A Dog's Got Personality: A Cross-Species Comparative Approach to Personality Judgments in Dogs and Humans

Samuel D. Gosling
University of Texas at Austin

Virginia S. Y. Kwan
Princeton University

Oliver P. John
University of California, Berkeley

This research offers a blueprint for how a cross-species comparative approach can be realized empirically. In a single design, parallel procedures and instruments were used in 2 species, dogs (*Canis familiaris*) and humans (*Homo sapiens*), to test whether personality differences exist and can be judged in dogs as accurately as in humans. Personality judgments of humans and dogs were compared on 3 accuracy criteria: internal consistency, consensus, and correspondence. Results showed that, on all 3 criteria, judgments of dogs were as accurate as judgments of humans. These findings are consistent with the evolutionary continuity hypothesis and suggest an important conclusion not widely considered by either personality or animal researchers: Personality differences do exist and can be measured in animals other than humans.

"A dog's got personality and personality goes a long way." So said Jules Winnfield, a gangster from the movie *Pulp Fiction*. Fictional gangsters are not the only ones concerned with personality in nonhuman animals. Robert Fagen, a professor of Biometry, used the personality traits "irascible, irritable, manipulative, and grumpy" to describe Suzy June, a brown bear he had observed for several years (Aschenbach, 1995). Although scientists such as Fagen are beginning to apply personality constructs to animals, a systematic empirical evaluation of such personality judgments has yet to be performed. In this report, we focus on personality judgments of domestic dogs, directly comparing them with personality judgments of humans.

Animal models have played a central role in much psychological science (Domjan & Purdy, 1995). Yet, although scientists widely accept the idea that the anatomy and physiology of humans show considerable continuity to other mammals, most have been reluctant to ascribe emotions and personality traits to animals. As

a result, there are few systematic studies on animals and cross-disciplinary bridges between personality psychologists and animal-behavior researchers are virtually nonexistent (Gosling, 2001). Most questions about animal personality remain untested, and a field of animal personality remains to be developed.

However, there is nothing in evolutionary theory to suggest that only physical traits are subject to selection pressures. Indeed, Darwin (1872/1998) argued explicitly that emotions exist in both human and nonhuman animals, including primates, cats, and dogs. Similarly, personality traits like Neuroticism, Extraversion, and Agreeableness may not be as uniquely human as once was thought (Buss, 1988; Gosling & John, 1999; Hogan, 1996). In this article, we examine a core issue for animal-personality research: whether differences in personality traits exist and can be measured in animals.

To show that personality traits exist in nonhuman animals, assessments of animal personality must be proven accurate. A small number of studies have reported reliabilities, but systematic efforts to examine the accuracy of assessment methods have been rare (see Gosling, 2001, for a review of the studies available). This dearth of studies may reflect the challenges of assessing creatures that cannot speak to us and may not have the mental or physical faculties (e.g., conceptions of self, autobiographical memory) required by the tests of personality developed for humans. However, there is one method that may be suited for use in animals—judgments made by informants who are well acquainted with the target individuals. It should be noted that this *personality-judgment approach* (Funder, 1999) is far from a method of last resort. As Hofstee (1994) noted about personality in humans, "The averaged judgment of knowledgeable others provides the best available point of reference for both the definition of personality structure in general and for assessing someone's personality in particular" (p. 149). Thus, many human-personality researchers consider judg-

ments by knowledgeable informants or observers the sine qua non of personality (Funder, 1995; Kenny, 1994; McCrae, 1982; Wiggins, 1973), making such judgments a logical candidate for studies of animals. We thus propose to extend the personality-judgment approach to research on animals; specifically, we test whether personality traits exist and can be judged in one particular nonhuman species, dogs.

Introducing a Cross-Species Comparative Approach to Personality

How should such research be performed? The logic of comparative research suggests that multiple species should be examined, preferably using at least one well-studied species. Such cross-species designs are useful because they provide researchers with a benchmark against which to evaluate results for a newly examined species. Given that we know more about the personality of humans than of any other species, humans should be one of the species included.

What should the other species be? Animal-personality research is most likely to be conducted in research facilities, zoos, farms, and animal shelters, on species such as rhesus monkeys, chimpanzees, sheep, and dogs. We focus on domestic dogs (*Canis familiaris*) because (a) they are widely owned, making them a readily available source of subjects, (b) they can safely and naturally engage in wide array of behaviors, (c) their behavioral repertoire is well understood by a large pool of observers, and (d) they can travel safely to research sites without specialist care. Together, these reasons make them a manageable species for research purposes (cf. lions, elephants), especially in the nascent stages of research on animal personality.

Domestic dogs are found throughout the world, living alongside humans, an association traced as far back as 14,000 to 20,000 years ago. The evolutionary origins of domestic dogs are still subject to debate (Nowak, 1999). Some authorities have suggested that modern dogs were derived from one or a few wolf subspecies and then spread throughout the world in association with people (Nowak, 1979). Others have suggested that domestication occurred at different times and different locations with humans domesticating whichever local species of canid happened to be around (Coren, 1994). In any event, many of the modern morphological and behavioral characteristics associated with modern domestic dogs have been selected within the context of human–dog relationships since domestication (Hare, Brown, Williamson, & Tomasello, 2002).

Criteria for Evaluating Accuracy

If we are going to use human judgments as data for animal-personality research, these judgments need to be subjected to the same rigorous accuracy evaluations as judgments of human personality (Funder, 1995). Although the specific criteria and how they are labeled vary somewhat from theorist to theorist, human personality judgments have been evaluated with respect to three major accuracy criteria: internal consistency, consensus, and correspondence (e.g., Kruglanski, 1989; Robins & John, 1997). Internal consistency reflects the degree to which judgments about an individual's personality are consistent across observations or items thought to reflect the same behavioral dimension (Robins & John,

1997; Wiggins, 1973). Consensus implies agreement among independent informants or observers, and is often considered the hallmark of accuracy (e.g., Block, 1961; Funder, 1995; Kenny, 1994; Kruglanski, 1989; McCrae, 1982). Correspondence refers to the extent to which judgments predict an external criterion for "reality" (Kruglanski, 1989); perhaps the most valuable external criterion is independent observations of behaviors (Funder, 1995; Kenny, 1994). In the present research, we use these three accuracy criteria, developed in human research, to evaluate the accuracy of personality judgments of dogs. For each criterion, we compare the dog findings with parallel findings from humans. If personality traits do not exist in dogs, then judgments of dogs should fall short on these criteria when compared with judgments of humans.

The most knowledgeable informants about dogs are typically their owners, who have extensively observed the animals in their care, across both situations and time. Therefore, we focused our judgment approach on how the personality of dogs is judged by their owners. To estimate *consensus* between the owner and an independent judge, we obtained judgments of the dogs by a second person familiar with the target animals (a "peer"). *Internal consistency* was studied in the judgments of both informants (owners and peers). Finally, we tested the *correspondence* of the owner's judgments by comparing them with the dog's behavior rated by independent observers in a local dog park.

Judgments of dogs, just like judgments of people, may be based on physical and appearance characteristics, rather than actual behavior. Thus, we also took photographs of the dogs, permitting us to obtain personality assessments on the basis of appearance alone; these data allowed us to test the extent to which assessments of dogs are affected by breed and other appearance stereotypes. In addition, we controlled for such other background characteristics as sex and age that can influence personality judgments.

To implement a cross-species comparative approach, we obtained personality judgments not only of the dogs but also of their human owners. In this cross-species design, we used parallel procedures, instruments, and constructs for both dogs and humans, allowing us to compare dog findings directly with human findings within the same study.

Lessons Learned From Previous Research

Studies of animal personality are isolated and few and far between. Reports are scattered across a multitude of disciplines and journals, ranging from veterinary medicine and zoology to agricultural science and psychology (e.g., Capitanio, 1999; Fairbanks, 2001; King & Figueredo, 1997; Sinn, Perrin, Mather, & Anderson, 2001). With few systematic attempts to assess personality in dogs and other nonhuman species, comprehensive evaluations of the accuracy of personality judgments have not been possible. Nonetheless, a number of useful lessons can be learned by surveying the limitations of the few studies that do exist. These lessons have guided the design of the present research.

Previous studies of personality in dogs have not been comprehensive in their coverage of relevant traits (e.g., Murphy, 1995) and behavioral domains (e.g., Cattell & Korth, 1973), even though the importance of examining a broad array of trait dimensions is now widely understood in the human literature (John & Srivastava, 1999). Another important reason for examining multiple traits in the same study is to address discriminant validity. Are personality

ratings on multiple dimensions independent and valid, or do they simply reflect a single evaluative dimension, such as "like–dislike"? To address this possibility, human research on personality judgments has assessed multiple traits simultaneously, permitting an evaluation of each dimension while controlling for judgments of the other dimensions. The same now needs to be done with personality dimensions in dogs.

In some cases, previous studies of dogs have assessed personality dimensions using only one item (e.g., Murphy, 1995; Slabbert & Odendaal, 1999). Single items are problematic because they provide less reliable measurements than do multiple indicators and do not permit the assessment of internal consistency (Gosling, Rentfrow, & Swann, in press). Thus, personality dimensions should be estimated using multiple items.

Moreover, research on dogs has focused on only one accuracy criterion at a time, rather than systematically testing all the important aspects of accuracy. For example, there may be internal consistency, or there may be consensus, but if judgments reflect only shared stereotypes about breed, then there may be little correspondence. Similarly, if consensus is low, correspondence might be severely underestimated. Thus, all three accuracy criteria need to be tested in the same study.

Most previous studies of dogs have been done in applied contexts where the behavior of dogs is of interest in and of itself (e.g., Slabbert & Odendaal, 1999; Svartberg, 2002). By focusing on one species, rather than taking the comparative approach proposed here, these studies make it difficult to evaluate the broader significance of the findings. Should a consensus correlation of .40 be considered large or small? Some kind of benchmark is needed against which the findings from dogs can be evaluated. The accuracy of personality judgments has been studied most extensively in humans, making them the most appropriate benchmark species.

One limitation of some of the previous research is that it has examined dog breeds globally, rather than judgments of individual dogs (e.g., Coren, 1998; Hart & Miller, 1985). These studies have shown good consensus among experts in judging breed-typical characteristics; however, they do not address the accuracy of judgments of individual dogs. To illustrate this point, consider that judges might well agree about the stereotypical traits associated with the peoples of England, Hong Kong, and Germany, yet relying on these national stereotypes may be of little value for judging the personalities of particular individuals. Indeed, such national stereotypes may even lead judges to overlook the individuating behaviors of the individuals. Nonetheless, the previous work showing personality differences among dog breeds is important for interpreting findings on differences among individual dogs; researchers must ensure that personality judgments indeed reflect real behaviors, rather than impressions based on breed differences or other stereotypes and biases (e.g., based on sex, age, size, or appearance).

In conclusion, to achieve a more complete understanding and evaluation of dog-personality judgments, research should: (a) include a comprehensive set of trait dimensions, (b) measure each dimension with multiple items, (c) evaluate multiple accuracy criteria, (d) assess two or more species in a comparative design, and (e) assess judgments of individuals (rather than groups or breeds) while controlling for the potentially biasing effects of stereotypes, such as those related to sex, age, and physical appearance. This kind of comprehensive species-comparative design has

not yet been implemented but is urgently needed if the field of animal personality is to emerge from its infancy.

Overview of Studies

Dog owners and their dogs were recruited in a local dog park to participate in three studies. In Study 1, each owner provided personality judgments of their dog as well as of their own personality. The owners also identified another person (a "peer") who was familiar with both the target dog and the owner and could thus judge both their personalities. This cross-species design allowed us to compare (a) the internal consistency of dog and human personality judgments and (b) the consensus between owner and peer in judging either dog or human personality. In Study 2, owners brought their dogs to a field-testing enclosure located at a dog park, where the dogs' behaviors were observed and rated by three independent observers in an observational field-testing session. To assess the correspondence criterion of accuracy, we tested how well the owners' personality judgments of their dogs predicted the behavior ratings obtained in the field-testing sessions. In Study 3, photographs of the dogs (taken at the dog park) were rated by a new set of observers. These photo-based ratings allowed us to examine the effects of breed and appearance characteristics.

Personality Judgments by Owner and Peer Informants: The Five-Factor Model (FFM) as a Framework

Which personality dimensions should be assessed? Although several factor-analytic studies have identified dimensions of dog personality (e.g., Cattell & Korth, 1973; Svartberg & Forkman, 2002), no single model has been adopted by the field. Using the FFM as an organizing framework, Gosling and John (1999) summarized the structural findings from the factor analytic studies of dogs and 11 other species. One of the most striking findings to emerge was that Conscientiousness did not appear as an independent personality dimension in dogs; in fact, Conscientiousness appeared only in humans and humans' closest relatives, namely chimpanzees (King & Figueredo, 1997). This pattern of findings suggests that Conscientiousness appeared as a distinct dimension of personality relatively recently in evolutionary history, long after canid ancestors diverged from hominid ancestors (Gosling & John, 1999).

These review findings, along with the results of a series of comprehensive trait studies (Gosling & John, 1998), suggest that at this point, a four-dimensional model is the most promising for personality traits in dogs. These four dimensions represent canine analogs of four of the five human FFM factors: Energy (analogous to human Extraversion), Affection (analogous to human Agreeableness), Emotional Reactivity (analogous to human Neuroticism), and Intelligence (analogous to human Openness/Intellect). Thus, we focused on these four dimensions, which currently provide the most reasonable, albeit provisional, model for organizing personality judgments of dogs. With no evidence for a separate Conscientiousness dimension in any species other than humans and chimpanzees, it would make little sense to assess this trait in dogs.

Study 1: Internal Consistency and Consensus

Method

Participants. Seventy-eight owners (67% women) rated their own personality and their dog's (50% female) personality, and returned their judgments by mail; peer informants rated the personality of owners and dogs and returned their judgments directly to the experimenter by mail. The owners reported the age and sex of the dogs.

Personality judgments. Judgments of humans were made using a standard FFM instrument, the Big Five Inventory (BFI; see John & Srivastava, 1999). For the dog judgments, the BFI was adapted slightly. Two experts reviewed each item on the human instrument to determine whether it was applicable to canine targets. Most items could be applied to canine behavior after minor editing. Care was taken to retain the original sense of the items. For example, the item "Is original, comes up with new ideas" was changed to "Is original, comes up with new ways of doing things." Only one item ("Has few artistic interests") could not be translated to a canine form, and was therefore omitted from both human and canine BFIs (the canine version of the BFI is available from the authors).

The judges indicated the degree to which each item was characteristic of the target (dog or human owner) on a scale ranging from 1 (*disagree strongly*) to 5 (*agree strongly*). There were eight items for Extraversion/Energy (e.g., "Is full of energy"), nine for Agreeableness/Affection (e.g., "Is cooperative"), eight for Neuroticism/Emotional Reactivity (e.g., "Can be tense"), and nine for Openness/Intelligence (e.g., "Is curious about many different things"). The variances did not vary across human and dog targets, with mean standard deviations of .76 and .78, respectively.

Results and Discussion

Internal consistency: Are personality judgments of dogs consistent across items? We computed Cronbach's coefficient alpha across the items on each BFI scale, and these values are shown in Table 1 for both human and dog targets. For humans, the alphas averaged .82 for the owners' self-judgments and .85 for the peers' judgments of the owners; these values are similar to previous research. How do the values for the dogs compare? The mean alpha was .83 for the owners' judgments of their dogs and .82 for the peers' judgments of the same dogs—values quite similar to those for human targets even though the BFI was derived in research on humans. In short, the personality judgments showed substantial internal consistencies for both species.

Consensus: Do owner judgments agree with peer judgments? To assess consensus we computed unit-weighted scale scores for humans and dogs for each of the four BFI dimensions. To provide

a human comparison standard, Table 1 shows the correlations between the owners' self-judgments and how they were judged by the peer informants. These human consensus correlations were strong, averaging .55, and quite similar to previous research on human personality (e.g., Funder, Kolar, & Blackman, 1995; McCrae & Costa, 1987). What about the dogs? Table 1 shows that the four consensus correlations for canine targets were all significant and averaged .62. That is, they were substantial in size and at least as large as those for human targets, suggesting that owners can judge the personality of their dogs with substantial consensus.

Could these substantial consensus correlations be due to nonbehavioral characteristics of the animals? That is, one possible artifactual explanation for consensus is that instead of making judgments on the basis of behavior, informants might have made judgments on the basis of such nonbehavioral variables as the dogs' sex or age. To find out, we estimated consensus using partial correlations. The partial correlations were essentially unchanged; averaged across the BFI scales, the partial consensus correlations controlling for sex and age were .62 and .61, respectively—virtually the same as the mean zero-order correlation of .62. Thus, consensus cannot be attributed simply to judges relying on sex or age stereotypes about dogs.

A second possible explanation for the consensus findings is that judges did not actually discriminate among the FFM dimensions but formed only a single impression (such as like–dislike), and then made their BFI judgments on some inferential basis (e.g., Borkenau, 1992). If so, each of the four dimensions should be highly correlated with the other three. Thus, when controlling for the discriminant correlations by partialing out the other three BFI scales, there should be no unique consensus left—the partial consensus correlations should all no longer be significant and approach zero. This was not the case, however, as the partial consensus correlations in Table 1 show. They averaged .56 for dogs, as compared with .52 for humans. Thus, our consensus findings cannot be explained by the intercorrelations among the BFI scales, neither for the humans nor the dogs.

Study 2: Correspondence

Probably the most important evidence for accuracy is whether personality judgments can predict external criteria. In the case of personality traits, the definitive test is whether personality judg-

Table 1

Internal Consistency and Consensus: Personality Judgments of Humans and Dogs

| BFI scales | Internal consistency (Cronbach's α) | | | | Consensus (owner–peer correlations) | |
| | Owner judgments of | | Peer judgments of | | Target judged | |
	Human (self)	Own dog	Human owner	Dog	Human owner	Dog
Extraversion	.83	.77	.84	.81	.66* (.61*)	.76* (.76*)
Agreeableness	.81	.84	.84	.83	.47* (.43*)	.55* (.43*)
Neuroticism	.80	.89	.86	.86	.45* (.43*)	.57* (.51*)
Openness	.83	.81	.84	.75	.58* (.60*)	.55* (.47*)
M	.82	.83	.85	.82	.55* (.52*)	.62* (.56*)

Note. Numbers shown in parentheses are partial correlations remaining after all discriminant correlations have been controlled. BFI = Big Five Inventory.
* $p < .05$.

ments predict behaviors. How should such behaviors be assessed? Funder, Furr, and Colvin (2000) observed that most measures of social behavior in psychological research have two major limitations: (a) from an ecological point of view, the behaviors examined are often intrinsically uninteresting (e.g., response latencies to stimuli), and (b) measures typically focus on an extremely small number (usually just one) of the many specific behaviors an individual actually emits. Thus, Funder et al. (2000) recommended focusing on behavioral assessments that are "psychologically meaningful and relevant to individuals in behavioral interaction, but that would also require a minimum of subjective interpretation on the part of the coders" (p. 454). This midlevel analysis is above that of narrow molecular units but below the level of broad cross-situational behavioral trends. Thus, in the context of the behavior of dogs, rather than attempting to measure the frequency of head shakes, the velocity of a scamper, or the amplitude of a tail wag, a midlevel approach would require that observers rate, for example, whether a dog performed *shy* behaviors or *nervous* behaviors in a specific situation.

The midlevel approach is well-suited to the study of dog behavior because it permits the measurement of situation-specific behaviors that are a priori related to the traits under study while retaining the breadth that allows multiple traits to be assessed in a naturalistic setting. We adopted this approach in the present research because we wanted to capture behavior relevant to all four of the trait domains we had examined in Study 1. Thus, correspondence in this study reflects the degree to which the owners' personality judgments of their dogs correspond with trait-relevant dog behaviors rated by independent observers in a field-testing session.

Method

For the behavioral field-testing session, owners brought the target animals to a fenced-off region of a dog park. To elicit a broad array of behaviors, each animal performed several tasks: *getting acquainted* (owner went for short walk with the three observers), *express energy* (owner instructed to run with dog), *show affection* (owner encouraged to elicit affection from dog), *obedience* (performed basic obedience tricks, "sit" and "stay"), *social anxiety/stress* (stranger took dog for short walk and dog watched its owner walk another dog), and *problem solving* (dog shown small edible dog biscuit placed under plastic cup). Three independent observers who did not know the dog beforehand observed the dog's behavior across all tasks and then rated its behavior in this field-testing session on three behavioral markers for each of the dimensions under study (e.g., *nervous* for Neuroticism, *shy* [reversed] for Extraversion; the full rating instrument is available from the authors). Unit-weighted scale scores were computed from these markers for each of the four dimensions.

Results and Discussion

Correspondence: Do owner judgments predict behaviors in a field-testing session? To test this accuracy criterion, we examined the correlations between owners' personality judgments and independent behavior ratings in the field session. Significant correspondence correlations would be particularly impressive because the owner personality judgments and the behavior ratings represent different kinds of information (Block & Block, 1980): They were obtained from different sources (owner vs. independent observers), instruments (adapted human BFI vs. behavioral ratings), situations

(everyday behaviors vs. responses in a specific field-testing session), and time (separated by several weeks). In other words, these two reports shared little method variance, and significant findings would provide strong evidence for the existence of personality traits and the accuracy of the owner judgments.

There are surprisingly few trait-behavior studies on humans that could serve as an appropriate benchmark comparison for our correspondence correlations in dogs. The closest equivalent in the human literature would seem to be the so-called zero-acquaintance studies, which show rather small correspondence correlations for most FFM traits (e.g., Norman & Goldberg, 1966; Watson, 1989). However, unlike our study of dogs, most of these studies provide very little behavioral information about the targets. One exception is a study by Paulhus and Bruce (1992), in which personality judgments were made about humans in a similar task, namely a single-session group interaction, in which four to six observers rated the behavior of each human subject who had also provided self-judgments on the FFM dimensions. In this study, the mean correlation between personality judgments and behavior ratings was .22, a finding typical in the human literature. For comparison purposes, Table 2 presents these human correspondence correlations.

What about the dogs? Their correspondence correlations were at least as strong as the human ones. As shown in Table 2, all four correspondence correlations for the dogs were significant, and averaged .27. As in the case of consensus, correspondence could not be explained by nonbehavioral attributes; after controlling for the dogs' sex and age, the mean correspondence correlations were .28 and .27 (both *p*s < .05), respectively. Moreover, even when the effects of the other three scales were partialed out (i.e., controlling for all discriminant correlations), correspondence remained significant, averaging .26 (*p* < .05), still as strong as the values found in the human literature (comparable partial correlations were not reported in Paulhus & Bruce, 1992).

Study 3: Ruling Out Breed and Appearance Effects

The correspondence correlations obtained in Study 2 seem impressive. However, human observers were needed to translate the

Table 2

Correspondence: Zero-Order Correlations Between Owners' Personality Judgments and Behavior Ratings in the Field-Testing Session (and Partial Correlations After Controlling for Discriminant Correlations)

	Target judged	
Big Five scales	Human[a]	Dog
Extraversion	.35*	.32* (.34*)
Agreeableness	.01	.33* (.24*)
Neuroticism	.25*	.21* (.24*)
Openness	.27*	.23* (.20*)
M	.22*	.27* (.26*)

Note. Numbers shown in parentheses are partial correlations remaining after all discriminant correlations have been controlled.
[a] Data from Paulhus and Bruce (1992, Table 4), comparing self-reports with observer ratings. Comparable partial correlations were not available.
* *p* < .05.

behavioral conduct of the dogs in the field session into quantitative ratings of trait-relevant behaviors. Thus, it is possible that nonbehavioral variables influenced the behavior ratings. One possible type of confounding variable is represented by variables like the age and sex of the dogs; we ruled out this possibility in Study 2.

Another possibility involves commonly held beliefs about dog breeds. Dogs have been domesticated for thousands of years during which humans have applied selective breeding techniques to create many distinct breeds that differ systematically in their morphological and behavioral traits (Coren, 1994). Thus, certain breeds or groups of dogs can be associated with certain behaviors. One possibility, therefore, is that the observers based their ratings on commonly held beliefs about breed characteristics rather than on the actual behaviors of the individual dog they observed in the field sessions. For example, they might have rated a golden retriever as high on the behavioral items related to Affection and a pit bull as low, regardless of the individual dogs' performance on the behavioral tests. Or these shared breed stereotypes might have influenced how the observers interpreted the dogs' behaviors, leading observers to interpret ambiguous behaviors in ways consistent with breed stereotypes.

Only 51% of the dogs assessed were pure breeds. However, even for mixed-breed dogs it is often possible to make a reasonable guess about which breeds have been mixed. So breed stereotypes could affect observers' ratings of dogs of mixed heritage, too. Moreover, observers could rely on general appearance stereotypes not necessarily tied to specific breeds, such as large versus small size, smooth versus long hair, or pure versus mixed breed. Therefore, the goal of Study 3 was to examine the hypothesis that the correspondence correlations in Study 2 may be due to observers' use of appearance-based impressions. Thus, we recruited a new sample of observers to obtain appearance-based ratings of the likely behavior of each of the dogs rated in Study 2.

Method

Each of the dogs from Study 2 had been photographed alone. In subsequent laboratory sessions, the photographs were projected onto a screen and a new set of 6–11 independent judges rated their impressions of each dog. Using exactly the same behavior rating form as the observers in Study 2, the judges rated their impressions of the dogs' likely behavior on the three behavioral markers for each of the BFI dimensions. Thus, the design of Study 3 matched the design of Study 2, with the one difference that judges in Study 3 did not observe the dogs' actual behaviors in the field session and thus had to base their impressions purely on the physical appearance of the dogs.

Results and Discussion

We first tested whether the appearance-based impressions were systematic—that is, did the judges exposed only to the photos of the dogs show agreement in their impressions? Coefficient alpha reliability ranged from .69 (Openness) to .84 (Extraversion), with a mean of .76, showing that the judges did indeed agree about the likely behavior of the dogs.

Next, we tested whether photo-based ratings correlated with the behavior-based ratings. As shown in Table 3, the mean correlation was small but positive (.18), with significant correlations between the photo ratings and the behavior ratings for Agreeableness and Neuroticism. There are two alternative interpretations of these

Table 3

Correlations Between Photo Ratings and Field Behavior Ratings and Correspondence Correlations After Controlling for Appearance Stereotypes

BFI scales	Correlation between photo-based ratings and behavior ratings	Correspondence controlling for appearance
Extraversion	.17	.28*
Agreeableness	.20*	.33*
Neuroticism	.30*	.21*
Openness	.04	.23*
M	.18	.26*

Note. Correspondence correlations are partial correlations between the owners' judgments of their dogs and the field-test behavior ratings from Study 2, controlling for the appearance-based impressions from the photo judges. BFI = Big Five Inventory.
* *p* < .05.

correlations. One possibility is that the photo judges were using their knowledge about real breed differences to make valid inferences about the dogs' likely behaviors, and the dogs indeed manifested these real breed differences in the field-testing session. A second possibility is that both the photo judges and the behavior observers were biased by the same stereotypic beliefs about the dog breeds.

In either case, the question remains whether the significant correspondence correlations obtained in Study 2 could be explained by appearance-based impressions alone. We thus computed partial correlations between the owner judgments and the field behavior ratings from Study 2, controlling for the appearance-based impressions from the photo judges in this study. If the correspondence correlations could be explained by the field observers' reliance on breed or appearance-based stereotypes, then the partial correlations removing these effects should no longer be significant and should approach zero.

This was not the case. Instead, the partialed correspondence correlations in Table 3 (mean partial *r* = .26) were almost identical to the zero-order correlations (mean *r* = .27). Thus, the correspondence between owner judgments of personality and behavior ratings in the field session cannot be attributed to shared beliefs on the basis of the dogs' appearance. These findings further solidify the evidence for the accuracy of owners' judgments of their dogs' personalities.

General Discussion

Taken together, these studies show that personality traits can be judged in dogs with impressive levels of accuracy. Our two-species comparative approach made it possible for the first time to interpret the animal findings in direct comparison with findings based on human participants. We were able to show that the consistency, consensus, and correspondence of dog personality judgments were not only significant but also as substantial in size as those found for humans.

By taking an integrative approach to issues examined previously in piecemeal fashion, this research represents a departure from the past. As such, we suggest that this research can provide a blueprint, or prototype, for conducting species-comparative studies of per-

sonality. Specifically, this research examined a comprehensive array of dog-relevant personality dimensions, thus permitting comparisons across traits, and tests of discriminant validity with partial correlations. In addition, we included controls for nonbehavioral variables, such as sex and age, that could have influenced the judgments. We used multiple indicators to measure each dimension, allowing us to test the internal consistency of the judgments. Moreover, we evaluated personality judgments according to a comprehensive set of accuracy criteria. To provide a benchmark against which the accuracy criteria could be tested, we implemented the cross-species comparative approach, simultaneously examining two species, including the species (humans) and personality trait model (the FFM) upon which most research on personality judgment so far has been based. Finally, in an additional study, we controlled for the potentially biasing effects of breed and appearance-based stereotypes on personality judgments of individual dogs.

The findings emerging from this research are important because they suggest a conclusion not widely considered by either human-personality or animal-behavior researchers: Differences in personality traits do exist and can be measured in animals. This conclusion also converges with another newly emerging line of research indicating that behavioral traits in nonhuman animals are heritable (e.g., Weiss, King, & Enns, 2002). The time has come, we suggest, to extend Darwin's argument of cross-species continuity to the domain of personality. Below we discuss some directions that such extensions might take.

Limitations and Future Directions

Our findings suggest that the personality judgment approach can be applied successfully to nonhuman populations. Nonetheless, the boundary conditions of these effects, obtained here in a sample of domestic dogs and their owners in an urban dog park, still need to be established. Future research should examine whether these findings generalize to other contexts and populations. One extension would be to conduct studies with seeing-eye dogs or bomb-sniffing dogs whose personalities could be judged by their trainers or handlers. Other types of criteria could then be used to evaluate the correspondence criterion of judgmental accuracy, such as performance in standardized working-dog trials or subsequent on-the-job success as a working dog (e.g., Goddard & Beilharz, 1983; Slabbert & Odendaal, 1999; Svartberg, 2002). Such work is important because it could open the way for obtaining judgments by trainers and handlers in controlled studies of dog personality.

Even more important from the perspective of the cross-species comparative approach, this work needs to be extended to other species. There is some encouraging evidence that the findings for personality judgments of dogs may indeed generalize to some other species of mammals, at least on some accuracy criteria. In terms of correspondence, for example, the pioneering studies of Stevenson-Hinde (e.g., 1983) found that personality judgments of rhesus monkeys (e.g., on aggressiveness) corresponded with behavior observed in specific situations. In terms of consensus, a study of spotted hyenas showed consensus for traits related to four of the human FFM, such as active, aggressive, fearful, and imaginative (Gosling, 1998).

An additional limitation of the present work was our reliance on the midlevel assessment of behaviors (Funder et al., 2000), in which we used behavior ratings rather than detailed codings or frequency counts of specific acts. Our multidimensional approach necessitated the use of these midlevel assessments because fine-grained recordings of behaviors are extremely time-consuming and thus would have been prohibitive in a study of more than a few traits. Nonetheless, future studies should examine other kinds of correspondence criteria. Work in ethology suggests that naturalistic observation and coding techniques may be used to assess correspondence in nonhuman animals. For example, in piglets, behavioral records of vocalizations, nose contacts, and location in the pen can serve as behavioral markers for a personality dimension akin to human sociability (Forkman, Furuhaug, & Jensen, 1995). Studies of these new species and new criteria should be a top priority for the next generation of studies.

Implications and Conclusions

Broadly, this first cross-species personality study represents a necessary and fundamental step toward bridging the gap between personality and animal research. To illustrate the benefits of a comparative approach to personality, we briefly consider its potential impact on personality research in genetics and development.

Progress in personality-assessment procedures for nonhumans opens the way for new interdisciplinary partnerships between genetics and personality researchers. Although genetic studies on humans are essential, animal-personality research provides an avenue of research that offers important advantages (Gershenfeld & Paul, 1998). Compared with humans, many laboratory animals have brief intergenerational periods and are inexpensive to maintain. More extensive and intrusive manipulations are possible in studies of animals than would be permitted with humans, so animal studies could be used to test specific hypotheses derived from human research. In addition, transgenic methods and new cloning techniques (e.g., Wakayama, Perry, Zuccotti, Johnson, & Yanagimachi, 1998; Wilmut, Schnieke, McWhir, Kind, & Campbell, 1997) could provide novel opportunities for animal research to further our understanding of the genetic influences on personality (Flint et al., 1995). Among the many possibilities, one can foresee expanded twin studies in which, instead of using pairs of identical human twins, a large number of genetically identical cloned or inbred animals are raised in systematically varied environments to examine the interaction of genetic and environmental influences on personality. The present findings suggest that personality judgments can play a part in such studies.

The present findings also bode well for animal studies of personality change. Typically, the most useful information on personality change is derived from longitudinal studies. In some respects, animal studies provide an ideal situation in which to investigate personality development. Many captive animals are observed almost every day of their lives and biological, environmental, and social events that are hypothesized to influence personality change can be recorded, or even manipulated experimentally, to test hypotheses about environmental influences. For example, cross-fostering studies in rhesus monkeys have already shown that infants' responses to separation from their foster mothers is best predicted from their inherited levels of emotional reactivity, rather than from their foster mother's level of reactivity or care-taking style (Suomi, 1999).

In sum, animal studies provide unique opportunities to elucidate the dynamic interaction of biological, genetic, and environmental effects on personality, and to study personality change, links between personality and health, and even processes in personality perception. The use of accurate personality judgments by owners, handlers, and observers would make collecting animal personality data so much less cumbersome and time-consuming, thus opening up research opportunities with vast potential. Moreover, the present evidence for the existence of personality traits in mammals paves the way for personality researchers to incorporate animal studies into their research programs, and for animal researchers to include personality constructs in their studies. Thus, the present approach should help enrich both personality and animal research, and will lay the groundwork for fruitful collaborations between the two disciplines. Indeed, as our opening quote suggests, even in animals, personality goes a long way.

References

Aschenbach, J. (1995, May 7). Grouchy to gleeful, bears at play showcase personalities. *Los Angeles Times,* p. B1.

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research.* Springfield, IL: Thomas.

Block, J. H., & Block, J. (1980). The role of ego-control and ego resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 13, pp. 39–101). Hillsdale, NJ: Erlbaum.

Borkenau, P. (1992). Implicit personality theory and the five-factor model. *Journal of Personality, 60,* 295–327.

Buss, A. H. (1988). *Personality: Evolutionary heritage and human distinctiveness.* Hillsdale, NJ: Erlbaum.

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

Cattell, R. B., & Korth, B. (1973). The isolation of temperament dimensions in dogs. *Behavioral Biology, 9,* 15–30.

Coren, S. (1994). *The intelligence of dogs: Canine consciousness and capabilities.* New York: Free Press.

Coren, S. (1998). *Why we love the dogs we do: How to find the dog that matches your personality.* New York: Free Press.

Darwin, C. (1998). *The expression of emotions in man and animals.* Oxford, England: Oxford University Press. (Original work published 1872)

Domjan, M., & Purdy, J. E. (1995). Animal research in psychology: More than meets the eye of the general psychology student. *American Psychologist, 50,* 496–503.

Fairbanks, L. A. (2001). Individual differences in response to a stranger: Social impulsivity as a dimension of temperament in vervet monkeys (*Cercopithecus aethiops sabaeus*). *Journal of Comparative Psychology, 115,* 22–28.

Flint, J., Corley, R., DeFries, J. C., Fulker, D. W., Gray, J. A., Miller, S., & Colins, A. C. (1995, September). A simple genetic basis for a complex physiological trait in laboratory mice. *Science, 269,* 1432–1435.

Forkman, B., Furuhaug, I. L., & Jensen, P. (1995). Personality, coping patterns, and aggression in piglets. *Applied Animal Behaviour Science, 45,* 31–42.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652–670.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception.* San Diego, CA: Academic Press.

Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The Riverside Behavioral Q-sort: A tool for the description of social behavior. *Journal of Personality, 68,* 451–489.

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintance. *Journal of Personality and Social Psychology, 69,* 656–672.

Gershenfeld, H. K., & Paul, S. M. (1998). Towards a genetics of anxious temperament: From mice to men. *Acta Psychiatrica Scandinavica, 98*(Suppl. 393), 56–65.

Goddard, M. E., & Beilharz, R. G. (1983). Genetics of traits which determine the suitability of dogs as guide-dogs for the blind. *Applied Animal Ethology, 9,* 299–315.

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., & John, O. P. (1998, May). *Personality dimensions in dogs, cats, and hyenas.* Paper presented at the annual meeting of the American Psychological Society, Washington, DC.

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., Rentfrow, P. J., & Swann, W. B., Jr. (in press). A very brief measure of the Big Five personality domains. *Journal of Research in Personality.*

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002, November 22). The domestication of social cognition in dogs. *Science, 298,* 1634–1636.

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.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality, 8,* 149–162.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 163–179). New York: Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford Press.

King, J. E., & Figueredo, A. J. (1997). The five-factor model plus dominance in chimpanzee personality. *Journal of Research in Personality, 31,* 257–271.

Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106,* 395–409.

McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology, 43,* 293–303.

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81–90.

Murphy, J. A. (1995). Assessment of the temperament of potential guide dogs. *Anthrozoös, 8,* 224–228.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4,* 681–691.

Nowak, R. M. (1979). North American Quaternary *Canis. Monographs of the Museum of Natural History, University of Kansas, 6,* 1–154.

Nowak, R. M. (1999). *Walker's mammals of the world* (6th ed.). Baltimore: Johns Hopkins University Press.

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.

Robins, R. W., & John, O. P. (1997). The quest for self-insight: Theory and

research on the accuracy of self-perceptions. In R. Hogan, J. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 649–679). New York: Academic Press.

Sinn, D., Perrin, N., Mather, J. A., & Anderson, R. C. (2001). Early temperamental traits in an octopus (*Octopus bimaculoides*). *Journal of Comparative Psychology, 115,* 351–364.

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.

Stevenson-Hinde, J. (1983). Individual characteristics and the social situation. In R. A. Hinde (Ed.), *Primate social relationships: An integrated approach* (pp. 28–35). Sunderland, MA: Sinauer Associates.

Suomi, S. J. (1999). Behavioral inhibition and impulsive aggressiveness: Insights from studies with rhesus monkeys. In L. Balter & C. S. Tamis-Lemonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 510–525). Philadelphia: Taylor & Francis.

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

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

Wakayama, T., Perry, A. C. F., Zuccotti, M., Johnson, K. R., & Yanagi-machi, R. (1998, July 23). Full-term development of mice enucleated oocytes injected with cumulus cell nuclei, *Nature, 394,* 369–374.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57,* 120–128.

Weiss, A., King, J. E., & Enns, R. M. (2002). Subjective well-being is heritable and genetically correlated with dominance in chimpanzees (*Pan troglodytes*). *Journal of Personality and Social Psychology, 83,* 1141–1149.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment.* Reading, MA: Addison-Wesley.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., & Campbell, K. H. S. (1997, February 27). Viable offspring derived from fetal and adult mammalian cells. *Nature, 385,* 810–813.

ORDER FORM

Start my 2004 subscription to *Journal of Personality and Social Psychology*!   ISSN: 0022-3514

____ $190.00, APA MEMBER/AFFILIATE _____
____ $382.00, INDIVIDUAL NONMEMBER _____
____ $1,069.00, INSTITUTION   _____
In DC add 5.75% / In MD add 5% sales tax   _____
TOTAL AMOUNT ENCLOSED $ _____

Subscription orders must be prepaid. (Subscriptions are on a calendar year basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.

SEND THIS ORDER FORM TO:
American Psychological Association
Subscriptions
750 First Street, NE
Washington, DC 20002-4242

AMERICAN
PSYCHOLOGICAL
ASSOCIATION

Or call **(800) 374-2721**, fax **(202) 336-5568**.
TDD/TTY **(202) 336-6123**.
For subscription information, e-mail:
subscriptions@apa.org

○ **Send me a FREE Sample Issue**
○ **Check enclosed** (make payable to **APA**)

Charge my: ○ VISA   ○ MasterCard   ○ American Express

Cardholder Name _____
Card No. _____ Exp. Date _____

_____
Signature (Required for Charge)
BILLING ADDRESS: _____

_____
City _____ State _____ Zip _____
Daytime Phone _____
E-mail _____

SHIP TO:
Name _____
Address _____

_____
City _____ State _____ Zip _____
APA Member # _____ *PSPA14*