scores that load on each facet. Facet-level test-retest reliability correlations are presented alongside inter-rater reliability ICCs in Table 8.4. The first two columns of correlations contain inter-rater reliability ICCs for the long and then the short form; the second two columns to the right of the table contain test-retest reliability correlations for the long form and then the short form. The facet-level correlations are organized by the factors with which they are associated and are presented underneath the bolded factor labels and factor-related statistics.

The factor scores are the mean of the score raters assigned to items composing each factor, after the reverse keyed items have been rekeyed. Factor scores were calculated separately for the long form and the short form of the DPQ, because, as with the facets, the factors on the short form consist of subsets of items on the corresponding factors on the long form. Factor-level test-retest reliability correlations indicate the strength of the relationship between the paired factor-level ratings, which were calculated based on averaging the item-level scores that load on each factor. Factor-level test-retest correlations are presented alongside inter-rater reliability ICCs in bolded rows in Table 8.4.

I compared inter-rater reliability levels across the different facets and factors to address the question of whether differences in reliability exist between the facets or between the factors, and I did the same with test-retest reliability correlations. As displayed in Table 8.4, facet test-retest reliability correlations were generally high, ranging from .735 (Dominance over Other Dogs) to .923 (Non-social Fear) on the 75-item long form, and from .750 (Situational Aggression) to .936 (Nonsocial Fear) on the 45-item short form. Factor test-retest correlations ranged from .878 (Aggression towards

People) to .939 (Fearfulness) on the 75-item long form, and from .872 (Aggression towards People) to .929 (Fearfulness) on the 45-item short form. It should be noted that it is expected that test-retest reliability correlations, like the ICCs, across facets and factors are higher than those at the item level because facets and factors allow for aggregation of ratings, creating a composite analogous to measuring behavior at multiple instances instead of a single instant in time.

Table 8.4. Facet- and factor-level inter-rater and test-retest reliability

	Inter-Rater		Test-Retest	
Factor **Facet**	**Long form**	**Short form**	**Long form**	**Short form**
Factor 1 – Fearfulness	**.753**	**.777**	**.939**	**.929**
Facet 1 – Fear of People	.738	.767	.899	.880
Facet 2 – Nonsocial Fear	.732	.675	.923	.936
Facet 3 – Fear of Dogs	.503	.492	.859	.835
Facet 4 – Fear of Handling	.603	.640	.799	.804
Factor 2 – Aggression towards People	**.659**	**.697**	**.878**	**.872**
Facet 1 – General Aggression	.683	.642	.893	.870
Facet 2 – Situational Aggression	.468	.597	.804	.750
Factor 3 – Activity/Excitability	**.786**	**.745**	**.890**	**.884**
Facet 1 – Excitability	.725	.669	.819	.831
Facet 2 – Playfulness	.872	.774	.900	.832
Facet 3 – Active Engagement	.704	.558	.745	.786
Facet 4 – Companionability	.550	.459	.746	.786
Factor 4 – Responsiveness to Training	**.701**	**.666**	**.906**	**.907**
Facet 1 – Trainability	.689	.627	.859	.848
Facet 2 – Controllability	.679	.587	.893	.867
Factor 5 – Aggression towards Animals	**.685**	**.734**	**.903**	**.878**
Facet 1 – Aggression towards Dogs	.678	.749	.906	.882
Facet 2 – Prey Drive	.590	.539	.874	.796
Facet 3 – Dominance over Other Dogs	.669	.628	.735	.758
MEAN across Factors	**.720**	**.726**	**.906**	**.896**

Note. Mean correlations were computed using Fisher's r-to-z transformation.

SUMMARY OF FINDINGS

Test-retest reliability levels associated with the items, facets, and factors on the DPQ long and short forms were shown to be acceptable. Most DPQ test-retest reliability levels were comparable to or stronger than test-retest reliability levels found in previous dog personality research. The DPQ long form's item-level test-retest reliability ranged from .325 to .923 with a mean of .750, facet-level test-retest reliability ranged from .735 to .923 with a mean of .854, and factor-level test-retest reliability ranged from .878 to .939 with a mean of .906. The DPQ short form's item-level test-retest reliability ranged from .325 to .923 with a mean of .753, facet-level test-retest reliability ranged from .758 to .936 with a mean of .838, and factor-level test-retest reliability ranged from .872 to .929 with a mean of .896.

CHAPTER 9

Study 6: Predictive Validity

INTRODUCTION

Validity is an index of the degree to which a tool measures what it is intended to measure (e.g., Cronbach & Meehl, 1955). Predictive validity is a type of validity in which a measure is evaluated in terms of how well it predicts an independent measure of the same construct (e.g., Fearfulness). As discussed, a complete evaluation of a tool's validity involves specifying not only the things to which each construct is expected to relate (convergent validity), but also the things to which each construct should be unrelated (discriminant validity) (e.g., Campbell & Fiske, 1959). In Study 6, I investigated the relationship between each factor and facet measured on the DPQ and behaviors predicted to be related and unrelated to each factor and facet.

Under consideration as criteria were (1) a second, independently designed Rating of Individual Dogs, (2) a Test Battery, and (3) an Observational Test. The prediction of behaviors has been described as the definitive test of personality judgments (e.g., Gosling et al., 2003a), which suggested a Test Battery or Observational Test was the preferred criterion. I chose to use a Test Battery because Test Batteries allow for greater control of the test situation through control of the environment and the stimuli to which the dog is exposed (see Chapter 2).

For Study 6, a new Test Battery was designed to have subtests corresponding to each DPQ factor, and to as many DPQ facets as were amenable to assessment through behavioral testing. However, given that the facets of each factor are also related to each other, clear convergent and discriminant correlations at the facet level were not expected. Dogs' scores on each factor of the DPQ were predicted to correlate with (and predict) the dogs' behavioral responses to theoretically related—but not theoretically unrelated—components (or subtests) of the Test Battery. The same types of predictions were made for each facet. I predicted, for example, that dogs' Fearfulness as measured on the DPQ

would predict the dogs' fearful behavior or lack thereof during the Test Battery, but would not predict the dogs' excitable behavior during the Test Battery.

However, it must be noted that it is notoriously difficult to predict single, specific instances of behavior (e.g., Epstein, 1983). Personality can be thought of as useful in predicting an individual's *average* behavior across time and situations. The likelihood of a single snapshot or sample of an individual's behavior corresponding perfectly to the individual's average behavior is low. If the DPQ is an accurate measure of dogs' personality, it may still have low correlations with theoretically related behavioral tests performed in Study 6.

METHOD

Design

Test Battery with subtests

A Test Battery afforded the advantage that it could be built with subtests that correspond to the factors and facets of the DPQ. In order to elicit a broad array of behaviors corresponding to those related to the five factors of the DPQ, the Test Battery designed for use in the current study required that each dog be put through multiple subtests. These subtests were drawn from the existing dog personality research. Taylor and Mills (2006) discussed and summarized the behavioral temperament tests used for companion dogs, highlighting the 20 most commonly used tests. Thirteen of the 18 subtests in Study 6 (Subtests 1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 15, and 17, described in Appendix J and listed in Table 9.1) are drawn from this list. Subtests 12 and 13 are drawn from a single test ("Object Play") in Taylor and Mills' (2006) review. Subtests were selected to assess behavior thought related to the five DPQ factors (Fearfulness, Aggression towards People, Responsiveness to Training, Activity/Excitability, and Aggression towards Animals) and as many facets as possible while maintaining a Test Battery of reasonable length with subtests that are similar to dogs' experiences in everyday life. The situations created in the Test Battery were relatively normal, everyday occurrences in the kennel and daycare environment in which the Test Battery was

conducted, with exception of the Threatening Approach (Subtest 3), which was deemed an important subtest for assessing fear- and aggression-related behaviors.

Each subtest is predicted to elicit behaviors related to one or more facets and/or factors. Table 9.1 lists each subtest as a column heading, in the order in which the tests were conducted. The rows indicate the factors and their respective facets with text in factor rows bolded. Behavioral descriptions, discussed in more detail in the scoring section below, appear in the table when a specific subtest is expected to elicit behavior motivated by and thus related to the factor or facet listed in that row. A "+" indicates a positive correlation is predicted; a "-" indicates a negative correlation is predicted. It should be noted that some subtests were predicted to elicit behavior associated with a subset of a factor's facets (e.g., 1 of 4); these subtests are predicted to be associated with factor level scores in addition to the specified facet.

Test Battery scoring

Another consideration in designing the Test Battery was the scoring system, or how dogs' behavioral responses were assigned numerical values. Existing Test Batteries have employed various scoring systems. Some of these systems are driven by the goal of attaining objectivity by focusing on specific elements of behavior. For example, Ledger and Baxter (1997) used behavioral coding, attempting to categorize then count as many behavioral responses (e.g., whines) as possible, and Mahut (1958) attempted to describe dogs' behavior only in terms of what was observable (e.g., tail wags, steps). Some researchers have attempted to create a very simple coding system by placing dogs' behavior in one of two categories (e.g., pass or fail; Weiss, 2002). Other researchers have asked raters to subjectively interpret dogs' behavior by using Likert scales to rate the degree to which each dog showed certain types of behavior (e.g., nervous and shy behavior, Gosling et al., 2003a).

Taylor and Mills (2006) provided evaluative critiques of these scoring methods and others, advising against the most subjective scoring systems. They suggested a Likert scale labeled in terms of behavioral reactions (e.g., 1 = no aggression, 5 = biting) as used

by Netto and Planta (1997) may be a satisfactory combination of the behavioral detail attained in behavioral coding and the speed and ease of subjective ratings. Unfortunately, the dog personality literature presents too little reliability or validity data (see Chapter 2; Taylor & Mills, 2006) to argue strongly for or against any one method of scoring Test Batteries. So, in attempting to decide which method of scoring to use, I turned to the human personality literature. Measures of behavior in psychological research have been described as having two major limitations: (1) studies tend to focus narrowly on ecologically uninteresting but relatively easily specified and defined behaviors (e.g., response latencies), and (2) measures tend to focus on only a few, specific behaviors, while individuals emit many relevant and complex behaviors (Funder et al., 2000). Funder et al. (2000) recommended avoiding these limitations by assessing behaviors that are meaningful and relevant to the situation and that require only minimal interpretation. Rating dogs broadly on traits (e.g., aggression, excitability) pertinent to the study instead of using detailed behavioral coding to count the number of times a dog, for example, growls or jumps up to greet a person, might avoid these pitfalls. However, it still leaves open the question of whether to use a system such as Netto and Planta's (1997) with a Likert scale in which specific behaviors are associated with each rating, or a system such as Gosling et al.'s (2003a) with a Likert scale in which items ask, for example, "How excitable is this dog?" and responses vary from 1 = "not at all excitable" to 5 = "extremely excitable".

Taylor and Mills' (2006) review presented the issue that dogs of different breeds display dissimilar behaviors associated with some traits. It is likely inappropriate to assess every dog's level of a given trait by measuring the frequency with which or degree to which the dog displays one specific behavior, because the behavior may not be in the behavioral repertoire for a given breed. For example, if aggression is assessed based only on barking and growling, aggression level may be inaccurately assessed for dogs belonging to less vocal breeds (e.g., Basenjis). Thus, it may be problematic to restrict scoring of a Test Battery to a scale that enumerates specific behaviors.

In the free-response portion of Study 1 of this dissertation, participants indicated that they found it easier to rate their dogs on items assessing some personality traits (e.g., aggression, friendliness) if associated behaviors were not specified. Indeed, if behaviors were specified, participants indicated uncertainty about how to rate dogs who did not display all of the behaviors listed. If an owner thought her dog was extremely friendly and the dog tended to wag her tail a lot when greeting people, but the dog did not show a large amount of lip licking when greeting people, was that dog not actually an extremely friendly dog? It follows from participants' feedback that the specific behaviors dogs display in association with a given trait, while driven by the same underlying trait, differ from dog to dog. That is, not all extremely excitable dogs behave in the exact same way.

A Likert scale dependent on a list of behaviors could lead to inaccurate ratings because of behavioral differences associated both with breeds (Taylor & Mills, 2006) and because of raters' difficulty applying the scale. To avoid these issues, I used a scoring system in which ratings are made using 5-point Likert scale specifying a broader description of behavior (e.g., "not at all" to "extremely" for each behavioral description) rather than specific behaviors (e.g., "remains still" to "jumps, barks multiple times").

Funder et al. (2000) pointed out that individuals are likely to emit many behaviors in any given situation. For example, a dog may display behavior associated with Factor 1 (Fearfulness), Factor 2 (Aggression towards People), and/or Factor 3 (Activity/Excitability) when approached by an unfamiliar person during a subtest (subtests 3 and 4). So, although subtests were selected based on whether they were predicted to elicit behavior motivated by specific facets and factors on the DPQ, dogs' behavior was scored on multiple descriptions for each subtest (as shown on the scoring sheet presented in Appendix K). However, dogs' were rated on behavioral descriptions though to capture key elements of behavior related to each subtest situation, not non-central behavior. For example, dogs were rated on how fearful, confident, aggressive, and friendly they seemed when they were on-leash and near an unfamiliar male dog (Subtest 7), but not during a subtest in which the researcher instructed the dog in basic commands (Subtest 15). In total, the behavioral descriptions included fearful; confident; aggressive;

friendly; excitable; calm; engaged, alert (to the environment); easy to control; submissive; pushy, assertive; playful; interested in people; aloof or uninterested; affectionate; easy to train; and obedient.

Measuring many aspects of behavior predicted to be related to the five personality traits yielded multiple scores for each dog on each description, thus increasing the amount of information gathered about each dog. Behavioral scores can thus be examined separately or averaged across all assessments of each type of behavior (e.g., fearful, confident, aggressive on all subtest for which each was scored). Both of these yield useful information. Does a factor or facet of the DPQ predict a specific behavior in response to a specific subtest? Does a factor or facet of the DPQ correlate with an average of behavioral scores across subtests? For example, does the DPQ factor Fearfulness predict the average fearful score on the Test Battery? Table 9.1 indicates the specific behavioral descriptions that were predicted to correlate with each facet and/or factor and thus used to test each of these predictions.

Table 9.1. Predicted relationships between behavioral descriptions and DPQ factors and facets

	Subtest																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	Collar grab	**Walk on leash with stranger**	**Threaten-ing approach**	**Friendly approach**	**Affection or Petting from stranger**	**Restraint**	**Other dog (on leash)**	**Novel situation (room)**	**Novel object**	**Doll test**	**Prey drive**	**Engage-ment in play with tester**	**Tug-o-war/toy release**	**Train new task**	**Basic com-mands**	**Other dog (off leash)**	**Activity in free play**	**Reunion with owner**
Factor Facet																		
Factor 1 – Fearfulness	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**	**+fearful -confident**						**+fearful -confident**		
Facet 1 – Fear of People	+fearful -confident	+fearful -confident	+fearful -confident	+fearful -confident	+fearful -confident	+fearful -confident				+fearful -confident								
Facet 2 – Non-social Fear		+fearful -confident						+fearful -confident	+fearful -confident									
Facet 3 -- Fear/ Submission towards Dogs							+fearful -confident									+fearful -confident		
Facet 4 – Fear during Handling	+fearful -confident				+fearful -confident	+fearful -confident												
Factor 2 – Aggression towards People	**+aggressive -friendly**		**+aggressive -friendly**	**+aggressive -friendly**	**+aggressive -friendly -aloof**	**+aggressive -friendly -affection**				**+aggressive -friendly**								
Facet 1 – General Aggression towards People			+aggressive -friendly	+aggressive -friendly						+aggressive -friendly								
Facet 2 – Situational Aggression towards People	+aggressive -friendly		+aggressive -friendly	+aggressive -friendly	+aggressive -friendly -aloof	+aggressive -friendly -affection												

Table 9.1. (*Continued*)

	Subtest																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	Collar grab	**Walk on leash with stranger**	**Threaten-ing approach**	**Friendly approach**	**Affection or Petting from stranger**	**Restraint**	**Other dog (on leash)**	**Novel situation (room)**	**Novel object**	**Doll test**	**Prey drive**	**Engage-ment in play with tester**	**Tug-o-war/toy release**	**Train new task**	**Basic com-mands**	**Other dog (off leash)**	**Activity in free play**	**Reunion with owner**
Factor Facet																		
Factor 3 – Activity/ Excitability		**+excitable -calm +engaged, alert**	**+excitable -calm +playful +engaged, alert +interested in people**	**+excitable -calm +playful +engaged, alert +interested in people**	**+excitable -calm +playful -aloof +engaged, alert +interested in people**	**+excitable -calm +affection**	**+excitable -calm +playful -aloof**	**+excitable -calm +playful +engaged, alert**	**+excitable -calm +playful +engaged, Alert**	**+excitable -calm**	**+excitable -calm +playful +engaged**	**+excitable -calm +playful -aloof +engaged, alert +affection**		**+engaged +interested in people**		**+excitable -calm +playful -aloof**	**+excitable -calm +active -aloof +engaged, alert +playful**	**+excitable -calm -aloof +interested in people**
Facet 1 – Excitability		+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm	+excitable -calm				+excitable -calm	+excitable -calm +active	+excitable -calm
Facet 2 – Playfulness			+playful	+playful	+playful -aloof		+playful -aloof				+playful	+playful -aloof				+playful -aloof	-aloof	
Facet 3 – Active Engagement		+engaged, alert	+engaged, alert	+engaged, alert	+engaged, alert			+engaged, alert	+engaged, alert		+engaged	+engaged, alert		+engaged			+engaged, alert +active	-aloof
Facet 4 – Companion-ability			+interested in people	+interested in people	-aloof +playful +interested in people +affection	+affection	-aloof					+affection -aloof +playful		+interest in people			+playful	-aloof +interested in people
Factor 4 – Responsive-ness to Training		**+easy-to-control +sub-missive -pushy/ assertive**											**+easy to train +easy to control +sub-missive -pushy, assertive**	**+easy to train +obedient**	**+obedient**			**+easy to control +sub-missive -pushy, assertive**
Facet 1 – Trainability													+easy to train +easy to control	+easy to train +obedient	+obedient			
Facet 2 – Control-lability		+easy-to-control +sub-missive -pushy/ assertive											+easy to train +easy to control +sub-missive -pushy, assertive		+obedient			+easy to control +sub-missive -pushy, Assertive

Table 9.1. (*Continued*)

	Subtest																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Factor	**Collar grab**	**Walk on leash with stranger**	**Threaten-ing approach**	**Friendly approach**	**Affection or Petting from stranger**	**Restraint**	**Other dog (on leash)**	**Novel situation (room)**	**Novel object**	**Doll test**	**Prey drive**	**Engage-ment in play with tester**	**Tug-o-war/toy release**	**Train new task**	**Basic com-mands**	**Other dog (off leash)**	**Activity in free play**	**Reunion with owner**
Facet																		
Factor 5 – Aggression towards Animals							**+aggressive -friendly -playful -submissive +pushy, assertive**				**+aggressive**					**+aggressive -friendly -submissive +pushy, assertive**		
Facet 1 – Aggression towards Dogs							+aggressive -friendly -playful -aloof									+aggressive -friendly		
Facet 2 – Prey Drive											+aggressive							
Facet 3 – Dominance over Other Dogs							+aggressive -friendly -submissive -pushy, assertive +aloof									-submissive +pushy, assertive		

Note. The predicted direction, or valence, of the loadings is indicated by the "+" or "-" preceding each descriptor.

Participants

The Test Batteries were conducted at an Austin-area dog daycare and kennel, and the 125 (25 in pilot testing, 100 in Study 6) dogs and owners who participated were regular patrons of the facility. All participants were required to show evidence of current vaccinations (rabies, Bordatella) or to have records on file at the kennel before they were assessed using the behavioral Test Battery.

The diversity and how familiar kennel staff were with the 25 dogs who took part in pilot testing was only a minor concern, because the purpose of the pilot testing was to allow testers to streamline the test process, practice scoring, and to examine initial behavioral assessment score correlations among judges.

To ensure that the kennel was not a novel environment to any of the dogs and that the kennel staff who rated the dogs personalities using the DPQ had ample opportunity to become familiar with each dog in the 100-dog sample in Study 6, only dogs who had been to the kennel at least 10 times in the six month prior to testing were included. To attempt to ensure a relatively diverse group of 100 dogs, owners of dogs of a variety of breeds, sizes, and ages were invited to participate. The demographic data for Study 6 owners (Table 9.2) and their dogs (Table 9.3) are presented below.

Sample demographics

Both sexes of dog owners were represented in Study 6. Again, there were fewer men than women; however, men composed over a third of the sample (37%, or 37 men), a greater percentage than in previous studies. The average age of participants was 39.8 years, with a standard deviation of 12.7 years – similar to previous studies.

All of the owners and their dogs were Austin-area residents. The majority of participants identified themselves as Caucasian/White (N = 94). Other participants identified themselves as African American/Black (N = 1), Hispanic (N = 3), East Asian (N = 1), and Multiracial or Other (N = 1).

Of the 100 participants, only four (4%) indicated having a dog-related job or career; three reported working as trainers or behaviorists, and one as a volunteer with a local organization that rescues and re-homes dogs.

The mean number of dogs each participant had owned in his or her lifetime, not including current dogs but including childhood family dogs, was 5.0 (s.d. = 3.7). The mean number of dogs currently owned by participants in Study 6 was 1.5 (s.d. = .8).

Table 9.2. Demographic information for dog owner participants in Study 6

		Dog owners
Sex (dog owner)	Male	37 (37.0%)
	Female	63 (63.0%)
Age (dog owner)	Mean	39.8
	standard deviation	12.7
	Age not reported	0
Race	Caucasian/White	94 (94.0%)
	African-American/Black	1 (1.0%)
	Hispanic	3 (3.0%)
	East Asian	1 (1.0%)
	Native American	0
	Asian Indian	0
	Multiracial or Other	1 (1.0%)
Dog-related careers	Breeder and/or Exhibitor	0
	Trainer and/or Behaviorist	3 (3.0%)
	Dog rescue worker/volunteer	1 (1.0%)
	Veterinarian	0
	Groomer	0
	Kennel/Daycare owner or employee	0
	Assistance dog partner	0
Number of dogs owned	Past (mean; s.d.)	5.0 (3.7)
	Number not reported	0
	Currently (mean; s.d.)	1.5 (.8)
	Number not reported	0
Total number of participants		**100**

The demographic information for the 100 dogs in Study 6 was examined to confirm that the sample was diverse. The first dimension I examined was age. Study 6, the dogs' ages ranged from one to 12 years, and averaged 5.0 years (s.d. = 2.7).

Other demographic information about the dogs is presented in Table 9.3. As in the previous chapters, the number of dogs in each category is listed in the far right column, which is entitled "Number of dogs". The sample of 100 dogs included dogs of both sexes and a variety of breeds or lineages, though most of the dogs are reported to primarily play the role of being pets. Sixty percent (N = 60) of the dogs were male. All of the dogs were castrated, because all dogs over the age of one year are required to be castrated to attend the kennel's daycare program unless medical issues do not permit it.

Purebred dogs composed slightly more than half the sample (N =54, or 54%), with 43 breeds represented. The breeds most represented included the Labrador Retriever (N = 5) and Golden Retriever (N = 4). The sample also included seven mixed-breed dogs thought or known to be partially pit bull. Dogs of relatively rare breeds (e.g., Japanese Chin, Schipperke) were also included in the sample.

Information about whether the dogs were reported to have ever bitten a person, had any reported health issues, what role they played in their owners' lives (e.g., pet), and whether they were involved in any sports is also included in Table 9.3.

Only one dog was reported to have bitten someone (dogs believed to be aggressive are not permitted at the daycare facility for safety reasons), two dogs had unknown bite histories, and the remaining 97 were reported to have never bitten a person (intentionally, outside play, and outside work or sport requiring biting, e.g., Schutzhund training or competition). Twelve dogs were indicated to have health problems or disabilities, with two dogs having more than one health problem or disability. Owners who indicated their dogs had other disabilities listed their dogs as having, allergies, frequency ear infections, and a congenital deformation that resulted in a missing leg.

All of the dogs in this sample were considered pets or companions by their owners. None were working dogs, though one did Animal Assisted Therapy visits to a

local retirement community. As in the previous studies, the popularity of agility was also represented in this sample (N = 11).

Table 9.3. Demographic information about target dogs in Study 6

			Number of dogs
Sex	Male		60 (60.0%)
	Female		40 (40.0%)
Age	Mean (s.d.)		5 (2.7)
Castration Status	Males	Neutered	60 (100.0%)
		Intact	0
	Females	Spayed	40 (100.0%)
		Intact	0
Purebred	Unknown		0
	No (mixed)		46 (47.0%)
	Yes		54 (53.0%)
Bitten a Person	Unknown		2 (2.0%)
	No		97 (97.0%)
	Yes		1 (1.0%)
Disability and Health issues	Unilaterally deaf		1 (1.0%)
	Bilaterally deaf		1 (1.0%)
	Blind in one eye		1 (1.0%)
	Blind in both eyes		0
	Arthritic		2 (2.0%)
	Hip dysplasia		2 (2.0%)
	Elbow dysplasia		3 (3.0%)
	Other disabilities		4 (4.0%)
Job or Role	**Pet/Companion**		100 (100.0%)
	Assistance dog		0
	Guide dog		0
	Hearing ear dog		0
	Medical assistance dog		0
	Search and rescue		0
	Guard dog		0
	Animal Assisted Therapy		1 (1.0%)
	Dam or sire for breeding		0

Table 9.3. (*Continued*)

		Number of dogs
Sport	Obedience	1 (1.0%)
	Sledding	0
	Carting	0
	Frisbee	2 (2.0%)
	Earth dog	2 (2.0%)
	Show/Conformation	0
	Schutzhund	0
	Agility	11 (11.0%)
	Herding (competitive)	1 (1.0%)
	Flyball	0
	Hunting	2 (1.0%)
Total number of dogs		**100**

Note. The sum of the number of dogs who participate in sports will not equal the total number of dogs in the study; some dogs are not in sports, and others are in multiple. Similarly, participants may consider a dog to be both a pet/companion and to have another job or role (e.g., assistance dog, search and rescue dog).

Dogs with unknown histories (e.g., bite history) are typically dogs who were adopted from shelters or rescues.

Materials and Procedures

DPQ

After owners volunteered their dogs for the study and the dogs vaccination records were verified, the owners and two kennel staff rated the dog using the paper version of the long form of the DPQ; this form is presented in Appendix H and is the same form presented online in Studies 4 and 5. The kennel staff, one man and one woman, had both worked at the kennel (taking care of and thus directly interacting with the dogs by moving them from kennel to kennel, grooming them, feeding them, and so on) for at least six months. The same two staff members rated all 100 dogs in Study 6. They completed their ratings independently and before the Test Batteries were conducted.

The dogs' owners also filled out the long form of the DPQ; they completed the questionnaire while in the kennel waiting area, waiting to pick up their dogs on the day

that the dog was tested but prior to being reunited with their dogs (Subtest 18). The owners were not witness to any part of the Test Battery except Subtest 18.

Test Batteries

Two additional kennel staff, a female familiar to the dogs and a newly hired male novel to the dogs, conducted all of the Test Batteries. Tester 1, the familiar female, conducted Subtest 1 and 2, and held the dogs on leash for Subtests 3, 4, 5, 7, 10, and 18. Tester 2, the novel male, conducted Subtests 3-17. This division of testing allowed observation of the dogs with both familiar and unfamiliar people (e.g., Subtest 3, the threatening approach by a novel person).

Two subtests in the Test Battery required a second dog. For these tests, the same dog was always used. This dog was a six-year-old male of unknown lineage (possibly part Australian Cattle Dog). He was neutered, had never threatened or bitten a person or another dog, and had grown up spending three days per week in the reception area of a busy veterinary office and was thus accustomed to being around a large variety of dogs and people. He weighed about 55 pounds, was light to medium brown, and had dark brown eyes, erect ears, and a long tail.

The same two testers and test dog conducted every test and played the same roles in every test to reduce variability in how the tests were conducted across dogs. To minimize biases in testing and scoring, the testers worked independently from the kennel staff who provided DPQ ratings of each dog and were not privy to how each dog had been scored on the DPQ until all Test Batteries were completed.

All Test Batteries were conducted between 9a.m. and 3p.m. during July, 2007, with exception of Subtest 18 (the reunion with owner). Subtests 1 through 17 took approximately 20-25 minutes per dog. The final subtest, Subtest 18, was completed any time between 4p.m. and 7p.m., when dogs were picked up from the kennel by their owners.

Immediately after each Test Battery subtest, the two individuals who tested the dogs scored them using the score sheet presented in Appendix K. I, as a third independent

judge, also scored the dogs after each subtest. As a result, each dog had a total of three scores for every subtest.

Pilot testing

Minimizing variance from sources extraneous to the personality of the target dog (e.g., variance due to environmental differences, different testers) during the Test Battery served to strengthen the statistical relationship between behavioral tests and theoretically related ratings of dogs' personalities on the DPQ. In conducting the Test Battery, as many variables as possible were held constant. For example, when each dog was tested, only that dog and any other people or stimuli necessary to the subtest were in the test area, which should have served to minimize the degree to which target dogs are distracted or otherwise affected by uncontrolled stimuli (e.g., other dogs). The sequence of subtests, the testers, and as many other aspects of the test procedure as possible were held as constant as possible in an effort to have any variance in behavior the dogs display during the test be attributable to the dogs themselves (Taylor & Mills, 2006).

Consistency of testing was increased through pilot testing, or practicing, before conducting the test with any of the 100 dogs included in Study 6. The Test Battery was piloted with 25 dogs. The ICCs among behavioral scores of the initial 25 dogs were examined and found to range from .471 to .934 with a mean of .782. Pilot test dogs participated only in the Test Battery and were not rated on the DPQ.

The original research plan included videotaping all Test Batteries. However, pilot testing revealed that the majority of dogs (N = 18) were notably distracted and/or disturbed by the presence of a video camera, even if the camera stationary, as indicated by their barking, growling, or staring at the camera. As a consequence, Test Batteries were not videotaped.

ANALYSES

The degree to which facet and factor scores on the DPQ (derived from owners' and kennel staff's ratings of the dogs' personalities) predict the independent observers' ratings of behavior scored during the Test Battery served as a measure of the predictive

validity of the DPQ. To evaluate the convergent and discriminant validity of the dog personality questionnaire against the criteria of behavior elicited and rated during the Test Battery, I correlated the DPQ factor ratings (averaged across all three raters) with theoretically related and unrelated mean scores on the Test Battery (averaged across all three raters). Because the short form of the questionnaire (developed in Study 3) bases scores associated with each facet and factor on a subset of the items in the long form of the questionnaire, the convergent and discriminant validity of the two forms was evaluated separately. The long form collects more data, and so it was expected to have more accurate predictive validity (i.e., to have larger convergent correlations and smaller discriminant correlations).

Inter-rater reliability of DPQ factor and facet ratings

The DPQ factor and facet scores used in Study 6 predictive validity analyses are composite, or mean, scores across three raters who rated each dog independently; the same two kennel staff, and each dog's owner, rated all 100 dogs. In Study 4, the inter-rater reliability of the DPQ was examined at the item-, factor-, and facet-level and determined to be relatively high, comparable to previous dog personality studies and to human personality ratings. Prior to averaging scores across the three raters in the current study, I examined the inter-rater reliability at the factor- and facet-level to confirm that inter-rater reliability is high. The factor- and facet-level scores are the focus of these analyses because those were the scores used in the predictive validity analyses.

As previously discussed, reliability of a test is a necessary prerequisite to validity. In the current study, inter-rater reliability correlations were computed for the DPQ using ICC (a one-way random-effects model). This method was selected because there were a total of three raters, ruling out use of Pearson's *r*, and because, although two raters were constant for all 100 dogs, the third rater was different for every dog. As displayed in Table 9.4, inter-rater reliability correlations (ICCs) were relatively high – higher than those from Study 4. As an estimate of the scores reliability, Cronbach's alpha is also included in Table 9.4. For the long form, which was expected to be associated with

higher levels of reliability because it included more items per facet and factor, the factor-level ICCs ranged from .833 (Activity/Excitability) to .906 (Fearfulness), with a mean of .856 across all five factors. The facet-level ICCs associated with the long form ranged from .672 (Playfulness on the Activity/Excitability factor) to .874 (Nonsocial Fear on the Fearfulness factor). For the short form, the factor-level ICCs ranged from .769 (Aggression towards Animals) to .881 (Fearfulness), with a mean of .819 across all five factors. The facet-level ICCs associated with the short form ranged from .586 (Playfulness on the Activity/Excitability factor) to .834 (Fear of Handling on the Fearfulness factor). It was concluded that inter-rater reliability among these raters was high enough to support using the average of their scores of each dog in the following predictive validity analyses.

Table 9.4. Facet- and factor-level inter-rater reliability

Factor	Long form		Short form	
Facet	ICC	Alpha	ICC	Alpha
Factor 1 – Fearfulness	**.908**	**.971**	**.881**	**.960**
Facet 1 – Fear of People	.774	.918	.663	.858
Facet 2 – Nonsocial Fear	.874	.956	.821	.934
Facet 3 – Fear of Dogs	.845	.948	.770	.916
Facet 4 – Fear of Handling	.852	.946	.834	.938
Factor 2 – Aggression towards People	**.836**	**.950**	**.786**	**.924**
Facet 1 – General Aggression	.784	.926	.685	.881
Facet 2 – Situational Aggression	.775	.920	.713	.882
Factor 3 – Activity/Excitability	**.833**	**.946**	**.788**	**.927**
Facet 1 – Excitability	.844	.942	.766	.907
Facet 2 – Playfulness	.672	.869	.586	.824
Facet 3 – Active Engagement	.711	.894	.679	.875
Facet 4 – Companionability	.812	.935	.783	.917
Factor 4 – Responsiveness to Training	**.851**	**.946**	**.846**	**.944**
Facet 1 – Trainability	.817	.932	.823	.933
Facet 2 – Controllability	.839	.940	.791	.921
Factor 5 – Aggression towards Animals	**.837**	**.939**	**.769**	**.908**
Facet 1 – Aggression towards Dogs	.844	.944	.760	.906
Facet 2 – Prey Drive	.826	.935	.803	.926
Facet 3 – Dominance over other Dogs	.825	.934	.710	.880
MEAN across Factors	**.856**	**.950**	**.819**	**.933**

Note. Mean correlations are computed using Fisher's r-to-z transformation. Single-measure ICCs are reported

Inter-rater reliability of Test Battery behavioral scoring

The criteria against which the DPQ was evaluated were ratings on the Test Battery. High inter-rater reliability levels on the Test Battery support the claim that the dogs' scores are functions of the dogs themselves, not of the people rating the dogs (though not all other sources of variance can be eliminated). For that reason, and because I intended to use the mean of the three raters scores on each behavioral description for each Test Battery subtest, I examined inter-rater reliability among the three raters. The results are presented in Table 9.5.

The mean ICC, across all 118 scores on all 18 Subtests of the Test Battery, was .764 (computed using Fisher's r-to-z transformation). ICCs ranged from .566 ("Engaged,

Alert" on Subtest 8: Novel situation (room)) to .920 ("Affectionate" on Subtest 6: Restraint). Inter-rater reliability among these raters was high enough to support using the average of their scores of each dog in the following predictive validity analyses.

Table 9.5. Test Battery inter-rater reliability

	ICC	Alpha
Subtest 1: Collar Grab		
Fearful	.690	.870
Confident	.580	.805
Aggressive	.668	.858
Friendly	.660	.853
Subtest 2: Walk on leash with stranger/tester		
Fearful	.812	.928
Confident	.704	.878
Excitable	.726	.887
Calm	.755	.902
Engaged, Alert	.576	.802
Easy to control	.689	.872
Submissive	.732	.891
Pushy, Assertive	.743	.896
Subtest 3: Threatening approach		
Fearful	.754	.901
Confident	.683	.865
Aggressive	.676	.861
Friendly	.794	.920
Excitable	.794	.918
Calm	.666	.857
Playful	.729	.889
Engaged, Alert	.573	.800
Interested in people	.717	.883

Table 9.5. (*Continued*)

	ICC	Alpha
Subtest 4: Friendly approach		
Fearful	.718	.883
Confident	.683	.865
Aggressive	.733	.891
Friendly	.800	.918
Excitable	.725	.887
Calm	.763	.907
Playful	.752	.900
Engaged, Alert	.698	.875
Interested in people	.674	.861
Subtest 5: Affection/Petting from stranger		
Fearful	.706	.878
Confident	.669	.858
Aggressive	.687	.868
Friendly	.826	.934
Excitable	.755	.902
Calm	.739	.895
Aloof or Uninterested	.686	.867
Playful	.777	.912
Engaged, Alert	.716	.883
Interested in people	.784	.916
Affectionate	.811	.929
Subtest 6: Restraint		
Fearful	.758	.903
Confident	.786	.916
Aggressive	.752	.900
Friendly	.804	.924
Affectionate	.920	.932
Calm	.793	.920
Excitable	.764	.906

Table 9.5. (*Continued*)

	ICC	Alpha
Subtest 7: Other dog (on leash)		
Fearful	.804	.924
Confident	.753	.901
Aggressive	.844	.942
Friendly	.846	.943
Excitable	.773	.911
Calm	.842	.941
Aloof or Uninterested	.733	.891
Playful	.783	.916
Submissive	.746	.898
Pushy, Assertive	.815	.930
Subtest 8: Novel situation (room)		
Fearful	.838	.939
Confident	.756	.902
Excitable	.778	.913
Calm	.760	.904
Engaged, Alert	.566	.796
Subtest 9: Novel objects		
Fearful	.800	.922
Confident	.736	.892
Excitable	.779	.913
Calm	.793	.919
Engaged, Alert	.612	.824
Subtest 10: Doll test		
Fearful	.753	.901
Confident	.714	.882
Aggressive	.803	.924
Friendly	.818	.931
Excitable	.800	.923
Calm	.801	.924
Subtest 11: Prey drive		
Aggressive	.777	.912
Excitable	.753	.901
Calm	.739	.895
Playful	.882	.957
Engaged	.631	.836

Table 9.5. (*Continued*)

	ICC	Alpha
Subtest 12: Engagement in play with tester		
Affectionate	.861	.948
Excitable	.861	.949
Calm	.782	.915
Aloof or Uninterested	.729	.890
Playful	.838	.939
Engaged, Alert	.670	.860
Interested in people	.756	.902
Subtest 13: Tug-o-war, Toy release		
Easy to train	.848	.943
Easy to control	.812	.928
Submissive	.852	.945
Pushy, Assertive	.821	.933
Subtest 14: Train new task		
Engaged	.649	.847
Interested in people	.813	.929
Easy to train	.785	.916
Obedient	.828	.934
Subtest 15: Basic commands		
Obedient	.787	.917
Subtest 16: Other dog (off leash)		
Fearful	.849	.944
Confident	.811	.927
Aggressive	.848	.944
Friendly	.739	.894
Excitable	.684	.866
Calm	.789	.917
Aloof or Uninterested	.765	.907
Playful	.748	.899
Submissive	.741	.895
Pushy, Assertive	.786	.916
Subtest 17: Activity in free play		
Active	.777	.912
Excitable	.803	.924
Calm	.806	.926
Aloof	.720	.885
Playful	.775	.911
Engaged, Alert	.706	.877

Table 9.5. (*Continued*)

	ICC	Alpha
Subtest 18: Reunion with owner		
Excitable	.839	.939
Calm	.751	.901
Aloof or Uninterested	.757	.903
Interested in people	.680	.864
Easy to control	.807	.925
Submissive	.768	.908
Pushy, Assertive	.709	.879

Note. Single-measure ICCs are reported.

Predictive validity: Correlation between DPQ and Test Battery ratings

The relationships between facet and factor scores on the DPQ (derived from owners' and kennel staff's ratings of the dogs' personalities) and the independent observers' ratings of facet- and factor-relevant (and unrelated) behavior displayed during the Test Battery served as a measure of the predictive validity of the DPQ. The mean (across all three raters) score for each factor and facet was correlated with every mean (across three independent raters) behavioral assessment score on the Test Battery subtests.

Results of these predictive validity analyses are presented in Tables 9.6 to 9.10. Each table presents the discriminant and convergent validity correlations for a single factor and its facets. The columns present the long form's and then the short form's factors and facets. The rows list each subtest, in the order in which they were completed, and every behavioral description on which dogs were rated for each subtest.

Convergent correlations are those that were predicted to be high and are presented in bold in all five tables. Discriminant correlations are those that were predicted to be low and are presented in normal font (i.e., not bolded) in the tables. Correlations that are as predicted are underlined in the table. Correlations that were significant at the $p \leq .001$, or that are greater than ± .320, are considered "high". Correlations that were less than ± .190, and which were thus not significant ($p \geq .05$), were considered "low". Thus, the

correlations that are not underlined are those which were predicted to be high but were actually < ± .320, and those that were predicted to be low but were actually > ± .190.

Because specific predictions regarding correlations were made, Bonferroni adjustments are not necessarily mandated. However, a conservative cut-off of a Bonferroni-corrected p = .05 (± .400) for convergent correlations is also included. Correlations that were predicted to be convergent and were greater than or equal to ± .400 are marked with an asterisk in Tables 9.6 through 9.10.

Table 9.6. Fearfulness convergent and discriminant validity correlations

Factor 1: Fearfulness										
	Long form					**Short form**				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 1: Collar Grab										
Fearful	**.622***	**.507***	.555	.518	**.521***	**.612***	**.496***	.586	.405	**.465***
Confident	**-.599***	**-.481***	-.541	-.502	**-.495***	**-.593***	**-.467***	-.565	-.398	**-.461***
Aggressive	.493	.491	.438	.348	.392	.466	.377	.436	.241	.412
Friendly	-.113	-.151	-.061	-.070	-.116	-.127	-.079	-.092	.035	-.226
Subtest 2: Walk on leash with stranger/tester										
Fearful	**.537***	**.468***	**.465***	.492	.401	**.548***	**.482***	**.516***	.401	.368
Confident	**-.530***	**-.426***	**-.502***	-.467	-.389	**-.543***	**-.455***	**-.522***	-.412	-.361
Excitable	.233	.151	.342	.118	.132	.237	.198	.293	.095	.152
Calm	-.240	-.136	-.328	-.132	-.175	-.231	-.168	-.267	-.091	-.188
Engaged, Alert	.055	-.009	.141	.014	.005	.043	-.005	.091	.004	.031
Easy to control	-.033	.013	-.022	-.068	-.035	.003	.032	-.025	.003	.006
Submissive	.308	.168	.313	.388	.168	.305	.205	.288	.473	.077
Pushy, Assertive	-.228	-.113	-.233	-.263	-.154	-.235	-.144	-.234	-.371	-.054
Subtest 3: Threatening approach										
Fearful	**.620***	**.507***	.575	.447	.552	**.620***	**.499***	.603	.344	.509
Confident	**-.639***	**-.547***	-.558	-.495	-.557	**-.636***	**-.525***	-.598	-.378	-.515
Aggressive	.371	.387	.319	.157	.384	.353	.252	.329	.089	.402
Friendly	-.221	-.305	-.128	-.009	-.314	-.253	-.227	-.181	.025	-.371
Excitable	.158	.059	.259	.083	.090	.164	.107	.218	.060	.116
Calm	-.150	-.078	-.243	-.051	-.097	-.182	-.140	-.211	-.068	-.143
Playful	-.112	-.102	-.033	-.048	-.205	-.120	-.096	-.068	-.010	-.187
Engaged, Alert	.055	-.009	.141	.014	.005	.043	-.005	.091	.004	.031
Interested in people	-.026	-.157	.081	-.047	-.010	-.043	-.119	.058	-.050	-.054

Table 9.6. (*Continued*)

Factor 1: Fearfulness (*Continued)*

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 4: Friendly approach										
Fearful	**.458***	**.477***	.342	.367	.386	**.452***	**.377**	.404	.346	.331
Confident	**-.517***	**-.532***	-.408	-.399	-.430	**-.505***	**-.441***	-.458	-.396	-.342
Aggressive	.254	.314	.180	.175	.209	.213	.174	.197	.162	.154
Friendly	-.098	-.182	-.071	-.005	-.082	-.070	-.047	-.063	.006	-.101
Excitable	.120	.100	.225	.018	.028	.122	.210	.157	.011	.016
Calm	-.096	-.085	-.210	.053	-.040	-.101	-.152	-.159	.013	-.015
Playful	-.184	-.246	-.087	-.104	-.208	-.197	-.176	-.136	-.039	-.254
Engaged, Alert	.035	-.010	.140	-.082	.028	.065	.042	.095	-.077	.109
Interested in people	.000	-.050	.117	-.128	.012	.014	.037	.075	-.129	.025
Subtest 5: Affection/Petting from stranger										
Fearful	**.384**	**.465***	.264	.287	**.313**	**.397**	**.359**	.329	.324	**.279**
Confident	**-.420***	**-.482***	-.315	-.297	**-.349**	**-.418***	**-.377**	-.375	-.317	**-.288**
Aggressive	.242	.377	.107	.155	.219	.229	.231	.157	.160	.197
Friendly	-.093	-.228	-.002	.026	-.135	-.079	-.082	-.024	.022	-.150
Excitable	.187	.062	.285	.130	.116	.187	.132	.222	.136	.104
Calm	-.175	-.068	-.260	-.112	-.114	-.160	-.096	-.185	-.152	-.082
Aloof or Uninterested	.001	.077	-.067	-.008	.026	.012	.029	-.050	-.006	.066
Playful	-.032	-.113	.029	.044	-.083	-.042	-.040	-.006	.096	-.150
Engaged, Alert	.081	-.038	.144	.079	.059	.087	.019	.120	.112	.028
Interested in people	.068	-.052	.181	.060	.001	.080	.022	.169	.035	.012
Affectionate	-.059	-.186	.027	.083	-.141	-.079	-.082	.021	.051	-.216
Subtest 6: Restraint										
Fearful	**.601***	**.516***	.516	.413	**.582***	**.595***	**.460***	.546	.334	**.533***
Confident	**-.550***	**-.466***	-.479	-.385	**-.520***	**-.534***	**-.392**	-.500	-.313	**-.475***
Aggressive	.401	.445	.303	.263	.361	.405	.341	.344	.183	.402
Friendly	-.176	-.309	-.053	-.009	-.250	-.195	-.200	-.091	-.002	-.300
Affectionate	-.143	-.252	-.061	.015	-.199	-.162	-.164	-.084	.058	-.285
Calm	-.391	-.213	-.481	-.288	-.290	-.350	-.233	-.421	-.188	-.248
Excitable	.331	.228	.389	.226	.242	.297	.302	.316	.148	.181

Table 9.6. (*Continued*)

Factor 1: Fearfulness (*Continued)*

	Long form					**Short form**				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 7: Other dog (on leash)										
Fearful	**.570***	.459	.483	**.514***	.480	**.560***	.461	.485	**.469***	.402
Confident	**-.552***	-.458	-.476	**-.442***	-.493	**-.549***	-.462	-.474	**-.411***	-.421
Aggressive	.250	.355	.158	.138	.221	.237	.203	.191	.028	.299
Friendly	-.045	-.107	.041	.089	-.186	-.039	.034	.035	.070	-.219
Excitable	.235	.181	.322	.158	.103	.252	.277	.288	.145	.104
Calm	-.212	-.152	-.270	-.166	-.106	-.190	-.190	-.238	-.126	-.059
Aloof or Uninterested	-.060	-.060	-.106	-.082	.051	-.031	-.129	-.085	-.048	.130
Playful	-.032	-.139	.059	.058	-.108	-.047	-.050	.032	.051	-.161
Submissive	.394	.208	.443	.450	.211	.381	.281	.406	.459	.123
Pushy, Assertive	-.268	-.075	-.326	-.372	-.112	-.261	-.197	-.285	-.383	-.026
Subtest 8: Novel situation (room)										
Fearful	**.417***	.425	**.321**	.340	.343	**.414***	.368	**.375**	.345	.260
Confident	**-.392**	-.411	**-.272**	-.328	-.343	**-.390**	-.360	**-.316**	-.355	-.252
Excitable	.271	.244	.356	.098	.177	.278	.288	.344	.028	.193
Calm	-.125	-.068	-.210	-.016	-.089	-.116	-.104	-.168	.025	-.093
Engaged, Alert	.071	.041	.146	-.009	.034	.067	.105	.110	-.097	.064
Subtest 9: Novel objects										
Fearful	**.498***	.431	**.397**	.359	.501	**.487***	.393	**.415***	.345	.407
Confident	**-.476***	-.385	**-.382**	-.353	-.492	**-.455***	-.355	**-.376**	-.336	-.392
Excitable	.160	.092	.231	.063	.118	.171	.127	.205	.050	.142
Calm	-.184	-.084	-.258	-.076	-.163	-.193	-.116	-.218	-.075	-.182
Engaged, Alert	.080	.044	.116	-.009	.097	.100	.069	.064	.045	.132
Subtest 10: Doll test										
Fearful	**.571***	**.491***	.523	.444	.466	**.542***	**.415***	.545	.369	.396
Confident	**-.560***	**-.481***	-.505	-.427	-.473	**-.524***	**-.382**	-.519	-.360	-.405
Aggressive	.426	.432	.390	.235	.377	.407	.312	.390	.173	.390
Friendly	-.292	-.308	-.230	-.096	-.348	-.289	-.123	-.281	-.055	-.394
Excitable	.265	.142	.392	.165	.145	.268	.178	.351	.143	.162
Calm	-.192	-.118	-.290	-.116	-.089	-.170	-.123	-.240	-.087	-.079
Subtest 11: Prey drive										
Aggressive	.372	.281	.423	.143	.364	.340	.249	.370	.028	.372
Excitable	.126	.084	.158	.016	.141	.117	.127	.110	-.058	.161
Calm	-.249	-.130	-.269	-.147	-.266	-.238	-.167	-.254	-.094	-.218
Playful	-.062	-.140	-.029	.021	-.070	-.078	-.076	-.059	.075	-.156
Engaged	.019	.101	.058	-.126	.014	.012	.162	.020	-.180	.014

Table 9.6. (*Continued*)

Factor 1: Fearfulness (*Continued*)

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 12: Engagement in play with tester										
Affectionate	-.066	-.112	-.005	.022	-.140	-.087	-.061	-.038	.082	-.218
Excitable	.214	.052	.267	.228	.148	.188	.135	.256	.135	.073
Calm	-.218	-.113	-.277	-.159	-.158	-.205	-.164	-.233	-.094	-.149
Aloof or Uninterested	-.006	.091	-.077	-.001	-.003	.004	.038	-.042	-.020	.038
Playful	-.125	-.192	-.055	-.008	-.182	-.116	-.076	-.073	.001	-.195
Engaged, Alert	-.148	-.118	-.070	-.178	-.157	-.149	-.038	-.114	-.189	-.144
Interested in people	.049	-.069	.164	.014	.014	.042	.055	.105	-.026	-.013
Subtest 13: Tug-o-war, Toy release										
Easy to train	-.017	-.130	.012	.131	-.074	-.030	-.102	.016	.217	-.176
Easy to control	-.022	-.141	-.018	.175	-.083	-.035	-.091	-.029	.217	-.152
Submissive	.228	.109	.208	.332	.126	.233	.198	.205	.320	.071
Pushy, Assertive	-.148	-.008	-.119	-.324	-.063	-.153	-.126	-.114	-.348	.026
Subtest 14: Train new task										
Engaged	-.071	-.053	-.064	-.088	-.036	-.056	.010	-.074	-.079	-.037
Interested in people	.044	-.065	.070	.079	.052	.068	-.026	.065	.141	.044
Easy to train	-.016	-.192	.031	.136	-.044	-.046	-.109	.004	.155	-.157
Obedient	-.085	-.230	-.046	.094	-.112	-.078	-.152	-.031	.162	-.186
Subtest 15: Basic commands										
Obedient	-.121	-.237	-.100	.065	-.135	-.110	-.165	-.091	.143	-.192
Subtest 16: Other dog (off leash)										
Fearful	**.512***	.453	.438	**.419***	.428	**.511***	.440	.428	**.388**	.393
Confident	**-.517***	-.451	-.446	**-.437**	-.421	**-.523***	-.416	-.434	**-.447***	-.399
Aggressive	.314	.312	.218	.217	.329	.279	.158	.221	.197	.301
Friendly	-.276	-.326	-.106	-.176	-.364	-.291	-.197	-.141	-.210	-.379
Excitable	.174	.111	.247	.077	.117	.172	.190	.223	.027	.092
Calm	-.255	-.190	-.304	-.148	-.193	-.243	-.230	-.278	-.111	-.149
Aloof or Uninterested	-.014	.024	-.041	-.049	.026	.001	-.026	-.032	-.084	.113
Playful	-.069	-.131	-.029	.045	-.121	-.060	.012	-.055	.089	-.186
Submissive	.367	.185	.374	.427	.244	.378	.336	.351	.450	.136
Pushy, Assertive	-.350	-.145	-.369	-.369	-.276	-.370	-.282	-.359	-.436	-.156

Table 9.6. (*Continued*)

Factor 1: Fearfulness (*Continued*)

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 17: Activity in free play										
Active	.234	.126	.275	.159	.202	.252	.169	.268	.138	.210
Excitable	.206	.087	.298	.142	.130	.204	.157	.263	.109	.111
Calm	-.278	-.192	-.358	-.155	-.196	-.277	-.219	-.330	-.122	-.191
Aloof	-.164	-.061	-.270	-.092	-.088	-.144	-.170	-.220	-.077	.001
Playful	.031	-.062	.124	.049	-.037	.004	.027	.090	.085	-.159
Engaged, Alert	.109	.046	.199	-.064	.139	.112	.137	.149	-.092	.124
Subtest 18: Reunion with owner										
Excitable	.051	-.051	.133	-.009	.061	.051	.020	.116	-.031	.033
Calm	-.137	-.026	-.223	-.040	-.132	-.141	-.075	-.214	-.059	-.081
Aloof or Uninterested	-.092	-.016	-.213	.004	-.039	-.070	-.093	-.172	.084	-.012
Interested in people	.120	-.005	.219	.139	.017	.109	.144	.204	.058	-.050
Easy to control	.181	-.044	.204	.295	.140	.139	.013	.145	.293	.030
Submissive	.248	.093	.246	.311	.181	.253	.159	.227	.330	.129
Pushy, Assertive	-.239	-.115	-.209	-.302	-.185	-.252	-.175	-.193	-.360	-.125

Note. Bolded correlations are those that are predicted to be convergent. Other correlations are predicted to be discriminant. The cut-off for significance at p = .05 is ± .190; numbers below this are considered discriminant. The cut-off for significance at p = .01 is ± .250. The cut-off for significance at p = .001 is ± .320; numbers at or above .320 are considered convergent. The Bonferroni-corrected cut-off for significance at p = .05 is ± .400. Correlations that are in line with predictions are underlined. Correlations predicted to be convergent that are equal to or greater than .320 are both bolded and underlined. Correlations predicted to be discriminant that are equal to or less than .190 are underlined. Correlations predicted to be convergent that are equal to or greater than ± .400 are marked with asterisks.

Table 9.7. Aggression towards people convergent and discriminant validity correlations

Factor 2: Aggression towards People						
	Long form			**Short form**		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 1: Collar Grab						
Fearful	.422	.443	.335	.470	.445	.402
Confident	-.435	-.442	-.361	-.482	-.468	-.403
Aggressive	**.656***	.555	**.654***	**.645***	.527	**.627***
Friendly	**-.518***	-.457	**-.491***	**-.502***	-.439	**-.463***
Subtest 2: Walk on leash with stranger/tester						
Fearful	.337	.345	.277	.395	.393	.321
Confident	-.331	-.348	-.262	-.373	-.374	-.302
Excitable	-.051	.028	-.120	-.007	.016	-.026
Calm	.028	-.061	.110	-.022	-.063	.019
Engaged, Alert	-.029	-.001	-.052	-.024	.028	-.065
Easy to control	-.012	.037	-.062	-.039	-.037	-.033
Submissive	-.183	-.110	-.224	-.150	-.123	-.146
Pushy, Assertive	.185	.085	.254	.156	.090	.186
Subtest 3: Threatening approach						
Fearful	.378	.464	.234	.425	.476	.297
Confident	-.419	-.509	-.263	-.477	-.531	-.337
Aggressive	**.619***	**.557***	**.581***	**.593***	**.487***	**.574***
Friendly	**-.479***	**-.478***	**-.399**	**-.450***	**-.386**	**-.421***
Excitable	-.090	-.011	-.151	-.069	-.028	-.092
Calm	.073	.004	.129	.035	-.011	.069
Playful	-.245	-.243	-.205	-.249	-.207	-.239
Engaged, Alert	-.465	-.364	-.486	-.451	-.377	-.431
Interested in people	-.012	.037	-.062	-.039	-.037	-.033
Subtest 4: Friendly approach						
Fearful	.480	.465	.421	.499	.492	.409
Confident	-.462	-.454	-.399	-.478	-.467	-.396
Aggressive	**.507***	**.429***	**.505***	**.467***	**.403***	**.436***
Friendly	**-.586***	**-.474***	**-.600***	**-.535***	**-.417***	**-.539***
Excitable	-.205	-.125	-.251	-.161	-.120	-.166
Calm	.147	.051	.220	.115	.056	.146
Playful	-.500	-.456	-.459	-.476	-.412	-.441
Engaged, Alert	-.020	.014	-.053	.001	-.021	.021
Interested in people	-.337	-.301	-.319	-.316	-.334	-.239

Table 9.7. (*Continued*)

Factor 2: Aggression towards People (*Continued*)						
	Long form			**Short form**		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 5: Affection/Petting from stranger						
Fearful	.396	.369	.363	.416	.411	.341
Confident	-.460	-.472	-.377	-.490	-.506	-.382
Aggressive	**.448***	.398	**.430***	**.430***	.442	**.338**
Friendly	**-.559***	-.456	**-.570***	**-.526***	-.445	**-.498***
Excitable	-.124	-.048	-.177	-.062	-.022	-.086
Calm	.079	.010	.133	.026	-.017	.060
Aloof or Uninterested	**.380**	.303	**.392**	**.362**	.286	**.360**
Playful	-.373	-.378	-.304	-.336	-.330	-.277
Engaged, Alert	-.162	-.110	-.191	-.149	-.154	-.115
Interested in people	-.440	-.353	-.452	-.410	-.356	-.380
Affectionate	-.530	-.498	-.473	-.489	-.465	-.416
Subtest 6: Restraint						
Fearful	.389	.422	.295	.404	.414	.318
Confident	-.394	-.438	-.288	-.411	-.446	-.302
Aggressive	**.525***	.501	**.465***	**.512***	.503	**.423***
Friendly	**-.534***	-.449	**-.530***	**-.481***	-.386	**-.475***
Affectionate	**-.531***	-.489	**-.485***	**-.476***	-.439	**-.418***
Calm	-.028	-.085	.030	-.056	-.056	-.046
Excitable	-.049	.017	-.106	.009	.005	.011
Subtest 7: Other dog (on leash)						
Fearful	.386	.432	.279	.418	.441	.318
Confident	-.375	-.436	-.256	-.419	-.465	-.297
Aggressive	.495	.444	.468	.464	.465	.375
Friendly	-.448	-.393	-.426	-.423	-.376	-.384
Excitable	-.146	-.069	-.197	-.096	-.053	-.116
Calm	.188	.113	.229	.128	.068	.158
Aloof or Uninterested	.335	.241	.370	.292	.165	.349
Playful	-.340	-.305	-.316	-.276	-.229	-.265
Submissive	-.050	.038	-.128	.024	.093	-.043
Pushy, Assertive	.086	-.028	.186	.021	-.097	.124
Subtest 8: Novel situation (room)						
Fearful	.327	.343	.260	.368	.393	.275
Confident	-.323	-.340	-.257	-.367	-.397	-.270
Excitable	-.031	.036	-.092	.000	.026	-.024
Calm	.114	.003	.202	.087	-.027	.174
Engaged, Alert	.045	.089	-.007	.049	.041	.047

Table 9.7. (*Continued*)

Factor 2: Aggression towards People (*Continued*)

	Long form			Short form		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 9: Novel objects						
Fearful	.344	.395	.240	.410	.426	.316
Confident	-.282	-.362	-.160	-.350	-.412	-.228
Excitable	-.140	-.090	-.167	-.121	-.135	-.085
Calm	.134	.072	.173	.110	.100	.098
Engaged, Alert	-.085	-.117	-.039	-.093	-.205	.026
Subtest 10: Doll test						
Fearful	.506	.571	.360	.529	.563	.396
Confident	-.482	-.551	-.337	-.505	-.551	-.366
Aggressive	**.649***	**.682***	.510	**.637***	**.642***	.510
Friendly	**-.490***	**-.512***	-.388	**-.466***	**-.451***	-.390
Excitable	-.018	.047	-.077	.000	.001	.000
Calm	-.062	-.143	.029	-.066	-.094	-.028
Subtest 11: Prey drive						
Aggressive	.312	.315	.256	.331	.266	.326
Excitable	-.057	-.010	-.093	-.045	-.093	.008
Calm	.044	-.026	.107	.013	.001	.021
Playful	-.261	-.284	-.191	-.236	-.235	-.191
Engaged	-.200	-.104	-.262	-.177	-.104	-.209
Subtest 12: Engagement in play with tester						
Affectionate	-.455	-.443	-.389	-.413	-.351	-.390
Excitable	-.248	-.149	-.305	-.197	-.170	-.184
Calm	.105	.050	.140	.094	.101	.069
Aloof or Uninterested	.414	.345	.411	.381	.333	.351
Playful	-.521	-.478	-.476	-.482	-.448	-.420
Engaged, Alert	-.353	-.233	-.416	-.316	-.209	-.353
Interested in people	-.446	-.342	-.475	-.411	-.357	-.381
Subtest 13: Tug-o-war, Toy release						
Easy to train	-.367	-.318	-.354	-.314	-.246	-.314
Easy to control	-.333	-.282	-.327	-.298	-.243	-.290
Submissive	-.297	-.204	-.339	-.257	-.190	-.269
Pushy, Assertive	.324	.237	.356	.274	.208	.282
Subtest 14: Train new task						
Engaged	-.116	-.060	-.159	-.120	-.097	-.119
Interested in people	-.357	-.245	-.408	-.343	-.274	-.339
Easy to train	-.451	-.373	-.452	-.406	-.359	-.370
Obedient	-.456	-.374	-.462	-.418	-.335	-.413
Subtest 15: Basic commands						
Obedient	-.449	-.377	-.447	-.424	-.348	-.411

Table 9.7. (*Continued*)

Factor 2: Aggression towards People (*Continued*)						
	Long form			**Short form**		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 16: Other dog (off leash)						
Fearful	.350	.422	.223	.355	.421	.227
Confident	-.350	-.390	-.256	-.360	-.386	-.268
Aggressive	.499	.424	.495	.470	.430	.417
Friendly	-.314	-.356	-.224	-.298	-.340	-.202
Excitable	-.258	-.209	-.262	-.201	-.186	-.175
Calm	.137	.056	.191	.081	.014	.126
Aloof or Uninterested	.283	.248	.268	.242	.207	.227
Playful	-.393	-.371	-.348	-.351	-.346	-.288
Submissive	-.145	-.043	-.223	-.085	-.028	-.121
Pushy, Assertive	.130	.010	.227	.070	-.015	.133
Subtest 17: Activity in free play						
Active	-.121	-.006	-.216	-.081	-.047	-.096
Excitable	-.240	-.128	-.310	-.198	-.152	-.201
Calm	.128	.007	.225	.093	.029	.133
Aloof	.307	.228	.332	.251	.202	.246
Playful	-.278	-.242	-.267	-.245	-.198	-.240
Engaged, Alert	-.223	-.135	-.276	-.184	-.186	-.147
Subtest 18: Reunion with owner						
Excitable	-.395	-.273	-.449	-.353	-.285	-.346
Calm	.309	.188	.376	.260	.200	.265
Aloof or Uninterested	.331	.225	.380	.293	.253	.271
Interested in people	-.376	-.304	-.385	-.338	-.288	-.318
Easy to control	-.113	-.095	-.111	-.102	-.150	-.039
Submissive	-.210	-.160	-.223	-.205	-.217	-.155
Pushy, Assertive	.131	.098	.140	.132	.139	.102

Note. Bolded correlations are those that are predicted to be convergent. Other correlations are predicted to be discriminant. The cut-off for significance at p = .05 is ± .190; numbers below this are considered discriminant. The cut-off for significance at p = .01 is ± .250. The cut-off for significance at p = .001 is ± .320; numbers at or above .320 are considered convergent. The Bonferroni-corrected cut-off for significance at p = .05 is ± .400. Correlations that are in line with predictions are underlined. Correlations predicted to be convergent that are equal to or greater than .320 are both bolded and underlined. Correlations predicted to be discriminant that are equal to or less than .190 are underlined. Correlations predicted to be convergent that are equal to or greater than ± .400 are marked with asterisks.

Table 9.8. Activity/Excitability convergent and discriminant validity correlations

Factor 3: Activity/Excitability

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 1: Collar Grab										
Fearful	-.029	.100	.007	-.146	-.075	-.024	.102	-.085	-.142	.016
Confident	.096	-.048	.028	.200	.139	.096	-.058	.139	.179	.062
Aggressive	-.201	.030	-.101	-.204	-.364	-.188	.078	-.197	-.209	-.281
Friendly	.463	.318	.302	.243	.545	.425	.246	.331	.205	.476
Subtest 2: Walk on leash with stranger/tester										
Fearful	-.092	-.008	.008	-.175	-.117	-.095	.016	-.094	-.179	-.048
Confident	.085	-.004	.019	.184	.078	.096	-.039	.138	.194	.026
Excitable	**.653***	**.751***	.336	.419	.440	**.612***	**.641***	.295	.376	.429
Calm	**-.657***	**-.738***	-.369	-.439	-.421	**-.625***	**-.653***	-.336	-.365	-.426
Engaged, Alert	**.355**	.333	.177	**.339**	.230	**.293**	.258	.100	**.295**	.196
Easy to control	.150	.194	.114	.053	.078	.149	.186	.119	.069	.037
Submissive	.283	.159	.225	.172	.320	.288	.170	.190	.146	.344
Pushy, Assertive	-.197	-.078	-.137	-.069	-.319	-.218	-.102	-.131	-.053	-.357
Subtest 3: Threatening approach										
Fearful	-.024	.094	-.033	-.151	-.017	-.020	.094	-.059	-.155	.026
Confident	-.003	-.141	-.020	.111	.069	-.002	-.136	.026	.125	.018
Aggressive	-.190	.056	-.192	-.181	-.303	-.197	.102	-.262	-.217	-.268
Friendly	.381	.218	.347	.262	.359	.374	.130	.366	.296	.346
Excitable	**.568***	**.574***	.341	.409	.392	**.492***	**.460***	.258	.313	.383
Calm	**-.473***	**-.506***	-.257	-.337	-.319	**-.445***	**-.461***	-.231	-.232	-.341
Playful	**.377**	.324	**.327**	.285	.226	**.346**	.256	**.309**	.235	.217
Engaged, Alert	**.355**	.333	.177	**.339**	.230	**.293**	.258	.100	**.295**	.196
Interested in people	**.486***	.296	.225	.390	**.576***	**.422***	.188	.193	.361	**.519***
Subtest 4: Friendly approach										
Fearful	-.181	-.075	-.027	-.257	-.205	-.206	-.086	-.094	-.260	-.183
Confident	.160	.083	-.008	.236	.181	.184	.088	.068	.254	.145
Aggressive	-.220	-.060	-.045	-.189	-.381	-.194	-.029	-.074	-.131	-.356
Friendly	.437	.222	.322	.256	.550	.412	.148	.336	.222	.536
Excitable	**.630***	**.533***	.369	.439	.573	**.582***	**.423***	.309	.439	.532
Calm	**-.574***	**-.505***	-.292	-.450	-.493	**-.516***	**-.416***	-.249	-.405	-.429
Playful	**.508***	.334	**.364**	.427	.447	**.508***	.262	**.399**	.416	.444
Engaged, Alert	**.306**	.326	.080	**.267**	.240	**.268**	.266	.065	**.167**	.264
Interested in people	**.440***	.283	.214	.345	**.504***	**.401***	.188	.187	.327	**.495***

Table 9.8. (*Continued*)

Factor 3: Activity/Excitability (*Continued*)

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 5: Affection/Petting from stranger										
Fearful	-.107	.039	-.016	-.223	-.152	-.139	.023	-.091	-.227	-.146
Confident	.110	-.026	.044	.175	.165	.124	-.001	.086	.154	.147
Aggressive	-.232	-.066	-.049	-.241	-.359	-.238	-.074	-.109	-.198	-.338
Friendly	.507	.320	.344	.323	.567	.495	.279	.373	.267	.549
Excitable	**.580***	**.585***	.335	.324	.495	**.565***	**.500***	.348	.270	.512
Calm	**-.591***	**-.597***	-.388	-.327	-.466	**-.580***	**-.517***	-.402	-.262	-.493
Aloof or Uninterested	**-.529***	-.367	**-.352**	-.344	**-.556***	**-.530***	-.306	**-.380**	-.344	**-.541***
Playful	**.526***	.361	**.379**	.402	**.482***	**.508***	.252	**.410***	.379	**.480***
Engaged, Alert	**.415***	.336	.168	**.317**	.433	**.356**	.266	.134	**.235**	.401
Interested in people	**.572***	.377	.354	.428	**.596***	**.525***	.288	.353	.369	**.550***
Affectionate	**.508***	.261	.442	.375	**.508***	**.517***	.211	.461	.378	**.511***
Subtest 6: Restraint										
Fearful	.001	.122	-.027	-.121	-.004	-.004	.108	-.074	-.112	.028
Confident	-.017	-.153	.023	.088	.026	-.024	-.143	.052	.066	-.006
Aggressive	-.182	.044	-.105	-.197	-.325	-.169	.085	-.160	-.181	-.292
Friendly	.446	.297	.309	.252	.505	.455	.255	.378	.231	.493
Affectionate	**.472***	.259	.372	.340	**.494***	**.474***	.186	.435	.320	**.491***
Calm	**-.506***	**-.542***	-.296	-.287	-.388	**-.501***	**-.494***	-.264	-.255	-.418
Excitable	**.541***	**.531***	.299	.302	.487	**.504***	**.432***	.248	.269	.505
Subtest 7: Other dog (on leash)										
Fearful	.022	.157	-.014	-.152	.032	.005	.118	-.044	-.136	.039
Confident	-.009	-.164	.060	.178	-.050	-.014	-.144	.053	.167	-.071
Aggressive	-.207	-.005	-.083	-.183	-.377	-.191	.006	-.107	-.116	-.372
Friendly	.484	.317	.446	.244	.487	.469	.305	.352	.219	.506
Excitable	**.531***	**.475***	.383	.325	.432	**.475***	**.410***	.276	.285	.404
Calm	**-.568***	**-.510***	-.457	-.310	-.449	**-.539***	**-.473***	-.392	-.244	-.454
Aloof or Uninterested	**-.568***	-.365	**-.547***	-.372	**-.484***	**-.512***	-.280	**-.444***	-.347	**-.459***
Playful	**.548***	.420	**.449***	.411	.411	**.527***	.347	**.437***	.364	.412
Submissive	.330	.290	.227	.071	.397	.359	.285	.283	.057	.414
Pushy, Assertive	-.198	-.143	-.145	.059	-.355	-.226	-.140	-.210	.075	-.386
Subtest 8: Novel situation (room)										
Fearful	-.003	.049	.117	-.085	-.088	-.013	.024	.070	-.089	-.052
Confident	-.043	-.090	-.185	.056	.081	-.040	-.068	-.156	.083	.030
Excitable	**.555***	**.599***	.312	.372	.381	**.493***	**.522***	.205	.341	.333
Calm	**-.528***	**-.555***	-.340	-.346	-.352	**-.489***	**-.500***	-.265	-.297	-.335
Engaged, Alert	**.246**	.240	.154	**.200**	.155	**.203**	.196	.095	**.156**	.135

Table 9.8. (*Continued*)

Factor 3: Activity/Excitability (*Continued*)

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 9: Novel objects										
Fearful	.030	.100	.018	-.065	.017	.021	.088	-.009	-.046	.006
Confident	-.122	-.148	-.099	-.009	-.104	-.121	-.124	-.093	-.032	-.096
Excitable	**.520***	**.520***	.386	.353	.322	**.486***	**.452***	.314	.281	.353
Calm	**-.512***	**-.489***	-.385	-.351	-.333	**-.476***	**-.439***	-.311	-.257	-.362
Engaged, Alert	**.214**	.151	.170	**.233**	.117	**.193**	.150	.137	**.144**	.133
Subtest 10: Doll test										
Fearful	.050	.248	-.059	-.055	-.032	.059	.230	-.098	.006	-.020
Confident	-.049	-.229	.029	.044	.051	-.057	-.220	.084	.007	.012
Aggressive	-.115	.138	-.209	-.163	-.177	-.113	.118	-.245	-.087	-.175
Friendly	.195	-.008	.286	.102	.255	.181	-.069	.256	.120	.276
Excitable	**.532***	**.564***	.272	.313	.437	**.471***	**.474***	.175	.288	.401
Calm	**-.431***	**-.496***	-.211	-.255	-.317	**-.380**	**-.442***	-.115	-.213	-.297
Subtest 11: Prey drive										
Aggressive	.300	.461	.100	.284	.040	.298	.428	.086	.222	.080
Excitable	**.488***	**.464***	.320	.412	.295	**.425***	**.406***	.213	.340	.263
Calm	**-.556***	**-.535***	-.355	-.452	-.350	**-.500***	**-.464***	-.268	-.366	-.344
Playful	**.386**	.301	**.272**	.313	.300	**.375**	.207	**.273**	.353	.290
Engaged	**.441***	.297	.345	**.331**	.391	**.391**	.210	.250	**.302**	.401
Subtest 12: Engagement in play with tester										
Affectionate	.331	.142	.330	.168	.395	.317	.072	.367	.147	.385
Excitable	**.682***	**.591***	.496	.481	.516	**.596***	**.460***	.392	.409	.481
Calm	**-.598***	**-.556***	-.405	-.432	-.424	**-.559***	**-.494***	-.331	-.376	-.415
Aloof or Uninterested	**-.539***	-.377	**-.411***	-.399	**-.476***	**-.532***	-.337	**-.411***	-.350	**-.478***
Playful	**.564***	.381	**.463***	.414	**.487***	**.540***	.295	**.462***	.394	**.466***
Engaged, Alert	**.453***	.352	.300	**.264**	.462	**.424***	.286	.276	**.199**	.486
Interested in people	.516	.324	.353	.343	**.566***	.474	.233	.317	.288	**.573***
Subtest 13: Tug-o-war, Toy release										
Easy to train	.244	.054	.177	.146	.382	.257	.007	.263	.138	.387
Easy to control	.207	.028	.142	.141	.338	.223	-.007	.195	.166	.341
Submissive	.231	.064	.259	.061	.339	.220	.023	.255	.051	.349
Pushy, Assertive	-.189	-.005	-.221	-.064	-.312	-.184	.000	-.238	.000	-.334
Subtest 14: Train new task										
Engaged	**.195**	.127	.099	**.170**	.200	**.162**	.113	.104	**.087**	.172
Interested in people	**.472***	.290	.291	.359	**.510***	**.425***	.217	.313	.276	**.461***
Easy to train	.398	.192	.244	.285	.508	.385	.109	.285	.287	.487
Obedient	.274	.083	.161	.187	.416	.281	.036	.254	.166	.409

Table 9.8. (*Continued*)

Factor 3: Activity/Excitability (*Continued*)

	Long form					Short form				
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**	**Facet 4**
Subtest 15: Basic commands										
Obedient	.183	.001	.072	.090	.401	.199	-.040	.170	.078	.414
Subtest 16: Other dog (off leash)										
Fearful	.056	.180	-.031	-.172	.140	.043	.156	-.076	-.173	.170
Confident	-.101	-.202	-.010	.099	-.149	-.085	-.181	.036	.126	-.184
Aggressive	-.223	-.018	-.189	-.205	-.302	-.222	-.016	-.231	-.166	-.276
Friendly	.287	.172	.210	.200	.304	.254	.067	.155	.251	.303
Excitable	**.665***	**.556***	.497	.465	.516	**.610***	**.438***	.411	.442	.502
Calm	**-.628***	**-.554***	-.496	-.399	-.469	**-.578***	**-.456***	-.390	-.366	-.474
Aloof or Uninterested	**-.432***	-.285	**-.416***	-.303	-.342	**-.361**	-.191	**-.321**	-.232	-.336
Playful	**.584***	.427	**.430***	.403	.534	**.536***	.321	**.384**	.353	.528
Submissive	.346	.232	.155	.132	.515	.324	.169	.157	.122	.505
Pushy, Assertive	-.264	-.174	-.065	-.049	-.486	-.248	-.120	-.086	-.023	-.496
Subtest 17: Activity in free play										
Active	**.656***	**.626***	.369	**.482***	.503	**.604***	**.526***	.332	**.399**	.489
Excitable	**.714***	**.618***	.421	.527	.604	**.655***	**.488***	.386	.462	.581
Calm	**-.724***	**-.674***	-.452	-.545	-.529	**-.676***	**-.582***	-.401	-.456	-.520
Aloof	**-.583***	-.455	**-.376**	-.342	-.596	**-.537***	-.349	**-.317**	-.329	-.582
Playful	**.594***	.479	.370	.402	.553	**.556***	.330	.376	.417	.525
Engaged, Alert	**.413***	.322	.196	**.309**	**.425***	**.337**	.250	.104	**.214**	**.410***
Subtest 18: Reunion with owner										
Excitable	**.609***	**.460***	.359	.472	.569	**.547***	**.364**	.297	.419	.530
Calm	**-.566***	**-.459***	-.315	-.428	-.517	**-.513***	**-.378**	-.267	-.361	-.492
Aloof or Uninterested	**-.536***	-.338	-.301	**-.401***	**-.601***	**-.481***	-.260	-.240	**-.352**	**-.574***
Interested in people	**.497***	.290	.307	.362	**.568***	**.426***	.190	.219	.311	**.552***
Easy to control	-.063	-.146	-.125	-.104	.174	-.075	-.159	-.153	-.080	.188
Submissive	.197	.074	.159	.073	.303	.177	.070	.109	.045	.305
Pushy, Assertive	-.093	.028	-.069	.014	-.261	-.085	.015	-.054	.046	-.266

Note. Bolded correlations are those that are predicted to be convergent. Other correlations are predicted to be discriminant. The cut-off for significance at p = .05 is ± .190; numbers below this are considered discriminant. The cut-off for significance at p = .01 is ± .250. The cut-off for significance at p = .001 is ± .320; numbers at or above .320 are considered convergent. The Bonferroni-corrected cut-off for significance at p = .05 is ± .400. Correlations that are in line with predictions are underlined. Correlations predicted to be convergent that are equal to or greater than .320 are both bolded and underlined. Correlations predicted to be discriminant that are equal to or less than .190 are underlined. Correlations predicted to be convergent that are equal to or greater than ± .400 are marked with asterisks.

Table 9.9. Responsiveness to Training convergent and discriminant validity correlations

Factor 4: Responsiveness to Training						
	Long form			**Short form**		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 1: Collar Grab						
Fearful	-.065	-.163	.061	-.045	-.167	.104
Confident	.096	.229	-.076	.082	.240	-.121
Aggressive	-.307	-.293	-.207	-.309	-.274	-.235
Friendly	.136	.314	-.099	.161	.225	.033
Subtest 2: Walk on leash with stranger/tester						
Fearful	-.054	-.165	.082	-.061	-.165	.075
Confident	.062	.160	-.061	.062	.168	-.076
Excitable	-.090	.136	-.289	-.087	-.009	-.142
Calm	.092	-.112	.268	.117	.016	.184
Engaged, Alert	.039	.180	-.121	.021	.129	-.105
Easy to control	**-.018**	.078	**-.110**	**-.006**	.057	**-.074**
Submissive	**.230**	.273	**.099**	**.232**	.224	**.156**
Pushy, Assertive	**-.213**	-.201	**-.146**	**-.207**	-.173	**-.169**
Subtest 3: Threatening approach						
Fearful	-.121	-.171	-.024	-.086	-.152	.016
Confident	.205	.222	.111	.179	.222	.067
Aggressive	-.256	-.227	-.191	-.262	-.186	-.250
Friendly	.186	.354	-.057	.233	.292	.084
Excitable	-.094	.098	-.257	-.067	-.007	-.108
Calm	.120	-.045	.246	.117	.010	.192
Playful	-.042	.149	-.223	.014	.087	-.070
Engaged, Alert	.039	.180	-.121	.021	.129	-.105
Interested in people	.185	.378	-.083	.244	.289	.106
Subtest 4: Friendly approach						
Fearful	-.025	-.093	.054	.007	-.084	.103
Confident	.098	.163	-.007	.056	.156	-.074
Aggressive	-.206	-.178	-.158	-.155	-.143	-.113
Friendly	.265	.369	.059	.288	.305	.165
Excitable	.025	.300	-.269	.030	.210	-.178
Calm	.047	-.220	.305	.038	-.147	.227
Playful	.174	.347	-.070	.209	.252	.086
Engaged, Alert	.065	.173	-.070	.000	.102	-.112
Interested in people	.155	.338	-.091	.173	.302	-.029

Table 9.9. (*Continued*)

Factor 4: Responsiveness to Training (*Continued*)						
	Long form			**Short form**		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 5: Affection/Petting from stranger						
Fearful	-.013	-.036	.016	-.003	-.048	.047
Confident	.074	.113	.007	.050	.102	-.024
Aggressive	-.210	-.248	-.092	-.178	-.233	-.054
Friendly	.213	.372	-.030	.223	.299	.059
Excitable	-.010	.237	-.261	-.004	.144	-.164
Calm	.033	-.247	.310	.017	-.149	.193
Aloof or Uninterested	-.129	-.334	.131	-.120	-.242	.058
Playful	.203	.414	-.092	.252	.313	.093
Engaged, Alert	.144	.215	.017	.112	.140	.039
Interested in people	.141	.344	-.120	.145	.261	-.035
Affectionate	.137	.322	-.105	.135	.246	-.036
Subtest 6: Restraint						
Fearful	-.116	-.059	-.133	-.070	-.077	-.038
Confident	.208	.145	.197	.168	.163	.114
Aggressive	-.205	-.243	-.089	-.183	-.239	-.056
Friendly	.220	.362	-.007	.221	.299	.057
Affectionate	.222	.400	-.044	.212	.328	.008
Calm	.030	-.174	.230	-.002	-.088	.093
Excitable	-.063	.148	-.257	-.032	.027	-.085
Subtest 7: Other dog (on leash)						
Fearful	-.183	-.173	-.126	-.114	-.157	-.027
Confident	.159	.165	.093	.106	.155	.014
Aggressive	-.333	-.303	-.239	-.244	-.269	-.129
Friendly	.172	.297	-.020	.121	.214	-.025
Excitable	-.004	.165	-.177	-.015	.069	-.101
Calm	-.010	-.226	.213	.001	-.131	.144
Aloof or Uninterested	-.023	-.222	.192	.002	-.137	.153
Playful	.165	.344	-.082	.154	.224	.023
Submissive	.106	.249	-.080	.161	.254	.001
Pushy, Assertive	-.143	-.174	-.059	-.187	-.217	-.086
Subtest 8: Novel situation (room)						
Fearful	-.042	-.012	-.058	-.013	-.016	-.006
Confident	.080	.074	.056	.076	.101	.022
Excitable	-.227	.014	-.391	-.213	-.081	-.280
Calm	.174	-.028	.318	.189	.043	.279
Engaged, Alert	-.099	-.035	-.128	-.135	-.033	-.198

Table 9.9. (*Continued*)

Factor 4: Responsiveness to Training (*Continued*)

	Long form			Short form		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 9: Novel objects						
Fearful	-.032	.023	-.078	-.010	.011	-.030
Confident	.027	-.040	.088	.006	-.006	.017
Excitable	-.151	.029	-.281	-.151	-.048	-.208
Calm	.039	-.106	.175	.050	-.047	.138
Engaged, Alert	-.037	.083	-.147	-.054	.092	-.194
Subtest 10: Doll test						
Fearful	-.175	-.101	-.186	-.124	-.129	-.074
Confident	.139	.090	.139	.100	.114	.050
Aggressive	-.318	-.226	-.294	-.258	-.209	-.219
Friendly	.199	.230	.092	.194	.186	.132
Excitable	-.096	.157	-.322	-.071	.053	-.181
Calm	.157	-.012	.272	.152	.067	.190
Subtest 11: Prey drive						
Aggressive	-.253	-.026	-.393	-.239	-.127	-.275
Excitable	-.219	-.039	-.324	-.238	-.127	-.273
Calm	.057	-.104	.202	.087	.000	.149
Playful	.131	.305	-.099	.174	.200	.082
Engaged	.002	.219	-.222	-.038	.127	-.204
Subtest 12: Engagement in play with tester						
Affectionate	.178	.371	-.088	.221	.347	.004
Excitable	-.006	.165	-.180	.005	.044	-.039
Calm	.066	-.132	.246	.079	-.039	.178
Aloof or Uninterested	-.128	-.316	.114	-.103	-.241	.084
Playful	.061	.252	-.160	.084	.101	.034
Engaged, Alert	.083	.193	-.060	.013	.090	-.074
Interested in people	.177	.360	-.078	.173	.245	.032
Subtest 13: Tug-o-war, Toy release						
Easy to train	**.410***	**.402***	**.266**	**.405***	**.346**	**.322**
Easy to control	**.383**	**.356**	**.269**	**.407***	**.319**	**.355**
Submissive	**.325**	.304	**.226**	**.349**	.260	**.318**
Pushy, Assertive	**-.413***	-.315	**-.361**	**-.401***	-.284	**-.384**
Subtest 14: Train new task						
Engaged	-.005	.028	-.038	-.054	.003	-.096
Interested in people	.208	.303	.032	.191	.206	.106
Easy to train	**.339**	**.373**	.176	**.341**	**.303**	.259
Obedient	**.491***	**.450***	.351	**.521***	**.429***	.432
Subtest 15: Basic commands						
Obedient	**.501***	**.398**	**.421***	**.506***	**.383**	**.456***

Table 9.9. (*Continued*)

Factor 4: Responsiveness to Training (*Continued*)						
	Long form			**Short form**		
	Factor	**Facet 1**	**Facet 2**	**Factor**	**Facet 1**	**Facet 2**
Subtest 16: Other dog (off leash)						
Fearful	-.146	-.130	-.108	-.082	-.134	.004
Confident	.165	.151	.118	.130	.192	.016
Aggressive	-.322	-.253	-.274	-.210	-.182	-.166
Friendly	.162	.268	-.008	.145	.195	.036
Excitable	.006	.251	-.249	.040	.127	-.069
Calm	.084	-.144	.289	.053	-.022	.116
Aloof or Uninterested	-.134	-.274	.060	-.116	-.194	.012
Playful	.154	.312	-.066	.149	.180	.061
Submissive	.182	.227	.069	.213	.189	.161
Pushy, Assertive	-.203	-.202	-.129	-.242	-.185	-.215
Subtest 17: Activity in free play						
Active	.029	.167	-.125	.004	.036	-.032
Excitable	-.004	.190	-.202	.009	.065	-.055
Calm	.031	-.158	.215	.043	-.028	.106
Aloof	-.061	-.251	.158	-.040	-.153	.097
Playful	.127	.382	-.185	.174	.280	-.005
Engaged, Alert	.130	.256	-.049	.035	.132	-.084
Subtest 18: Reunion with owner						
Excitable	.033	.237	-.190	.028	.104	-.064
Calm	.005	-.182	.197	.006	-.070	.088
Aloof or Uninterested	-.002	-.218	.221	.055	-.103	.207
Interested in people	.089	.242	-.102	.098	.159	-.005
Easy to control	**.283**	.168	**.297**	**.280**	.203	**.261**
Submissive	**.219**	.267	**.088**	**.212**	.219	**.126**
Pushy, Assertive	**-.165**	-.156	**-.113**	**-.159**	-.155	**-.105**

Note. Bolded correlations are those that are predicted to be convergent. Other correlations are predicted to be discriminant. The cut-off for significance at p = .05 is ± .190; numbers below this are considered discriminant. The cut-off for significance at p = .01 is ± .250. The cut-off for significance at p = .001 is ± .320; numbers at or above .320 are considered convergent. The Bonferroni-corrected cut-off for significance at p = .05 is ± .400. Correlations that are in line with predictions are underlined. Correlations predicted to be convergent that are equal to or greater than .320 are both bolded and underlined. Correlations predicted to be discriminant that are equal to or less than .190 are underlined. Correlations predicted to be convergent that are equal to or greater than ± .400 are marked with asterisks.

Table 9.10. Aggression towards animals convergent and discriminant validity correlations

Factor 5: Aggression towards Animals								
	Long Form				**Short Form**			
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**
Subtest 1: Collar Grab								
Fearful	.031	.187	.097	-.223	.075	.203	.128	-.176
Confident	-.057	-.233	-.054	.169	-.083	-.247	-.068	.133
Aggressive	.344	.452	.170	.041	.327	.381	.185	.074
Friendly	-.304	-.536	.064	-.142	-.357	-.540	.012	-.211
Subtest 2: Walk on leash with stranger/tester								
Fearful	-.026	.154	-.010	-.181	.013	.202	-.002	-.147
Confident	.087	-.110	.022	.243	.084	-.142	.034	.239
Excitable	.065	-.396	.501	-.052	-.023	-.467	.459	-.148
Calm	-.086	.348	-.487	.044	-.033	.395	-.443	.087
Engaged, Alert	.023	-.202	.280	-.072	-.044	-.233	.217	-.122
Easy to control	.055	-.158	.138	.105	.025	-.182	.139	.053
Submissive	-.346	-.406	.148	-.435	-.377	-.359	.097	-.495
Pushy, Assertive	.437	.387	-.061	.527	.421	.330	-.031	.529
Subtest 3: Threatening approach								
Fearful	-.006	.169	.061	-.235	.091	.215	.090	-.114
Confident	-.083	-.196	-.141	.180	-.160	-.213	-.170	.074
Aggressive	.311	.382	.128	.091	.352	.312	.210	.151
Friendly	-.169	-.415	.079	-.016	-.267	-.437	.022	-.143
Excitable	.044	-.361	.380	.005	.000	-.379	.360	-.070
Calm	-.009	.340	-.386	.085	.028	.379	-.366	.129
Playful	-.034	-.311	.186	.024	-.130	-.391	.164	-.084
Engaged, Alert	.023	-.202	.280	-.072	-.044	-.233	.217	-.122
Interested in people	-.194	-.489	.189	-.114	-.251	-.497	.097	-.144
Subtest 4: Friendly approach								
Fearful	.117	.274	-.066	.036	.122	.229	-.021	.054
Confident	-.039	-.215	.041	.083	-.070	-.217	.007	.047
Aggressive	.355	.427	.086	.177	.322	.295	.130	.199
Friendly	-.331	-.512	.024	-.173	-.328	-.456	-.005	-.209
Excitable	-.013	-.418	.311	.026	-.075	-.433	.246	-.040
Calm	.016	.363	-.262	-.023	.050	.364	-.208	.011
Playful	-.131	-.394	.131	-.022	-.201	-.426	.087	-.103
Engaged, Alert	.042	-.247	.318	-.038	.038	-.238	.290	-.046
Interested in people	-.145	-.436	.232	-.118	-.157	-.420	.161	-.109

Table 9.10. (*Continued*)

Factor 5: Aggression towards Animals (*Continued)*

	Long Form				Short Form			
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**
Subtest 5: Affection/Petting from stranger								
Fearful	.137	.212	-.060	.124	.123	.175	-.031	.112
Confident	-.106	-.266	.072	-.029	-.121	-.245	.052	-.070
Aggressive	.326	.423	-.020	.240	.292	.325	.024	.235
Friendly	-.382	-.612	.168	-.331	-.408	-.593	.111	-.369
Excitable	-.095	-.403	.336	-.168	-.150	-.463	.281	-.190
Calm	.053	.375	-.337	.115	.104	.457	-.319	.151
Aloof or Uninterested	.247	.409	-.091	.180	.279	.394	-.008	.188
Playful	-.104	-.424	.177	.007	-.178	-.456	.134	-.087
Engaged, Alert	-.148	-.343	.174	-.149	-.173	-.331	.144	-.196
Interested in people	-.287	-.584	.239	-.256	-.332	-.548	.156	-.315
Affectionate	-.325	-.547	.163	-.278	-.383	-.527	.074	-.337
Subtest 6: Restraint								
Fearful	.112	.182	.129	-.100	.157	.181	.147	-.025
Confident	-.167	-.200	-.176	.064	-.199	-.192	-.201	.016
Aggressive	.343	.397	.104	.162	.377	.360	.179	.190
Friendly	-.352	-.562	.078	-.224	-.443	-.578	-.019	-.303
Affectionate	-.313	-.532	.100	-.200	-.414	-.552	.000	-.292
Calm	.022	.309	-.379	.166	.044	.342	-.337	.160
Excitable	-.066	-.351	.327	-.153	-.080	-.354	.278	-.150
Subtest 7: Other dog (on leash)								
Fearful	.072	.167	.147	-.180	.090	.168	.181	-.177
Confident	-.022	-.154	-.074	.183	-.056	-.163	-.107	.154
Aggressive	**.494***	**.505***	.161	**.283**	**.492***	**.426***	.198	**.327**
Friendly	**-.310**	**-.540***	.119	**-.209**	**-.374**	**-.508***	.050	**-.310**
Excitable	-.104	-.410	.253	-.090	-.120	-.381	.202	-.118
Calm	.096	.453	-.310	.099	.128	.455	-.281	.159
Aloof or Uninterested	**.204**	**.496***	-.191	.130	**.279**	**.514***	-.112	.197
Playful	**-.157**	**-.488***	.183	-.039	**-.217**	**-.544***	.151	-.101
Submissive	**-.312**	-.313	.128	**-.433***	**-.354**	-.336	.079	**-.452***
Pushy, Assertive	**.434***	.303	.028	**.503***	**.442***	.309	.039	**.508***
Subtest 8: Novel situation (room)								
Fearful	.043	.169	-.016	-.060	.041	.133	-.005	-.032
Confident	-.025	-.141	-.005	.092	-.032	-.100	-.026	.055
Excitable	.088	-.291	.424	-.025	.085	-.285	.394	-.033
Calm	-.071	.268	-.340	-.012	-.073	.263	-.332	.007
Engaged, Alert	.005	-.086	.122	-.043	.045	-.081	.129	.011

Table 9.10. (*Continued*)

Factor 5: Aggression towards Animals (*Continued)*

	Long Form				Short Form			
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**
Subtest 9: Novel objects								
Fearful	.013	.158	-.013	-.110	.009	.117	-.014	-.067
Confident	-.006	-.112	-.039	.136	-.010	-.062	-.039	.076
Excitable	.046	-.313	.443	-.103	.017	-.326	.400	-.130
Calm	-.070	.275	-.422	.072	-.061	.272	-.386	.079
Engaged, Alert	.027	-.136	.243	-.086	.021	-.118	.183	-.063
Subtest 10: Doll test								
Fearful	.121	.174	.177	-.126	.091	.125	.153	-.105
Confident	-.149	-.174	-.201	.100	-.138	-.146	-.185	.074
Aggressive	.341	.417	.152	.088	.315	.338	.173	.102
Friendly	-.220	-.293	-.076	-.057	-.307	-.296	-.139	-.162
Excitable	-.004	-.328	.371	-.106	-.080	-.361	.294	-.164
Calm	-.012	.221	-.264	.060	.027	.242	-.234	.101
Subtest 11: Prey drive								
Aggressive	**.415***	.111	**.588***	.039	**.411***	.049	**.571***	.089
Excitable	.105	-.205	.390	-.034	.065	-.235	.352	-.068
Calm	-.043	.250	-.339	.055	-.052	.239	-.324	.057
Playful	-.082	-.319	.117	.015	-.216	-.422	.061	-.106
Engaged	.004	-.294	.290	-.036	-.004	-.289	.266	-.053
Subtest 12: Engagement in play with tester								
Affectionate	-.315	-.448	-.004	-.170	-.405	-.474	-.104	-.228
Excitable	-.084	-.413	.376	-.182	-.144	-.446	.325	-.242
Calm	.065	.352	-.316	.139	.076	.357	-.271	.133
Aloof or Uninterested	.193	.413	-.167	.161	.224	.400	-.083	.162
Playful	-.088	-.428	.227	-.015	-.149	-.437	.179	-.098
Engaged, Alert	-.125	-.387	.156	-.045	-.117	-.374	.143	-.055
Interested in people	-.215	-.483	.239	-.216	-.244	-.472	.166	-.230
Subtest 13: Tug-o-war, Toy release								
Easy to train	-.314	-.327	-.073	-.206	-.370	-.280	-.157	-.269
Easy to control	-.240	-.263	-.027	-.177	-.315	-.238	-.116	-.249
Submissive	-.397	-.408	.035	-.407	-.379	-.335	-.026	-.385
Pushy, Assertive	.492	.455	.030	.470	.479	.378	.075	.477
Subtest 14: Train new task								
Engaged	-.187	-.220	.019	-.168	-.140	-.167	-.012	-.104
Interested in people	-.270	-.403	.107	-.249	-.293	-.369	.020	-.246
Easy to train	-.316	-.380	.036	-.279	-.356	-.342	-.045	-.316
Obedient	-.327	-.374	-.067	-.195	-.361	-.352	-.131	-.221
Subtest 15: Basic commands								
Obedient	-.386	-.344	-.151	-.242	-.406	-.294	-.205	-.272

Table 9.10. (*Continued*)

Factor 5: Aggression towards Animals (*Continued*)								
	Long Form				**Short Form**			
	Factor	**Facet 1**	**Facet 2**	**Facet 3**	**Factor**	**Facet 1**	**Facet 2**	**Facet 3**
Subtest 16: Other dog (off leash)								
Fearful	.015	.119	.090	-.179	.025	.094	.118	-.163
Confident	.019	-.058	-.121	.221	.019	-.042	-.139	.223
Aggressive	**.434***	**.461***	.162	.211	**.422***	**.383**	.221	.208
Friendly	**-.191**	**-.346**	-.089	.060	**-.292**	**-.374**	-.152	-.051
Excitable	.018	-.461	.419	.007	-.022	-.465	.388	-.069
Calm	-.004	.413	-.372	.013	.035	.432	-.342	.070
Aloof or Uninterested	.149	.353	-.072	.032	.221	.388	-.029	.109
Playful	-.151	-.448	.150	-.030	-.222	-.483	.116	-.125
Submissive	**-.434***	-.369	.057	**-.534***	**-.454***	-.323	.012	**-.573***
Pushy, Assertive	**.534***	.415	-.011	**.629***	**.536***	.364	.025	**.648***
Subtest 17: Activity in free play								
Active	-.082	-.422	.379	-.173	-.095	-.414	.336	-.192
Excitable	-.092	-.470	.407	-.177	-.106	-.448	.355	-.202
Calm	.025	.433	-.478	.161	.037	.430	-.436	.180
Aloof	.296	.519	-.113	.198	.353	.508	-.044	.265
Playful	-.097	-.416	.204	-.016	-.190	-.492	.160	-.105
Engaged, Alert	-.190	-.404	.170	-.167	-.173	-.345	.107	-.142
Subtest 18: Reunion with owner								
Excitable	-.160	-.513	.286	-.134	-.201	-.493	.225	-.198
Calm	.150	.445	-.279	.170	.182	.428	-.231	.223
Aloof or Uninterested	.356	.553	-.172	.341	.355	.501	-.090	.326
Interested in people	-.290	-.482	.196	-.309	-.290	-.399	.115	-.321
Easy to control	-.374	-.127	-.171	-.400	-.387	-.089	-.221	-.394
Submissive	-.344	-.334	.075	-.419	-.338	-.317	.030	-.387
Pushy, Assertive	.347	.210	.002	.456	.310	.169	.030	.394

Note. Bolded correlations are those that are predicted to be convergent. Other correlations are predicted to be discriminant. The cut-off for significance at p = .05 is ± .190; numbers below this are considered discriminant. The cut-off for significance at p = .01 is ± .250. The cut-off for significance at p = .001 is ± .320; numbers at or above .320 are considered convergent. The Bonferroni-corrected cut-off for significance at p = .05 is ± .400. Correlations that are in line with predictions are underlined. Correlations predicted to be convergent that are equal to or greater than .320 are both bolded and underlined. Correlations predicted to be discriminant that are equal to or less than .190 are underlined. Correlations predicted to be convergent that are equal to or greater than ± .400 are marked with asterisks.

The mean magnitude of the convergent and discriminant validity correlations associated with each factor and each facet, for both the long and the short forms of the DPQ, are presented in Table 9.11. The columns indicate the long form's and short form's convergent and discriminant validity. The rows are labeled by factors, with each factor being followed by its facets. For consistency with the immediately previous tables, the convergent validity correlations (under the column heading "Convergent") are presented in bold, and the discriminant validity correlations (under the column heading "Discriminant") are not bolded. The correlations that were predicted to be convergent and that were indeed high ($\geq .320$, or significant at $p \leq .001$) are underlined. The correlations that were predicted to be discriminant and were indeed low ($\leq .320$, or significant at $p \geq .001$) are also underlined.

It should be noted that these means indicate the strength of the average strength of the convergent and discriminant relationships, not their directionality; they were calculated using the absolute value of the correlations presented in Tables 9.6 through 9.10 so that positive and negative correlations would not cancel out.

It is apparent from examination of Table 9.11 that the convergent validity correlations were, for the most part, as predicted. That is, most of the correlations that fall under the "Convergent" headings and are bolded are also underlined. For the long form and the short form, this is true of all but the Active Engagement (Facet 3 of Activity/Excitability) and Controllability (Facet 2 of Responsiveness to Training) correlations.

Table 9.11. Mean factor- and facet-level predictive validity correlations

Factor	Long form		Short form	
Facet	Convergent	Discriminant	Convergent	Discriminant
Factor 1 – Fearfulness	**.528***	.168	**.522***	.165
Facet 1 – Fear of People	**.490***	.170	**.439***	.161
Facet 2 – Nonsocial Fear	**.438***	.235	**.459***	.226
Facet 3 – Fear of Dogs	**.454***	.175	**.429***	.163
Facet 4 – Fear of Handling	**.469***	.189	**.421***	.184
Factor 2 – Aggression towards People	**.539***	.272	**.510***	.257
Facet 1 – General Aggression	**.528***	.256	**.469***	.256
Facet 2 – Situational Aggression	**.513***	.272	**.469***	.231
Factor 3 – Activity/Excitability	**.524***	.197	**.481***	.195
Facet 1 – Excitability	**.562***	.203	**.478***	.163
Facet 2 – Playfulness	**.401***	.268	**.380**	.257
Facet 3 – Active Engagement	**.307**	.297	**.239**	.261
Facet 4 – Companionability	**.528***	.411	**.508***	.398
Factor 4 – Responsiveness to Training	**.335**	.120	**.340**	.113
Facet 1 – Trainability	**.396**	.204	**.357**	.153
Facet 2 – Controllability	**.221**	.150	**.252**	.106
Factor 5 – Aggression towards Animals	**.362**	.195	**.392**	.221
Facet 1 – Aggression towards Dogs	**.475***	.380	**.461***	.371
Facet 2 – Prey Drive	**.588***	.189	**.571***	.167
Facet 3 – Dominance over Other Dogs	**.443***	.171	**.479***	.192
MEAN across Factors	**.462***	.191	**.452***	.191

Note. Mean correlations are computed using Fisher's r-to-z transformation and the absolute values of the correlations. Bolded correlations are those that are predicted to be convergent. Other correlations are predicted to be discriminant. The item-level cut-offs for significance are used. Significance at p = .05 is ± .190; numbers below this are considered discriminant. The cut-off for significance at p = .01 is ± .250. The cut-off for significance at p = .001 is ± .320; numbers at or above .320 are considered convergent. The Bonferroni-corrected cut-off for significance at p = .05 is ± .400. Correlations that are in line with predictions are underlined. Correlations predicted to be convergent that are equal to or greater than .320 are both bolded and underlined. Correlations predicted to be discriminant that are equal to or less than .190 are underlined. Correlations predicted to be convergent that are equal to or greater than ± .400 are marked with asterisks.

The discriminant validity correlations are not as consistent with predictions. Of the factor-level correlations, only the mean discriminant validity correlations associated with Fearfulness and Responsiveness to Training were less than .190 (for both the long

and the short forms). Facet-level correlations are mixed. For both the long and the short form, Companionability (Facet 4 of Activity/Excitability) and Aggression towards Dogs (Facet 1 of Aggression towards Animals) are associated with correlations greater than the ±.320 cut-off, such that these correlations are considered "high".

Evaluation of the predictive validity of the DPQ relative to other instruments and dog personality assessments is limited by the fact that predictive validity, particularly discriminant validity, has not been extensively reported in the dog personality assessment literature (see Chapter 2). In the human literature, for example, trait-behavior correlations are commonly cited as rarely exceeding .30 (e.g., Gosling and Vazire, 2002; Mischel, 1968), which has been defended as a stronger effect size than it may appear (e.g., Funder & Ozer, 1983). However, as illustrated in the few convergent validity correlations presented in the existing dog literature (see Table 2.5 in Chapter 2) and in the current study of the DPQ's convergent validity correlations, trait-behavior correlations of > .30 are not entirely uncommon in dog personality research. Trait-behavior correlations presented in Table 2.5 indicate that previous studies have found convergent correlations ranging from $r < .001$ (associated with "attention/distraction" in Weiss & Greenberg, 1997; "affability" in Wilsson & Sundgren, 1998) to $r > .99$ (associated with "fear/submission" in Weiss & Greenberg, 1997). Mean convergent validity (non-weighted) across all traits examined and reported in Table 2.5 was .54, slightly higher than the mean convergent validity correlations associated with the DPQ, presented in Table 9.11.

SUMMARY OF FINDINGS

Results from Study 6 showed support for convergent validity. Many correlations between DPQ factors or facets and behaviors scored during the Test Battery that were predicted to be high were high, whether evaluated against the p = .001 cut-off of ±.320 or the more conservative, Bonferroni-corrected cut-off of ±.400. The mean convergent validity correlation for the DPQ long-form, at the factor level, was .462, and for the short form was .452.

Results from Study 6 showed less support for discriminant validity. Correlations predicted to be discriminant, or low, ranged from .001 to levels above the cut-off for convergence (e.g., .523). On average, however, the mean discriminant validity correlation across all five factors of the long form and of the short form were .191, just missing the .190 cut-off for being categorized as discriminant. The mean discriminant validity correlations are also noticeably lower than the mean convergent validity correlations.

CHAPTER 10

General Discussion

The overarching goal of the research recorded in this dissertation was to develop a dog personality assessment tool that could meet a diverse set of ten evaluative criteria (described in Chapter 3) pertaining to reliability, validity, sensitivity, order effects, the availability of psychometric properties to potential users, comprehensiveness and detail, wide applicability, ease of use, efficiency, and manageability or logistics. Evaluation of dog personality assessment methods against the ten evaluative criteria guided the selection of a "Rating of Individual Dogs" approach, or a questionnaire designed for rating the personality of dogs individually.

Review of the existing dog personality literature (see Chapter 2) drew attention to number of specific issues, including an overly narrow focus on specific breeds and lack of attention to discriminant validity in the existing literature. These trends helped guide the development of the DPQ.

The six studies that compose the development and evaluation of the DPQ were designed to meet specific sub-goals necessary to achieving the overarching research goal. The studies, their individual goals, and their results are summarized below.

STUDY 1: PILOT TESTING THE INITIAL ITEM POOL

In Chapter 4, I described Study 1, the goal of which was to begin the development of the DPQ in a way that would make it highly comprehensive and easily understood. The items in the initial item pool were drawn from a variety of sources. Item sources included the dog personality and temperament research literature, assessment tools used in applied settings (e.g., shelters), and nominations by dog experts. An initial pool of 1,284 descriptors was generated. Items were eliminated if redundant or if they did not fit the ten criteria (e.g., were applicable to very narrow contexts or only certain types of dogs, like guide dogs). By the end of this process, a 360-item online questionnaire was developed. That questionnaire was administered to a pilot sample of participants (N =

152) who gave feedback about how easily understood and how generally applicable the items were. The items were modified based on the participants' feedback.

STUDY 2: FACTOR STRUCTURE IDENTIFICATION

In Chapter 5, I described Study 2, which used EFA (PCA, varimax rotation) to determine the number of factors underlying the behaviors and descriptors in the 360-item questionnaire. Scrutiny of the results of EFA of the responses from 3,737 participants to the 360-item online questionnaire indicated the five-factor solution was the best fit. The fit was then confirmed using a split-sample method in which one half of the data was reanalyzed using EFA and the derived model was fit to the second half of the data.

I used the EFA and other criteria to guide further item elimination in creating a shorter, more manageable questionnaire that strove to simultaneously maximize manageability, ease of use, and the instrument's psychometric properties (e.g., internal consistency, validity). This new questionnaire contained 102 items, representing 15 facets within the five personality factors.

The repeatability of the five-factor model was again further supported using a split-sample procedure in which I divided the participant set into two randomly selected halves, repeated the EFA procedure on one half of the data (N = 1,868), then performed CFA to fit the model to the second half of the data (N = 1,869).

STUDY 3: FACTOR STRUCTURE CONFIRMATION

In Study 3 (described in Chapter 6), the goal was to confirm the five-factor structure model (found in Study 2) using the new 102-item questionnaire and a new online sample of participants. I used SEM to perform CFA on data collected from 2,556 new participants' ratings of their dogs on the 102-item questionnaire. Analyses confirmed that the best fit was associated with a five-factor model that included correlations between factors 1 and 2 (Fearfulness and Aggression towards People) and between factors 2 and 5 (Aggression towards People and Aggression towards Animals).

As in Study 2, the repeatability of the five-factor model was confirmed using a split-sample procedure in which I divided the participant set into two randomly selected

halves, repeated the EFA procedure on one half of the data (N = 1,278), then repeated the CFA procedure on the second half of the data (N = 1,278).

In addition to the CFA results, examination of the psychometric properties (e.g., content validity, construct validity, including discriminant and convergent validity; internal consistency) guided the creation of the two final forms of the DPQ: the long form with five items per facet (or 75 items in total) and the short form with three items per facet (or 45 items in total).

STUDY 4: INTER-RATER RELIABILITY

The specific goal of Study 4, described in Chapter 7, was to assess the DPQ's inter-rater reliability—the degree to which the ratings (by item, facet, or factor) that different observers ascribed to a target were correlated. Ninety-nine pairs of participants in which both people were familiar with the same dog rated the dog using the DPQ online. Inter-rater reliability at the item, facet, and factor levels of both the long form and the short form indicated that inter-rater reliability on the DPQ was comparable to previous inter-observer reliability levels found in the dog personality literature (reviewed in Chapter 2). The item-level mean for the long form was .555, and for the short form was .547. However, the item-level inter-rater reliability of the DPQ was highly variable, ranging from .240 to .839. The facet-level inter-rater reliability ICCs ranged from .459 (Companionability on the short form) to .872 (Playfulness on the long form), and the factor-level ICCs inter-rater reliability ICCs ranged from .659 (Aggression towards People on the long form) to .786 (Activity/Excitability on the long form).

STUDY 5: TEST-RETEST RELIABILITY

In Study 5, described in Chapter 8, I addressed test-retest reliability by administering the online questionnaire twice, with approximately four to six weeks between administrations, to 100 participants who rated the same dog both times. I then examined how well the paired ratings correlated for each item, facet, and factor of both the long and short DPQ forms. The DPQ was found to have test-retest reliability rates generally comparable to those found in human personality rating studies. Mean item-level

test-retest reliability levels were .750 for the long form and .753 for the short form, or slightly higher than the averages reported in previous dog personality studies (see Chapter 2). Again, the reliability levels were variable, with item-level test-retest reliability ranging from .325 to .923. Facet-level test-retest reliability levels ranged from .735 to .923 on the long form, and from .750 to .936 on the short form. Factor-level test-retest reliability levels ranged from .878 to .939 on the long form, and from .872 to .929 on the short form.

STUDY 6: PREDICTIVE VALIDITY

The goal of Study 6, described in Chapter 9, was to address how well participants' ratings of their dogs on the DPQ predicted the dogs' behavior on a Test Battery designed to elicit behaviors relevant to the factors and facets of the DPQ. Behaviors assessed on the Test Battery were predicted to be related to or unrelated to specific factors and facets measured on the DPQ; these predictions served as the basis for evaluating the convergent and discriminant validity of the DPQ. The mean convergent validity correlations on the DPQ were, at the item, facet, and factor level, markedly higher than the discriminant validity correlations. Across all of the factors, the mean convergent validity correlation for the long form was .462, and for the short form was .452. The mean discriminant validity correlation across all factors was .191 for both forms.

EVALUATION AGAINST TEN CRITERIA

The development of the DPQ, from the choice of methodology (Rating of Individual Dogs) to the final development of a short and long form, to their validation, was guided by ten criteria. These criteria were developed to help ensure my research goals could be met by the selected assessment method, and that the final DPQ would be able to meet the needs of a wide variety of groups seeking to assess personality in dogs. Whether and how well the final DPQ met these ten criteria is reviewed below.

Reliability

The first criterion against which the long and short forms of the DPQ were evaluated was reliability. Three aspects of reliability were addressed: internal consistency (Study 3), inter-rater reliability (Study 4), and test-retest reliability (Study 5).

Internal consistency is reliability across items within a scale, or whether items that are purported to measure a single construct yield consistent scores. Study 3 results indicated that the internal consistency levels associated with the scales on both forms were comparable or slightly higher than those found in other studies of dog personality (reviewed in Chapter 2).

Both forms of the DPQ were also shown to have acceptable inter-rater and test-retest reliability at the item, facet, and factor levels. Each of these types of reliability is discussed in more detail above, in the summary of Study 4 (inter-rater reliability) and Study 5 (test-retest reliability).

Based on these studies, the DPQ was found to consist of internally consistent scales and to be a reliable test of dog personality across raters and across time.

Validity

The second criterion against which the DPQ was evaluated was validity. Three aspects of validity were addressed: content validity (during item development, Study 1, Study 2, Study 3); construct validity (Study 2, Study 3), and predictive validity (Study 6).

Content validity is the extent to which the set of items or subtests within an assessment tool represents all facets of the construct being measured. The DPQ's content validity was addressed at multiple stages of its development. First, a large number of diverse behavioral descriptions were culled from the literature. Second, expert judges and dog owner participants reviewed the items piloted in the initial assessment tool. Third, all stages of DPQ item pool revision were completed with content validity as a criterion. The final long and short forms consist of items representing five factors, and 15 facets, of dog personality.

Construct validity is the extent to which the items or subtests within an assessment tool measure the broad construct (e.g., personality trait) they were intended to

measure. The DPQ's construct validity was examined using factor analyses in Study 2 and Study 3 to evaluate the factors underlying the instrument's items, and by examining convergent and discriminant correlations associated with the questionnaire's items, facets, and factors. Items were also eliminated from preliminary versions of the DPQ if they did not appear to be univocal, or to clearly load on a single factor (i.e., in Study 3, Appendices E and F). Ultimately, the convergent and discriminant validity associated with the DPQ indicated that the DPQ factors were relative unique and their items loaded univocally, with only a few exceptions. As predicted and as confirmed by the CFA model fit described in Study 3, the factors Aggression towards People and Aggression towards Animals were correlated, as were Fearfulness and Aggression towards People.

Predictive validity is extent to which the results of an assessment tool are consistent with results from another. The DPQ's predictive validity was evaluated in Study 6, in terms of both convergent and discriminant validity. As discussed in the summary of Study 6, evidence indicated that the DPQ had predictive validity, though a clearer case can be made for convergent than discriminant validity based on the predicted correlations tested in this study.

Based on these studies, the DPQ was been found to be a valid measure of dog personality. It addresses as many facets of dog personality as are applicable to a general population, has an underlying structure in which items load univocally on factors unless theoretically supported relationships indicate otherwise, and it has been shown to predict real-life dog behavior.

Sensitivity

The third criterion, sensitivity, guided the development of the DPQ. A tool is sensitive if it is able to differentiate not only among dogs whose personalities and behaviors are very different from one another, but also among dogs whose personalities or behaviors differ relatively minimally, but still meaningfully. That is, the tool should be able to distinguish between dogs who are pervasively or generally aggressive and dogs who are aggressive in only a few specific situations, and between dogs who are a little bit

fearful and dogs who are extremely fearful. Two specific aspects of the DPQ that permit it to be a sensitive tool are its assessment of different facets of personality (e.g., General Aggression and Situational Aggression on the Aggression towards People factor) and its use of a Likert scale for rating so that the degree to which a descriptor is fitting can be incorporated into the scoring process.

Minimal order effects

The fourth criterion, minimal order effects, also guided the DPQ's development. This criterion emphasizes the importance of addressing the effect of each item on subsequent items. The impact of order effects on the factor structure derived in Study 2 and confirmed in Study 3, and on the selection of items retain in the final DPQ, was minimized by presenting items in Studies 1-3 in randomized order such that items were unlikely to appear in the same order for any two participants. I also sought to minimize order effects in the final long and short form of the DPQ by presenting the items so that items on the same facets and factors are not grouped together. However, order effects present in the final 75-item and 45-item instruments have not been evaluated.

Availability of psychometric properties to the public

The fifth criterion, availability of psychometric properties (e.g., reliability, validity) to the public, is readily fulfilled by the DPQ. The goal of this criterion is to allow potential users of the tool to evaluate its effectiveness and suitability for their purposes. To meet this criterion and educate potential users of the DPQ, the pertinent results of this dissertation will be made readily available.

Comprehensiveness and detail

The sixth criterion, comprehensiveness and detail, must be balanced with other criteria (e.g., wide applicability, ease of use). The DPQ is comprehensive in that its final forms includes items assessing all 15 facets of personality that were uncovered in Study 2 and is detailed in that these items address specific aspects of behavior (e.g., aggression when fearful, fear during grooming). However, even the longer 75-item form is far from

including every aspect of dog behavior. The DPQ sacrifices some comprehensiveness and detail for wide applicability. The items included in the DPQ are intended to apply to any population of domestic dogs, so behaviors specific to specialized groups (e.g., guide dogs) are excluded. The DPQ is likely still useful in predicting success in guide dog or other work, but as a general personality assessment, not an assessment of specific behavior. Furthermore, the 45-item form is necessarily less comprehensive and detailed than the 75-item form. The shorter questionnaire sacrifices comprehensiveness and detail in favor of brevity, manageability, and ease of use.

Wide applicability

The seventh criterion against which the DPQ was evaluated was wide applicability, or whether it is useable across a range of dogs and contexts. The general applicability of items included in the DPQ was addressed in Study 1 and Study 2. Items that more than 20% of the participants in Study 2 indicated were not applicable to their dogs (i.e., described situations that were too rare or their dog had never experienced) were removed.

Ease of use

The eighth criterion against which the DPQ was evaluated was ease of use. So that the DPQ can be generally useful, it should be easy to administer, to score, and to apply the results without much expertise or training. One aspect of making the questionnaire easy to administer is having items that are easy to interpret. In Study 1, 25 items were reworded based on participants' indications that those items were difficult to interpret. In later studies, lay dog owners completed the DPQ form with no further instructions than are included on the form. Scoring the DPQ requires reverse coding some items, then finding the mean of the items that load onto each facet or factor. The resulting scores are readily interpretable with respect to the factor labels (e.g., final Fearfulness scores will range from 1-5, with 5 being the most fearful).

Efficiency

The ninth criterion that guided the DPQ's development was efficiency. To be efficient, the tool should require a minimum of time to administer to collect a maximum amount of information about the dog. The long form of the DPQ requires less than ten minutes for a dog owner to complete, and, as discussed above, has relatively high content validity, comprehensiveness, and detail in balance with the criterion of wide applicability and manageability.

Manageability

The final criterion that guided the development of the DPQ was manageability. The goal was to design a test that could be administered with minimal monetary resources, physical space, and time, and with few test administrators. Generally, the DPQ meets this criterion. It can be presented online or as a paper-and-pencil form, raters can complete it with minimal guidance or with only the instructions provided on the form, and it takes approximately ten minutes to complete the long form. However, use of the DPQ is likely to be manageable for any group that is able to access people familiar with the dogs they aim to assess.

GENERAL EVALUATION OF THE DPQ

Strengths and major findings

The structured and systematic development of the DPQ fostered a number of technical and psychometric strengths in the instrument itself, and also some advances in basic and theoretical understanding of animal personality.

One strength of the process used to develop the DPQ is that it utilized a very large sample that was, in many ways, diverse. Dogs included in the studies varied in age, breed, and other demographic characteristics described for each study. Excluding dogs whose data were not ultimately used in creating the DPQ (e.g., dogs in the piloting of Study 6, dogs whose ratings were eliminated during data refinement), a total of 6,743 dogs were evaluated during the development of the DPQ (151 in Study 1; 3,737 in Study 2; 2,556 in Study 3; 99 in Study 4; 100 in Study 5; and 100 in Study 6). This is a strength

of the DPQ not only for reasons of statistical power, but also because the canine populations that drove the selection of items and derivation of factors and facets were likely to have been diverse in terms of personality.

A second strength of the DPQ is that many of its psychometric properties have been evaluated. This is an important strength, because the psychometric properties of dog personality assessments are frequently not assessed, and data about the tools' psychometrics are not available to people who might want to use the tools. The DPQ's reliability and validity have both been assessed. Its reliability, including internal consistency, inter-rater reliability, and test-retest reliability have all been evaluated in independent tests. The DPQ's predictive validity. both convergent and discriminant (the latter of which has typically been neglected), has also been evaluated.

Results of the evaluation of the DPQ's psychometric properties have a number of substantive implications that go beyond the basic utility of the DPQ as a personality assessment tool. First, the DPQ's high reliability and convergent validity, and even the less consistent discriminant validity findings, support the notion that non-human animals can be conceptualized in terms of personality, and that people can assign meaningful ratings to non-human animals' personalities. Dogs have individual personalities that come through in the rating process, despite other variables that may impact how the dog is rated. For example, it can be argued that each person who rates a dog affects his or her own ratings of the dog, such that each person brings his or her own personality and biases into play when rating a dog (Kwan, Gosling, & John, 2008). However, despite the effect of the individual rater, inter-rater reliability for the DPQ is high. Different people agree when they rate a dog's personality.

Furthermore, people's ratings of dogs' personalities and personality-driven behaviors, or specific DPQ items, group to form interpretable personality factors when analyzed through factor analysis. The question may remain whether the better DPQ factor solution is the four- or the five-factor solution, as discussed below. But the issue of whether there are four or five factors, whether Aggression should be one coherent factor or two factors separated by the target of the aggressive behavior, is a relatively trivial one

when compared to the question of whether dog personality can be understood in terms of factors at all. Each scale in the final form of the DPQ was found to be relatively cohesive. Each of the five factors has high internal consistency even though each includes a variety of items that represent different facets and different types of behavior. As discussed in Gosling et al. (2003a), internal consistency of scales is necessary to supporting the notion that personality traits exist in non-human animals.

The ultimate criterion in determining whether a personality assessment tool is useful is whether it predicts behavior. The DPQ showed relatively high convergent validity, particularly at the factor level, indicating that factor scores did correlate to the behaviors they were expected to predict. Discriminant validity correlations were more mixed, with some being unexpectedly high. Although evidence for discriminant validity was mixed, the overall predictive validity evidence has an important implication: rating of a non-human animal on a personality scale can, indeed, predict that animal's real world behavior. Dog personality ratings on the DPQ do correspond to behavior, further supporting the claim that non-human animals can be usefully and meaningfully characterized in terms of personality and, more specifically, personality traits.

A second substantial finding that arose from the development of the DPQ pertains to the number of factors or traits that best characterize dog personality. As mentioned above, the best model to fit the personality ratings that formed the basis for the DPQ is somewhat ambiguous. There may be four factors, or there may be five factors. Conceptually, however, the structure is the same for either factor solution. Dog personality can be characterized in terms of the factors of Fearfulness, Aggression (towards People and Animals), Activity/Excitability, and Responsiveness to Training. These labels are all descriptive, and, in particular, the Responsiveness to Training label is one that reflects that way that humans and dogs interact. Responsiveness to Training could also be described in terms of impulse control, attentiveness, and learning from the environment; reconceptualizing it in such terms might make the Responsiveness to Training factor easier to relate to factors derived in the study of other animals (e.g., humans') personality structures.

The four- or five-factor solution derived in the DPQ's development has noticeably fewer factors than the number suggested by the review of dog personality literature presented in Chapter 2. In the review of dog personality literature, dog behavior experts categorized the personality and behavioral descriptions from previous studies under various headings, resulting in seven dog personality trait categories: Reactivity, Fearfulness, Activity, Sociability, Responsiveness to Training, Submissiveness, and Aggression. It was suggested that Reactivity and Fearfulness might represent different facets of a larger Fearfulness or Neuroticism factor. In the DPQ analyses, behaviors related to Reactivity and Fearfulness in the literature review did indeed combine to form one broader factor, which I labeled Fearfulness. An Activity factor was also found in the DPQ analyses, labeled here as Activity/Excitability. The DPQ Activity/Excitability factor includes Companionability, some items of which (e.g., "Dog seeks companionship from people") relate to the literature review's Sociability factor. Items related to the Sociability factor (e.g., friendliness towards people, friendliness towards other animals) also load on the DPQ's Aggression towards People and Aggression towards Animals items; in the DPQ's factor solution, aggression and friendliness form opposite ends of a spectrum. The Responsiveness to Training factor from the literature review was also derived in the DPQ analyses. Submissiveness, seen as a separate factor by the experts who performed the sorting task in Chapter 2, was not found to be a separable factor in the DPQ analyses. Items related to submissive behavior when greeting people (e.g., "Dog exhibits submissive behavior [e.g., rolls over, avoids eye contact, yawns, licks lips] when greeting unfamiliar people") were found to load onto the Fearfulness facet, items related to dominance over the owner (e.g., "Dog is dominant over owner") were found to load onto the Responsiveness to Training factor. Items related to submissive or dominant behavior with other dogs (e.g., "Dog is assertive or pushy with other dogs [e.g., if in a home with other dogs, when greeting]") were found to load onto a separable facet of the Aggression towards Animals factor in the DPQ's five-factor model. The Aggression trait category from the literature review can be seen either as a single factor, as in the four-factor model of the DPQ, or divided by the target of the aggressive behavior to create Aggression

towards People and Aggression towards Animals, which were shown to be moderately correlated DPQ factors.

How do the DPQ's four conceptual factors (Fearfulness, Aggression, Activity/Excitability, and Responsiveness to Training) relate to the personality factors found in human and other non-human personality research? The four factors appear, broadly, to confirm Gosling and John's (1999) finding that dogs' personalities can be conceptualized in terms of traits paralleling some aspect of the human five-factor model traits. Neuroticism can be seen as parallel to the DPQ's Fearfulness factor. Agreeableness can be seen as parallel to the DPQ's Aggression factor (or factors, in the five-factor model). Extraversion, a large component of which is activity level in many non-human animal studies, is parallel to the DPQ's Activity/Excitability factor. In further support of this parallel, the Activity/Excitability factor also contains a facet labeled as Companionability, which includes behaviors such as seeking company and interaction with others. And, finally, Gosling and John (1999) suggested that dogs' personalities may include a factor that can be conceptualized as a combination of Openness and Conscientiousness. They indicated that this factor included learning and obedience ability (e.g., Coren, 1998) and also trainability (e.g., Hart & Hart, 1985). The Responsiveness to Training factor of the DPQ is very similar, and might also be adequately characterized as a blend of Openness and Conscientiousness. Gosling and John (1999) found Dominance and Territoriality to be a separate factor, however, whereas the analyses for the DPQ found dominance and related behaviors to be subsumed into the Fearfulness, Responsiveness to Training, and Aggression towards Animals factors, as discussed above.

Another interesting insight into the structure of canine personality, particularly with respect to how dogs can be meaningfully characterized, came from owners' feedback when they were rating their dogs. In Study 1, owners were asked to indicate when items were difficult to interpret or it was difficult to apply an item to their dog. If 5% or more of the participants flagged an item as difficult to interpret or use, then the item was presumed to be difficult to understand. At this threshold, 25 of the 360 items in

Study 1 were thought to be difficult for dog owners to understand or use. Seven of these 25 items, or 28% of the items, were from a pool of items modeled after the human Conscientiousness factor, drawn from the BFI (e.g., John, 1990). Of all of the items in the original pool of 1,284 items, these seven were the only ones modeled after human Conscientiousness items, or, indeed, posited to assess Conscientiousness. Because the items were thought to be difficult for dog owners to understand or use, an attempt was made to revise them. However, the items proved quite difficult to reword, in part because providing example behaviors to illustrate them was challenging. For example, "Dog does things efficiently" was reworded to "Dog achieves tasks (e.g., fetches objects) quickly and easily." However, the finding that it was difficult for participants to interpret traditional Conscientiousness-related items with respect to their dogs is consistent with Gosling and John's (1999) suggestion that Conscientiousness does not appear as a separate personality dimension in species other than humans, chimpanzees, and possibly other closely related apes. If it does appear in species such as dogs, it may appear only in conjunction with what might be described as the human Openness factor, as suggested by Gosling and John (1999), and by the DPQ's Responsiveness to Training factor.

Limitations and weaknesses

Although it can be argued that the DPQ meets the majority of the 10 criteria that guided its development and contributes to understanding canine personality in general, it is not a perfect instrument. One issue is that some DPQ items are weak in terms of inter-rater and/or test-retest reliability. Further study of the aspects of dog personality assessed by DPQ items with weak reliability might help researchers understand these aspects of dog personality. Perhaps the items have low reliability correlations because they are not clear to dog owners, or the items might use behavior descriptions that dog owners would not typically use. A better understanding of these aspects of personality and how people describe their dogs' personalities might facilitate the development of items with higher reliability.

A second potential weakness of the DPQ relates to its five-factor structure. Researchers might argue that a stronger case can be made for a four-factor structure in which Aggression is collapsed (instead of divided, as it is in all of the five-factor solutions). In terms of content, however, it is arguable whether this difference in structure would significantly impact the DPQ. Furthermore, division of aggression into Aggression towards People and Aggression towards Animals, as it appears in the five-factor model, is of practical use to people assessing dog personality. The simple division of Aggression into items pertaining to people and items to animals is a meaningful and useful one because of how potential users of the DPQ (e.g., lay people, shelter workers, dog trainers) are likely to want to discuss canine aggression. So, although dogs rated on the DPQ as Aggressive towards People also tended to have been rated as Aggressive towards Animals (according to the discriminant validity correlations presented in Study 6), maintaining two separate factors is of practical utility.

On a larger scale, the DPQ is limited in what it assesses. That is, the DPQ does not purport to assess everything that might be argued to be related to individual differences and personality. Personality and personality traits were defined much as Allport (1937) defined them when beginning to assess human personality using the lexical approach—as predispositions to respond in a particular way to a broad range of situations. So, the DPQ was designed with the goal of measuring personality-driven behavioral differences, and over-arching personality characteristics (e.g., anxiousness), in dogs. This definition of personality excludes the domains of physical appearance or attractiveness, and also interaction effects, such as how two individuals get along together. Although physical appearance and attractiveness are relatively stable across time and situation, and they may indirectly impact behavior, they themselves are not psychological predispositions. The DPQ does not ask how cute a dog is, or whether his or her ears are erect or pendulous, although these characteristics might affect whether or not, for example, a person seeking to adopt a dog from a shelter wishes to adopt the dog. A dog's appearance can probably be better communicated through a photograph than through a series of ratings or measurements, which would serve as a poor substitute.

Similarly, the DPQ does not directly address issues of how different individuals' personalities and related behavioral patterns interact. That is, a dog might respond differently (e.g., more playfully) to people who are themselves energetic and playful, and a dog may be better behaved and more under control with people who communicate clearly with the dog. The DPQ attempts to assess a more general trait of the dog, not his or her response to particular traits in others. Thus, the DPQ can provide general personality information about an individual dog, but does not purport to capture all of the nuances of the dog as an individual and a social animal.

Another weakness related to the development of the DPQ is the dog owner demographic used in its development. That is, the majority of the dog owners who took part in the DPQ's developmental phase, or Studies 1, 2, and 3, were Caucasian women. A total of 5,572 women were included in Studies 1, 2, and 3, composing 86.5% of the participants across those three studies. The average age in the studies was just over 43 years of age (Study 1 mean age = 44 with a standard deviation of 12; Study 2 mean age = 43 with a standard deviation of 12; Study 3 mean age = 43 with a standard deviation of 12.6). Although the average age for all three samples was approximately the same, it should be noted that the standard deviation is about 12 years for each study, and that the range across all three studies was 18 to 84 years of age. Because the demographic of the dog owners who took part in the developmental stages of the DPQ is so uniform in terms of sex, age, and race, the question of how or whether the DPQ would differ if developed using a different demographic (e.g., younger dog owners, men with working dogs) remains to be addressed.

The demographic of the DPQ's samples also evokes the issue of the self-selection bias, and particularly the question of online sampling, in psychological research. Studies 1 through 5 utilize web-based data collection and self-selected participants. As discussed, web-based studies and online self-selected Internet samples are becoming more and more popular as the Internet becomes more widespread and accessible, but these methods have also been a target of concern. It is important to address and understand the concerns pertinent to the methods utilized in the studies. Specific issues that have been addressed

include the impact of the integrity of the data (McGraw et al., 2000), the correspondence between surveys conducted online with those conducted through traditional postal mail (McCabe et al., 2006), the effects of recruitment strategies (Buchanan et al., 2005), and self-selection (Walsh et al., 1992). In each case, the data have argued against these concerns. McGraw et al. (2000) and McCabe et al. (2006) found that data collected through the Internet were the same as data collected through traditional means. Buchanan et al. (2005) found no effects of how Internet samples were recruited. Walsh et al. (1992) found that a self-selected sample and a randomly selected sample did not differ significantly in demographic information; the participants in the self-selected sample, however, tended to provide longer free-responses and left fewer items blank, suggesting that they might be more interested in or care more about their participation. Walsh and colleagues' finding was probably not replicated in the DPQ's online studies, in which the samples were composed of primarily Caucasian women. They did not represent a random sample of dog owners. However, the participants who took part in the DPQ did tend to provide lengthy free-responses, follow-up questions, and indicate they cared about the subject matter. The sample demographics may have been very similar to the demographics that would have been found in a randomly selected sample, if that sample were randomly selected from dog owners invested enough to fill out such a lengthy questionnaire. The skewed demographics of the sample that provided the basis for the DPQ's development may be a weakness, but only if it lead to an idiosyncratic factor solution or the development of items that are readily interpretable by only a specific group of dog owners. These are issues that could be addressed through the validation of the DPQ in a more demographically diverse sample.

In addition to the potential issues of data integrity, self-selection, and so on, Gosling and his colleagues (2004) culled from the literature what they term "preconceptions" that researchers may have about web-based research, and particularly about people who participate in web-based research. Their analyses found that three of these six preconceptions, that Internet samples are socially maladjusted, isolated, or depressed; that Internet data do not generalize across the various formats in which data

can be presented on the internet; and that internet participants are unmotivated, are unsupported by data. Gosling et al. also found evidence suggesting that another preconception, that internet data are unique and different from traditionally collected data, is likely untrue but they believe more data are needed. Mixed evidence was found for whether internet samples are demographically diverse. Evidence supported the preconception that internet data can be compromised by the anonymity of the participants, but Gosling and his colleagues describe multiple means by which researchers can reduce this risk through various means (e.g., removing data from people who respond to the questionnaire multiple times).

Web-based data collection also affords many advantages over traditional methods of data collection. For example, online data collection tends to be very efficient, because many participants can be reached and can complete the study at once without requiring an experimenter to administer the study. Online data collection also removes the necessity of entering data and, with it, the risk of data entry mistakes. In addition, using the internet enables researchers to reach people outside of the typical subject pool (e.g., Gosling et al., 2004), including people who are disabled, geographically distant, elderly, or in a specific and rare population.

For purposes of the current set of studies, I deemed the advantages of online data collection to greatly outweigh the potential risks. The goal of creating a questionnaire that is widely applicable requires piloting and evaluating the questionnaire in a diverse population. Internet respondents (dog owners) had the potential to vary in terms of age, sex, occupation, type and extent of experience with dogs, their geographic region, etc. However, the demographics of the current sample were not particularly diverse, which likely reflects a self-selection bias due to participants' interest in the study and investment in their dogs, not due to the web-based nature of the study. The use of the internet to develop the DPQ likely facilitated its development, allowing me to reach a large, specialized group of people who are both knowledgeable about dogs and their dog in particular, and interested enough in dog behavior to complete a lengthy questionnaire.

Future directions

Future studies and analyses are indicated by some of the limitations and weaknesses in the current series of studies, and also by questions that have been elucidated through the process of developing the DPQ. One question is related to the fundamental issue of where personality comes from. How much does biology (e.g., genetics) determine adult behavior and personality? How much of what is seen as personality is behavior that is learned through interaction with the environment? Because of the nature of people's relationships with dogs, which often include socializing puppies in order to increase their friendliness towards people and training them to behave in certain ways, this question is of particular relevance when assessing dogs' personalities. If personality is conceptualized as behavioral patterns and tendencies that remain consistent across time and situation, it is possible that training and early experiences can shape dogs' personalities. In fact, the importance of early experiences and training may be part of why personality assessments performed with puppies have not been very accurate in predicting adult dogs' behavior (see the literature review in Chapter 2; e.g., Goddard & Beilharz, 1986; Hennessy et al., 2001; Wilsson & Sundgren, 1998). The issue of the power of experience and training in shaping dogs' personalities suggests multiple research questions. For example, if puppies are raised in specific and controlled environments and trained in way to increase friendliness, playfulness, and so on, will all of the puppies end up the same? Evidence from genetic studies (e.g., Saetre et al., 2004) indicates that the answer is no, because personality traits, particularly those related to Fearfulness, or possibly to Svartberg and colleagues' (e.g., Svartberg, 2002; Svartberg et al., 2005) broad Boldness/Shyness dimension, have genetic components. So, how much can a carefully controlled rearing environment impact behavior? And how long-lasting are the effects of training? If a puppy or an adult dog is, for example, reinforced for being friendly towards strangers or playing vigorously with toys, will that reinforced behavior become a long-term and enduring behavior? How much can training shape personality?

A second future direction for the DPQ is evaluation of the DPQ using a more diverse sample of dog owners. As discussed in terms of limitations and weaknesses, the

DPQ's online studies have relatively limited demographic diversity in terms of owner age, sex, and race, although the samples are diverse in terms of a number of other variables (e.g., experience with dogs, geographic location). In order to determine the effect that the dog owners' demographics of the original DPQ study samples on the development of the DPQ, similar studies would need to be repeated using different demographics. However, these issues could also be addressed by examining the DPQ's predictive validity when the DPQ is used by different dog owners. Are the scores of men who rate their dogs on the DPQ as predictive as women's? Are there differences depending related to the race of the dog owner? Such differences might indicate that the DPQ's structure and development was affected by idiosyncrasies of the initial, or Study 1, 2, and 3, samples.

A number of future directions for DPQ research involve examining the validity of the DPQ. The third possible future direction is validation of the DPQ in additional contexts. For example, it is mentioned above that predicting adult personality and behavior from assessments of puppies is notoriously difficult. How well do DPQ scores assigned to puppies predict adult dogs' DPQ scores and other personality assessments (e.g., behavioral test results)? Another issue is predicting dogs' behavior further in the future than was assessed in the current validation study (Study 6). That is, how do DPQ scores predict dogs' behavior two, three, or five years from now? Also, the test battery used in Study 6 was designed to include relatively normal situations. Do scores on the DPQ predict how the dogs will respond to unusual circumstances or contexts that may not be clearly related to items on the DPQ?

The discriminant correlations found in the validation of the DPQ bring up a fourth potential future direction. A number of correlations predicted to be discriminant were actually quite large. At least three explanations could make sense of these unexpectedly large correlations. First, the predictions may have been poorly made; perhaps these correlations are reasonable and replicable. Second, the DPQ may simply have items or factors with low discriminant validity. Or third, discriminant and convergent validity might be better conceptualized as graded, as described next.

A fifth question related to the validation of the DPQ is related to how discriminant and convergent validity correlations should be conceptualized. In Study 6, correlations were predicted to be either discriminant or convergent. The categorizing of correlations as either discriminant or convergent indicates that they should be either rather low (discriminant) or rather high (convergent), and implies that no correlations are expected to be in the middle range. However, as Cronbach and Meehl (1955) argue, this black and white image of construct validity is neither the most intuitive nor the most accurate. A more graded conceptualization of validity may be more appropriate. Indeed, in the case of the DPQ, it is logical to expect moderate correlations between the Fearfulness on the DPQ and tests expected to be related to Aggression towards People and vice versa, and also to expect moderate correlations between the Aggression towards People on the DPQ and tests expected to be related to Aggression towards Animals. These correlations are logically expected because the DPQ's Fearfulness and Aggression towards People factors, and the DPQ's Aggression towards People and Aggression towards animal factors are correlated. Re-examination of the Study 6 discriminant and convergent validity data, conceptualized in terms of low, moderate, and high correlations, may give a more accurate idea of the DPQ's predictive validity.

A sixth potential future direction, which also related to the DPQ's validation, is whether different raters are associated with different levels of predictive accuracy. In Study 6, for example, the dogs' owners and kennel staff who were also familiar with the dogs rated them on the DPQ. The dogs' owners likely spend more time with the dogs and see them in a wider variety of contexts, so they might be argued to know the dogs' personalities better. However, the kennel staff are familiar with the dogs in the kennel environment, where the behavioral assessments were conducted. Do the kennel staff members' ratings better predict the dogs' behavior in the kennel environment, because the kennel staff are more familiar with the dogs in that environment? Or are the dog owners' ratings better predictors, because the dog owners are familiar with the dogs more generally and in more contexts? The Study 6 data could be reanalyzed to gain insight into these questions, or additional studies could be conducted.

A seventh potential future direction is how much people's ratings of their and other dogs are shaped by their own personalities. Given that the DPQ had relatively high inter-rater reliability and that scores on the DPQ predicted behavioral test results, the ratings were assessments of the target dogs themselves. However, the ratings are also impacted by a number of others factors, one of which is likely the raters' personalities. Do raters project their own personalities or those of specific others onto the dogs they rate such that more neurotic people will rate dogs as more fearful, or less fearful if they contrast themselves with the dogs? Or is there another systematic way in which raters' personalities shape the ratings they assign to dogs? Using the DPQ, these questions can be addressed. Findings might be expected to replicate those reported in Kwan, Gosling, and John (2008): raters do project their own personalities onto the dogs whom they rate, but less than they project their own personalities onto other people whom they rate.

Eighth, just as people's own personalities and experiences shape the ratings they assign to a dog's personality, people's experiences and cultures might shape how they conceptualize specific species or all animals and their general behavior. If the DPQ had been developed in a different culture, one that did see dogs as companion animals and use them in practical and functional working capacities, then the DPQ might have developed differently. How likely would people be to rate a dog as highly affectionate or intelligent if those people live in a culture in which dogs are used as a food source or are looked down upon as dirty? If developed in a different culture, the DPQ might have had a different factor structure, leading to the selection of different items for inclusion in the final form of the DPQ. Examination of factor structures derived from using the DPQ in a variety of cultures might have interesting implications for dog personality's structure or for people's perception of animals as related to the role they play in a culture.

Finally, the DPQ is a broad and general tool for assessing personality in dogs. This is one of its strengths and, indeed, allows it to meet the criteria of wide applicability. However, it might also be a weakness. Future studies or applications of the DPQ might benefit from supplementing the DPQ in ways that are useful to the specific context in which it is being used. For example, a dog shelter that uses the DPQ during surrender of

dogs to be re-homed might want to include supplemental questions asking whether the dog has any specific and idiosyncratic behaviors (e.g., specific phobias, aggression towards very specific targets). A group assessing dogs for service as search and rescue animals might be interested in dogs' abilities and tendencies to search by smell and their physical and mental endurance. Such detailed information is not included in the DPQ, but could lead to the development of supplemental DPQ forms to meet the needs of groups with specific needs and goals.

As a general personality assessment tool, the DPQ will, hopefully, fit the needs of a wide variety of groups (e.g., shelters, guide dog programs, researchers) seeking to evaluate dog personality and use personality to predict other factors (e.g., dogs' suitability for specific homes, as guide dogs, propensity for health problems). Accurate evaluation and good research depend on accurate, reliable, valid measurement. The DPQ should provide a solid foundation for a broad array of research, have uses in applied settings, and open up avenues for increased understanding of how people rate and evaluate personality in non-human animals.

Appendix A. Study 1: 51 broad, content-based categories for sorting items

Activity
Adaptability
Affection
Affiliation
Aggression
Alertness
Anxiety
Attachment
Attention
Barking
Body sensitivity
Boldness-Shyness
Chasing
Competence
Conscientiousness
Cooperation
Curiosity
Destructiveness
Digging/Burying
Distractability/Focus
Emotionality
Excitability
Extraversion
Fearfulness
Fetch/Retrieve
Following
Food-related behavior
Forgiving/Forgetting unpleasant experiences
Hardness
Hearing sensitivity
Intelligence
Jealousy
Marking behavior
Mounting behavior
Nervous aggression
Nervousness
Nose ability
Obedience
Opportunistic
Pain sensitivity
Playfulness
Recovery
Self-sufficiency
Separation-related behavior
Sociability
Submission
Suspicion
Temperament
Trainability
Tug-o-War behavior
Willingness

Appendix B. Study 2: Loading of 352 personality items on five varimax-rotated factors

Item text	Loading of each item on each factor				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1: Fearfulness					
[192] Dog is fearful.	**.721**	.178	-.062	.028	.117
[194] Dog is timid.	**.712**	.077	-.125	-.011	-.011
[269] Dog is nervous.	**.697**	.157	.030	.085	.168
[135] Dog is shy.	**.662**	.163	-.177	-.104	-.013
* [193] Dog is confident.	**-.653**	-.061	.228	-.058	-.088
[211] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when near crowds of people.	**.651**	.254	-.034	-.008	.060
[9] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar men.	**.625**	.310	-.044	-.088	-.077
*[122] *Dog is anxious.*	**.622**	.139	.076	.164	.190
[197] Dog attempts to flee from novel objects or situations.	**.619**	.113	-.036	.098	.047
*[270] *Dog is anxious.*	***.618***	*.164*	*.085*	*.151*	*.183*
[10] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar women.	**.614**	.328	-.075	-.098	-.079
[5] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when an unfamiliar person visits the home.	**.607**	.328	-.044	-.077	-.125
[226] Dog exhibits fearful behavior (e.g., attempts to flee, tucks tail, trembles) when cornered by a person.	**.595**	.249	.008	.034	.050
[209] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar objects (e.g., canes, wheelchairs, umbrellas).	**.586**	.148	.046	.092	.046
[3] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when head or collar is reached for by an unfamiliar person.	**.585**	.325	-.013	-.044	.010
[13] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar children.	**.582**	.248	-.058	-.092	.044
[345] Dog behaves fearfully (e.g., raises hackles, flees) towards unfamiliar people.	**.582**	.421	-.005	-.068	.003
[207] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards highly active, excited children (e.g., running, yelling).	**.576**	.168	-.046	-.053	.061
[219] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards loud or heavy traffic.	**.568**	.077	.033	.078	.037
* [240] Dog remains calm in stressful situations.	**-.562**	-.163	-.073	-.220	-.274
[245] Dog is fearful when walking near loud, heavy traffic.	**.557**	.060	.015	.103	.052
[214] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when an unfamiliar person approaches the home or yard.	**.538**	.276	-.004	.003	-.111
* [64] Dog adapts easily to new situations and environments.	**-.538**	-.203	.136	-.168	-.225
* [195] Dog is quick to recover after a frightening experience.	**-.525**	-.080	.132	-.195	-.133
[224] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards moving trucks, cars, or motorcycles.	**.518**	.082	.048	.066	-.004
[241] Dog is startled by sudden loud noises (e.g., a slamming door, car horns).	**.515**	.061	.117	.165	.157
[17] While on leash, dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar large dogs.	**.511**	.063	.051	.025	.070

Appendix B. (*Continued)*

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 1: Fearfulness (*Continued*)					
* [271] Dog is able to adapt to various types of situations.	**-.511**	-.212	.152	-.245	-.184
* [65] Dog adapts easily to stressful environments (e.g., kennels, new places).	**-.510**	-.149	.061	-.165	-.249
[123] Dog often appears anxious (e.g., has tight facial muscles, holds ears back tightly, darts and pulls leash).	**.507**	.229	.085	.175	.179
[343] Dog shows uncertainty or caution towards specific object, animal, or person.	**.503**	.190	.056	-.030	.155
[7] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when an unfamiliar person approaches a family member.	**.501**	.333	.022	.015	-.107
[227] Dog exhibits fearful behavior (e.g., attempts to flee, tucks tail, trembles) when cornered by another dog.	**.500**	-.051	.073	.058	-.049
[221] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when examined by the vet.	**.499**	.167	.029	.100	.110
[133] Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a doorframe).	**.494**	.093	-.014	.098	.099
[220] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) during visits to the veterinary office.	**.492**	.153	-.005	.084	.146
[225] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards moving bicycles, joggers, or skateboarders.	**.490**	.141	.023	.052	-.045
[198] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when threatened by a dog (e.g., barked, growled, or lunged at).	**.490**	-.057	.099	.040	-.148
* [273] Dog is able to relax and overcome frightening experiences.	**-.490**	-.091	.093	-.231	-.150
[11] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar men.	**.487**	.333	-.085	-.022	-.149
* [134] Dog is bold.	**-.486**	.144	.313	.083	.150
* [298] After being startled, dog is quick to recover.	**-.483**	-.123	.094	-.221	-.120
[222] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when walking on a leash.	**.480**	.143	-.059	.057	-.006
[215] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when scolded or punished.	**.477**	.013	.093	.147	.100
[174] Dog can seem tense.	**.477**	.195	.036	.112	.299
[348] Dog tends to over-react to events or stimuli in the environment.	**.476**	.148	.169	.284	.238
[205] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when it is restrained.	**.474**	.235	-.007	.139	.034
[18] While off leash, dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar large dogs.	**.473**	.044	.083	.032	.063
[344] Dog is suspicious of people, things, or situations.	**.467**	.446	.043	-.043	.198
[19] While on leash, dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar small dogs.	**.467**	.067	.015	.041	.043
[14] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar children.	**.457**	.252	-.085	-.035	.026
[208] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards toddlers.	**.447**	.226	-.019	-.078	.080

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 1: Fearfulness (*Continued*)					
[20] While off leash, dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar small dogs.	**.435**	.056	.019	.050	.056
* [268] Dog reacts appropriately to various situations.	**-.433**	-.333	.050	-.277	-.226
[12] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar women.	**.430**	.327	-.096	-.002	-.137
[340] Dog is submissive.	**.425**	-.167	-.006	.000	-.185
* [158] Dog enjoys going new places.	**-.425**	-.109	.252	-.098	-.054
[6] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when a familiar person visits the home.	**.425**	.319	-.103	.003	-.133
[202] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when it is bathed.	**.420**	.025	.017	.231	.107
* [213] Dog is relaxed when greeting an unfamiliar woman.	**-.420**	-.403	-.053	-.018	-.132
[4] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when head or collar is reached for by a familiar person.	**.415**	.258	-.069	.118	-.051
[341] Dog is cautious, careful.	**.411**	.098	-.107	-.195	.049
[210] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards specific objects (e.g., the toaster, garden hose, shoes).	**.407**	.062	.074	.134	.037
[352] Dog is sensitive, easily upset by corrections.	**.402**	-.082	-.001	-.034	.103
[204] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when it is groomed (e.g., brush coat, brush teeth).	**.391**	.101	.002	.216	.019
[201] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when its ears are examined.	**.380**	.139	.008	.180	.028
[21] While on leash, dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar large dogs.	**.368**	.112	-.022	.106	.078
* [172] Dog is happy-go-lucky, carefree.	**-.365**	-.264	.255	.032	-.313
[8] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when a familiar person approaches a family member.	**.363**	.345	-.053	.077	-.136
[218] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards some toys.	**.357**	.114	-.006	.091	.048
[223] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards family members.	**.348**	.232	-.080	.095	-.032
[203] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when its toe nails are trimmed.	**.344**	.051	.101	.176	.068
[332] Dog exhibits submissive behavior (e.g., rolls over, avoids eye contact, yawns, licks lips) when greeting dogs.	**.339**	-.079	.122	.072	-.240
[22] While off leash, dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar small dogs.	**.339**	.106	-.047	.085	.045
[244] Dog is fearful of fireworks.	**.336**	-.011	.009	.069	.144
[2] Dog exhibits submissive behavior (e.g., rolls over, avoids eye contact, yawns, licks lips) when greeting unfamiliar people.	**.331**	-.093	.017	.096	-.104
[1] Dog exhibits submissive behavior (e.g., rolls over, avoids eye contact, yawns, licks lips) when greeting familiar people.	**.325**	.035	.052	.069	.006

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 1: Fearfulness (*Continued*)					
[331] Dog avoids pushy dogs.	**.315**	.094	.031	.143	.050
[310] Dog exhibits separation-related problems.	**.310**	-.158	-.085	-.122	-.073
[304] When left or about to be left alone, dog becomes restless or agitated (e.g., paces).	**.304**	.012	.205	.247	.085
[296] Dog behaves consistently across different situations.	**.301**	.045	.189	.276	.092
* [307] When alone or about to be left alone, dog shakes, shivers, or trembles.	**-.298**	-.209	-.008	-.233	-.128
[283] Dog is sensitive to mild pain.	**.292**	.092	.104	.201	.011
[239] Dog easily gets over unpleasant experiences (e.g., painful toe nail clippings).	**.280**	.003	.054	.060	.039
* [325] Dog is aloof towards unfamiliar men.	**-.278**	-.165	.060	-.149	-.133
[216] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) during or just before car rides.	**.275**	.247	-.175	-.204	.048
[336] Dog urinates when stressed or threatened.	**.275**	.025	-.009	.125	.021
[243] Dog is fearful of thunderstorms.	**.264**	.118	.086	.176	.013
[284] Dog is sensitive to moderate pain.	**.264**	-.062	-.031	.083	.161
[306] When alone or about to be left alone, dog salivates excessively.	**.255**	-.014	.037	.057	.083
[15] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards unfamiliar cats.	**.246**	.109	.049	.223	.016
[327] Dog is aloof towards familiar men.	**.236**	.051	.004	.100	-.079
[282] Dog reacts (e.g., yelps, cries) when a person steps on its paw or tail.	**.229**	.224	-.183	-.025	-.002
[312] Dog tends to be independent.	**.226**	.008	.108	.093	.107
* [200] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards small wild animals (e.g., squirrels and birds).	**-.216**	.167	-.081	.025	.074
[16] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) towards familiar cats.	**.203**	.102	.000	.076	-.095
[217] Dog exhibits fearful behavior (e.g., flees, tucks tail, trembles) when going up or down stairs.	**.201**	.064	.006	.081	-.017
[303] Dog will not eat when left alone.	**.189**	-.016	-.022	.119	.016
[234] Dog is a fussy or picky eater.	**.180**	.020	.077	.084	.028
[331] Dog avoids pushy dogs.	**.141**	.055	-.035	.066	-.053
Factor 2: Aggression towards People					
[48] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar women.	.166	**.716**	.040	-.039	.209
[47] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar men.	.167	**.697**	.060	-.069	.227

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 2: Aggression towards People (*Continued*)					
[111] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when directly approached by an unfamiliar person.	.169	**.693**	.080	-.061	.244
[45] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when an unfamiliar person visits the home.	.179	**.677**	.057	-.069	.224
[27] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when head or collar is reached for by an unfamiliar person.	.139	**.673**	.034	-.033	.183
[112] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when stared at by an unfamiliar person.	.143	**.633**	.055	-.041	.174
[38] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when cornered by an unfamiliar woman.	.135	**.625**	.066	-.046	.215
[37] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when cornered by an unfamiliar man.	.143	**.616**	.068	-.074	.219
[49] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards familiar men.	.144	**.605**	-.021	.061	.019
[43] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when an unfamiliar person approaches a family member.	.125	**.605**	.097	-.048	.218
[46] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when a familiar person visits the home.	.122	**.603**	.004	.043	.065
[50] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards familiar women.	.130	**.603**	-.035	.096	-.010
* [320] Dog is friendly.	-.269	**-.588**	.192	.040	-.178
[113] Dog has a tendency to attack (or attempt to attack) people.	.101	**.572**	-.009	.020	.193
[97] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar children.	.178	**.567**	.007	-.044	.229
[75] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when threatened by a person.	.039	**.545**	.100	-.019	.263
[44] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when a familiar person approaches a family member.	.108	**.543**	.030	.096	.057
[118] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) during visits to the veterinary office.	.055	**.540**	.019	.113	.116
* [321] Dog is friendly towards unfamiliar men.	-.353	**-.540**	.073	.194	-.163
[77] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when an unfamiliar person approaches the home or yard.	.145	**.536**	.130	-.074	.296
[267] Dog shows aggression when nervous or fearful.	.234	**.535**	.105	.095	.329
[119] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when examined by the vet.	.067	**.525**	.015	.112	.085
[84] Dog is aggressive.	.052	**.516**	.042	.105	.457
[99] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards highly active, excited children (e.g., running, yelling).	.174	**.507**	.041	.012	.271
[105] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards crowds of people.	.160	**.502**	.044	.063	.154
[107] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards family members.	.041	**.497**	-.016	.188	.037
* [85] Dog is not aggressive.	-.065	**-.490**	-.031	-.088	-.431

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 2: Aggression towards People (*Continued*)					
* [322] Dog is friendly towards unfamiliar women.	-.322	**-.485**	.095	.152	-.125
[127] Dog is bad-tempered.	.094	**.482**	-.048	.181	.128
* [212] Dog is relaxed when greeting an unfamiliar man.	-.422	**-.473**	-.032	.053	-.126
[40] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when cornered by a familiar child.	.137	**.468**	.005	.043	.196
[26] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when head or collar is reached for by a familiar person.	.083	**.467**	-.016	.184	-.032
[36] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when cornered by a family member.	.068	**.453**	.048	.203	.038
* [354] Dog is generally trusting.	-.395	**-.452**	.122	-.077	-.142
[98] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards familiar children.	.102	**.452**	.011	.027	.144
[108] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards moving bicycles, joggers, or skateboarders.	.122	**.451**	.106	.073	.300
[95] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when it is restrained.	.135	**.450**	.001	.222	.138
[101] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar objects (e.g., canes, wheelchairs, umbrellas).	.169	**.450**	.103	.102	.167
[100] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards toddlers.	.164	**.443**	.017	-.012	.207
* [86] Dog is amiable.	-.212	**-.431**	.089	-.111	-.071
* [289] Dog is playful with unfamiliar people.	-.338	**-.412**	.318	.190	-.207
[83] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when stared at by a family member.	.081	**.398**	.044	.195	-.002
[90] Dog displays aggression that is sudden and without apparent reason	.131	**.392**	.011	.137	.170
[117] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when someone takes away or attempts to take away stolen items (e.g., socks, tissues, people food).	.056	**.387**	.041	.291	.099
[94] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when it is groomed (e.g., brush coat, brush teeth).	.108	**.382**	.053	.221	.044
[78] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when removed from furniture.	.041	**.372**	.017	.244	.082
[96] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when its paws are handled.	.057	**.370**	.044	.219	.088
[82] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when directly approached by a family member.	.044	**.370**	-.025	.143	-.016
[91] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when its ears are examined.	.096	**.366**	.019	.202	.035
[106] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when scolded or punished.	.044	**.363**	.014	.261	.059
[92] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when it is bathed.	.072	**.361**	.030	.232	-.048

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 2: Aggression towards People (*Continued*)					
[116] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when someone takes away or attempts to take away delicious items (e.g., pig ears, bones).	.147	**.353**	.110	.121	.133
* [323] Dog is friendly towards familiar men.	-.138	**-.339**	-.067	-.120	-.206
[93] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when its toe nails are trimmed.	.061	**.338**	.052	.286	.147
* [324] Dog is friendly towards familiar women.	-.229	**-.337**	.145	.015	.012
[299] Dog behaves erratically.	.051	**.329**	.106	.237	.083
* [79] Dog exhibits aggressive behavior (e.g., bares teeth, growls, snaps) when disturbed while sleeping (e.g., told to move, petted, tripped over).	-.162	**-.325**	.147	.017	.021
[25] When on leash, dog barks at bicycles, children running, or joggers.	.273	**.321**	.103	**.320**	.132
[326] Dog is aloof towards unfamiliar women.	.068	**.315**	.017	.236	.135
[115] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when approached while eating.	.181	**.307**	.195	.155	.236
[328] Dog is aloof towards familiar women.	.267	**.304**	-.157	-.126	.048
[139] Dog likes to chase bicycles, joggers, and skateboarders.	.094	**.295**	.009	.246	.119
[132] Dog is responsive to petting, handling.	.186	**.262**	-.179	-.006	-.027
[102] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards specific objects (e.g., the toaster, garden hose, shoes).	.072	**.259**	.224	.181	.241
* [116] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when someone takes away or attempts to take away delicious items (e.g., pig ears, bones).	-.108	**-.241**	.218	-.142	-.024
[323] Dog is friendly towards familiar men.	.115	**.229**	.147	.145	.096
Factor 3: Activity/Excitability					
	-.048	.084	**.678**	-.097	-.042
[54] Dog is active during play with people.	-.094	.016	**.654**	-.072	-.075
[63] Dog is active.	-.064	.079	**.650**	-.150	-.085
[188] Dog is very excitable during play with toys.	.112	.055	**.633**	.133	-.020
[291] Dog enjoys playing with toys.	-.035	.057	**.624**	-.109	-.141
[292] Dog initiates play (e.g., by bringing toys).	-.047	.079	**.612**	-.126	-.129
[52] Dog is active in the home.	.005	.096	**.605**	-.031	-.015
[228] Dog chases after thrown objects (e.g., sticks, balls, or toys).	-.035	.030	**.587**	-.132	-.077
[56] Dog is highly active in the yard.	.019	.118	**.586**	-.009	.004
[294] Dog plays boisterously.	-.050	.069	**.575**	.144	.032
[356] Dog is interested in playing tug of war with people.	-.035	.071	**.569**	.052	-.035
[53] Dog is active when off leash.	-.080	.062	**.568**	-.096	-.024
[288] Dog is playful with familiar people.	-.134	-.105	**.548**	-.020	-.097
[58] Dog seeks constant activity.	.047	.067	**.548**	.152	.022
[178] Dog is very excitable during play with other dogs.	.046	.020	**.544**	.229	-.029
[229] Dog carries objects such as toys, balls, sticks, or leash in mouth.	.015	.038	**.540**	-.063	-.085

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 3: Activity/Excitability (*Continued*)					
[60] Dog is boisterous.	-.046	.100	**.540**	.225	.094
[184] Dog is very excitable when with people.	.122	-.023	**.527**	.319	.051
[358] Dog likes to grasp and shake toys.	.003	.110	**.521**	.044	.002
[157] Dog is curious.	-.202	.043	**.519**	-.167	-.001
[346] Dog is interested in what is going on around it.	-.081	-.013	**.519**	-.273	.031
[55] Dog is active during play with other dogs.	-.148	.016	**.510**	-.022	-.224
[230] Dog retrieves objects (e.g., balls, sticks, toys).	-.024	.040	**.509**	-.205	-.083
[121] Dog is very alert.	-.018	.076	**.493**	-.262	.094
[357] Dog is interested in playing tug of war with other dogs.	-.086	.054	**.490**	.036	-.172
[183] Dog is very excitable when visitors arrive.	.082	-.024	**.485**	.264	.184
[293] Dog enjoys play fighting or playing roughly.	-.037	.081	**.471**	.176	.044
[361] Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden.	-.053	-.005	**.449**	-.292	.051
[61] Dog is lethargic.	.036	-.041	**-.446**	.219	.030
[51] Dog is active (e.g., jumps, sniffs a lot) when walking on a leash.	.062	.038	**.442**	.232	.135
[179] Dog is very excitable when doorbell rings or there is a knock at the door.	.151	.061	**.432**	.083	.247
[190] Dog is very excitable when on a walk.	.157	.048	**.427**	.345	.197
[180] Dog is very excitable when owner returns home.	.189	-.025	**.424**	.135	.112
[308] Dog becomes wildly excited when owner returns home.	.168	.009	**.420**	.196	.088
[177] Dog is very excitable when meeting other dogs.	.034	.019	**.415**	.285	.090
[126] Dog is extremely demanding of attention.	.142	-.005	**.413**	.264	.191
[67] Dog frequently demands affection.	.118	-.054	**.409**	.172	.132
[246] Dog is clever.	-.126	.043	**.399**	-.353	.066
[342] Dog is alert, watchful (e.g., monitors yard for squirrels, monitors noises around the home).	.026	.139	**.398**	-.123	.226
[249] Dog is able to act on own initiative (e.g., brings owner leash or toy without being told to do so).	-.107	.040	**.398**	-.302	-.004
[233] Dog follows people around.	.108	-.113	**.393**	.037	.051
[242] Dog is interested in sounds (e.g., pays attention to sounds, looks for source).	-.105	.030	**.390**	-.187	.066
[311] Dog seeks companionship from people.	-.069	-.261	**.381**	-.049	.016
* [295] Dog gets bored in play quickly.	.045	-.014	**-.379**	.175	.156
[69] When in the home, dog follows owner/family member from room to room.	.154	-.057	**.372**	.030	.102
[66] Dog often expresses affection.	-.057	-.173	**.369**	-.148	.000
[181] Dog is very excitable when handled, petted, or groomed.	.135	.014	**.366**	.318	.015
[360] Dog appears to remember an object when it is out of sight (e.g., in your pocket, behind your back).	-.090	-.026	**.364**	-.325	.061
* [191] Dog is very vocal.	.042	.140	**.363**	.130	.217
[57] Dog tends to be calm.	-.226	-.184	**-.362**	-.301	-.191
[252] Dog actively explores new environments.	-.354	-.073	**.358**	-.113	.035
[68] Dog seeks affection from family members.	.013	-.121	**.356**	-.054	.041

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 3: Activity/Excitability (*Continued*)					
[171] Dog works at task (e.g., getting treats out of Kongs, shredding toys) until entirely finished.	-.082	-.037	**.345**	-.133	.046
* [73] Dog is undemanding.	-.072	-.057	**-.343**	-.204	-.216
[285] Dog is playful with children.	-.259	-.247	**.343**	.065	-.228
[141] Dog likes to chase children who are running.	.067	.178	**.337**	.214	.070
[182] Dog is very excitable just before being taken for a walk.	.102	-.035	**.333**	.168	.189
[187] Dog is very excitable when around squirrels, birds, or other small animals.	.126	.046	**.329**	.176	.302
[254] Dog becomes agitated (e.g., whines, jumps up, tries to intervene) when owner shows affection for another dog or animal.	.179	.054	**.326**	.223	.266
* [319] Dog is aloof.	.112	.208	**-.322**	.011	.046
[161] Dog chews up and destroys most toys.	-.043	.036	**.321**	.256	.084
* [173] Dog seems subdued, depressed.	.290	.037	**-.304**	.166	.051
[72] Dog generally prefers to be with owner/family member (rather than alone).	.052	-.159	**.296**	-.052	.060
[258] Dog appears jealous when owner pays attention to another pet.	.196	.016	**.295**	.189	.288
[274] Dog is interested in scents.	-.052	-.045	**.293**	-.074	.106
* [315] Dog is aloof or indifferent towards familiar dogs.	.107	.066	**-.290**	-.067	.169
[130] Dog barks at outside noises when in the home.	.087	.137	**.288**	.020	.256
* [300] Dog is not demanding or needy (e.g., is content to play alone).	-.192	.027	**-.286**	-.216	-.207
[350] Dog loves to be praised.	-.061	-.139	**.281**	-.244	.039
[186] Dog is very excitable when around bicyclists, joggers, or motorcycles.	.165	.233	**.280**	.188	.234
[176] Dog is very excitable during car rides.	.159	.090	**.277**	.273	.153
[309] Dog hates to be left alone.	.251	-.069	**.277**	.236	.127
[59] Dog is restless.	.256	.110	**.277**	.274	.116
[253] Dog becomes agitated (e.g., whines, jumps up, tries to intervene) when owners show affection for another person.	.222	.144	**.267**	.234	.143
[275] Dog has a good sense of smell.	-.090	-.009	**.262**	-.207	.072
[347] Dog's behavior varies from situation to situation (e.g., dog is quiet when others are quiet but more excited when invited to play).	.068	-.020	**.260**	-.119	.060
[257] Dog appears jealous when owner pays attention to another person.	.190	.120	**.253**	.247	.177
* [316] Dog is aloof or indifferent towards unfamiliar dogs.	.108	.006	**-.252**	-.169	-.009
[144] Dog chases tail.	.060	.084	**.250**	.140	-.097
[260] Dog sniffs frequently on walks.	-.006	.004	**.232**	.157	.117
* [231] When walking on leash, dog tends to walk behind you.	.131	.025	**-.210**	.019	-.074
* [71] When resting, dog prefers to be alone, rather than in the company of family members.	.034	.136	**-.207**	.099	.040
[165] Dog buries or tries to bury favorite toys and other objects inside the home (e.g., under rugs, cushions, clothing).	.079	.094	**.178**	.119	.060
[70] Dog prefers to sleep near owner/family member.	.036	-.083	**.178**	-.079	.076

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 3: Activity/Excitability (*Continued*)					
[148] Dog is systematic in its behaviors (e.g., routinely sniffing perimeter of yard).	.016	.054	**.155**	.000	.145
[264] Dog mounts (or attempts to mount) non-family dogs.	-.020	.097	**.121**	.119	.095
Factor 4: Responsive to Training					
*[280] Dog is responsive to training, readily trained.	-.138	-.107	.270	**-.631**	-.020
*[353] Dog is easy to train.	-.108	-.077	.236	**-.616**	-.032
*[247] Dog learns readily.	-.149	-.057	.318	**-.607**	.048
*[278] Dog is willing and able to react to signals and cues from the handler.	-.077	-.119	.215	**-.596**	.016
[351] Dog is slow to respond to corrections.	.011	.115	-.093	**.589**	.040
*[145] Dog is good at tasks it has been trained to do.	-.134	-.089	.250	**-.583**	.026
[154] Dog ignores commands.	.051	.098	-.011	**.583**	.093
*[169] Dog is attentive to actions and words of its owner.	-.041	-.107	.216	**-.581**	.034
[250] Dog is slow to learn new tricks or tasks.	.119	.027	-.302	**.560**	-.067
*[277] Dog leaves food or objects alone when told to do so.	-.048	-.112	-.005	**-.552**	-.055
*[168] Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).	-.249	-.071	.011	**-.535**	-.089
*[276] When off leash, dog comes immediately when called.	-.049	-.093	-.011	**-.532**	-.070
*[159] Dog enjoys learning new things.	-.180	-.021	.436	**-.526**	-.005
*[155] Dog is eager to please handler.	-.003	-.135	.272	**-.518**	-.011
*[359] Dog is willing to complete work or task without a reward.	-.120	-.088	.107	**-.515**	-.104
*[251] Dog uses what it has learned.	-.171	-.092	.256	**-.499**	.025
*[170] Dog is able to focus on a task in the absence of distractions.	-.136	-.082	.115	**-.488**	-.004
*[146] Dog is able to control impulses (e.g., resists chasing a squirrel when told to sit or come).	-.095	-.090	-.096	**-.486**	-.170
[279] Dog is disobedient.	.046	.104	.094	**.478**	.124
*[153] Dog is in tune wishes or moods of owner even without being given direct commands.	-.086	-.061	.174	**-.475**	.022
*[152] Dog achieves tasks (e.g., fetches objects) quickly and easily.	-.098	-.014	.456	**-.470**	-.001
*[248] Dog is intelligent.	-.121	-.033	.327	**-.445**	.085
[162] Dog is destructive.	.086	.041	.230	**.437**	.050
[281] Dog is quick to sneak out through open doors, gates.	.034	.031	.166	**.431**	.091
[335] Dog resists getting off the couch or moving out of the way of others.	.020	.132	.014	**.393**	.066
*[272] Dog is able to concentrate when emotionally aroused (e.g., nervous, fearful).	-.313	-.062	.085	**-.388**	-.137
[163] Dog chews inappropriate objects.	.056	-.008	.196	**.370**	-.032
*[131] Dog is responsive to physical corrections.	-.095	-.181	.055	**-.365**	-.041
[167] Dog is easily distracted by interesting sights, sounds, and smells.	.162	-.001	.305	**.358**	.160
*[156] Dog is eager to earn rewards (e.g., in training).	-.067	-.080	.349	**-.358**	.099
*[297] After being excited, dog is quick to recover.	-.313	-.122	-.067	**-.356**	-.143

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 4: Responsive to Training (*Continued*)					
*[355] Dog is willing to let go of toy when playing (e.g., during tug of war).	-.036	-.118	-.103	**-.341**	-.105
[232] When walking on leash, dog tends to pull ahead.	.079	.005	.336	**.338**	.199
[337] Dog is dominant over owner.	.052	.240	.122	**.328**	.118
[160] Dog destroys household objects by chewing, digging, or scratching when owner is at home.	.070	.065	.169	**.326**	.012
[302] When alone or about to be left alone, dog is destructive (e.g., chews/scratches doors, floors, windows, curtains).	.126	.030	.186	**.324**	-.007
[256] Dog becomes agitated when another dog receives food or a treat.	.132	.041	.160	**.312**	.305
[235] Dog tries to steal food.	.036	-.019	.085	**.300**	.139
[236] Dog persistently begs for food.	.068	-.027	.146	**.287**	.203
*[149] Dog's behavior is predictable.	-.194	-.256	-.022	**-.281**	-.067
*[349] Dog was easily housetrained.	-.121	-.083	-.007	**-.279**	-.067
*[175] Dog is highly predictable.	-.211	-.256	-.049	**-.279**	-.075
*[147] Dog keeps living area clean (e.g., tends to eliminate in one area).	-.044	-.023	.026	**-.278**	-.016
*[150] Dog keeps itself clean.	-.061	.025	.069	**-.262**	.028
*[301] When alone or about to be left alone, dog whines, barks, or howls.	.198	.038	.212	**.262**	.041
[334] Dog is physically pushy with people.	-.018	.122	.245	**.261**	.214
[128] Dog barks excessively.	.146	.167	.227	**.259**	.198
*[151] Dog is orderly (e.g., tends to keep and play with toys in a specific area).	-.002	.038	-.122	**-.250**	-.040
[305] When left or about to be left alone, dog urinates or defecates.	.142	.054	.034	**.206**	.023
*[196] Dog exhibits less fear of objects (e.g., vacuums, brushes) or situations (e.g., pet stores, kennels) after repeated exposure to them.	-.044	.028	.047	**-.165**	-.025
[261] Dog mounts (or attempts to mount) adults.	.022	.088	.066	**.162**	.037
[263] Dog mounts (or attempts to mount) objects (e.g., toys, blankets, furniture).	.077	.112	.107	**.150**	-.004
[166] Dog digs holes in dirt or sand.	-.011	-.006	.139	**.139**	.025
[266] Dog relentlessly mounts other dogs.	-.022	.117	.102	**.137**	.097
[262] Dog mounts (or attempts to mount) children.	.053	.116	.059	**.127**	.076
*[355] Dog is willing to let go of toy when playing (e.g., during tug of war).	-.036	-.118	-.103	**-.341**	-.105
Factor 5: Aggression towards Animals					
* [313] Dog is friendly towards unfamiliar dogs.	-.201	-.199	.137	.091	**-.651**
[30] While on leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar large dogs.	.090	.255	.025	.039	**.651**
[32] While on leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar small dogs.	.096	.264	.026	.042	**.644**
* [318] Dog exhibits friendly behavior towards unfamiliar dogs (e.g., sniffing, tail wagging, licking).	-.170	-.182	.164	.091	**-.634**

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 5: Aggression towards Animals (*Continued*)					
[31] While off leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar large dogs.	.098	.255	-.013	.031	**.633**
[41] Dog has a tendency to attack (or attempt to attack) male dogs.	.014	.229	-.037	.016	**.626**
[42] Dog has a tendency to attack (or attempt to attack) female dogs.	.051	.200	-.027	.050	**.623**
[33] While off leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar small dogs.	.079	.260	-.012	.035	**.610**
[80] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when threatened by a dog (e.g., barked, growled, or lunged at).	-.017	.223	.006	-.015	**.594**
* [286] Dog is playful with unfamiliar dogs.	-.209	-.111	.290	.114	**-.590**
[34] While on leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards familiar large dogs.	.092	.265	-.004	.089	**.537**
[35] While off leash, dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards familiar small dogs.	.052	.280	-.042	.066	**.497**
[339] Dog is often dominant over other dogs.	-.192	.211	.054	.034	**.489**
[329] Dog is assertive or pushy with other dogs.	-.099	.195	.160	.125	**.483**
* [314] Dog is friendly towards familiar dogs.	-.118	-.168	.186	-.038	**-.468**
[28] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards unfamiliar cats.	.056	.186	.059	.084	**.467**
* [317] Dog exhibits friendly behavior towards familiar dogs (e.g., sniffing, tail wagging, licking).	-.126	-.102	.213	-.011	**-.465**
[81] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards other dogs in household.	.096	.172	-.027	.090	**.425**
[110] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards small wild animals (e.g., squirrels, birds).	.077	.157	.134	.099	**.424**
* [287] Dog is playful with familiar dogs.	-.141	-.031	.388	-.020	**-.423**
[104] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) when walking on a leash.	.115	.371	.028	.077	**.414**
[114] Dog guards food or treats from other dogs	.081	.107	.052	.106	**.405**
* [74] Dog willingly shares toys with other dogs.	-.065	-.131	.004	-.186	**-.397**
[330] Dog guards food and feeding area from other dogs.	.096	.113	.037	.077	**.393**
[338] Dog is dominant.	-.226	.226	.097	.058	**.367**
[29] Dog exhibits aggressive behavior (e.g., bares teeth, growls, lunges) towards familiar cats.	.059	.133	.033	.106	**.365**
[333] Dog exhibits assertive behaviors (e.g., stands erect, ears forward, direct stare, tail up) when greeting dogs.	-.124	.120	.089	.000	**.354**
[136] Dog likes to chase cats.	.049	.055	.195	.140	**.338**
[255] Dog tends to be jealous.	.195	.093	.305	.179	**.326**
[137] Dog likes to chase squirrels, birds, and other small animals.	.083	.036	.268	.112	**.316**
[143] Dog kills other animals (e.g., squirrels, rabbits, birds).	-.034	.057	.042	.041	**.308**
[129] Dog barks loudly and persistently when doorbell rings or mail is delivered.	.156	.185	.266	.028	**.305**
[142] Dog catches other animals (e.g., squirrels, rabbits, birds).	-.042	.054	.082	.050	**.290**
[120] Dog barks when an unfamiliar person (e.g., delivery person) approaches the home.	.118	.205	.238	-.106	**.286**

Appendix B. (*Continued*)

Item text	Loading of each item on each factor				
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Factor 5: Aggression towards Animals (*Continued*)					
[185] Dog is very excitable when it encounters cats.	.084	.043	.239	.219	**.279**
[24] When at home, dog barks at bicycles, children running, or joggers.	.166	.267	.199	.127	**.271**
[259] Dog frequently urine marks when outside.	-.043	.072	.096	.055	**.228**
[138] Dog likes to chase cars and motorcycles.	.058	.210	.160	.142	**.222**
[237] Dog is very food-motivated.	.004	-.068	.157	.041	**.213**
[238] Dog gulps food.	.037	-.033	.099	.194	**.197**
[189] Dog is very excitable at feeding times.	.084	-.051	.151	.163	**.190**
[125] Dog often requests to be fed.	.048	-.055	.107	.182	**.183**
* [265] Dog is mounted by other dogs (apart from appropriate mating).	.121	-.100	.089	.120	**-.155**
* [290] Dog enjoys playing alone.	-.066	.075	.106	-.100	**-.138**
[124] Dog displays a strong attachment to a particular member of the household.	.123	.035	.118	-.104	**.130**
[164] Dog buries or tries to bury objects (e.g., toys, bones) outside.	.050	.038	.086	.087	**.088**

Note. The highest factor loading of each trait is in boldface type. Items with an asterisk next to them load negatively on the factor under which they are listed. The numbers presented before each item are assigned solely for the purpose of identifying and specifying each item (e.g., in "Source item number(s)" column of Appendix C) and had no influence on the order in which the items were administered. The bolded horizontal lines that form the borders between some items demarcate the .3 and .4 cut-offs indicated as the cut-off for significant loadings by Floyd and Widaman (1995). [a] The item "Dog is anxious" was accidentally included twice in the questionnaire. Loadings for both instances appear here and are italicized, though the item is counted once towards the total of 352 items in the table.

Appendix C. Study 2: Item list resulting from item reduction phase

FACTOR 1: Fearfulness	**Loading direction**	**Source item number(s)[a]**
Facet 1: Fear of people (10 items)		
Dog behaves fearfully towards unfamiliar people.	+	(5, 7, 9, 10, 214, 345)
Dog behaves fearfully towards familiar people.	+	(6, 11, 12, 223)
Dog behaves fearfully towards children.	+	(13, 14, 207, 208)
Dog behaves fearfully in response to perceived threats from people (e.g., being cornered, having collar reached for).	+	(3, 4, 226)
Dog behaves fearfully when near crowds of people.	+	(211)
Dog is shy.	+	135
Dog is cautious, careful.	+	341
Dog behaves fearfully towards moving bicycles, joggers, skateboarders.	+	(225)
Dog is relaxed when greeting people.	-	213
Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting people.	+	(1, 2)
Facet 2: Non-social fear (8 items)		
Dog adapts easily to new situations and environments.	-	64, (65, 240, 271)
Dog is confident.	-	193
Dog is anxious	+	122, 270, (123, 174, 192, 194, 269)
Dog is quick to recover after being startled or frightened.	-	(195, 273, 298)
Dog is fearful of loud noises (e.g., heavy traffic, car horns, slamming doors, fireworks).	+	(219, 241, 244, 245)
Dog attempts to flee from novel objects or situations.	+	197, (209, 343)
Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a door frame).	+	133
Dog easily gets over unpleasant experiences (e.g., painful toe nail clippings).	-	239
Facet 3: Fear/submission towards dogs (5 items)		
Dog behaves fearfully towards other dogs.	+	(17, 18, 19, 20, 21, 22)
Dog behaves fearfully when threatened by other dogs (e.g., growled or lunged at, cornered).	+	(198, 227)
Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.	+	(332, 340)
Dog avoids other dogs.	+	331
Dog is bold.	-	134
Facet 4: Fear of Handling (6 items)		
Dog behaves fearfully when groomed (e.g., nails trimmed, brushed, bathed, ears cleaned).	+	(201, 202, 203, 204, 206)
Dog behaves fearfully during visits to the veterinarian.	+	(220, 221)
Dog exhibits fearful behaviors when restrained.	+	(205)
Dog is sensitive (and reactive) to pain.	+	(282, 283, 284)
Dog is easily upset when corrected, scolded, or punished.	+	(215, 352)
When alone or about to be left alone, dog shakes, shivers, or trembles.	+	307

Appendix C. (*Continued*)

FACTOR 2: Aggression towards People	Loading direction	Source item number(s)[a]
Facet 1: General Aggression (8 items)		
Dog behaves aggressively towards unfamiliar people.	+	(27, 37, 38, 43, 45, 47, 48, 105, 111, 112, 113)
Dog behaves aggressively towards familiar people.	+	(44, 46, 49, 50, 113)
Dog behaves aggressively when a person (e.g., visitor, delivery person) approaches the house or yard.	+	(77)
Dog is friendly towards unfamiliar people.	-	(212, 289, 320, 321, 322, 323, 324, 326, 328)
Dog behaves aggressively towards children.	+	(40, 97, 98, 99, 100)
Dog shows aggression when nervous or fearful.	+	267
Dog behaves aggressively towards moving bicycles, joggers, skateboarders.	+	(108)
Dog behaves aggressively towards people with unfamiliar objects (e.g., canes, wheelchairs, umbrellas).	+	(101)
Facet 2: Situational Aggression (7 items)		
Dog behaves aggressively in response to perceived threats from people (e.g., being cornered, having collar reached for).	+	(26, 36, 83)
Dog behaves aggressively during visits to the veterinarian.	+	(118, 119)
Dog displays aggression that is sudden and without apparent reason.	+	90, (127)
Dog behaves aggressively when restrained or handled (e.g., groomed).	+	(91, 92, 93, 94, 95, 96)
Dog aggressively guards coveted items (e.g., stolen item, treats, food bowl).	+	(115, 116, 117)
Dog behaves aggressively when scolded or punished.	+	(106)
Dog behaves aggressively if disturbed or moved when resting.	+	(78, 79)
FACTOR 3: Activity/Excitability		
Facet 1: Excitability (9 items)		
Dog becomes very excited when owner returns home.	+	(180, 308)
Dog is active (e.g., jumps, sniffs a lot) when walking on a leash.	+	51
Dog is very excitable when visitors arrive.	+	183, (179)
Dog is boisterous.	+	60
Dog is very excitable around other dogs.	+	(177, 178)
Dog seeks constant activity.	+	58
Dog is very excitable when around squirrels, birds, or other small animals.	+	187
Dog tends to be calm.	-	57
Dog is very excitable just before being taken for a walk.	+	182
Facet 2: Playfulness (7 items)		
Dog retrieves objects (e.g., balls, toys, sticks).	+	(228, 230)
Dog gets bored in play quickly.	-	295
Dog enjoys playing with toys.	+	291, (188, 229, 292, 358)
Dog is interested in playing tug-o-war with people or dogs.	+	(356, 357)
Dog is active during play with other dogs.	+	55
Dog is playful with familiar people.	+	288
Dog is aloof or indifferent towards other dogs.	-	(315, 316)

Appendix C. (*Continued*)

FACTOR 3: Activity/Excitability (*Continued*)	**Loading direction**	**Source item number(s)[a]**
Facet 3: Active engagement (6 items)		
Dog is very alert.	+	121
Dog is curious.	+	157, (242, 252, 346)
Dog is lethargic	-	61, (52, 53, 56, 62)
Dog works at tasks (e.g., getting treats out of a Kong, shredding toys) until entirely finished.	+	171
Dog is very watchful (e.g., monitors for squirrels, attends to noises).	+	342
Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden.	+	361
Facet 4: Companionability (5 items)		
Dog is affectionate.	+	(66, 67, 68)
Dog follows people around.	+	233, (69)
Dog seeks companionship from people.	+	311, (70, 71, 72)
Dog loves to be praised.	+	350
Dog is aloof.	-	319
FACTOR 4: Responsiveness to Training		
Facet 1: Trainability (7 items)		
Dog is slow to respond to corrections.	-	351, (131)
Dog is attentive to owner's actions and words.	+	(153, 155, 169, 278)
Dog is willing to complete work, task, or training without a reward.	+	(359)
Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).	+	168, (170)
Dog ignores commands.	-	154
Dog is slow to learn new tricks or tasks.	-	250, (145, 247, 251, 280, 353)
Dog is intelligent.	+	248
Facet 2: Controllability (7 items)		
Dog is destructive.	-	162, (160, 163)
When off leash, dog comes immediately when called.	+	276, (146, 279)
Dog leaves food or objects alone when told to do so.	+	277, (235)
Dog is quick to sneak out through open doors, gates.	-	281, (335)
Dog is willing to let go of toys when playing (e.g., during tug-o-war).	+	355
When walking on leash, dog tends to pull ahead.	-	232
Dog is dominant over owner.	-	337
FACTOR 5: Aggression towards Animals		
Facet 1: Aggression towards dogs (5 items)		
Dog behaves aggressively toward dogs.	+	(30, 31, 32, 33, 34, 35)
Dog responds aggressively when threatened by another dog (e.g., growled or lunged at, cornered).	+	(80)
Dog has a tendency to attack (or attempt to attack) other dogs.	+	(41, 42)
Dog is friendly towards other dogs.	-	(313, 314, 317, 318)
Dog is playful with other dogs.	-	(286, 287)

Appendix C. (*Continued*)

FACTOR 5: Aggression towards Animals (*Continued*)	**Loading direction**	**Source item number(s)[a]**
Facet 2: Prey drive (6 items)		
Dog behaves aggressively towards cats.	+	(28, 29)
Dog likes to chase cats.	+	136
Dog likes to chase squirrels, birds, or other small animals.	+	137
Dog catches and kills other animals (e.g., squirrels, rabbits).	+	110, (142, 143)
Dog likes to chase bicycles, joggers, and skateboarders.	+	139
Dog is very excitable around cats.	+	(185)
Facet 3: Dominance over dogs (6 items)		
Dog guards food or treats from other dogs.	+	(114, 330)
Dog willingly shares toys with other dogs.	-	74
Dog is dominant over other dogs.	+	339
Dog is assertive or pushy with other dogs (e.g., if in a home with other dogs, when greeting).	+	(81, 329)
Dog frequently urine marks.	+	(259)
Dog jumps up on (e.g., in play) and/or mounts other dogs (outside appropriate mating).	+	(264, 266)

Notes. [a]Item numbers correspond to numbers assigned to each item in Studies 1 and 2 for purpose of identifying each item; see Appendix B to locate the item associated with each number. Item numbers not in parentheses indicate that the new item (in the questionnaire derived in Study 2) is an exact quotation of that item. Item numbers in parentheses indicate that the new item is used to address the content of all of the listed items, is a composite of those items, or is a modified version of the item.

Appendix D. Study 3: CFA model item loadings on five dog personality factors

Factor 1: Fearfulness

Item	Loading
1	1.382
2	**.467**
3	.906
4	1.286
5	1.428
6	1.461
7	.968
8	.961
9*	.907
10	.594
11*	1.099
12*	1.180
13	1.302
14*	.830
15	1.136
16	1.164
17	.937
18*	.525
19	.819
20	.974
21	**.426**
22	.577
23*	.857
24	1.000
25	1.145
26	.836
27	**.495**
28	.727
29	**.325**

Appendix D. (*Continued*)

Factor 2: Aggression towards People

Item	Loading
30	1.108
31	**.338**
32	1.230
33*	1.034
34	.733
35	1.264
36	.884
37	.819
38	1.000
39	.591
40	**.448**
41	**.532**
42	.622
43	.345
44	.444

Factor 3: Activity/Excitability

Item	Loading
45	**.443**
46	.761
47	.769
48	1.308
49	.931
50	1.360
51	.615
52*	.836
53	**.487**
54	1.444
55*	1.045
56	1.261
57	1.434
58	.820
59	.627
60*	.686
61	**.499**
62	.506
63*	.794
64	.879
65	.606
66	.803

Appendix D. (*Continued*)

Factor 3: Activity/Excitability (*continued*)

Item	Loading
67	1.000
68	.627
69	**.472**
70	**.323**
71*	.754

Factor 4: Responsiveness to Training

Item	Loading
72	1.119
73*	.655
74*	.902
75*	1.063
76	1.359
77	.693
78*	**.313**
79	1.000
80*	1.423
81*	.861
82	1.060
83*	.549
84	.834
85	.627

Appendix D. (*Continued*)

Factor 5: Aggression towards Animals

Item	Loading
86	1.689
87	1.538
88	1.412
89*	1.331
90*	.827
91	.792
92	.639
93	.543
94	.672
95	**.475**
96	.552
97	1.000
98*	.839
99	1.108
100	1.089
101	**.457**
102	**.330**

Note**.** Factors 1 and 2, 2 and 5 are correlated in this model. Items with loadings ≤ .500 are bolded; these items were considered for elimination form the questionnaire based on their loadings. Item numbers are from items as they were numbered in Study 3. Reverse-coded items were rekeyed before the model was fit so that all items should load positively in the model; those items that are reverse coded and were rekeyed are indicated by an asterisk after the item number.

Appendix E. Study 3: Convergent validity – Correlations of items that load on the same factors and facets

Factor 1: Fearfulness

	Facet 1: Fear of people										**Facet 2: Non-social fear**								**Facet 3: Fear/ submission to dogs**					**Facet 4: Fear of handling**					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	***9**	**10**	***11**	***12**	**13**	***14**	**15**	**16**	**17**	***18**	**19**	**20**	**21**	**22**	***23**	**24**	**25**	**26**	**27**	**28**	**29**
Facet 1: Fear of people																													
1	**1**																												
2	**.425**	**1**																											
3	**.496**	**.287**	**1**																										
4	**.529**	**.346**	**.401**	**1**																									
5	**.667**	**.356**	**.457**	**.515**	**1**																								
6	**.642**	**.365**	**.398**	**.481**	**.626**	**1**																							
7	**.374**	**.208**	**.216**	**.345**	**.373**	**.461**	**1**																						
8	**.467**	**.269**	**.390**	**.395**	**.497**	**.408**	**.264**	**1**																					
***9**	**.387**	**.228**	**.259**	**.290**	**.368**	**.270**	**.102**	**.283**	**1**																				
10	**.236**	**.170**	**.134**	**.206**	**.219**	**.292**	**.173**	**.179**	**.087**	**1**																			
Facet 2: Non-social fear																													
***11**	.471	.302	.312	.369	.523	.472	.272	.371	.355	.160	**1**																		
***12**	.514	.342	.350	.434	.543	.592	.335	.391	.323	.251	**.555**	**1**																	
13	.412	.289	.303	.398	.468	.458	.308	.316	.348	.201	**.509**	**.529**	**1**																
***14**	.363	.240	.271	.347	.436	.380	.261	.305	.244	.141	**.468**	**.492**	**.425**	**1**															
15	.320	.194	.214	.314	.403	.375	.294	.347	.144	.183	**.338**	**.347**	**.347**	**.384**	**1**														
16	.482	.303	.321	.435	.537	.516	.340	.423	.292	.250	**.469**	**.507**	**.418**	**.403**	**.404**	**1**													
17	.338	.242	.237	.358	.373	.355	.310	.314	.198	.174	**.336**	**.361**	**.356**	**.346**	**.398**	**.437**	**1**												
***18**	.196	.144	.144	.205	.208	.173	.116	.180	.104	.060	**.235**	**.231**	**.178**	**.328**	**.143**	**.181**	**.198**	**1**											
Facet 3: Fear/submission to dogs																													
19	.387	.252	.315	.334	.389	.390	.225	.339	.218	.192	.321	.446	.347	.281	.233	.341	.244	.186	**1**										
20	.314	.209	.261	.303	.314	.345	.213	.289	.156	.220	.252	.399	.301	.300	.279	.334	.226	.172	**.517**	**1**									
21	.150	.109	.094	.128	.145	.230	.135	.142	.089	.403	.099	.234	.112	.116	.116	.184	.122	.027	**.221**	**.269**	**1**								
22	.241	.172	.237	.192	.254	.284	.182	.176	.090	.112	.244	.279	.244	.213	.152	.195	.152	.128	**.465**	**.278**	**.076**	**1**							
***23**	.296	.179	.189	.239	.323	.457	.300	.225	.072	.216	.285	.539	.259	.289	.263	.356	.208	.126	**.225**	**.281**	**.213**	**.151**	**1**						
Facet 4: Fear of handling																													
24	.239	.171	.185	.313	.261	.217	.187	.247	.164	.135	.264	.255	.281	.268	.238	.259	.240	.184	.223	.223	.089	.147	.129	**1**					
25	.397	.225	.328	.372	.395	.376	.273	.283	.216	.161	.350	.339	.330	.317	.300	.312	.254	.178	.283	.252	.083	.218	.209	**.429**	**1**				
26	.351	.272	.312	.449	.378	.323	.240	.324	.261	.180	.363	.354	.380	.324	.272	.365	.308	.185	.297	.247	.106	.182	.130	**.323**	**.358**	**1**			
27	.116	.081	.108	.190	.130	.140	.221	.137	.027	.100	.098	.133	.150	.144	.203	.150	.195	.120	.148	.193	.071	.108	.126	**.241**	**.211**	**.169**	**1**		
28	.180	.121	.136	.279	.220	.220	.218	.164	.087	.153	.216	.261	.254	.251	.232	.245	.221	.133	.176	.202	.117	.141	.160	**.215**	**.197**	**.255**	**.217**	**1**	
29	.137	.127	.141	.185	.206	.151	.117	.173	.144	.123	.208	.202	.282	.184	.157	.208	.187	.114	.158	.126	.084	.111	.069	**.145**	**.177**	**.220**	**.068**	**.142**	**1**

Appendix E. (*Continued*)

Factor 2: Aggression towards People

	Facet 1: General aggression towards people								**Facet 2: Situational aggression towards people**								
	30	**31**	**32**	***33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**
Facet 1: General aggression towards people																	
30	**1**																
31	**.299**	**1**															
32	**.548**	**.194**	**1**														
***33**	**.593**	**.211**	**.456**	**1**													
34	**.446**	**.262**	**.321**	**.379**	**1**												
35	**.487**	**.309**	**.384**	**.360**	**.414**	**1**											
36	**.414**	**.251**	**.373**	**.275**	**.334**	**.383**	**1**										
37	**.472**	**.228**	**.366**	**.338**	**.330**	**.443**	**.439**	**1**									
Facet 2: Situational aggression towards people																	
38	.479	.286	.361	.382	.388	.545	.328	.383	**1**								
39	.373	.266	.238	.290	.307	.421	.272	.284	**.426**	**1**							
40	.291	.258	.206	.180	.260	.382	.249	.235	**.279**	**.211**	**1**						
41	.236	.231	.182	.164	.246	.373	.258	.213	**.393**	**.419**	**.234**	**1**					
42	.172	.152	.176	.087	.185	.270	.212	.161	**.241**	**.162**	**.206**	**.204**	**1**				
43	.195	.319	.131	.093	.188	.295	.197	.157	**.341**	**.198**	**.245**	**.311**	**.249**	**1**			
44	.171	.220	.136	.099	.166	.310	.176	.152	**.285**	**.232**	**.250**	**.364**	**.313**	**.310**	**1**		
45	.056	.012	.049	.020	-.006	-.003	.016	.017	**.030**	**-.003**	**-.018**	**.005**	**.040**	**.006**	**.024**	**1**	
46	.051	.045	.065	.029	.055	.066	.107	.111	**.065**	**.027**	**.029**	**.084**	**.105**	**.052**	**.054**	**.154**	**1**

Appendix E. (*Continued*)

Factor 3: Activity/Excitability

	Facet 1: Excitability							**Facet 2: Playfulness**							**Facet 3: Active engagement**						**Facet 4: Companionability**				
	47	**48**	**49**	**50**	**51**	***52**	**53**	**54**	***55**	**56**	**57**	**58**	**59**	***60**	**61**	**62**	***63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**
Facet 1: Excitability																									
47	**1**																								
48	**.316**	**1**																							
49	**.394**	**.328**	**1**																						
50	**.189**	**.381**	**.265**	**1**																					
51	**.238**	**.203**	**.301**	**.175**	**1**																				
***52**	**.292**	**.421**	**.302**	**.471**	**.180**	**1**																			
53	**.336**	**.198**	**.223**	**.128**	**.213**	**.173**	**1**																		
Facet 2: Playfulness																									
54	.110	.171	.050	.279	.045	.116	.016	**1**																	
***55**	.041	.110	.023	.145	-.043	.061	-.020	**.378**	**1**																
56	.138	.232	.129	.281	.068	.141	.025	**.565**	**.372**	**1**															
57	.160	.240	.177	.258	.090	.146	.072	**.417**	**.254**	**.544**	**1**														
58	.065	.204	.220	.224	.065	.081	.020	**.180**	**.207**	**.271**	**.240**	**1**													
59	.176	.219	.141	.182	.060	.063	.032	**.249**	**.208**	**.361**	**.273**	**.264**	**1**												
***60**	.157	.161	.399	.063	.076	.110	.061	**.033**	**.121**	**.126**	**.115**	**.317**	**.134**	**1**											
Facet 3: Active engagement																									
61	.111	.179	.064	.252	.117	.117	.081	.163	.143	.159	.147	.199	.162	.066	**1**										
62	.131	.219	.117	.214	.137	.082	.042	.156	.093	.219	.185	.219	.240	.101	**.303**	**1**									
***63**	.067	.195	.074	.305	.057	.221	.000	.264	.294	.280	.176	.257	.208	.141	**.326**	**.227**	**1**								
64	.095	.151	.068	.190	.078	.093	.072	.288	.213	.288	.228	.141	.167	.052	**.155**	**.199**	**.165**	**1**							
65	.144	.158	.142	.192	.366	.120	.121	.090	.044	.105	.093	.088	.092	.021	**.410**	**.227**	**.146**	**.119**	**1**						
66	.084	.140	.043	.158	.074	.037	.062	.399	.265	.339	.241	.143	.194	.052	**.200**	**.248**	**.226**	**.439**	**.149**	**1**					
Facet 4: Companionability																									
67	.077	.078	.041	.034	-.034	-.066	.069	.089	.128	.137	.098	.143	.277	.097	.106	.155	.087	.050	.050	.103	**1**				
68	.136	.112	.081	.142	.016	.054	.068	.119	.071	.131	.120	.072	.174	.039	.092	.129	.068	.106	.082	.092	**.278**	**1**			
69	.136	.111	.074	.084	-.009	-.004	.067	.126	.119	.196	.132	.126	.264	.076	.088	.171	.076	.118	.043	.128	**.440**	**.410**	**1**		
70	.039	.051	.019	.036	-.032	-.071	.093	.108	.156	.132	.083	.117	.216	.064	.132	.155	.099	.131	.066	.139	**.426**	**.244**	**.289**	**1**	
71	-.069	-.112	-.070	-.066	.025	-.053	.009	-.119	-.202	-.158	-.084	-.144	-.210	-.294	-.018	-.099	-.202	-.078	.070	-.106	**-.313**	**-.174**	**-.311**	**-.184**	**1**

Appendix E. (*Continued*)

Factor 4: Responsiveness to Training

	Facet 1: Trainability							**Facet 2: Unruliness**						
	***72**	**73**	**74**	**75**	***76**	***77**	**78**	***79**	**80**	**81**	***82**	**83**	***84**	***85**
Facet 1: Trainability														
***72**	**1**													
73	**.433**	**1**												
74	**.364**	**.388**	**1**											
75	**.322**	**.349**	**.321**	**1**										
***76**	**.525**	**.464**	**.402**	**.387**	**1**									
***77**	**.374**	**.297**	**.295**	**.277**	**.339**	**1**								
78	**.249**	**.278**	**.242**	**.246**	**.221**	**.524**	**1**							
Facet 2: Unruliness														
***79**	.217	.232	.165	.171	.261	.061	.066	**1**						
80	.402	.441	.340	.420	.563	.251	.178	**.216**	**1**					
81	.359	.339	.326	.286	.344	.236	.198	**.215**	**.325**	**1**				
***82**	.264	.245	.214	.218	.367	.127	.057	**.307**	**.424**	**.272**	**1**			
83	.202	.193	.188	.199	.187	.093	.069	**.134**	**.170**	**.289**	**.158**	**1**		
***84**	.233	.198	.157	.247	.290	.052	.065	**.176**	**.283**	**.193**	**.278**	**.145**	**1**	
***85**	.338	.273	.224	.201	.343	.157	.109	**.192**	**.270**	**.271**	**.213**	**.200**	**.217**	**1**

Appendix E. (*Continued*)

Factor 5: Aggression towards Animals																	
	Facet 1: Aggression towards dogs					**Facet 2: Prey drive**						**Facet 3: Dominance over dogs**					
	86	**87**	**88**	***89**	***90**	**91**	**92**	**93**	**94**	**95**	**96**	**97**	***98**	**99**	**100**	**101**	**102**
Facet 1: Aggression towards dogs																	
86	**1**																
87	**.617**	**1**															
88	**.773**	**.553**	**1**														
***89**	**.706**	**.508**	**.664**	**1**													
***90**	**.443**	**.299**	**.411**	**.662**	**1**												
Facet 2: Prey drive																	
91	.271	.234	.258	.200	.104	**1**											
92	.194	.189	.183	.113	-.003	**.646**	**1**										
93	.176	.180	.160	.102	.001	**.372**	**.525**	**1**									
94	.215	.217	.211	.167	.101	**.389**	**.374**	**.438**	**1**								
95	.182	.151	.184	.119	.003	**.190**	**.274**	**.263**	**.152**	**1**							
96	.161	.144	.163	.089	-.009	**.647**	**.726**	**.454**	**.297**	**.239**	**1**						
Facet 3: Dominance over dogs																	
97	.272	.324	.254	.263	.189	.175	.136	.140	.169	.138	.110	**1**					
***98**	.318	.305	.282	.358	.315	.169	.117	.125	.189	.118	.088	**.509**	**1**				
99	.439	.460	.362	.324	.172	.184	.158	.138	.230	.078	.107	**.320**	**.308**	**1**			
100	.424	.402	.369	.306	.122	.200	.202	.164	.184	.160	.176	**.306**	**.287**	**.561**	**1**		
101	.148	.176	.097	.114	.125	.115	.104	.087	.134	.077	.071	**.127**	**.124**	**.190**	**.122**	**1**	
102	.107	.097	.096	-.007	-.143	.111	.170	.155	.126	.157	.156	**.123**	**.060**	**.252**	**.269**	**.165**	**1**

Appendix F. Study 3: Discriminant validity – Correlations of items that load on different factors and facets Fearfulness and Aggression towards People

Factor 1: Fearfulness

	Facet 1: Fear of people										**Facet 2: Non-social fear**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	***9**	**10**	***11**	***12**	**13**	***14**	**15**	**16**	**17**	***18**
Factor 2: Aggression towards People																		
Facet 1: General aggression towards people																		
30	.428	.230	.320	.322	.341	.223	.198	.263	.355	.010	.280	.208	.262	.212	.124	.218	.185	.114
31	.199	.248	.159	.208	.180	.104	.084	.152	.172	.029	.127	.135	.165	.115	.067	.127	.120	.108
32	.323	.179	.226	.240	.258	.177	.178	.206	.302	.014	.236	.164	.223	.155	.142	.179	.163	.108
***33**	.640	.336	.426	.401	.530	.461	.311	.349	.409	.073	.419	.382	.332	.299	.189	.339	.221	.154
34	.326	.197	.596	.287	.313	.210	.134	.259	.263	.046	.251	.224	.234	.200	.126	.212	.188	.111
35	.354	.199	.318	.370	.316	.205	.171	.290	.292	.053	.312	.273	.331	.256	.157	.226	.215	.198
36	.219	.127	.218	.219	.225	.094	.093	.337	.252	-.001	.200	.120	.200	.150	.110	.179	.142	.091
37	.339	.186	.262	.277	.281	.175	.161	.293	.279	.039	.249	.193	.243	.206	.161	.234	.242	.155
Facet 2: Situational aggression towards people																		
38	.323	.234	.270	.432	.274	.182	.167	.247	.250	.036	.245	.197	.242	.184	.137	.177	.206	.153
39	.245	.201	.214	.251	.232	.113	.110	.194	.208	.031	.208	.139	.185	.163	.072	.139	.170	.170
40	.163	.169	.187	.172	.171	.079	.073	.149	.149	.019	.184	.133	.181	.145	.115	.114	.103	.141
41	.168	.162	.154	.217	.170	.076	.093	.201	.167	.053	.203	.116	.186	.168	.120	.135	.165	.146
42	.115	.112	.112	.153	.121	.045	.049	.124	.108	.029	.134	.027	.150	.098	.097	.102	.112	.122
43	.110	.131	.089	.173	.125	.037	.041	.128	.126	.033	.137	.084	.149	.123	.075	.098	.108	.124
44	.091	.115	.104	.158	.106	.063	.077	.144	.115	.019	.131	.068	.139	.126	.121	.109	.132	.121
45	.029	.027	.002	.059	.039	.037	.067	.044	.082	.038	.023	-.008	.070	.003	.089	.063	.074	-.021
46	.046	.000	.032	.103	.088	.006	-.050	.053	.191	.019	.035	.010	.119	.028	.052	.085	.090	-.017

Appendix F. (*Continued*)
Fearfulness and Aggression towards People

Factor 1: Fearfulness											
	Facet 3: Fear/submission towards dogs					**Facet 4: Fear during handling**					
	19	**20**	**21**	**22**	***23**	**24**	**25**	**26**	**27**	**28**	**29**
Factor 2: Aggression towards People											
Facet 1: General aggression towards people											
30	.196	.096	-.034	.162	-.006	.154	.256	.260	.050	.102	.071
31	.125	.101	.029	.086	-.001	.108	.149	.154	.098	.073	.090
32	.186	.101	-.031	.143	-.035	.142	.252	.215	.058	.088	.053
***33**	.254	.157	.023	.217	.161	.147	.346	.287	.062	.107	.102
34	.203	.147	.012	.171	.032	.146	.229	.282	.086	.082	.095
35	.351	.217	-.025	.193	.008	.237	.268	.364	.134	.124	.104
36	.167	.082	-.033	.074	-.069	.164	.190	.221	.091	.074	.082
37	.216	.143	.015	.100	-.017	.189	.222	.252	.092	.096	.114
Facet 2: Situational aggression towards people											
38	.217	.100	-.001	.171	-.018	.218	.236	.326	.138	.110	.091
39	.155	.104	-.024	.127	-.015	.237	.332	.290	.116	.067	.089
40	.164	.077	-.016	.122	-.045	.125	.104	.185	.068	.038	.086
41	.165	.113	.028	.102	-.023	.366	.219	.329	.159	.095	.108
42	.120	.066	.003	.086	-.086	.149	.089	.167	.113	.074	.074
43	.106	.060	-.003	.085	-.059	.140	.106	.189	.077	.090	.111
44	.116	.086	-.011	.111	-.048	.184	.136	.188	.109	.082	.057
45	.019	.052	.048	.007	-.034	.037	.056	.034	.081	.102	.087
46	.043	.075	.050	-.047	-.124	.109	.090	.083	.086	.051	.047

Appendix F. (*Continued*)
Fearfulness and Activity/Excitability

Factor 1: Fearfulness

	Facet 1: Fear of people										**Facet 2: Non-social fear**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	***9**	**10**	***11**	***12**	**13**	***14**	**15**	**16**	**17**	***18**
Factor 3: Activity/Excitability																		
Facet 1: Excitability																		
47	-.020	-.034	-.016	.037	-.013	-.073	-.083	.040	.268	-.017	.022	-.014	.080	.003	.051	.022	.060	.007
48	-.020	-.034	-.011	-.040	-.054	-.140	-.193	-.022	.208	-.049	-.009	-.079	.054	-.042	-.057	-.026	-.022	-.050
49	.035	.007	.015	.070	.046	-.040	-.076	.078	.264	.053	.039	.019	.133	.026	.050	.078	.087	.010
50	.022	.000	.033	.016	.007	-.052	-.088	.021	.136	.032	-.017	-.042	.122	.010	-.010	.020	.035	-.021
51	.020	.013	.033	.081	.036	-.001	-.019	.039	.122	.027	.023	.019	.110	.011	.051	.039	.078	-.007
***52**	.146	.093	.131	.141	.143	.044	-.125	.140	.402	.026	.216	.182	.296	.144	.061	.127	.113	.047
53	.025	.031	.005	.039	.008	.012	-.018	.034	.127	.023	.049	.013	.120	.024	.080	.063	.067	-.011
Facet 2: Playfulness																		
54	.010	-.016	-.020	-.066	-.064	-.061	-.040	-.026	.039	-.012	-.072	-.085	-.025	-.080	-.073	-.046	-.021	-.047
***55**	-.029	-.052	-.054	-.084	-.079	-.079	-.082	-.055	.028	-.049	-.067	-.101	-.103	-.101	-.091	-.097	-.089	-.065
56	-.015	-.032	-.030	-.069	-.067	-.099	-.068	-.028	.061	-.029	-.072	-.119	-.061	-.122	-.082	-.034	-.038	-.065
57	-.054	-.061	-.075	-.055	-.086	-.116	-.101	-.045	.084	-.035	-.068	-.133	-.050	-.115	-.058	-.028	-.015	-.069
58	-.076	-.076	-.106	-.089	-.114	-.162	-.118	-.091	-.024	-.013	-.164	-.192	-.146	-.154	-.130	-.078	-.105	-.095
59	-.144	-.184	-.144	-.126	-.146	-.161	-.088	-.074	-.067	-.040	-.175	-.191	-.114	-.171	-.073	-.083	-.052	-.092
***60**	-.036	-.048	-.060	-.054	-.037	-.114	-.134	-.026	.076	-.025	-.059	-.061	-.089	-.063	-.054	-.027	-.036	-.064
Facet 3: Active engagement																		
61	.007	-.029	.009	.029	-.026	-.028	.019	-.036	.039	-.054	-.084	-.160	.004	-.114	-.029	-.057	-.042	-.049
62	-.148	-.131	-.100	-.112	-.168	-.200	-.127	-.142	-.048	-.127	-.232	-.301	-.113	-.255	-.146	-.184	-.144	-.151
***63**	-.019	-.032	-.008	-.056	-.053	-.122	-.082	-.028	.038	-.093	-.078	-.125	-.087	-.133	-.103	-.088	-.101	-.073
64	-.078	-.042	-.072	-.054	-.105	-.115	-.086	-.073	-.005	-.040	-.096	-.135	-.028	-.093	-.074	-.075	-.051	-.045
65	.089	.034	.072	.074	.064	.034	.130	.034	.102	.018	.016	-.053	.091	-.017	.063	.018	.053	-.017
66	-.046	-.044	-.036	-.059	-.122	-.108	-.065	-.095	-.026	-.057	-.131	-.138	-.063	-.146	-.101	-.100	-.084	-.073
Facet 4: Companionability																		
67	-.105	-.093	-.115	-.112	-.131	-.125	-.047	-.047	-.109	.010	-.145	-.121	-.115	-.127	-.013	-.043	-.044	-.109
68	-.050	-.050	-.090	-.063	-.045	-.035	-.015	-.027	-.026	.050	-.055	-.028	.014	-.055	.008	-.008	.000	-.069
69	-.173	-.164	-.173	-.124	-.161	-.186	-.072	-.078	-.101	-.016	-.166	-.140	-.104	-.146	-.027	-.096	-.043	-.072
70	-.083	-.065	-.105	-.047	-.089	-.065	.003	-.073	-.114	.030	-.100	-.104	-.047	-.120	-.013	-.057	-.024	-.093
71	.134	.116	.118	.126	.142	.197	.179	.054	.062	.044	.110	.103	.147	.130	.080	.088	.096	.082

Appendix F. (*Continued*)
Fearfulness and Activity/Excitability

	Facet 3: Fear/submission towards dogs					Facet 4: Fear during handling					
	19	**20**	**21**	**22**	***23**	**24**	**25**	**26**	**27**	**28**	**29**
Factor 1: Fearfulness											
Factor 3: Activity/Excitability											
Facet 1: Excitability											
47	.056	.085	.051	-.015	-.121	.081	.038	.068	.035	.075	.033
48	.009	.005	.031	-.075	-.306	.012	-.054	.032	-.025	-.034	.029
49	.072	.063	.078	-.188	-.150	.092	-.001	.126	.015	.061	.042
50	.025	.049	.072	-.017	-.223	.013	-.032	.105	-.021	-.002	.057
51	.057	.073	.021	.010	-.121	.086	.060	.091	.081	.082	.041
***52**	.159	.123	.068	.039	-.119	.125	.102	.225	-.010	.034	.094
53	.025	.038	.004	.006	-.087	.094	.072	.083	.068	.079	.019
Facet 2: Playfulness											
54	.006	.008	.067	.001	-.060	-.063	-.014	-.049	-.035	-.030	-.019
***55**	-.050	-.034	.011	-.082	-.054	-.123	-.093	-.085	-.070	-.059	-.072
56	-.023	-.025	.090	-.102	-.124	-.079	-.071	-.031	-.031	-.019	-.022
57	-.028	-.003	.049	-.065	-.192	-.036	-.035	-.046	-.017	-.033	.031
58	-.243	-.133	.045	-.389	-.141	-.115	-.119	-.084	-.087	-.069	-.045
59	-.058	-.043	.039	-.112	-.139	-.099	-.129	-.063	-.047	-.036	-.050
***60**	-.099	-.048	-.021	-.413	-.059	-.034	-.090	-.045	-.036	-.070	-.072
Facet 3: Active engagement											
61	-.028	-.020	-.026	-.070	-.199	-.026	.002	-.015	.058	-.013	-.040
62	-.113	-.089	-.039	-.134	-.316	-.112	-.138	-.105	-.034	-.066	-.040
***63**	-.055	-.042	-.022	-.095	-.137	-.106	-.095	-.040	-.066	-.090	-.049
64	-.062	-.052	.016	-.053	-.138	-.079	-.075	-.075	-.064	-.023	-.062
65	.053	.017	-.020	.014	-.176	.047	.077	.064	.055	.067	.012
66	-.034	-.039	-.004	-.014	-.125	-.108	-.071	-.075	-.047	-.029	-.093
Facet 4: Companionability											
67	-.090	-.055	.026	-.088	-.039	-.098	-.103	-.096	-.004	.025	-.014
68	.011	.056	.065	-.031	-.041	-.019	-.063	-.016	.044	.090	.075
69	-.062	-.031	.006	-.106	-.077	-.033	-.134	-.095	.007	.052	-.001
70	-.050	-.022	.006	-.060	-.013	-.056	-.066	-.073	.038	.112	-.056
71	.087	.029	.003	.159	.025	.072	.111	.104	.035	.032	.065

Appendix F. (*Continued*)
Fearfulness and Responsiveness to Training

Factor 1: Fearfulness

	Facet 1: Fear of people										**Facet 2: Non-social fear**								**Facet 3: Fear/submission towards dogs**				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	***9**	**10**	***11**	***12**	**13**	***14**	**15**	**16**	**17**	***18**	**19**	**20**	**21**	**22**	***23**
Responsiveness to Training																							
Facet 1: Trainability																							
***72**	-.042	-.046	-.055	-.069	-.101	-.034	.041	-.105	-.153	-.006	-.106	-.086	-.110	-.122	-.068	-.113	-.089	-.076	-.074	-.055	.008	-.026	.040
73	-.027	-.036	-.023	-.041	-.062	-.004	.083	-.063	-.144	-.034	-.101	-.095	-.062	-.139	-.051	-.082	-.053	-.095	-.039	-.034	-.027	.031	.010
74	-.061	-.019	-.039	-.115	-.096	-.060	-.005	-.106	-.103	-.046	-.130	-.102	-.084	-.132	-.101	-.136	-.103	-.103	-.070	-.098	-.051	-.023	-.027
75	-.155	-.078	-.086	-.163	-.241	-.155	-.053	-.198	-.303	-.110	-.295	-.270	-.244	-.258	-.204	-.242	-.190	-.112	-.153	-.146	-.090	-.002	-.124
***76**	-.030	-.055	-.030	-.086	-.102	-.039	.062	-.091	-.169	-.018	-.090	-.075	-.120	-.107	-.072	-.109	-.080	-.070	-.090	-.065	-.015	.014	.039
***77**	-.063	-.048	-.035	-.077	-.131	-.126	-.050	-.075	-.067	-.125	-.149	-.172	-.121	-.148	-.121	-.171	-.126	-.057	-.061	-.077	-.065	-.062	-.133
78	-.071	-.102	-.053	-.093	-.098	-.103	-.001	-.097	-.131	-.118	-.176	-.247	-.125	-.150	-.103	-.145	-.118	-.075	-.061	-.089	-.087	-.033	-.185
Facet 2: Controllability																							
***79**	-.008	-.038	-.001	-.042	-.035	.022	.108	-.050	-.148	-.043	-.045	-.020	-.093	-.037	.025	-.051	-.036	-.020	-.017	-.037	-.093	.059	.091
80	-.001	-.012	-.005	-.053	-.063	-.006	.058	-.073	-.165	-.013	-.100	-.069	-.076	-.093	-.060	-.099	-.068	-.059	-.055	-.069	-.014	.057	.014
81	-.047	-.035	-.037	-.065	-.068	-.003	.075	-.107	-.137	-.018	-.116	-.060	-.088	-.111	-.058	-.083	-.071	-.072	-.046	-.061	-.024	.002	.047
***82**	.029	-.008	.009	-.060	-.021	.033	.086	-.046	-.111	-.032	-.017	.021	-.074	-.040	-.047	-.048	-.057	.017	-.016	-.046	-.054	.044	.129
83	-.029	-.009	-.012	-.031	-.039	.012	.035	-.093	-.147	-.013	-.076	.004	-.046	-.076	-.050	-.050	-.070	-.023	-.041	-.048	-.019	-.021	.085
***84**	-.033	.001	-.018	-.058	-.061	.026	.078	-.049	-.205	-.033	-.071	-.035	-.117	-.054	-.050	-.076	-.070	-.037	-.031	-.083	-.008	.057	.111
***85**	-.074	-.059	-.063	-.123	-.102	-.028	-.002	-.142	-.175	-.007	-.097	-.030	-.181	-.093	-.067	-.102	-.101	-.079	-.116	-.089	-.005	-.065	.116

Appendix F. (*Continued*)
Fearfulness and Responsiveness to Training

Factor 1: Fearfulness

	Facet 3: Fear/submission towards dogs					Facet 4: Fear during handling					
	19	**20**	**21**	**22**	***23**	**24**	**25**	**26**	**27**	**28**	**29**
Responsiveness to Training											
Facet 1: Trainability											
***72**	-.074	-.055	.008	-.026	.040	-.133	-.068	-.118	-.013	.029	-.110
73	-.039	-.034	-.027	.031	.010	-.089	-.027	-.085	.025	.052	-.093
74	-.070	-.098	-.051	-.023	-.027	-.149	-.065	-.120	-.093	-.054	-.094
75	-.153	-.146	-.090	-.002	-.124	-.185	-.139	-.207	-.051	-.058	-.130
***76**	-.090	-.065	-.015	.014	.039	-.121	-.069	-.116	-.056	-.014	-.120
***77**	-.061	-.077	-.065	-.062	-.133	-.145	-.070	-.080	-.048	-.096	-.119
78	-.061	-.089	-.087	-.033	-.185	-.120	-.067	-.124	.008	-.041	-.078
Facet 2: Controllability											
***79**	-.017	-.037	-.093	.059	.091	-.052	.035	-.087	.035	.026	-.109
80	-.055	-.069	-.014	.057	.014	-.108	-.055	-.086	-.048	.030	-.116
81	-.046	-.061	-.024	.002	.047	-.147	-.052	-.124	-.057	.023	-.104
***82**	-.016	-.046	-.054	.044	.129	-.063	-.010	-.095	-.063	.004	-.088
83	-.041	-.048	-.019	-.021	.085	-.066	-.045	-.075	-.016	-.012	-.078
***84**	-.031	-.083	-.008	.057	.111	-.139	-.093	-.103	-.036	-.057	-.072
***85**	-.116	-.089	-.005	-.065	.116	-.142	-.120	-.137	-.068	-.012	-.131

Appendix F. (*Continued*)
Fearfulness and Aggression towards Animals

Factor 1: Fearfulness

	Facet 1: Fear of people										Facet 2: Non-social fear								Facet 3: Fear/submission towards dogs				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	***9**	**10**	***11**	***12**	**13**	***14**	**15**	**16**	**17**	***18**	**19**	**20**	**21**	**22**	***23**
Aggression towards Animals																							
Facet 1: Aggression towards dogs																							
86	.146	.106	.150	.150	.152	.045	.047	.105	.182	.010	.181	.102	.211	.123	.080	.085	.113	.073	.292	.030	-.167	.243	-.141
87	.102	.047	.084	.152	.080	-.015	.044	.088	.121	-.013	.120	.010	.136	.077	.065	.051	.081	.066	.153	-.034	-.204	.164	-.188
88	.141	.108	.146	.125	.137	.049	.031	.124	.169	.012	.177	.117	.186	.113	.066	.094	.102	.066	.263	.035	-.141	.210	-.115
***89**	.191	.116	.202	.151	.197	.132	.092	.125	.185	.020	.254	.212	.219	.180	.099	.104	.116	.117	.395	.113	-.141	.439	-.019
***90**	.126	.096	.167	.120	.153	.157	.105	.098	.102	.007	.223	.222	.202	.180	.116	.089	.092	.119	.343	.127	-.089	.512	.080
Facet 2: Prey drive																							
91	.040	.057	.071	.079	.068	.008	.009	.061	.096	-.009	.061	.003	.093	.030	.057	.034	.069	.013	.065	.024	-.043	.028	-.103
92	.027	.024	.034	.059	.028	-.014	-.033	.030	.102	.007	.006	.004	.060	.010	.023	.011	.060	.013	.027	.017	-.019	-.028	-.106
93	.036	.012	.044	.065	.029	.009	-.011	.014	.099	.021	.012	.004	.085	-.006	.023	.025	.048	.003	.033	.046	.001	-.002	-.113
94	.003	.047	.051	.054	.021	-.006	-.012	-.006	.023	.000	.004	-.006	.061	-.013	-.023	.007	.015	.020	.024	-.037	-.076	.034	-.143
95	.135	.085	.117	.152	.133	.060	.024	.190	.235	.014	.131	.089	.157	.097	.069	.133	.116	.030	.118	.086	.023	-.001	-.089
96	.032	.029	.032	.063	.045	.003	-.020	.056	.119	.026	.041	.028	.089	.026	.058	.046	.084	-.011	.061	.051	.027	-.005	-.106
Facet 3: Dominance over dogs																							
97	.098	.086	.124	.149	.103	.032	.027	.074	.107	.028	.126	.047	.124	.086	.072	.090	.097	.087	.134	.043	-.047	.138	-.096
***98**	.086	.054	.122	.103	.101	.038	.015	.061	.105	.002	.164	.063	.124	.089	.064	.064	.064	.072	.126	.026	-.099	.173	-.064
99	-.032	-.024	.010	-.012	-.028	-.148	-.070	-.040	.053	-.131	.007	-.198	.013	-.049	-.051	-.075	-.034	-.005	-.061	-.207	-.339	.018	-.378
100	.044	.021	.066	.041	.040	-.073	-.091	.014	.164	-.056	.080	-.082	.085	-.006	-.022	.003	.032	.039	.021	-.086	-.187	.012	-.306
101	.015	.031	.020	.080	.023	-.019	.033	-.002	.012	.018	.018	-.016	.080	.028	.043	.032	.057	.045	.018	-.019	-.080	.060	-.099
102	.000	-.020	.041	-.010	-.015	-.090	-.113	.015	.112	.002	-.019	-.056	.014	-.053	-.013	.010	.011	-.008	-.051	-.012	-.072	-.102	-.191

Appendix F. (*Continued*)
Fearfulness and Aggression towards Animals

Factor 1: Fearfulness

	Facet 4: Fear during handling					
	24	**25**	**26**	**27**	**28**	**29**
Aggression towards Animals						
Facet 1: Aggression towards dogs						
86	.143	.127	.206	.039	.078	.054
87	.103	.089	.134	.034	.066	.029
88	.136	.122	.198	.010	.047	.048
***89**	.162	.187	.190	.053	.083	.076
***90**	.136	.170	.125	.051	.081	.070
Facet 2: Prey drive						
91	.061	.082	.113	.013	.013	.024
92	.048	.064	.073	.007	.019	-.009
93	.087	.068	.062	.077	.073	.023
94	.023	.011	.055	.015	.018	.003
95	.129	.105	.151	.060	.068	.068
96	.053	.044	.091	.011	.025	.026
Facet 3: Dominance over dogs						
97	.129	.105	.138	.074	.070	.048
***98**	.096	.110	.127	.021	.031	.035
99	.010	.011	.038	-.058	-.021	-.005
100	.052	.023	.103	-.060	.016	.035
101	.054	.053	.070	.074	.100	.043
102	.012	-.020	.054	-.041	-.025	.044

Appendix F. (*Continued*)
Aggression towards People and Activity/Excitability

	Facet 1: General aggression towards people								Facet 2: Situational aggression towards people								
Factor 2: Aggression towards People	**30**	**31**	**32**	***33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**
Factor 3: Activity/Excitability																	
Facet 1: Excitability																	
47	.080	.044	.129	-.067	.044	.100	.124	.082	.038	.045	.067	.088	.109	.066	.061	.281	.281
48	.058	.044	.088	-.024	.061	.079	.119	.084	.048	.032	.063	.080	.097	.077	.044	.143	.265
49	.088	.050	.063	-.021	.075	.127	.147	.133	.068	.079	.079	.101	.089	.074	.023	.125	.295
50	.070	.052	.043	.012	.084	.073	.090	.081	.055	.038	.033	.074	.078	.091	.066	.122	.246
51	.093	.069	.131	.011	.094	.121	.178	.125	.086	.048	.064	.090	.114	.058	.067	.095	.251
***52**	.145	.104	.146	.134	.176	.218	.187	.161	.150	.130	.130	.177	.140	.138	.123	.095	.240
53	.083	.039	.117	.025	.050	.058	.063	.057	.047	.035	.052	.043	.126	.030	.067	.338	.283
Facet 2: Playfulness																	
54	.004	.018	.042	-.009	-.014	-.043	.010	.000	-.009	-.003	-.046	-.004	-.056	-.012	-.057	.053	.036
***55**	-.015	-.008	.013	-.003	-.040	-.040	-.043	-.026	-.031	-.038	-.045	-.055	-.080	-.044	-.065	.021	-.003
56	-.013	.003	.016	-.030	-.005	-.035	-.011	-.008	.000	-.013	-.009	.006	-.040	-.025	-.019	.107	.095
57	.011	.024	-.001	-.059	-.026	-.031	.007	.027	.005	.030	-.038	.024	-.028	.031	-.011	.102	.137
58	-.086	-.040	-.078	-.091	-.074	-.121	-.023	-.024	-.088	-.067	-.093	-.082	-.052	-.092	-.089	.031	.090
59	-.109	-.137	-.093	-.178	-.125	-.093	-.035	-.045	-.063	-.062	-.087	-.059	-.042	-.060	-.058	.125	.102
***60**	-.029	-.007	-.021	-.078	-.035	-.003	.059	.007	-.039	-.022	-.027	.006	-.013	-.026	-.035	.020	.127
Facet 3: Active engagement																	
61	.068	.006	.096	.049	.015	.025	.092	.040	.018	-.007	-.007	-.001	.050	-.024	-.015	.120	.116
62	-.075	-.026	-.030	-.131	-.059	-.085	-.007	-.053	-.041	-.034	-.060	-.019	-.008	-.054	-.050	.040	.175
***63**	-.017	-.011	-.013	.007	-.025	-.033	.024	.002	-.021	-.051	-.031	-.025	-.036	-.017	-.057	.063	.100
64	-.031	-.032	.013	-.054	-.044	-.050	-.001	-.026	-.030	-.032	-.063	-.020	.037	-.010	-.030	.039	.087
65	.174	.040	.243	.145	.092	.102	.177	.124	.118	.057	.036	.061	.097	.031	.033	.133	.191
66	-.040	-.042	.032	-.045	-.032	-.029	-.002	-.044	-.017	-.017	-.070	-.036	.020	-.018	-.044	.027	.066
Facet 4: Companionability																	
67	-.076	-.061	-.050	-.160	-.114	-.125	-.038	-.049	-.091	-.070	-.108	-.122	-.032	-.132	-.130	.224	.018
68	-.055	-.035	-.035	-.109	-.095	-.066	-.042	-.025	-.086	-.062	-.082	-.062	-.003	-.016	-.031	.202	.081
69	-.136	-.063	-.088	-.241	-.141	-.116	-.079	-.094	-.126	-.105	-.095	-.055	-.020	-.080	-.064	.161	.058
70	-.086	-.078	-.056	-.118	-.116	-.098	-.072	-.079	-.075	-.058	-.121	-.121	-.024	-.117	-.110	.203	.005
71	.132	.069	.099	.195	.125	.095	.043	.053	.134	.073	.110	.067	.077	.089	.094	-.078	-.011

Appendix F. (*Continued*)
Aggression towards People and Responsiveness to Training

	Factor 2: Aggression towards People																
	Facet 1: General aggression towards people								Facet 2: Situational aggression towards people								
	30	31	32	*33	34	35	36	37	38	39	40	41	42	43	44	45	46
Factor 4: Responsiveness to Training																	
Facet 1: Trainability																	
***72**	-.080	-.123	-.047	-.012	-.093	-.160	-.120	-.110	-.112	-.079	-.152	-.157	-.154	-.237	-.145	-.002	-.125
73	-.014	-.070	.023	-.001	-.067	-.084	-.052	-.053	-.068	-.068	-.100	-.145	-.107	-.200	-.113	.057	-.091
74	-.013	-.072	-.014	.005	-.027	-.089	-.065	-.076	-.073	-.041	-.080	-.106	-.167	-.151	-.119	.024	-.096
75	-.090	-.085	-.049	-.107	-.103	-.132	-.133	-.153	-.119	-.094	-.103	-.137	-.084	-.112	-.079	-.035	-.193
***76**	-.075	-.132	-.006	.003	-.087	-.122	-.137	-.104	-.089	-.076	-.135	-.154	-.162	-.178	-.126	-.038	-.186
***77**	-.024	-.063	-.005	-.035	-.029	-.069	-.007	-.052	-.046	-.011	-.074	-.060	-.080	-.094	-.080	-.044	-.026
78	-.009	-.072	-.011	-.044	-.035	-.056	-.014	-.029	-.030	-.033	-.060	-.052	-.012	-.052	-.046	.029	-.008
Facet 2: Controllability																	
Face	-.062	-.104	-.034	.004	-.071	-.071	-.094	-.086	-.062	-.092	-.103	-.122	-.126	-.122	-.066	-.037	-.145
Face	-.013	-.079	.037	.029	-.048	-.079	-.094	-.080	-.038	-.030	-.103	-.113	-.081	-.120	-.085	-.008	-.191
Face	-.031	-.139	-.026	.003	-.051	-.142	-.104	-.076	-.111	-.099	-.136	-.150	-.272	-.196	-.149	-.003	-.147
Face	-.010	-.095	-.011	.049	-.036	-.052	-.127	-.072	-.063	-.043	-.086	-.078	-.168	-.116	-.095	-.071	-.245
Face	-.119	-.137	-.101	-.036	-.061	-.124	-.110	-.099	-.124	-.124	-.100	-.137	-.207	-.169	-.121	-.035	-.120
Face	-.087	-.068	-.084	-.011	-.082	-.103	-.131	-.108	-.045	-.068	-.073	-.108	-.120	-.115	-.085	-.097	-.490
Face	-.157	-.251	-.127	-.056	-.121	-.216	-.192	-.156	-.180	-.166	-.239	-.252	-.208	-.284	-.216	-.053	-.132

Appendix F. (*Continued*)
Aggression towards People and Aggression towards Animals

	Facet 1: General aggression towards people								Facet 2: Situational aggression towards people								
	30	**31**	**32**	***33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**
Factor 5: Aggression towards Animals																	
Facet 1: Aggression towards dogs																	
86	.323	.172	.313	.201	.279	.436	.310	.251	.291	.222	.319	.207	.246	.151	.181	-.007	.065
87	.264	.146	.304	.183	.171	.370	.254	.214	.283	.177	.206	.168	.249	.133	.167	.020	.062
88	.273	.175	.256	.161	.259	.406	.280	.231	.260	.224	.355	.203	.212	.155	.155	-.007	.060
***89**	.278	.111	.274	.272	.257	.369	.218	.212	.259	.206	.251	.195	.165	.101	.157	-.038	.011
***90**	.170	.065	.164	.191	.154	.230	.099	.093	.162	.124	.172	.113	.094	.077	.118	-.041	-.059
Facet 2: Prey drive																	
91	.177	.089	.203	.087	.163	.195	.227	.140	.141	.114	.094	.108	.185	.108	.094	-.004	.151
92	.084	.039	.114	.031	.103	.113	.179	.111	.068	.058	.042	.060	.114	.062	.026	.031	.177
93	.083	.044	.136	.037	.096	.094	.166	.106	.090	.038	.034	.049	.097	.019	.045	.051	.216
94	.059	.032	.091	.031	.085	.104	.140	.083	.078	.053	.050	.030	.153	.017	.057	-.053	.105
95	.231	.140	.214	.134	.196	.218	.617	.283	.189	.167	.130	.192	.140	.138	.122	.036	.203
96	.067	.058	.098	.016	.096	.105	.156	.090	.064	.052	.070	.072	.112	.079	.048	.057	.200
Facet 3: Dominance over dogs																	
97	.162	.105	.180	.106	.163	.234	.170	.144	.194	.122	.134	.131	.593	.124	.187	.023	.096
***98**	.155	.074	.163	.122	.163	.194	.138	.143	.161	.117	.125	.116	.412	.113	.165	.004	.093
99	.178	.075	.192	.088	.114	.166	.157	.107	.144	.116	.153	.093	.249	.112	.119	.008	.070
100	.167	.084	.195	.082	.140	.206	.167	.153	.174	.102	.167	.118	.260	.146	.125	.030	.100
101	.071	.063	.075	.050	.048	.102	.097	.081	.100	.061	.063	.074	.088	.087	.071	.003	.121
102	.026	.033	.058	-.025	.056	.064	.116	.072	.038	.069	.075	.044	.101	.074	.041	.028	.155

Appendix F. (*Continued*)
Activity/Excitability and Responsiveness to Training

Factor 3: Activity/Excitability																				
	Facet 1: Excitability							**Facet 2: Playfulness**							**Facet 3: Active engagement**					
	47	**48**	**49**	**50**	**51**	***52**	**53**	**54**	***55**	**56**	**57**	**58**	**59**	***60**	**61**	**62**	***63**	**64**	**65**	**66**
Factor 4: Responsiveness to Training																				
Facet 1: Trainability																				
***72**	-.126	-.107	-.172	-.038	-.099	-.154	-.099	.160	.173	.098	.009	.115	.114	-.030	.143	.066	.157	.080	.034	.166
73	-.088	-.090	-.169	-.011	-.103	-.150	-.047	.147	.115	.123	.034	.051	.107	-.055	.217	.139	.128	.096	.081	.179
74	-.104	-.023	-.087	.060	-.073	-.087	-.045	.203	.198	.153	.062	.098	.106	-.009	.180	.090	.179	.137	.080	.197
75	-.147	-.075	-.245	-.004	-.110	-.195	-.101	.162	.143	.088	.014	.104	.077	-.125	.136	.094	.134	.157	.018	.203
***76**	-.163	-.142	-.206	-.040	-.145	-.173	-.104	.193	.212	.120	.029	.047	.055	-.021	.097	.008	.149	.123	.022	.170
***77**	-.019	.063	-.042	.130	-.004	.023	-.064	.276	.201	.188	.126	.162	.158	.041	.266	.218	.287	.214	.098	.313
78	-.026	.044	-.043	.106	.049	-.070	-.037	.174	.095	.140	.089	.117	.170	-.002	.344	.270	.203	.145	.191	.215
Facet 2: Controllability																				
***79**	-.137	-.268	-.233	-.217	-.129	-.247	-.086	-.027	.046	-.093	-.094	-.111	-.036	-.105	.028	-.071	-.042	-.131	-.014	.003
80	-.156	-.094	-.225	.014	-.172	-.134	-.100	.206	.171	.107	.035	.028	.092	-.120	.104	.024	.113	.115	.028	.173
81	-.118	-.099	-.128	-.038	-.094	-.182	-.078	.110	.110	.080	.055	.099	.068	-.032	.143	.044	.094	.051	.090	.135
***82**	-.183	-.200	-.220	-.169	-.226	-.203	-.181	.096	.124	.052	-.025	-.052	.015	-.107	-.025	-.077	.002	-.016	-.073	.059
83	-.119	-.101	-.121	-.131	-.075	-.149	-.063	-.028	-.040	-.049	-.138	.067	.023	-.004	.028	.010	.008	-.028	-.047	.041
***84**	-.272	-.232	-.291	-.174	-.273	-.222	-.279	.001	.037	-.045	-.117	-.054	-.021	-.111	-.047	-.061	-.032	-.033	-.112	-.039
***85**	-.160	-.146	-.130	-.117	-.087	-.197	-.127	.004	.107	.024	-.043	.090	.052	.029	-.011	-.020	.056	.032	-.049	.083

Appendix F. (*Continued*)
Activity/Excitability and Responsiveness to Training

Factor 3: Activity/Excitability					
	Facet 4: Companionability				
	67	**68**	**69**	**70**	**71**
Factor 4: Responsiveness to Training					
Facet 1: Trainability					
***72**	.147	.053	.065	.216	-.135
73	.208	.153	.166	.312	-.117
74	.138	.068	.100	.217	-.108
75	.074	.022	.071	.162	-.059
***76**	.091	.032	.061	.187	-.130
***77**	.034	.013	.019	.135	-.095
78	.094	.036	.068	.169	-.045
Facet 2: Controllability					
***79**	.008	-.028	.005	.071	.003
80	.124	.080	.077	.209	-.081
81	.084	.020	.039	.145	-.030
***82**	.027	-.008	.004	.064	-.057
83	.044	-.013	.010	.069	-.026
***84**	.048	-.049	-.008	.066	-.010
***85**	.104	-.008	.059	.138	-.123

Appendix F. (*Continued*)
Activity/Excitability and Aggression towards Animals

Factor 3: Activity/Excitability																				
	Facet 1: Excitability							**Facet 2: Playfulness**							**Facet 3: Active engagement**					
	47	**48**	**49**	**50**	**51**	***52**	**53**	**54**	***55**	**56**	**57**	**58**	**59**	***60**	**61**	**62**	***63**	**64**	**65**	**66**
Factor 5: Aggression towards Animals																				
Facet 1: Aggression towards dogs																				
86	.093	.093	.207	.051	.183	.176	.105	-.066	-.058	-.064	-.034	-.201	-.075	.029	.022	-.038	-.035	-.017	.127	-.058
87	.074	.081	.136	.067	.166	.134	.097	-.038	-.052	-.022	.014	-.102	-.027	.007	.074	.000	-.002	.019	.146	.020
88	.090	.079	.205	.060	.170	.192	.101	-.042	-.033	-.053	-.022	-.186	-.066	.068	.031	-.063	-.031	-.004	.083	-.031
***89**	.040	.015	.034	-.024	.108	.152	.059	-.073	-.056	-.109	-.074	-.371	-.125	-.129	-.037	-.144	-.057	-.035	.075	-.035
***90**	-.035	-.128	-.177	-.127	.007	.030	.027	-.124	-.167	-.227	-.187	-.664	-.254	-.339	-.122	-.186	-.178	-.100	-.032	-.107
Facet 2: Prey drive																				
91	.108	.093	.169	.083	.399	.118	.113	-.060	-.072	-.028	-.011	-.007	-.038	.065	.040	.052	-.014	.004	.193	-.015
92	.131	.164	.233	.164	.501	.169	.136	-.001	-.064	.033	.065	.098	.026	.097	.072	.095	.060	.052	.239	.033
93	.160	.175	.216	.142	.756	.144	.166	.055	-.040	.061	.077	.084	.036	.079	.105	.152	.057	.087	.368	.090
94	.021	.083	.078	.030	.417	.052	.075	-.088	-.131	-.074	-.055	-.027	-.048	-.019	.046	.082	-.007	.035	.185	.026
95	.174	.212	.239	.199	.267	.266	.116	.031	-.033	.036	.081	.056	.038	.131	.096	.054	.076	.038	.196	.046
96	.178	.170	.280	.188	.499	.203	.150	.011	-.059	.055	.100	.070	.039	.089	.051	.104	.036	.045	.223	-.004
Facet 3: Dominance over other dogs																				
97	.089	.093	.072	.050	.129	.147	.097	-.062	-.100	-.051	-.031	-.077	-.047	-.040	.034	.006	-.026	.027	.093	.039
***98**	.045	.043	.037	.051	.112	.143	.114	-.048	-.068	-.027	-.027	-.191	-.081	-.059	-.008	-.027	-.006	.043	.065	.026
99	.057	.147	.094	.072	.136	.105	.109	-.042	-.063	-.005	.033	-.007	-.032	.007	.086	.079	.018	.052	.164	.057
100	.137	.252	.236	.167	.168	.235	.139	-.012	-.046	.035	.062	.038	.009	.088	.074	.064	.038	.035	.121	.011
101	.042	.021	.061	.002	.096	.005	.075	-.068	-.121	-.103	-.056	-.097	-.075	-.074	.016	.013	-.039	-.017	.056	-.001
102	.137	.228	.244	.165	.148	.167	.084	-.015	-.045	.033	.091	.179	.041	.149	.036	.091	.082	.053	.068	.061

Appendix F. (*Continued*)
Activity/Excitability and Aggression towards Animals

Factor 3: Activity/Excitability					
	Facet 4: Companionability				
	67	**68**	**69**	**70**	**71**
Factor 5: Aggression towards Animals					
Facet 1: Aggression towards dogs					
86	-.086	-.025	-.052	-.059	.094
87	-.040	-.025	-.025	-.031	.089
88	-.044	-.039	-.060	-.053	.059
***89**	-.138	-.053	-.115	-.116	.123
***90**	-.175	-.093	-.142	-.132	.143
Facet 2: Prey drive					
91	-.066	-.051	-.075	-.073	.059
92	-.053	-.007	-.054	-.054	.035
93	-.041	.001	-.023	-.034	.025
94	-.048	-.098	-.072	-.064	.049
95	-.043	.011	-.048	-.065	.014
96	-.047	.007	-.030	-.074	.050
Facet 3: Dominance over other dogs					
97	-.042	-.017	-.015	-.038	.057
***98**	-.071	-.065	-.061	-.061	.071
99	-.039	-.022	-.023	-.010	.089
100	-.042	.002	-.037	-.043	.047
101	-.036	-.022	-.029	-.035	.113
102	-.011	.064	-.006	-.072	-.017

Appendix F. (*Continued*)
Responsiveness to Training and Aggression towards Animals

Factor 4: Responsiveness to Training														
	Facet 1: Trainability							**Facet 2: Unruliness**						
	***72**	**73**	**74**	**75**	***76**	***77**	**78**	***79**	**80**	**81**	***82**	**83**	***84**	***85**
Factor 5: Aggression towards Animals														
Facet 1: Aggression towards dogs														
86	-.126	-.046	-.046	-.130	-.103	-.070	-.025	-.039	-.085	-.111	-.091	-.122	-.098	-.164
87	-.097	-.024	-.044	-.062	-.055	-.002	.033	-.043	-.056	-.052	-.076	-.103	-.085	-.121
88	-.127	-.076	-.072	-.147	-.108	-.059	-.048	-.067	-.100	-.105	-.102	-.109	-.122	-.164
***89**	-.095	-.044	-.058	-.134	-.051	-.047	-.064	.007	-.060	-.071	-.005	-.097	-.057	-.126
***90**	-.086	-.039	-.081	-.105	-.036	-.115	-.098	.080	-.027	-.057	.036	-.062	.017	-.088
Facet 2: Prey drive														
91	-.136	-.096	-.096	-.096	-.158	-.033	-.004	-.116	-.168	-.140	-.169	-.139	-.177	-.126
92	-.093	-.105	-.066	-.096	-.131	.043	.022	-.136	-.140	-.084	-.181	-.068	-.199	-.064
93	-.039	-.061	-.054	-.074	-.088	.044	.069	-.098	-.127	-.038	-.174	-.039	-.242	-.034
94	-.065	-.052	-.037	-.004	-.107	.009	.033	-.126	-.120	-.088	-.177	-.040	-.128	-.027
95	-.150	-.108	-.104	-.172	-.148	-.009	-.023	-.162	-.141	-.117	-.206	-.114	-.199	-.150
96	-.110	-.125	-.101	-.127	-.146	.001	-.003	-.156	-.151	-.117	-.196	-.119	-.202	-.113
Facet 3: Dominance over other dogs														
97	-.115	-.067	-.154	-.056	-.107	-.052	-.043	-.106	-.049	-.181	-.148	-.147	-.135	-.126
***98**	-.120	-.054	-.120	-.091	-.076	-.030	-.042	-.111	-.039	-.191	-.094	-.245	-.116	-.102
99	-.050	-.035	.012	.041	-.065	.027	.058	-.038	-.031	-.039	-.080	-.091	-.103	-.177
100	-.122	-.094	-.044	-.053	-.107	-.003	.013	-.118	-.074	-.086	-.155	-.113	-.163	-.158
101	-.086	-.009	-.048	.009	-.074	-.058	-.027	.005	-.048	-.029	-.051	-.050	-.087	-.085
102	-.120	-.109	-.082	-.056	-.151	.007	-.030	-.164	-.130	-.107	-.174	-.092	-.184	-.108

Appendix G. Study 3: Items retained and removed in creating the final DPQ forms

ITEMS RETAINED		ITEMS REMOVED	
FEARFULNESS			
Facet 1			
***1**	Dog behaves fearfully towards unfamiliar people.	**2**	Dog behaves fearfully towards familiar people.
3	Dog behaves fearfully towards children.	**4**	Dog behaves fearfully in response to perceived threats from people (e.g., being cornered, having collar reached for).
5	Dog behaves fearfully when near crowds of people.	**7**	Dog is cautious, careful.
***6**	Dog is shy.	**8**	Dog behaves fearfully towards moving bicycles, joggers, skateboarders.
***9**	Dog is relaxed when greeting people.	**10**	Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting people.
Facet 2			
***11**	Dog adapts easily to new situations and environments.	**15**	Dog is fearful of loud noises (e.g., heavy traffic, car horns, slamming doors, fireworks).
***12**	Dog is confident.	**16**	Dog attempts to flee from novel objects or situations.
***13**	Dog is anxious	**18**	Dog easily gets over unpleasant experiences (e.g., painful toe nail clippings).
14	Dog is quick to recover after being startled or frightened.		
17	Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a door frame).		
Facet 3			
***19**	Dog behaves fearfully towards other dogs.		
20	Dog behaves fearfully when threatened by other dogs (e.g., growled or lunged at, cornered).		
***21**	Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.		
***22**	Dog avoids other dogs.		
23	Dog is bold.		

Appendix G. (*Continued*)

ITEMS RETAINED		ITEMS REMOVED	
FEARFULNESS (*continued*)			
Facet 4			
***24**	Dog behaves fearfully when groomed (e.g., nails trimmed, brushed, bathed, ears cleaned).	**29**	When alone or about to be left alone, dog shakes, shivers, or trembles.
***25**	Dog behaves fearfully during visits to the veterinarian.		
***26**	Dog exhibits fearful behaviors when restrained.		
27	Dog is sensitive (and reactive) to pain.		
28	Dog is easily upset when corrected, scolded, or punished.		
AGGRESSION TOWARDS PEOPLE			
Facet 1			
***30**	Dog behaves aggressively towards unfamiliar people.	**31**	Dog behaves aggressively towards familiar people.
32	Dog behaves aggressively when a person (e.g., visitor, delivery person) approaches the house or yard.	**36**	Dog behaves aggressively towards moving bicycles, joggers, skateboarders.
***33**	Dog is friendly towards unfamiliar people.	**37**	Dog behaves aggressively towards people with unfamiliar objects (e.g., canes, wheelchairs, umbrellas).
34	Dog behaves aggressively towards children.		
***35**	Dog shows aggression when nervous or fearful.		
Facet 2			
***38**	Dog behaves aggressively in response to perceived threats from people (e.g., being cornered, having collar reached for).	**40**	Dog displays aggression that is sudden and without apparent reason.
39	Dog behaves aggressively during visits to the veterinarian.	**43**	Dog behaves aggressively when scolded or punished.
***41**	Dog behaves aggressively when restrained or handled (e.g., groomed).		
***42**	Dog aggressively guards coveted items (e.g., stolen item, treats, food bowl).		
44	Dog behaves aggressively if disturbed or moved when resting.		

Appendix G. (*Continued*)

ITEMS RETAINED		ITEMS REMOVED	
ACTIVITY & EXCITABILITY			
Facet 1			
47	Dog is very excitable when visitors arrive.	**45**	Dog becomes very excited when owner returns home.
***48**	Dog is boisterous.	**46**	Dog is active (e.g., jumps, sniffs a lot) when walking on a leash.
49	Dog is very excitable around other dogs.	**50**	Dog seeks constant activity.
***50**	Dog seeks constant activity.	**53**	Dog is very excitable just before being taken for a walk.
***52**	Dog tends to be calm.	**51**	Dog is very excitable when around squirrels, birds, or other small animals.
Facet 2			
***54**	Dog retrieves objects (e.g., balls, toys, sticks).	**58**	Dog is active during play with other dogs.
***55**	Dog gets bored in play quickly.	**59**	Dog is playful with familiar people.
***56**	Dog enjoys playing with toys.		
57	Dog is interested in playing tug-o-war with people or dogs.		
60	Dog is aloof or indifferent towards other dogs.		
Facet 3			
61	Dog is very alert.	**65**	Dog is very watchful (e.g., monitors for squirrels, attends to noises).
***62**	Dog is curious.		
***63**	Dog is lethargic		
***64**	Dog works at tasks (e.g., getting treats out of a Kong, shredding toys) until entirely finished.		
66	Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden.		
Facet 4			
***6**	Dog is affectionate.		
68	Dog follows people around.		
***69**	Dog seeks companionship from people.		
70	Dog loves to be praised.		
***71**	Dog is aloof.		

Appendix G. (*Continued*)

ITEMS RETAINED		ITEMS REMOVED	
RESPONSIVENESS TO TRAINING			
Facet 1			
***72**	Dog is slow to respond to corrections.	**74**	Dog is willing to complete work, task, or training without a reward.
73	Dog is attentive to owner's actions and words.	**78**	Dog is intelligent.
***75**	Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).		
***76**	Dog ignores commands.		
77	Dog is slow to learn new tricks or tasks.		
Facet 2			
79	Dog is destructive.	**83**	Dog is willing to let go of toys when playing (e.g., during tug-o-war).
***80**	When off leash, dog comes immediately when called.	**85**	Dog is dominant over owner.
***81**	Dog leaves food or objects alone when told to do so.		
***82**	Dog is quick to sneak out through open doors, gates.		
84	When walking on leash, dog tends to pull ahead.		
AGGRESSION TOWARDS ANIMALS			
Facet 1			
***86**	Dog behaves aggressively toward dogs.		
87	Dog responds aggressively when threatened by another dog (e.g., growled or lunged at, cornered).		
88	Dog has a tendency to attack (or attempt to attack) other dogs.		
***89**	Dog is friendly towards other dogs.		
***90**	Dog is playful with other dogs.		

Appendix G. (*Continued)*

ITEMS RETAINED		ITEMS REMOVED	
AGGRESSION TOWARDS ANIMALS (*Continued*)			
Facet 2			
***91**	Dog behaves aggressively towards cats.	92	Dog likes to chase cats.
***93**	Dog likes to chase squirrels, birds, or other small animals.		
94	Dog catches and kills other animals (e.g., squirrels, rabbits).		
***95**	Dog likes to chase bicycles, joggers, and skateboarders.		
96	Dog is very excitable around cats.		
Facet 3			
97	Dog guards food or treats from other dogs.	**101**	Dog frequently urine marks.
***98**	Dog willingly shares toys with other dogs.		
***99**	Dog is dominant over other dogs.		
***100**	Dog is assertive or pushy with other dogs (e.g., if in a home with other dogs, when greeting).		
102	Dog jumps up on (e.g., in play) and/or mounts other dogs (outside appropriate mating).		

Note. Items marked with an asterisk were retained for the short-, or 45-item, form.

Appendix H. DPQ Long (75-item) form and scoring sheet

Dog Personality Questionnaire (DPQ)

Here are a number of personality traits and behavioral descriptions that may or may not apply to your dog. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate your dog based on his or her general, overall behavior.

Disagree strongly	Disagree moderately	Disagree slightly	Neither agree nor disagree	Agree slightly	Agree moderately	Agree strongly
1	2	3	4	5	6	7

1._____ Dog is relaxed when greeting people.

2._____ Dog behaves aggressively if disturbed or moved when resting.

3._____ Dog is aloof or indifferent towards other dogs.

4._____ Dog is destructive.

5._____ Dog behaves aggressively toward dogs.

6._____ Dog is anxious

7._____ Dog loves to be praised.

8._____ Dog responds aggressively when threatened by another dog (e.g., growled or lunged at, cornered).

9._____ Dog is bold.

10._____ Dog is lethargic

11._____ When off leash, dog comes immediately when called.

12._____ Dog is shy.

13._____ Dog behaves aggressively towards unfamiliar people.

14._____ Dog will work to obtain an object or reward (e.g., ball, treat) that is hidden.

15._____ Dog likes to chase squirrels, birds, or other small animals.

16._____ Dog gets bored in play quickly.

17._____ Dog behaves aggressively when restrained or handled (e.g., groomed).

18._____ Dog is quick to sneak out through open doors, gates.

19._____ Dog is confident.

20._____ Dog is dominant over other dogs.

21._____ Dog avoids other dogs.

22._____ Dog catches and kills other animals (e.g., squirrels, rabbits).

23._____ Dog behaves aggressively when a person (e.g., visitor, delivery person) approaches the house or yard.

24._____ Dog is easily startled by unexpected contact with objects (e.g., tripping, brushing against a door frame).

25._____ Dog works at tasks (e.g., getting treats out of a Kong, shredding toys) until entirely finished.

26._____ Dog is very excitable around cats.

27._____ Dog is boisterous.

28._____ Dog behaves fearfully during visits to the veterinarian.

29._____ When walking on leash, dog tends to pull ahead.

30._____ Dog behaves fearfully when near crowds of people.

31._____ Dog enjoys playing with toys.

32._____ Dog is easily upset when corrected, scolded, or punished.

33._____ Dog is friendly towards unfamiliar people.

34._____ Dog is playful with other dogs.

35._____ Dog seeks companionship from people.

36._____ Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.

Appendix H. (*Continued*)

Disagree strongly	Disagree moderately	Disagree slightly	Neither agree nor disagree	Agree slightly	Agree moderately	Agree strongly
1	2	3	4	5	6	7

37. _____ Dog is attentive to owner's actions and words.
38. _____ Dog adapts easily to new situations and environments.
39. _____ Dog likes to chase bicycles, joggers, and skateboarders.
40. _____ Dog is curious.
41. _____ Dog guards food or treats from other dogs.
42. _____ Dog is sensitive (and reactive) to pain.
43. _____ Dog behaves aggressively in response to perceived threats from people (e.g., being cornered, having collar reached for).
44. _____ Dog is aloof.
45. _____ Dog is slow to learn new tricks or tasks.
46. _____ Dog is interested in playing tug-o-war with people or dogs.
47. _____ Dog behaves fearfully towards unfamiliar people.
48. _____ Dog is very alert.
49. _____ Dog willingly shares toys with other dogs.
50. _____ Dog is slow to respond to corrections.
51. _____ Dog behaves aggressively during visits to the veterinarian.
52. _____ Dog jumps up on (e.g., in play) and/or mounts other dogs (outside appropriate mating).
53. _____ Dog seeks constant activity.
54. _____ Dog behaves fearfully towards children.
55. _____ Dog is very excitable when visitors arrive.
56. _____ Dog leaves food or objects alone when told to do so.
57. _____ Dog has a tendency to attack (or attempt to attack) other dogs.
58. _____ Dog is quick to recover after being startled or frightened.
59. _____ Dog retrieves objects (e.g., balls, toys, sticks).
60. _____ Dog is friendly towards other dogs.
61. _____ Dog exhibits fearful behaviors when restrained.
62. _____ Dog aggressively guards coveted items (e.g., stolen item, treats, food bowl).
63. _____ Dog is affectionate.
64. _____ Dog ignores commands.
65. _____ Dog behaves aggressively towards cats.
66. _____ Dog behaves fearfully when threatened by other dogs (e.g., growled or lunged at, cornered).
67. _____ Dog follows people around.
68. _____ Dog shows aggression when nervous or fearful.
69. _____ Dog tends to be calm.
70. _____ Dog behaves fearfully towards other dogs.
71. _____ Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).
72. _____ Dog is very excitable around other dogs.
73. _____ Dog behaves aggressively towards children.
74. _____ Dog behaves fearfully when groomed (e.g., nails trimmed, brushed, bathed, ears cleaned).
75. _____ Dog is assertive or pushy with other dogs (e.g., if in a home with other dogs, when greeting).

Appendix H. (*Continued*)

SCORING KEY FOR DPQ LONG FORM

Factor Facet	Item number on long form
Factor 1 – Fearfulness	
Facet 1 – Fear of People	**R1, 12,** 30, **47,** 54
Facet 2 – Nonsocial Fear	**6, R19,** 24, **R38,** R58
Facet 3 – Fear of Dogs	R9, **21, 36,** 66, **70**
Facet 4 – Fear of Handling	**28,** 32, 42, **61, 74**
Factor 2 – Aggression towards People	
Facet 1 – General Aggression	**13,** 23, **R33, 68,** 73
Facet 2 – Situational Aggression	2, 17, **43, 51, 62**
Factor 3 – Activity/Excitability	
Facet 1 – Excitability	**27, 53,** 55, **R69**, 72
Facet 2 – Playfulness	R3, **R16, 31,** 46, **59**
Facet 3 – Active Engagement	**R10**, 14, **25, 40,** 48
Facet 4 – Companionability	7, **35, R44, 63,** 67
Factor 4 – Responsiveness to Training	
Facet 1 – Trainability	37, R45, **R50, R64, 71**
Facet 2 – Controllability	R4, **11, R18,** R29, **56**
Factor 5 – Aggression towards Animals	
Facet 1 – Aggression towards Dogs	**5,** 8, **R34,** 57, **R60**
Facet 2 – Prey Drive	**15,** 22, 26, **39, 65**
Facet 3 – Dominance over Other Dogs	**20,** 41, **R49,** 52, **75**

Note. An R in front a item indicates that the item is reverse coded. Bolded item numbers indicate items that also appear on the short form of the DPQ, though numbered differently.

Appendix I. DPQ Short (45-item) form and scoring sheet

Dog Personality Questionnaire (DPQ)

Here are a number of personality traits and behavioral descriptions that may or may not apply to your dog. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate your dog based on his or her general, overall behavior.

Disagree strongly	Disagree moderately	Disagree slightly	Neither agree nor disagree	Agree slightly	Agree moderately	Agree strongly
1	2	3	4	5	6	7

1. _____ Dog is relaxed when greeting people.
2. _____ Dog behaves aggressively toward dogs.
3. _____ Dog is anxious
4. _____ Dog is lethargic
5. _____ When off leash, dog comes immediately when called.
6. _____ Dog is shy.
7. _____ Dog behaves aggressively towards unfamiliar people.
8. _____ Dog likes to chase squirrels, birds, or other small animals.
9. _____ Dog gets bored in play quickly.
10. _____ Dog is quick to sneak out through open doors, gates.
11. _____ Dog is confident.
12. _____ Dog is dominant over other dogs.
13. _____ Dog avoids other dogs.
14. _____ Dog works at tasks (e.g., getting treats out of a Kong, shredding toys) until entirely finished.
15. _____ Dog is boisterous.
16. _____ Dog behaves fearfully during visits to the veterinarian.
17. _____ Dog enjoys playing with toys.
18. _____ Dog is friendly towards unfamiliar people.
19. _____ Dog is playful with other dogs.
20. _____ Dog seeks companionship from people.
21. _____ Dog behaves submissively (e.g., rolls over, avoids eye contact, licks lips) when greeting other dogs.
22. _____ Dog adapts easily to new situations and environments.
23. _____ Dog likes to chase bicycles, joggers, and skateboarders.
24. _____ Dog is curious.
25. _____ Dog behaves aggressively in response to perceived threats from people (e.g., being cornered, having collar reached for).
26. _____ Dog is aloof.
27. _____ Dog behaves fearfully towards unfamiliar people.
28. _____ Dog willingly shares toys with other dogs.
29. _____ Dog is slow to respond to corrections.
30. _____ Dog behaves aggressively during visits to the veterinarian.
31. _____ Dog seeks constant activity.
32. _____ Dog leaves food or objects alone when told to do so.
33. _____ Dog retrieves objects (e.g., balls, toys, sticks).
34. _____ Dog is friendly towards other dogs.
35. _____ Dog exhibits fearful behaviors when restrained.
36. _____ Dog aggressively guards coveted items (e.g., stolen item, treats, food bowl).

Appendix I. (*Continued*)

Disagree strongly	Disagree moderately	Disagree slightly	Neither agree nor disagree	Agree slightly	Agree moderately	Agree strongly
1	2	3	4	5	6	7

37. _____ Dog is affectionate.

38. _____ Dog ignores commands.

39. _____ Dog behaves aggressively towards cats.

40. _____ Dog shows aggression when nervous or fearful.

41. _____ Dog tends to be calm.

42. _____ Dog behaves fearfully towards other dogs.

43. _____ Dog is able to focus on a task in a distracting situation (e.g., loud or busy places, around other dogs).

44. _____ Dog behaves fearfully when groomed (e.g., nails trimmed, brushed, bathed, ears cleaned).

45. _____ Dog is assertive or pushy with other dogs (e.g., if in a home with other dogs, when greeting).

Appendix I. (*Continued*)

Scoring Key for DPQ Short Form

Factor Facet	Item number on short form
Factor 1 – Fearfulness	
Facet 1 – Fear of People	R1, 6, 27
Facet 2 – Nonsocial Fear	3, R11, R22
Facet 3 – Fear of Dogs	13, 21, 42
Facet 4 – Fear of Handling	16, 35, 44
Factor 2 – Aggression towards People	
Facet 1 – General Aggression	7, R18, 40
Facet 2 – Situational Aggression	25, 30, 36
Factor 3 – Activity/Excitability	
Facet 1 – Excitability	15, 31, R41
Facet 2 – Playfulness	R9, 17, 33
Facet 3 – Active Engagement	R4, 14, 24
Facet 4 – Companionability	20, R26, 37
Factor 4 – Responsiveness to Training	
Facet 1 – Trainability	R29, R38, 43
Facet 2 – Controllability	5, R10, 32
Factor 5 – Aggression towards Animals	
Facet 1 – Aggression towards Dogs	2, R19, R34
Facet 2 – Prey Drive	8, 23, 39
Facet 3 – Dominance over Other Dogs	12, R28, 45

Note. An R in front a item indicates that the item is reverse coded. Bolded item numbers indicate items that also appear on the short form of the DPQ, though numbered differently.

Appendix J. Study 6: Test Battery subtest descriptions form

TEST BATTERY (SUBTESTS & PROCEDURES)

When each dog arrives at the kennel, he or she will be placed in a familiar, indoor kennel in a room familiar to the dog. The dog will remain there until his or her test period begins. Then, the dog will be released from the kennel into the room.

Subtest 1: Collar grab
A researcher will reach for the dog's neck/collar and put a second collar on the dog; this collar will be attached to a leash for use in Subtest 2. The dogs' behavior will be rated on 5-point scales assessing how fearful, confident, aggressive, and friendly the dog appeared. Notes of other significant behaviors (e.g., not seeming to react to notice the researcher) will be made.

Subtest 2: Walk on leash with stranger
The same researcher will walk the dog approximately 25 feet down a hallway familiar to the dog to a 10*15-foot room that is familiar to the dog. The dog's on-leash behavior during will be assessed and rated on 5-point scales assessing how fearful, confident, excitable, calm, engaged/alert to the environment, easy-to-control, submissive, and pushy/assertive the dog appeared.

Subtest 3: Threatening approach
The researcher will stand still but ignore the dog. A second researcher (unfamiliar to the dog) will approach the dog. The researcher will exhibit mildly threatening behavior, targeted towards the dog, including staring directly at the dog and looming (leaning slightly) over the dog. The dog's behavior will be carefully observed for signs of arousal, fear, and aggression, including raised hackles, panting, pulling back of lips and ears, backing away, and barking (among others). The test will be ceased if the dog exhibits aggressive behavior readily. The researcher will not come within range (such that the dog could touch or bite the researcher) unless the dog exhibits signs of friendly behavior (e.g., lip licking, tail wagging, whining, body wiggling/wagging). The dogs' behavior will be rated on 5-point scales assessing how fearful, confident, aggressive, friendly, excitable, calm, playful, engaged/alert to the environment, and interested in people the dog appears. Notes of other significant behaviors (e.g., being distracted by a bird) will be made.

Subtest 4: Friendly approach
If Subtest 3 indicates that the dog is friendly to the approach of an unfamiliar person who is mildly threatening, a second approach test will be conducted. The same stranger will leave, then re-approach the dog, talking in a friendly voice ("Hi! It's so nice to see you! Who's a good doggie?"). If the dog shows mild to moderate discomfort (e.g., fear, stress, aggression) during the approach in subtest 3 or 4, the test will be conducted using an artificial hand and so that the researcher will be kept at a distance from the dog. If the dog shows more than moderate discomfort (e.g., by attempting to bite), the test will be ended.

The dog's behavior during the friendly approach will be rated using a 5-point scale to indicate the degree to which the dog appears how fearful, confident, aggressive, friendly, excitable, calm, playful, engaged/alert to the environment, and interested in people.

Subtest 5: Affection/Petting from stranger
The researcher will kneel next to the dog, then reach towards the dog's body to pet the dog in a minimally threatening fashion. The dogs' behavior will be rated on 5-point scales assessing how fearful, confident, aggressive, friendly, excitable, calm, aloof/uninterested, playful, engaged/alert to the environment, interested in people, and affectionate the dog appears. Notes of other significant behaviors (e.g., ignoring the petting) will be made.

Subtest 6: Restraint
The researcher will hold the dog down on its side or back, as if the dog were undergoing an inspection at the veterinarian's office. Pressure will begin as light but elevate to moderate over a 45 second period regardless of whether or not the dog struggles. The dogs' behavior will be rated on 5-point scales assessing how fearful, confident, aggressive, friendly, calm, and excitable the dog appears; notes of other significant behaviors (e.g., whining) will be made.

Subtest 7: Other dog (on leash)
The target dog will be exposed to another dog, the test dog (a non-aggressive, male, neutered dog who is highly socialized). Each dog will be held on leash by a different person. The dog being assessed will be held by the same researcher who served as the stranger in subtest 3 and conducted all tests following subtest 3. The test dog will be walked by the researcher who conducted subtests 1 and 2.

The test dog will be walked slowly around the room, and the target dog's behavior will be observed. The dog will be rated (using a 5-point scale) on the degree to which he or she appears fearful, confident, aggressive, friendly, excitable, calm, aloof/uninterested, playful, submissive, and pushy/assertive.

Subtest 8: Novel situation test (room)
The researcher will walk the dog out of the familiar room in which subtests 3-7 were conducted and into a room that is novel to the dog. The researcher will then drop the dog's leash and let the dog explore the room. The researcher will not respond if the dog solicits attention. The dog's behavior will be observed and rated on 5-point scales assessing how fearful, confident, excitable, calm, and engaged/alert to the environment the dog appears.

Subtest 9: Novel object (Remote controlled car, Jack-in-the-box)
The researcher will allow the dog to continue dragging his or her leash. The researcher will retrieve a remote controlled car and set it on the floor. After 20 seconds, regardless of the dog's response, the researcher will begin to drive the car around the room. After 1 minute of exposure to the car, the car will be removed and a Jack-in-the-box will be placed on the floor. The dog will be allowed to inspect closed, and which will then suddenly open. The dog's behavior will be observed and rated on 5-point scales assessing the degree to which the dog appears fearful, confident, excitable, calm, and engaged/alert to the environment.

Subtest 10: Doll test (Doll as child)
The dog will be held relatively stationary by the researcher while a second researcher approaches, holding a large (toddler-sized) doll by the hand, as if the doll were walking. The dog's behavior during the walk and when he/she reaches the dog will be observed and rated on 5-point scales assessing how fearful, confident, aggressive, friendly, excitable, and calm the dog appears.

Subtest 11: Prey drive
The dog will be exposed to a furry ball pulled on a string, then tossed. The dog's behavior will be observed and rated on 5-point scales assessing how aggressive, excitable, calm, playful, and engaged in the environment the dog appears.

Subtest 12: Engagement in play with tester
The researcher will excitedly ask the dog to play with a new squeaky toy. The researcher can nudge the dog with the toy, talk to the dog, clap his hands, kneel down, and run around to get the dog's interest. The dog's behavior will be rated on 5-point scales assessing how affectionate, excitable, calm, aloof or uninterested, playful, engaged/alert to the environment, and interested in people the dog appears.

Subtest 13: Tug-o-war or toy release
The researcher will attempt to engage the dog in a game of tug-o-war with a rope toy. At the end of the tug-o-war game (up to 20 seconds of pulling), the researcher will tell the dog to release the toy and/or try to retrieve the object from the dog. Researchers will observe the dog for unwillingness to release the toy. The dog's behavior will be rated on 5-point scales assessing how easy to train, easy to control, submissive, and pushy/assertive the dog appears.

Subtest 14: Train new task
The researcher will attempt to teach the dog to hit a square block (paired with other shapes of blocks) in order to receive a reward. The reward will be a food reward, toy, or praise, depending on what appears to motivate the dog and on what the owner reports the dog enjoys. The researcher will point to the block, lure the dog to touch the block, then give the dog a reward. The researcher will spend up to 4 minutes using luring, shaping, and positive reinforcement (plus verbal corrections) to teach the dog. The dog's behavior will be observed and rated on 5-point scales assessing how engaged in the environment and task, interested in people, easy to train, and obedient the dog appears.

Subtest 15: Basic commands
The researcher will give the dog basic commands (e.g., "sit," "come," "down") to get the dog to come, sit at a door, and then be released to go outside. The researcher will to give the dog that command without a food reward or other reward present, then with a reward if the dog does not perform without the reward. The dogs' behavior will be rated on 5-point scales assessing how obedient the dog appears.

Subtest 16: Other dog (off leash)
The target dog will be released into a play yard, which is typically used at the kennel where the test is conducted and which is familiar to the dog. In the play yard will be another dog, the test dog used in subtest (a non-aggressive, male, neutered dog who is highly socialized). Both dogs will simply be loose in the play yard, as during any other play period, for 4 minutes. Their interactions and the target dog's behavior will be observed. The target dog will be rated, using a series of 5-point scales, on of fearful, confident, aggressive, friendly, excitable, calm, aloof/uninterested, playful, submissive, and pushy/assertive the dog appears.

Subtest 17: Activity in free-play
The test dog will be removed from the play yard, and the target dog's activity level when alone will be observed. The dog will be rated on a 5-point scale on how active, excitable, calm, aloof, playful, and engaged/alert to the environment the dog appears.

Subtest 18: Reunion with owner
The researcher will walk the dog back to the owner. The dog's behavior during the walk and when he/she reaches the owner will be observed and rated on 5-point scales assessing how excitable, calm, aloof, interested in people, easy to control, submissive, and pushy/assertive the dog appears.

Appendix K. Study 6. Dog behavioral assessment scoring sheet

DATE: _____

YOUR NAME: _____

TESTER'S NAME: _____

SCORE SHEET – DOG BEHAVIORAL ASSESSMENT

DOG'S NAME: _____

BREED: _____ **SEX:** Male Female

NOTES:

Please rate the dog's behavior on each subtest using the traits or behaviors listed below that subtest. (Circle one.)

Subtest 1: COLLAR GRAB

FEARFUL	1	2	3	4	5
	not at all				extremely
CONFIDENT	1	2	3	4	5
	not at all				extremely
AGGRESSIVE	1	2	3	4	5
	not at all				extremely
FRIENDLY	1	2	3	4	5
	not at all				extremely

Subtest 2: WALK ON LEASH WITH STRANGER

FEARFUL	1	2	3	4	5
	not at all				extremely
CONFIDENT	1	2	3	4	5
	not at all				extremely
EXCITABLE	1	2	3	4	5
	not at all				extremely
CALM	1	2	3	4	5
	not at all				extremely
ENGAGED, ALERT (environment)	1	2	3	4	5
	not at all				extremely
EASY TO CONTROL	1	2	3	4	5
	not at all				extremely
SUBMISSIVE	1	2	3	4	5
	not at all				extremely
PUSHY, ASSERTIVE	1	2	3	4	5
	not at all				extremely

Appendix K. (*Continued*)

Subtest 3: THREATENING APPROACH

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
ENGAGED, ALERT (environment)	1 not at all	2	3	4	5 extremely
INTERESTED IN PEOPLE	1 not at all	2	3	4	5 extremely

Subtest 4: FRIENDLY APPROACH

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
ENGAGED, ALERT (environment)	1 not at all	2	3	4	5 extremely
INTERESTED IN PEOPLE	1 not at all	2	3	4	5 extremely

Appendix K. (*Continued*)

Subtest 5: AFFECTION/PETTING FROM STRANGER

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
ALOOF or UNINTERESTED	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
ENGAGED, ALERT (environment)	1 not at all	2	3	4	5 extremely
INTERESTED IN PEOPLE	1 not at all	2	3	4	5 extremely
AFFECTIONATE	1 not at all	2	3	4	5 extremely

Subtest 6: RESTRAINT

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
AFFECTIONATE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely

Appendix K. (*Continued*)

Subtest 7: OTHER DOG (ON-LEASH)

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
ALOOF or UNINTERESTED	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
SUBMISSIVE	1 not at all	2	3	4	5 extremely
PUSHY, ASSERTIVE	1 not at all	2	3	4	5 extremely

Subtest 8: NOVEL SITUATION (ROOM)

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
ENGAGED, ALERT (environment)	1 not at all	2	3	4	5 extremely

Appendix K. (*Continued*)

Subtest 9: NOVEL OBJECT (Remote controlled car, Jack-in-the-box)

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
ENGAGED, ALERT (environment)	1 not at all	2	3	4	5 extremely

Subtest 10: DOLL TEST (doll as child)

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely

Subtest 11: PREY DRIVE

AGGRESSIVE	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
ENGAGED (environment)	1 not at all	2	3	4	5 extremely

Appendix K. (*Continued*)

Subtest 12: ENGAGEMENT IN PLAY WITH TESTER

AFFECTIONATE	1	2	3	4	5
	not at all				extremely
EXCITABLE	1	2	3	4	5
	not at all				extremely
CALM	1	2	3	4	5
	not at all				extremely
ALOOF or UNINTERESTED	1	2	3	4	5
	not at all				extremely
PLAYFUL	1	2	3	4	5
	not at all				extremely
ENGAGED, ALERT (environment)	1	2	3	4	5
	not at all				extremely
INTERESTED IN PEOPLE	1	2	3	4	5
	not at all				extremely

Subtest 13: TUG-O-WAR OR TOY RELEASE

EASY TO TRAIN	1	2	3	4	5
	not at all				extremely
EASY TO CONTROL	1	2	3	4	5
	not at all				extremely
SUBMISSIVE	1	2	3	4	5
	not at all				extremely
PUSHY, ASSERTIVE	1	2	3	4	5
	not at all				extremely

Subtest 14: TRAIN NEW TASK

ENGAGED (environment)	1	2	3	4	5
	not at all				extremely
INTERESTED IN PEOPLE	1	2	3	4	5
	not at all				extremely
EASY TO TRAIN	1	2	3	4	5
	not at all				extremely
OBEDIENT	1	2	3	4	5
	not at all				extremely

Subtest 15: BASIC COMMANDS

OBEDIENT	1	2	3	4	5
	not at all				extremely

Appendix K. (*Continued*)

Subtest 16: OTHER DOG (OFF-LEASH)

FEARFUL	1 not at all	2	3	4	5 extremely
CONFIDENT	1 not at all	2	3	4	5 extremely
AGGRESSIVE	1 not at all	2	3	4	5 extremely
FRIENDLY	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
ALOOF or UNINTERESTED	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
SUBMISSIVE	1 not at all	2	3	4	5 extremely
PUSHY, ASSERTIVE	1 not at all	2	3	4	5 extremely

Subtest 17: ACTIVITY IN FREE PLAY

ACTIVE	1 not at all	2	3	4	5 extremely
EXCITABLE	1 not at all	2	3	4	5 extremely
CALM	1 not at all	2	3	4	5 extremely
ALOOF	1 not at all	2	3	4	5 extremely
PLAYFUL	1 not at all	2	3	4	5 extremely
ENGAGED, ALERT (environment)	1 not at all	2	3	4	5 extremely

Appendix K. (*Continued*)

Subtest 18: REUNION WITH OWNER

EXCITABLE	1	2	3	4	5
	not at all				extremely
CALM	1	2	3	4	5
	not at all				extremely
ALOOF	1	2	3	4	5
	not at all				extremely
INTERESTED IN PEOPLE	1	2	3	4	5
	not at all				extremely
EASY TO CONTROL	1	2	3	4	5
	not at all				extremely
SUBMISSIVE	1	2	3	4	5
	not at all				extremely
PUSHY, ASSERTIVE	1	2	3	4	5
	not at all				extremely

REFERENCES

Allport, G. W. (1937). *Personality: A Psychological Interpretation,* Holt, New York.

Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology, 44,* 1-26.

Beaudet, R., Chalifoux, A., & Dallaire, A. (1994). Predictive value of activity level and behavioral evaluation on future dominance in puppies. *Applied Animal Behaviour Science, 40,* 273-284.

Bentler, P. M. (1995). EQS: Structural equation models. *Psychological Bulletin, 107,* 238-246.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Bentler, P. M. & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods and Research, 16,* 78-117.

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research.* Springfield, IL: Charles C. Thomas.

Bogg, T. & Roberts, B. W. (2004). Conscientiousness and health behaviors: A meta-analysis. *Psychological Bulletin, 130,* 887-919.

Bollen, K. A. (1989). *Structural equation models with latent variables.* New York, NY: Wiley.

Bollen, K. A. & Long, J. S. (1992). Tests for structural equation models. *Sociological Methods and Research, 21,* 123-131.

Bradshaw, J. W. S., & Goodwin, D. (1998). Determination of behavioural traits of pure-bred dogs using factor analysis and cluster analysis; a comparison of studies in the USA and UK. *Research in Veterinary Science, 66,* 73-76.

Browne, M. W. & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research, 21,* 30-258.

Buirski, P., Plutchik, R., & Kellerman, H. (1978). Sex differences, dominance, and personality in the chimpanzee. *Animal Behaviour, 26,* 123-129.

Buss, A. H. (1995). *Personality: Temperament, social behavior, and the self.* Needham Heights, MA: Allyn & Bacon.

Campbell, W. E. (1972). A behavior test for puppy selection. *Modern Veterinary Practice, 12,* 29-33.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Carmines, E. G., & Melver, J. (1981). Analyzing models with unobserved variables: Analysis of covariance structures. In G. Bohrnstedt & E. Borgatta (Eds.), *Social measurements: Current issues* (pp. 65-116). Beverly Hills, CA: Sage.

Cattell, R. B. (1943). The description of personality, 1. Foundations of trait measurement. *Psychological Review, 50,* 559-594.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245-76.

Cattell, R. B., & Korth, B., (1973). The isolation of temperament dimensions in dogs. *Behavioral Biology, 9,* 15-30.

Comrey, A. L. (1973). *A first course in factor analysis.* New York, NY: Academic Press.

Coren, S. (1995). *The intelligence of dogs: A guide to the thoughts, emotions, and inner lives of our canine companions.* New York, NY: Free Press.

Coren, S. (1998). *Why we love the dogs we do: How to find a dog that matches your personality.* New York, NY: Free Press.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Dunayer, J. (2001). *Animal equality: Language and liberation.* Derwood, MD: Ryce.

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

Dugatkin, L. A. (2004). *Principles of animal behavior.* New York, NY: Norton.

Epstein, S. (1983). Aggregation and beyond: Some basic issues on the prediction of behavior. *Journal of Personality, 51,* 235-241.

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

Funder, D. C., & Ozer, D. J. (1983). Behavior as a function of a situation. *Journal of Personality and Social Psychology, 44,* 107-112.

Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The Riverside Behavioral Q-sort: A tool for description of social behavior. *Journal of Personality, 68,* 451-489.

Gilquin, G., & Jacobs, G. M. (2006). Elephants who marry mice are very unusual: The use of the relative pronoun *who* with nonhuman animals. *Society and Animals, 14,* 79-105.

Goddard, M. E., & Beilharz, R. G. (1982/83). Genetics of traits which determine the suitability of dogs as guide-dogs for the blind. *Applied Animal Ethology, 9,* 299-315.

Goddard, M. E., & Beilharz, R. G. (1984a). A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science, 12,* 253-265.

Goddard, M. E., & Beilharz, R. G. (1984b). The relationship of fearfulness to, and the effects of sex, age and experience on exploration and activity in dogs. *Applied Animal Behaviour Science, 12,* 267-278.

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

Goddard, M. E., & Beilharz, R. G. (1986). Early prediction of adult behavior in potential guide dogs. *Applied Animal Behaviour Science, 15,* 247-260.

Goldberg, L. R. (1982). From Ace to Zombie: Some explorations in the language of personality. In: Spielberger, C. D. and Butcher, J. N. (Eds), *Advances in Personality Assessment, 1* (pp. 203-234). Hillsdale, NJ: Erlbaum.

Goldberg, L. R. (2006). Doing it all bass-ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality, 40,* 347-358.

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

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment, 68,* 532-560.

Gosling, S. D. (1998). Personality dimensions in spotted hyenas (*Crocuta crocuta*). *Journal of Comparative Psychology, 112,* 107-118.

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

Gosling, S. D., & Bonnenburg, A. V. (1998). An integrative approach to personality research in anthrozoology: ratings of six species of pets and their owners. *Anthrozoös, 11,* 148-156.

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

Gosling, S. D., & Vazire, S. (2002). Are we barking up the right tree? Evaluating a comparative approach to personality. *Journal of Research in Personality, 36,* 607-614.

Gosling, S. D., Kwan, V. S. Y., & John, O. P. (2003a). A dog's got personality: A cross-species comparative approach to evaluating personality judgments. *Journal of Personality and Social Psychology, 85,* 1161-1169.

Gosling, S. D., Lilienfeld, S. O., & Marino, L. (2003b). Personality. In: Maestripieri, D. (Ed.), *Primate Psychology: The Mind and Behavior of Human and Nonhuman Primates* (pp. 254-288). Cambridge, MA: Harvard University Press.

Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology, 74,* 1337-1349.

Griffin, D., & Gonzalez, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin, 118,* 430-439

Guadagnoli, E. & Velicer, W. F. (1988). The relationship of sample size to the stability of component patterns. *Psychological Bulletin, 103,* 265-275.

Hart, B. L. (1995). Analyzing breed and gender differences in behaviour. In: Serpell, J. (Ed.), *The domestic dog: its evolution, behaviour and interactions with people* (pp. 65-78). Cambridge, England: Cambridge 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, 186,* 1181-1185.

Hart, B. L., & Miller, M. F. (1985). Behavioral profiles of dog breeds: A quantitative approach. *Journal of the American Veterinary Medical Association, 186,* 1175-1180.

Hart, B. L., Murray, S. R. J., Hahs, M., Cruz, B., & Miller, M. F. (1983). Breed-specific behavioral profiles of dogs: Model for a quantitative analysis. In A. H. Katcher & A. M. Beck (Eds.), *New perspectives on our lives with companion animals* (pp. 47-56). Philedelphia, PA: University of Pennsylvania Press.

Hart, L. A. (2000). Methods, standards, guidelines, and considerations in selecting animals for animal-assisted therapy: Part A: Understanding animal behavior, species, and temperament as applied to interactions with specific populations. In: Fine, A. H. (Ed.), *Handbook on animal-assisted therapy: Theoretical foundations and guidelines for practice* (pp. 81-97). San Diego, CA: Academic Press.

Heller, D., Watson, D., & Ilies, R. (2004). The role of person versus situation in life satisfaction: A critical examination. *Psychological Bulletin, 130,* 574-600.

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.

Hetzel, R. D. (1996). A primer on factor analysis with comments on patterns of practice and reporting. In B. Thompson (Ed.), *Advances in social science methodology, 4* (pp. 175-206). Greenwich, CT: JAI Press.

Hoffman, M. (1999). *Lend me an ear: The temperament, selection, and training of the hearing ear dog.* Wilsonville, OR: Doral Publishing,

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-185.

Hsu, Y., & Serpell, J. A. (2003). Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *Journal of the American Veterinary Medical Association, 223,* 1293-1300.

Humane Society of the United States (HSUS). (January 5, 2007). U. S. Pet Ownership Statistics. Retrieved March 26, 2007, from The Humane Society of the United States website: http://www.hsus.org/pets/issues_affecting_our_pets/pet_overpopulation_and_ownership_statistics/us_pet_ownership_statistics.html

Humphrey, E. S. (1934). "Mental Tests" for shepherd dogs: An attempted classification and evaluation of the various traits that go to make up "temperament" in the German Shepherd Dog. *Journal of Heredity, 25,* 128-136.

James, W. T. (1951). Social organization among dogs of different temperaments, terriers and beagles, reared together. *Journal of Comparative and Physiological Psychology, 44,* 71-77.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of Personality Theory and Research* (pp. 6-100), New York, NY: Guilford Press.

John, O. P., & Gosling, S. D. (2000). Personality traits. In: Kazdin, A. E. (Ed.), *Encyclopedia of Psychology, 6* (pp. 140-144). Washington, DC: American Psychological Association.

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Kennedy (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 719). New York: Guilford Press.

John, O. P. & Benet-Martínez, V. (2000). Measurement, scale construction, and reliability. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social Psychology* (pp. 339-369). Cambridge, England: Cambridge University Press.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2*(3), 171-203.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29,* 589-611.

Jones, A. C., & Gosling, S. D. (2005). Temperament and personality in dogs (*Canis familiaris*): A review and evaluation of past research. *Applied Animal Behaviour Science, 95,* 1-53.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the programs and applications.* Chicago, IL: SPSS.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Keeler, C. E., (1947). Coat color, physique, and temperament; materials for the synthesis of hereditary behavior trends in the lower mammals and man. *Journal of Heredity, 38,* 271-277.

Klag, S., Creed, P., & O'Callaghan, F. (2006). Development and initial validation of an instrument to measure perceived coercion to enter treatment for substance abuse. *Psychology of Addictive Behaviors, 20,* 463-470.

Klages, L. (1926). *The Science of Character* (Translated 1932). London, England: Allen & Unwin.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Kwan, V. S. Y., Gosling, S. D., & John, O. P. (2008). Anthropomorphism as a special case of social perception: A cross-species social relations model analysis of humans and dogs. *Social Cognition, 26,* 129-142.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cut-off criteria. What did they really say? *Organizational Research Methods, 9,* 202-220.

Ledger, R. A. (2003). Aggressive behaviour in dog breeds re-homed from rescue shelters. In: Ferrante V., & the Scientific Committee (Eds.), Proceedings of the 37th International Congress of the ISAE, Abano Terme, Italy. Fondazione Iniziative Zooprofilattiche e Zootechnicie, Brescie, Italy, p. 184.

Ledger, R., & Baxter, M. (1996). A validated test to assess the temperament of dogs. In I.J.H. Duncan, T.M. Widowski & D.B. Haley, (Eds.), Proceedings of the 30th International Congress of the ISAE, Guelph, Canada. Col. C.K. Centre for the Study of Animal Welfare, Canada, p. 111.

Ledger, R. A., & Baxter, M. R. (1997). The development of a validated test to assess the temperament of dogs in a rescue shelter. In D.S. Mills, S.E. Heath & L.J. Harrington (Eds.), Proceedings of the First International Conference on Veterinary Behavioral Medicine, Birmingham, UK. Universities Federation for Animal Welfare, United Kingdom, pp. 87-92.

Ledger, R. A., Baxter, M., & McNicholas, J. (1995). Temperament testing dogs in a rescue shelter: Improving owner-dog compatibility. In: Rutter, S. M., Rushen, J., Randle, H. D., & Eddison, J. C. (Eds.), Proceedings of the 29th International Congress of the ISAE, Exeter, UK. Universities Federation for Animal Welfare, United Kingdom, pp. 101-102.

Lester, D. (1983). Body build and temperament in dogs. *Perceptual and Motor Skills, 56,* 590.

Lindberg, S., Strandberg, E., & Swensen, L. (2004). Genetic analysis of hunting behaviour in Swedish Flatcoated Retrievers. *Applied Animal Behaviour Science, 88,* 289-298.

Lipsey, M.W., & Wilson, D.B. (1996). *Practical Meta-Analysis*. Newbury Park, CA: Sage.

MacCallum, R. C. (1990). The need for alternative measures of fit in covariance structure modeling. *Multivariate Behavioral Research, 25,* 157-162.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84-99.

Mahut, H. (1958). Breed differences in the dog's emotional behavior. *Canadian Journal of Psychology, 12,* 35-44.

McCabe, S. E., Couper, M. P., Cranford, J. A., & Boyd, C. J. (2006). Comparison of Web and mail surveys for studying secondary consequences associated with substance abuse: Evidence for minimal mode effects. *Addictive Behaviors, 31,* 162-168.

McCrae, R. R., Costa Jr., P. T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., Sanz, J., Sanchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78,* 173-186.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-48

McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of Web-delivered experiments: Can you trust the data? *Psychological Science, 11,* 502-506.

Mischel, W. (1968). *Personality and Assessment.* New York, NY: Wiley.

Mischel, W., & Peake, P. K., (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological Review, 89,* 730-755.

Morris, D., (2002). *Dogs: The Ultimate Dictionary of over 1,000 Breeds.* North Pomfret, VT: Trafalgar Square Publishing.

Murphy, J. A. (1995). Assessment of the temperament of potential guide dogs. *Anthrozoos, 8,* 224-228.

Murphy, J. A. (1998). Describing categories of temperament in potential guide dogs for the blind. *Applied Animal Behaviour Science, 58,* 163-178

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

Paulhus, D. L., & Bruce, M. N., (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63,* 816-824.

Pavlov, I. P., (1906). The Scientific Investigation of the Psychical Faculties or Processes in the Higher Animals. *Science, 24,* 613-619.

Pervin, L. & John, O. P. (1997). *Personality: Theory and research* (7th ed.). New York, NY: Wiley.

Podberscek, A. L., & Serpell, J. A. (1996). The English Cocker Spaniel: Preliminary findings on aggressive behavior. *Applied Animal Behaviour Science, 47,* 75-89.

Recklitis, C. J., Parsons, S. K., Shih, M.-C., Mertens, A., Robinson, L. L., & Zelter, L. (2006). Factor structure of the Brief Symptom Inventory—18 in adult survivors of childhood cancer: Results from the Childhood Cancer Survivor Study. *Psychological Assessment, 18,* 22-32.

Registration Statistics. Retrieved July 26, 2004, from http://www.akc.org/breeds/reg_stats.cfm

Reuterwall, C., & Ryman, N., (1973). An estimate of the magnitude of additive genetic variation of some mental characters in Alsatian dogs. *Hereditas, 73,* 277-284.

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

Roll, A., & Unshelm, J., (1997). Aggressive conflicts amongst dogs and factors affecting them. *Applied Animal Behavior Science, 52,* 229-242.

Rooney, N. J., Bradshaw, J. W. S., & Almey, H. (2004). Attributes of specialist search dogs-A questionnaire survey of UK dog handlers and trainers. *Journal of Forensic Science, 49,* 1-7.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Royce, J. R., (1955). A Factorial Study of Emotionality in the Dog. *Psychological Monographs, 69,* 1-27.

Ruefenacht, S., Gebhardt-Henrich, S., Miyake, T., & Gaillard, C., (2002). A behavior test on German Shepherd dogs: heritability of seven different traits. *Applied Animal Behaviour Scinece, 79,* 113-132.

Saetre, P., Strandberg, E., Sundgren, P.-E., Pettersson, U., Jazin, E., & Bergström, T. F. (2006). The genetic contribution to canine personality. *Genes, Brain and Behavior, 5*, 240–248

Scott, J. P., & Bielfelt, S. W., (1976). Analysis of the puppy testing program. In: Pfaffenberger, C. J., Scott, J. P., Fuller, J. L., Ginsburg, B. E., Bielfelt, S. W. (Eds.), *Guide Dogs for the Blind: Their Selection, Development and Training* pp. 39-75). Amsterdam, The Netherlands: Elsevier.

Scott, J. P., & Fuller, J. L., (1965). *Genetics and the social behavior of the dog.* Chicago, IL: University of Chicago Press.

Seksel, K., Mazurski, E. J., & Taylor, A., (1999). Puppy socialization programs: short and long term behavior effects. *Applied Animal Behaviour Science, 62,* 335-349.

Serpell, J. A., (1983). *The personality of the dog and its influence on the pet-owner bond.* In: Katcher, A. H., & Beck, A. M. (Eds.), New perspectives on our lives with companion animals (pp. 57-63). Philadelphia, PA: University of Pennsylvania Press.

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

Shrout, P.E. & Fleiss, J.L. (1979) Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin, 2*, 420-428.

Simms, L. J. & Watson, D. (2007). The construct validation approach to personality scale construction. In: Robins, R. W., Fraley, R. C., & Krueger, R. F. (Eds). *Handbook of Research Methods in Personality Psychology* (pp. 240-258) New York, NY: Guilford Press.

Slabbert, J. M., & Odendaal, J. S. J. (1999). Early prediction of adult police dog efficiency—a longitudinal study. *Applied Animal Behaviour Science, 64,* 269-288.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Lawrence Erlbaum Publishers.

Stephen, J. M., & Ledger, R. A. (2003). Owners are reliable observers of their own dog's behaviour. In V. Ferrante, & the Scientific Committee (Eds.), Proceedings of the 37th International Congress of the ISAE, Abano Terme, Italy. Fondazione Iniziative Zooprofilattiche e Zootechnicie, Brescie, Italy. p. 190.

Stephen, J. M., Ledger, R. A., & Stanton, N. (2001). Comparison of the perceptions of temperament in dogs by different members of the same household. In J.P. Garner, J.A. Mench, S.P. Heekin, (Eds.), Proceedings of the 35th International Congress of the ISAE, Davis, California, USA. Center For Animal Welfare, UC Davis, p. 113.

Stubbs, C. J., & Cook, M., (1999). Personality, anal character, and attitudes toward dogs. *Psychological Reports, 85,* 1089-1092.

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

Svartberg, K. (2005). A comparison of behaviour in test and in everyday life: evidence of three consistent boldness-related personality traits in dogs. *Applied Animal Behaviour Science, 91,* 103-128.

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

Svartberg, K., Tapper, I., Temrin, H., Radesater, T., & Thorman, S. (2005). Consistency of personality in dogs. *Animal Behaviour, 69,* 283-291.

Tabachnick, B. G., & Fidell, L.S. (2001). Using multivariate statistics (4^{th} ed.). New York, NY: Harper and Row.

Taylor, K. D. & Mills, D. S. (2006). The development and assessment of temperament tests for adult companion dogs. *Journal of Veterinary Behavior, 1,* 94-108.

Topál, J., Miklósi, Á, Csányi, V., & Dóka, A., (1998). Attachment behavior in dogs (Canis familiaris): A new application of Ainsworth's (1969) Strange Situation test. *Journal of Comparative Psychology, 112,* 219-229.

Tortora, D. F., (1983). *The Right Dog for You.* New York, NY: Simon and Schuster,

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

Vazire, S., Gosling, S. D., Dickey, A. S., & Shapiro, S. J. (2007). Measuring personality in non-human animals. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology.* New York, NY: Guilford Press.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate behavioral research, 25,* 1-28.

Wahlgren, K., & Lester, D., (2003). The big four: personality in dogs. *Psychological Reports, 92,* 828.

Walsh, W. & Betz, N. (1985). *Tests and assessments.* Upper Saddle River, NY: Prentice-Hall.

Walsh, J. P. & Kiesler, S., Sproull, L. S., & Hesse, B. W. (1992). Self-selected and randomly selected respondents in a computer network survey. *Public Opinion Quarterly, 56,* 241-244.

Weiss, E. (2002). Selecting shelter dogs for service dog training. *Journal of Applied Animal Welfare Science, 5,* 43-62.

Weiss, E., & Greenberg, G., (1997). Service dog selection tests: effectiveness for dogs from animal shelters. *Applied Animal Behaviour Science, 53,* 297-308.

Wilcox, B., & Walkowicz, C., (1995). *The atlas of dog breeds of the world.* 5th ed. Neptune, NJ: TFH Publications.

Wilsson, E., & Sundgren, P.-E., (1997). The use of a behavior test for the selection of dogs for service and breeding, I: Method of testing and evaluating test results in the adult dog, demands on different kinds of service dogs, sex and breed differences. *Applied Animal Behaviour Science, 53,* 279-295.

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.

Worthington, R. L. & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34,* 806-838.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate behavioral research, 17,* 253-269.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432-442.

VITA

Amanda Claire Jones, born on March 8, 1980, in Garden City, Michigan. Amanda graduated from Chelsea High School in Chelsea, Michigan, in June, 1998. She completed her undergraduate education at the Residential College at the University of Michigan, Ann Arbor. Amanda also attended The Language Schools at Middlebury College in Middlebury, Vermont, and studied at Albert-Ludwigs-Universität Freiburg im Breisgau, Germany. In December, 2001, she graduated with a Bachelor of Arts in Linguistics and in German. In Fall of 2003, Amanda began her graduate work at the University of Texas at Austin, working towards a doctorate in Social and Personality Psychology. While in graduate school, Amanda authored the article "Sensory development in puppies (*Canis lupus familiaris*): Implications for improving canine welfare," in *Animal Welfare*. She also co-authored "Temperament and personality in dogs (*Canis familiaris*): A review and evaluation of past research," in *Applied Animal Behaviour Science*; "Are we dog's best friend? Predicting canine cortisol response from human affiliative and punitive behaviors," in *Proceedings of the International Veterinary Behavior Meeting, American College of Veterinary Behaviorists*; and "Interspecies hormonal interactions between man and the domestic dog (*Canis familiaris*)," in *Hormones and Behavior*. Amanda was also a recipient of numerous awards, including Best Graduate Research poster at the annual meeting of the Association for Research in Personality (2006) and fellowships, including the National Science Foundation Graduate Research Fellowship (2002).

Permanent address: 1902 Blue Crest Drive, Unit A; Austin, Texas 78704

This dissertation was typed by the author.