



OPEN ACCESS

EDITED BY

Clara Mancini,
The Open University, United Kingdom

REVIEWED BY

Janice Lauren Baker,
Veterinary Tactical Group, United States
Emma Kathryn Grigg,
University of California, Davis, United States
Susan Hazel,
University of Adelaide, Australia

*CORRESPONDENCE

Ariella Y. Moser
✉ ariellamoser0@gmail.com

RECEIVED 27 June 2024

ACCEPTED 01 October 2024

PUBLISHED 16 October 2024

CITATION

Moser AY, Welch M, Brown WY,
McGreevy P and Bennett PC (2024) Methods
of behavioral testing in dogs: a scoping
review and analysis of test stimuli.
Front. Vet. Sci. 11:1455574.
doi: 10.3389/fvets.2024.1455574

COPYRIGHT

© 2024 Moser, Welch, Brown, McGreevy and
Bennett. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Methods of behavioral testing in dogs: a scoping review and analysis of test stimuli

Ariella Y. Moser^{1,2*}, Mitchell Welch³, Wendy Y. Brown¹,
Paul McGreevy⁴ and Pauleen C. Bennett²

¹School of Environmental and Rural Science, University of New England, Armidale, NSW, Australia, ²Anthrozoology Research Group, School of Psychology and Public Health, La Trobe University, Bendigo, VIC, Australia, ³School of Science and Technology, University of New England, Armidale, NSW, Australia, ⁴Sydney School of Veterinary Science, University of Sydney, Sydney, NSW, Australia

Background: Behavioral testing is widely used to measure individual differences in behavior and cognition among dogs and predict underlying psychological traits. However, the diverse applications, methodological variability, and lack of standardization in canine behavioral testing has posed challenges for researchers and practitioners seeking to use these tests. To address these complexities, this review sought to synthesize and describe behavioral testing methods by creating a framework that uses a “dog-centric” perspective to categorize the test stimuli used to elicit responses from dogs.

Methods: A scoping review was conducted to identify scientific literature that has reported behavioral testing to assess psychological traits in dogs. Five online databases were systematically searched. Following this, an inductive content analysis was conducted to evaluate and summarize the behavioral testing methods in the literature.

Results: A total of 392 publications met the selection criteria and were included in the analysis, collectively reporting 2,362 behavioral tests. These tests were individually evaluated and categorized. Our content analysis distinguished 29 subcategories of behavioral testing stimuli that have been used, grouped into three major categories: human-oriented stimuli; environmental stimuli; and motivator-oriented stimuli.

Conclusion: Despite the methodological heterogeneity observed across behavioral testing methods, our study identified commonalities in many of the stimuli used in test protocols. The resulting framework provides a practical overview of published behavioral tests and their applications, which may assist researchers in selecting and designing appropriate tests for their purposes.

KEYWORDS

behavioral assessment, behavioral testing, canine, dog cognition, dog personality, qualitative analysis, scoping review, temperament testing

1 Introduction

Behavioral testing offers an empirical lens to reveal individual differences in the behavior and cognition of animals (1–3), including dogs (4, 5). Domestic dogs exhibit substantial variation in their behavioral tendencies and cognitive abilities. This can be seen, for example, in the diversity of behavioral phenotypes that characterize dog breeds (6–9). Even within a given breed, individual dogs exhibit extensive variation in their behavior and cognition (9–11).

Accordingly, the attributes of individual dogs need to be considered and assessed to predict their behavior as companions and co-workers.

A behavioral test is a standardized protocol that presents a stimulus designed to elicit a measurable response in a subject. In many cases, a battery of behavioral tests, employing a series of stimuli, is used [e.g., (12, 13)]. The responses of different dogs undergoing the same protocol can then be compared to reveal individual differences. The premise of behavioral testing is that these responses reflect underlying traits and, in doing so, predict behavior beyond the testing context (5). In animal behavior research, “traits” are the inter-individual differences in behavior that are relatively stable across time and contexts (14). The term is most often used in relation to personality [e.g., (15)], but has also been used to describe perceived cognitive abilities or tendencies [e.g., (16)]. As these are types of psychological difference that we attempt to infer from dogs’ behavior, here we will refer to these as “psychological traits”.

Historically, canine behavioral tests have been used as a research tool in fields that include psychology [e.g., (17, 18)], neurophysiology [e.g., (19, 20)], animal science [e.g., (21, 22)] and ethology [e.g., (23, 24)], to answer questions about the biological and environmental bases of individual differences in the behavior of humans and animals. Beyond the scientific research context, canine behavioral tests have had many practical applications. For example, they are used to assess working dog candidates for various roles [for review, see (25, 26)], to determine dogs’ suitability for adoption from shelters [for review, see (27–29)], and to determine dogs’ breeding suitability according to breed club standards [e.g., (30, 31)]. Unsurprisingly, given this diverse range of applications, there is a great deal of heterogeneity in the behavioral testing literature concerning the studied populations, characteristics, interpretations, and methodologies that have been used.

Behavioral tests are often designed to reveal a certain trait (e.g., laterality, sociability), or super-trait (e.g., boldness) but may do so only imperfectly. Despite efforts to devise frameworks that define and categorize canine traits (25, 32–34), canine science has so far resisted the uptake of standard terminology or shared definitions. This may, in part, reflect the disparity of reasons for undertaking canine behavioral research (35). An additional problem is that we cannot measure psychological traits directly. Instead, we can measure only behavioral and physiological responses and then infer their meaning. Such inferences are primarily made in two ways: (1) *a priori* expectations that a test is designed to measure a certain trait (e.g., use of a loud sound intending to reveal ‘noise sensitivity’); and (2) *post hoc* interpretations, often from grouping correlated variables with factor analysis and using a descriptive label for each factor (e.g., cowering, vocalization, and distance from and latency to approach an unfamiliar human might be labelled ‘fearfulness’). Unfortunately, any attempt to interpret behavior is vulnerable to subjectivity and there is no consensus on the emergent terminology nor on the number of psychological dimensions that exist in dogs (4, 5, 36).

There is also currently no established methodological standard to adhere to when conducting behavioral tests. Protocols for tests and test batteries may be entirely or partly original or they may be adapted, rearranged, or replicated from previously established tests. In addition, data collection methods differ considerably and can range from subjective ratings of behavior, ratings of behavior on a scale, and behavior coding, to physiological measures or data from sensors (e.g., accelerometers, infrared beam breakers). This diversity in

methodology gives rise to inconsistency among the tests’ degrees of standardization, reliability, and validity [for detailed discussions on assessing these qualities, see (1, 4, 5, 29, 36)]. Given the many diverse aims of behavioral tests and continuously evolving methods arising from nascent ideas and emergent technologies, much of this inconsistency has been unavoidable.

The current state of flux concerning terminology and methodology in canine behavioral testing has frustrated attempts to unify or compare findings (4, 5, 36, 37). That said, there are similarities among many behavioral testing protocols when these are considered from the perspective of the dog being tested. From this “dog-centric” perspective, behavioral tests tend to employ stimuli that share important features even when the exact methodology, intended purposes, or interpretations differ. For example, tests in which a dog is released to roam free in an empty testing area have recently been used to measure traits including activity level [e.g., (38)], independence [e.g., (39)], arousal and anxiety [e.g., (40)], or exploration tendency [e.g., (41)]. Irrespective of the intention of the testers, the experience of an empty room is likely to be similar for the participating dogs. For efficiency in discussion, it is practical to consider these tests together, but also to acknowledge the numerous possible contributors to behavior that might be reflected in a single test. As such, a “dog-centric” perspective may help us to put aside the subjective interpretations of behaviors, and instead consider the measures themselves. By identifying the methodological similarities and categorizing behavioral tests based on stimulus attributes, we offer an approach to consolidate the research across such a disparate field.

The current review sought to provide a practical starting point for researchers and practitioners seeking to select or design canine behavioral tests. To do so, we investigated the behavioral test stimuli used to measure canine psychological traits in the scientific literature. We aimed to create a framework that could parsimoniously describe the stimuli used in behavioral tests from a dog-centric perspective and, from this, review various methodological options and practical considerations for their application.

2 Methods

This review used a scoping review search method with a content analysis of the behavioral testing methods from the articles identified. We followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Review (PRISMA-ScR) guidelines (42).

2.1 Eligibility criteria

We included peer-reviewed articles published in English that used behavioral testing to measure variation in psychological traits among dogs. Behavioral tests in this case included any standardized protocol in which a dog was presented with a stimulus and their responses recorded, and which did not require extensive specific training (> 1 week of training) for the dog to complete. Tests that analyzed the variation between individual dogs or groups of dogs were included and we excluded tests that analyzed at only a general species level (e.g., comparing dogs and wolves, or determining if dogs possess a cognitive

ability). Finally, only tests that intended to measure psychological traits were included. The inclusion and exclusion criteria are specified in Table 1. In cases where an article used both in-scope test(s) and out-of-scope test(s), the article was included but only the in-scope test(s) were analyzed.

2.2 Information sources and search strategy

Search terms were initially chosen based on the terms used in titles, keywords, and abstracts of an initial selection ($n = 71$) of articles in the research area that were known to the research team. Following this, synonyms, plurals, and alternate spellings were added. Preliminary searches were conducted to check if known articles were included. To capture articles with different purposes and to retrieve a comprehensive sample of publications, the search terms were broad. However, since some reports use behavioral tests as part of their methods but do not mention this explicitly in their title, abstract, or keywords, it is acknowledged that the list of retrieved articles is likely not exhaustive.

Five databases (SCOPUS, Web of Science, CAB abstracts, PsycINFO, and Medline) were searched from their inception dates. The searches were first conducted on the 6th of January 2023 and updated on the 18th of March 2024. The following terms were used to search the titles, abstracts, and keywords of peer-reviewed journal publications: (dog OR dogs OR puppy OR puppies) AND (test OR task OR assessment OR measure* OR score* OR procedure OR protocol) AND (cognit* OR behavio* OR personality OR temperament OR character OR “problem solving” OR “problem-solving” OR fearfulness) AND (trait* OR characteristic OR factor OR ability OR differences OR suited OR suitability OR selection OR problems OR stability). Irrelevant subject areas were excluded from

the search results (e.g., engineering, physics and astronomy, dentistry, economics).

Citations from the database searches were imported into the Covidence Systematic Review Software (43). Duplicates were automatically and manually removed. According to the exclusion criteria listed above, titles and abstracts were screened initially for relevance, followed by full-text screening of the remaining articles. The screening process was conducted by one reviewer (AM). A second reviewer was not deemed necessary to achieve our objective of a comprehensive, but not exhaustive, selection of publications, since high sensitivity and specificity of included articles were not necessary for the aims of this scoping review.

2.3 Data charting

All data were charted in a spreadsheet created for this purpose, which is supplied in the Supplementary materials. The data were extracted by manually reading each original article and recording the relevant features. The Supplementary materials present the list of categories and definitions for each feature.

For each article in the spreadsheet, the citation, author names, and year of publication were charted. We then categorized the general purpose of the behavioral test(s) in the article (e.g., research, shelter assessment, working dog assessment), extracted the terms used by the authors to describe the measured outcomes—referred to here as “trait descriptors”—and categorized these as relating primarily to the domains of behavior, cognition, or both.

Following this, methodological information was recorded, including the total number of participants, their age groups, and sources of studied dogs (e.g., companion, shelter), as well as the types of measures used (e.g., behavioral coding, hormonal assays). For each article, we counted the number of unique tests used (i.e., excluding

TABLE 1 Inclusion and exclusion criteria used to select articles.

Criteria	Inclusion	Exclusion
Article type	Experimental study published in a peer-reviewed journal	<ul style="list-style-type: none"> - Review article - Not published in a peer-reviewed journal (e.g., book, conference proceeding, thesis) - Retracted publication
Language	Article available in English	<ul style="list-style-type: none"> - Article available exclusively in a language other than English
Population	Participants were domestic dogs (<i>Canis familiaris</i>)	<ul style="list-style-type: none"> - Participants were exclusively a species other than domestic dogs (<i>Canis familiaris</i>)
Methods	Reports a standardized behavioral test (i.e., measures were taken from direct observation of dogs responding to stimuli at a specific time point(s) using a protocol that is consistent across dogs)	<ul style="list-style-type: none"> - Outcomes are only responses to a questionnaire - Measure is from general observation of behavior, not at a specific time point(s) - Test methods are not described or referenced - Test measures performance of behavior(s) that required extensive protracted training (> 1 week)
Analysis	Individual differences were measured (i.e., scores from individual dogs or groups of dogs are differentiated)	<ul style="list-style-type: none"> - Species-wide abilities or trends measured (e.g., only differentiations between species are reported)
Purpose of test	To measure or infer consistent tendencies or psychological traits	<ul style="list-style-type: none"> - To measure physical or physiological differences (e.g., visual or olfactory acuity) - To measure transient psychological states (e.g., fear during a surgical procedure) - To measure the outcome of a short-term intervention, including training, interaction, pharmacological, nutraceutical, or hormonal interventions - To measure clinical dysfunction or impairment - To measure food preference or palatability - To measure interactions or relationships within specific relationships (e.g., attachment style)

repeated tests from the count). Articles using more than one test were considered to have used a test “battery” and, if applicable or available, the name or a brief description of the battery was recorded so that it could be referenced and compared across different articles.

We then recorded the names (or if not available, brief descriptions) of each behavioral test in a given article and whether each test was reported as having been adapted or replicated from another article in the dataset or was described in the first instance in this dataset. We charted the methodological detail for each test by extracting the relevant segments of the reporting article’s methods section. In cases where a test battery was replicated in its entirety and referenced as such, the relevant data were duplicated from the data charted for the original, referenced article.

Finally, we charted whether or not reliability (test–retest, intra-rater, or inter-rater) and validity (construct or criterion) metrics were reported in the article. This was recorded simply as presence or absence.

To report the number and type of trait descriptors used, we made a list of all the terms authors used to describe the outcomes measured by the test(s) that had been recorded in the spreadsheet, then counted the number of times that each trait descriptor (e.g., sociability, frustration) appeared. We then condensed the list of trait descriptors by manually combining all terms with the same word-stem and meaning (e.g., aggressiveness, aggression, and aggressivity). To avoid subjective interpretation, synonyms that did not share a word-stem were not combined (e.g., boldness was not combined with confidence) and specific terms were not combined with general terms (e.g., aggression towards dogs was not combined with aggression).

2.4 Content analysis

As a method of qualitative analysis, inductive content analysis can be used to understand and synthesize text-based data (44). This method involves reading the text, creating codes to describe each piece of its content, and then grouping these codes to synthesize the data into broad categories. The steps to conducting this analysis are (1) reading and familiarization with content, (2) first-round coding, (3) second-round coding, (4) redefining subcategories, and (5) synthesis and interpretation (44).

In the current study, the aim was to synthesize and present the patterns of canine behavioral testing methods in a framework that enabled us to discuss a large number of protocols. To do this, content analysis was used to investigate the test stimuli in each behavioral testing protocol. A stimulus was defined as any object or event used to elicit a behavioral response from a participating dog. We sought to describe and categorize the major stimulus or stimuli of each test procedure according to those which were most relevant to the responses measured in the test. Potential stimuli that were incidental, such as objects, sounds, or people that may have been in the area but were not the focus or intention of the test, were not coded.

We used a slightly modified process of content analysis appropriate for the breadth of methodological data that were analyzed. The first- and second-round coding processes were carried out with an initial subset of 20% of the included articles, which determined the first iteration of codes. Then, in the data charting spreadsheet, every article was coded by recording the presence or absence of each stimulus code. These data are supplied in the [Supplementary materials](#). A test could be coded with more than one code in cases when more than one distinct stimulus was

presented in a test simultaneously or successively. Throughout this process, the codes were redefined when appropriate. In this way, the codes were created iteratively and were developed and modified throughout the process until the test stimuli were considered to be adequately and parsimoniously described. Finally, the codes derived from this process were grouped into overarching categories for synthesis and discussion and each code was labelled and defined as a subcategory.

3 Results

[Figure 1](#) shows the flowchart of the citation retrieval, screening, and inclusion process. From 4,430 unique citations, 392 were included in the review analyses. A complete list of the included articles is presented in the [Supplementary materials](#). The articles were published between 1948 and 2024, with a majority (61.48%; 241/392) published in the most recent decade, from 2015 to 2024.

Most articles employed a battery of tests, with a mean of 6 unique tests per article (not including repeated tests) ($SD=6.02$), ranging from 1 to 43. Overall, the reviewed articles included 2,362 behavioral tests in total. These included 326 that were adapted or modified and 982 that were replicated from other tests previously reported in the dataset. A further 1,054 tests were either described for the first time, replicated a test not previously appearing in the dataset or the scientific literature, or used a test without reference to its source in the text.

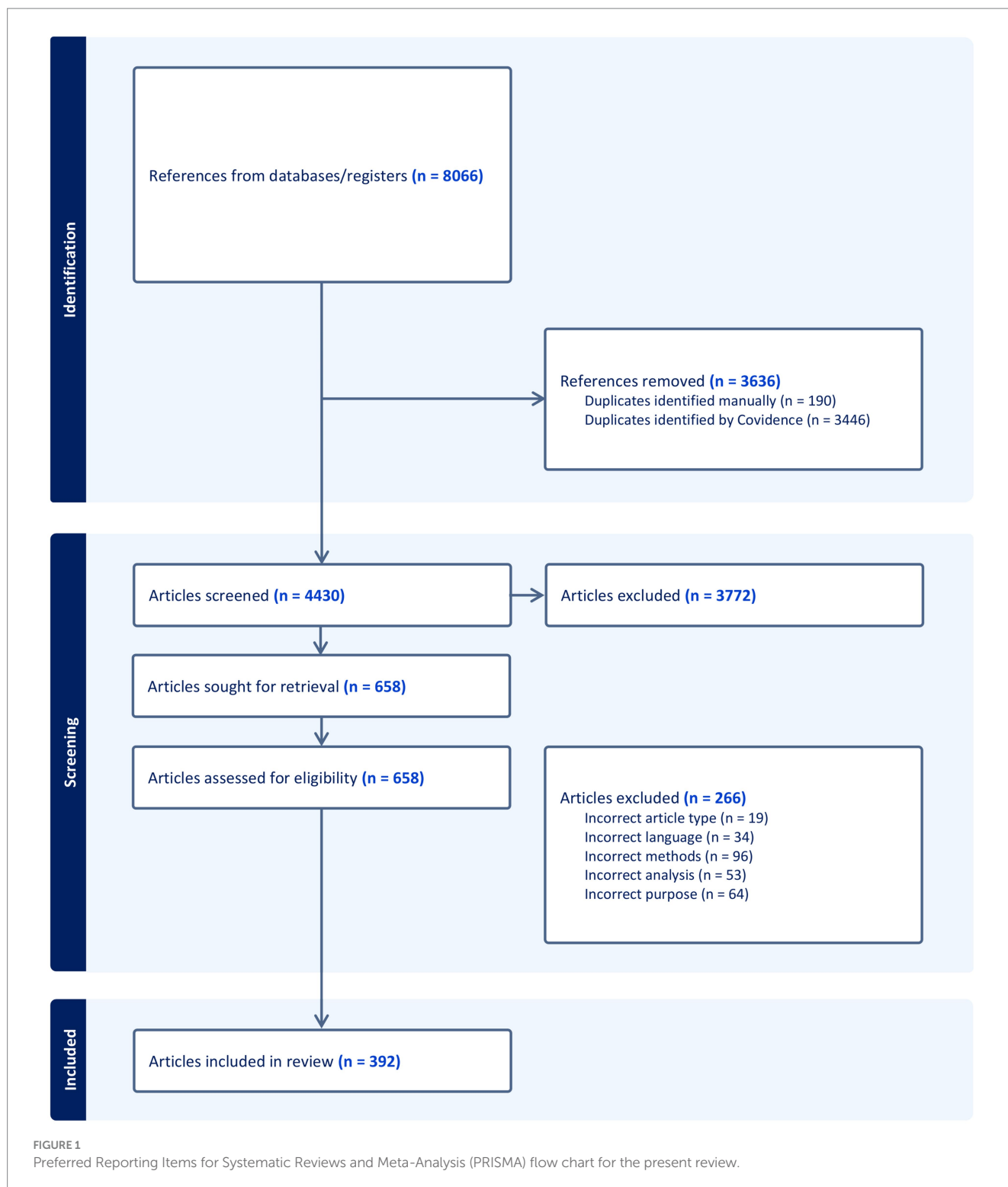
3.1 Characteristics of included articles

The complete data including all references and charted article characteristics are accessible in the [Supplementary materials](#).

3.1.1 Purpose of tests

Articles using behavioral tests for the primary purpose of research were most common in this dataset (68.62%; 269/392). These were articles that used behavioral tests with the primary aim of answering a research question to advance scientific understanding or to provide an instrument for research. The remaining articles reported tests used for purposes with specific applications separate from basic research. This included assessing behavior generally, often referred to as “temperament testing,” for companionship or breeding purposes (6.12%; 24/392). Another application was shelter testing (7.91%; 31/392), which sought to assess shelter dogs’ suitability for adoption as a companion, often with the aim of predicting potentially problematic behaviors post-adoption. Other articles assessed aggression specifically (3.83%; 15/392), for example, to determine if a companion dog is a risk to the community. Another important applied purpose was assessing dogs’ suitability to be used for various working roles, including assistance (5.87%; 23/392), military (2.81%; 11/392), detection (2.04%; 8/392), and other working roles (2.81%; 11/392).

In terms of the traits that were measured, 64.29% (269/392) of articles sought to measure traits related to behavior and 34.18% (154/392) of articles sought to measure traits relating to cognition, with an overlap of articles that measured traits related to both domains. The traits that the tests intended to measure were described with a wide range of terms, with 390 unique descriptors used among the reviewed articles. The most frequently used trait descriptors



were: aggression ($n=58$), fear ($n=48$), sociability ($n=41$), playfulness ($n=31$), curiosity ($n=16$), problem-solving ($n=15$), fearlessness ($n=14$), inhibitory control ($n=14$), chase-proneness ($n=13$), and anxiety ($n=11$). Thirty-three articles did not label any traits and instead described the findings using direct behavioral descriptions (e.g., number of yawns, distance from stimulus) or a test performance score.

3.1.2 Populations

The sample size used in each article ranged from 8 to 89,352 dogs (mean = 839.62, median = 70, SD = 5831.84). The distribution of sample size included outliers with very high sample sizes; the six highest sample sizes (>10,000 dogs) used data from the Swedish Dog Mentality Assessment project [described first in (15)] and the latest adaptation of this test (13).

Companion dogs (pet or family dogs) were used in 54.08% (233/392) of articles. Other populations, in order of most commonly to least commonly used, included candidate working dogs (dogs bred or selected to potentially perform a working role) (16.33%; 64/392), dogs in shelters (12.76%; 50/392), laboratory dogs (dogs owned by a research institution) (9.18%; 36/392), working dogs (dogs currently performing a working role) (9.18%; 36/392), kennel dogs (privately-owned dogs not kept for human companionship) (3.06%; 12/392), and free-ranging dogs (1.5%; 6/392).

Adult dogs (1 to 9 years old) were used in most articles (83.42%; 327/392) and, in many cases, juvenile (16 weeks to 12 months old) (25.26%; 99/392) and senior dogs (>9 years old) (29.08%; 114/392) were included in the same sample as adults. Puppies (<16 weeks old) were always differentiated as a separate sample from other age groups and generally administered different versions of behavioral tests, and were reported in 16.84% (66/392) of articles.

3.1.3 Measures

We identified several different types of measures used to collect data from behavioral tests. Subjective rating, for which traits or behaviors were rated based on the rater's interpretation of the behaviors (e.g., a rating of "fearfulness") was used in 19.90% (78/392) of articles. Behavior rating, for which specific behaviors were defined and rated on a scale (e.g., a rating of jumping when greeting) was used in 32.14% (12/392) of articles. Behavior coding, for which specific behaviors were defined and coded based on presence/absence, duration, latency, or a similarly objective parameter, was most common, used in 64.03% (251/392) of articles. Hormonal concentrations were measured from biological samples (e.g., saliva, urine, blood) in 7.40% (9/392) of articles. Cardiac measures [e.g., heart rate, heart rate variability (HRV)] were used in 3.32% (13/392) of articles, accelerometer recordings of activity were reported in 0.77% (3/392) of articles, and other technologies (e.g., MRI imagining) were used in 0.51% (2/392) articles.

3.1.4 Psychometric reporting

Psychometric indices relating to reliability and validity were reported in a minority of the articles. Indices of test reliability were reported in 39.54% (155/392) of articles, which included intra-rater reliability (2.30%; 9/392), inter-rater reliability (32.40%; 127/392), and test-retest reliability (10.46%; 41/392). Indices of test validity were reported in 32.91% (129/392) of articles. Construct validity was reported in 20.15% (79/392) of articles to indicate how well the test reflected a certain construct in comparison to other measures (e.g., a questionnaire measuring similar traits). Criterion validity was reported in 14.54% (57/392) of articles to describe how well the test measured or predicted an outcome (e.g., completion of training, success of adoption).

3.2 Content analysis of methods

The inductive content analysis process produced 29 codes to describe test stimuli; the codes assigned to each article are available in the [Supplementary materials](#). The stimulus codes, hereafter referred to as subcategories, were grouped into three broad categories, labelled as human-oriented stimuli, environmental stimuli, and motivator-oriented stimuli.

3.2.1 Human-oriented stimuli

This category describes test protocols that sought to elicit a response towards a human ([Table 2](#); [Supplementary Table S1](#)). These test stimuli were the most commonly used in the current sample of the literature and have been included in test batteries for almost every applied purpose. This emphasizes the importance placed on dogs' behavior towards humans specifically, which may help them to be safe and effective as companions and working dogs. Tests using human-oriented stimuli generally seek to measure traits described as sociability, aggression, fearfulness, playfulness, and obedience.

While the contents of [Table 2](#) are mostly self-explanatory, several points are of particular importance. First, by far the largest subcategory is Human Interaction (see [Figure 2](#)). This is a necessarily broad category to embrace the sheer scale of variation in tests involving human interactions. This category includes all protocols that involved a person interacting directly with a dog but that did not constitute the other, more specific, interaction categories (Physical Manipulation, Obedience Cues, Playful Encounter, or Hostile Encounter). These interactions include a spectrum of behavior from normal and sociable [e.g., calling the dog over and patting ([45](#))], to the unusual and unexpected [e.g., wearing a hooded cape and approaching the dog while crouching and widening the cape ([15](#))]. Different intensities of unusual behaviors may reveal different degrees of sociability, aggression, or fearfulness.

A second point of interest is that, in contrast to the stimuli that involve direct interactions, the subcategory Indirect Human Encounter refers to situations that allow testers to observe whether a dog chooses prosocial or antisocial behaviors while limiting the influence of the human leading the interaction. Additionally, Indirect Human Encounters often allow for the ability to manipulate the intensity of human behavior without it being directed towards or threatening the dog specifically [e.g., the "disgruntled stranger" test ([46](#))]. This stimulus type was sometimes used to determine the safety of the dog with strangers [e.g., indicating if they will respond aggressively to people moving past ([47](#))] but was more often used to indicate whether they seek out human contact [e.g., ([23](#))], often as an initial step before beginning an interaction.

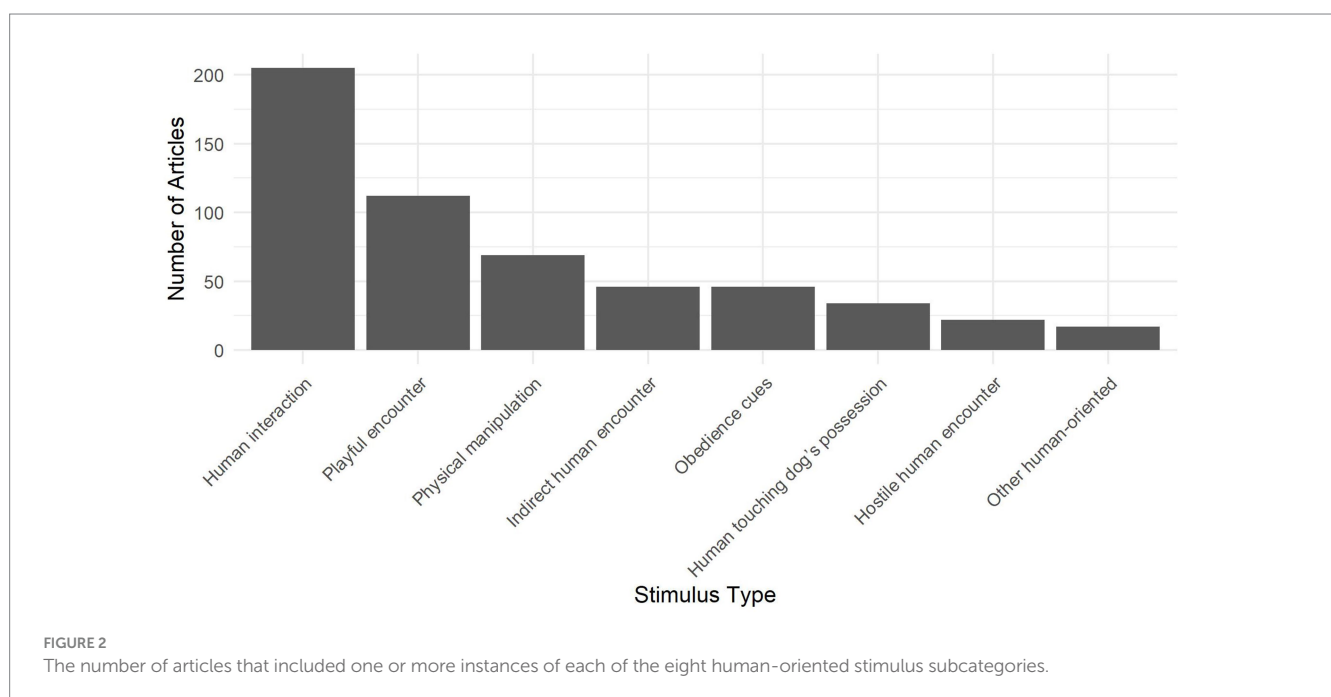
Third, Playful Encounters were used to reveal dogs' motivation to play with humans and/or objects. This has been tested not only because it is a desirable trait for companion dogs ([48](#)), but also because it is critical for working dogs. In detection and military dogs ([49](#), [50](#)), for example, playfulness may be required to effectively train these dogs to carry out their roles.

Fourth, within each subcategory, the intensity of interactions and behaviors was often manipulated according to the purpose of testing. For example, if testers sought to determine if there was any potential for undesirable traits, such as aggression or fearfulness, human behavior that was particularly intense or challenging was often used to reveal this. In addition, in tests seeking to measure aggression, stimuli that could cause conflict, such as a Physical Manipulation or a Human Touching a Dog's Possession, were used routinely [e.g., ([51](#), [52](#))].

Studies have reported reasonable-to-excellent psychometric qualities for tests using human-oriented stimuli. Test batteries including human-oriented stimuli have found good test-retest reliability [e.g., ([53](#), [54](#))]. Several studies have also established construct validity using owner reports of traits including aggression ([55](#)), fearfulness ([56](#)), sociability ([57](#)), activity-impulsivity ([58](#)), and

TABLE 2 Subcategories of human-oriented test stimuli and their descriptions.

Subcategory	Description
Indirect human encounter	One or more human(s) are in the area but are ignoring the participant dog and do not engage with them directly for the entire duration of the test.
Physical manipulation	A human physically manipulates the participant dog by grabbing, holding, and/or moving parts of the dog (i.e., more than merely patting or stroking.)
Obedience cues	A human provides verbal and/or non-verbal cues to perform a specific behavior, either after a brief training period (during the test) or with the assumption that the dog has learned that behavior previously. For example, “sit” or “come.”
Playful encounter	A human engages with the dog in a way that invites play, for example by playfully activating/engaging with an object.
Hostile human encounter	A human directly engages with the participant dog and/or handler in a way that is intentionally and explicitly aggressive, for example striking towards, or yelling at.
Human touching dog's possession	The dog is given a possession, for example a food bowl or toy, and then a human reaches towards, touches, and/or moves the object. This may be done with an artificial hand (excluding playful encounter).
Human interaction	One or more human(s) interact directly with the dog in a manner not described in the previous categories (physical manipulation, obedience cues, a playful encounter, hostile encounter, or touching a possession). This is broad and includes a spectrum of behaviors from those that would be considered typical in companion dog-human interactions (e.g., approaching, petting, placing on leash) as well as unusual interactions (e.g., approaching and running away, wearing a sheet over body and banging broom on the ground).
Other human-oriented	A human-oriented stimulus in which there is no direct interaction and which does not constitute one of the other categories.



other traits (13, 54). In addition, aggression tests have been able to differentiate dogs with a bite history from those without (59). In terms of criterion validity, evidence has been found for shelter tests using human stimuli to predict some behaviors of dogs following adoption, such as overall friendliness or fear (60), although they may not accurately predict undesirable behaviors, such as aggression (60–63). In addition, tests using human stimuli have been useful in predicting dogs' success in assistance roles (64) and military roles (65).

3.2.2 Environmental stimuli

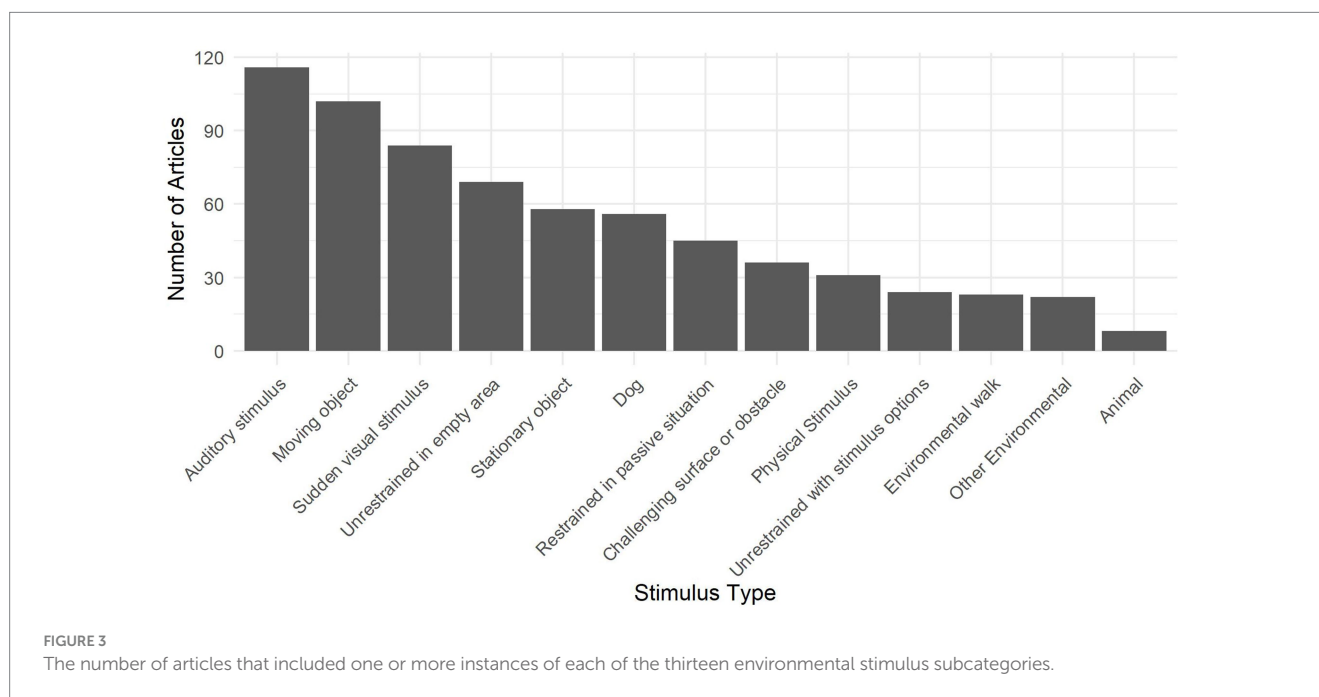
The Environmental stimuli category includes stimuli that involve a setting, context, or the presentation of distinct objects or sensory

stimuli (Table 3; Supplementary Table S2). These have been routinely used in test batteries for applied purposes, especially shelter tests, general temperament tests, and tests for working suitability, to observe whether dogs respond to the environment in ways that are adaptive and conducive to being successful in a companionship or working role. Tests using environmental stimuli tend to measure traits such as fearfulness, anxiety, boldness, reactivity, and activity level.

The environmental stimulus subcategories are described in Table 3 and some additional points are of note. Firstly, the most commonly used stimuli in this category were Auditory Stimuli, Moving Objects, and Sudden Visual Stimuli (see Figure 3). These subcategories share a common theme of the stimulus presentations being particularly salient or startling. These challenging stimulus types were often used

TABLE 3 Subcategories of environmental test stimuli and their descriptions.

Subcategory	Description
Unrestrained in empty area	The participant dog is let off leash with no human direction or interference in a testing area (not their usual home area), with no other purposeful test stimuli (note: If a human is present, they are not the focal stimulus and the dog had encountered them previously).
Unrestrained in area with stimulus options	The participant dog is let off leash with no human direction or interference in a testing area with multiple environmental stimuli.
Restrained in passive situation	The participant dog is restrained, such as with a leash or crate, in a testing area with no purposeful stimuli presentations.
Sudden visual stimulus	A visual stimulus is made to appear or change suddenly (i.e., quickly and unexpectedly).
Stationary object	A stationary object is presented.
Moving object	An object that moves, either on its own (e.g., robotic) or with human intervention (e.g., dragged on wheels), is presented. The object may become still after a period of moving.
Dog	One or more conspecific(s) is/are presented.
Animal	One or more non-human, interspecific animal(s) is/are presented.
Auditory stimulus	A loud or otherwise salient auditory stimulus is created in the vicinity of the participant dog.
Challenging surface or obstacle	An obstacle or surface (e.g., metal grate, stairs, wobbly surface) is presented.
Environmental walk	The participant dog is walked on leash by a person through a complex and naturalistic (i.e., non-controlled) environment.
Physical stimulus	A physical stimulus is applied to the participant dog.
Other environmental	An environmental stimulus which does not constitute one of the other categories.



to potentially provoke behaviors that could reflect underlying problematic behaviors or undesirable traits, such as fearfulness and/or aggression (52, 66).

Secondly, the Stationary Object and Moving Object stimuli were often used with the intention of testing dogs' response to novelty and, in particular, fear of novelty [e.g., (38, 67, 68)]. However, it was not reported how the novelty of these stimuli was ensured for most dogs. Nevertheless, many dogs predisposed to fear show fearful responses to objects that are familiar and so these may still be informative test stimuli even when novelty cannot be ensured.

In addition were objects that were used as proxies for humans, which are intended to gauge dogs' potential aggressive responses to

humans without risk of harm [e.g., (69)]. Although these tests are intended to measure a dog's potential response to a human, rather than environmental stimuli, it is debatable whether dogs genuinely respond to these objects as humans. Nevertheless, their widespread use suggests that they have been useful to reveal extreme reactions and assess the risk of human contact.

Thirdly, contexts in which the dog's free behavior could be observed, such as tests in which a dog is unrestrained in an area, with or without stimulus options, were often referred to in the animal behavior literature as open field tests [e.g., (6, 70, 71)] and arena tests [e.g., (24, 72, 73)]. These tests were also commonly predicated on the novelty of the area for the dog. In addition to observing a dog's fear or

boldness in a novel context, these tests were also used to observe a dog's general activity and behavioral preferences, such as independently exploring the space, or being passive. However, a complicating factor in these tests was the presence or absence of a human; the presence of a human was often considered incidental in these tests, and sometimes not even reported. However, it has been demonstrated that the presence or absence of a human, especially a familiar guardian, has an impact on a dog's exploratory or anxiety behavior (74). Conversely, in some cases, versions of these tests where no human was present were specifically referred to as "isolation" tests [e.g., (63, 75)].

Given the widespread importance of novelty for many of the environmental stimulus subcategories, determining test–retest reliability could be problematic in some cases. Some researchers have used different but comparable stimuli at each test time in an attempt to maintain novelty (49, 76). In cases when the time between testing is prolonged and there are not multiple instances of repeated testing, habituation to the stimulus is less likely and using the same stimuli may be appropriate [e.g., (54)].

Finally, studies using environmental stimuli have demonstrated construct validity by assessing concurrence between test behaviors and reports from owners and handlers, in particular for fearfulness (54, 56, 77). Others have found a link between test responses and physiological markers, such as salivary cortisol concentration (78, 79). Studies have also found criterion validity in predicting the likelihood of behavioral problems after adoption (80) or to become assistance dogs (81), detection dogs (82), military dogs (50), or guide dogs (83).

3.2.3 Motivator-oriented stimuli

Motivator-oriented stimuli describe testing paradigms in which the participating dog is expected to respond to a test stimulus in an attempt to reach a resource (Table 4; Supplementary Table S3), which is typically defined as a physical target that the dog has a desire to attain. Food is considered an intrinsic motivator for all dogs and was therefore used as a motivator most commonly, while play objects were also used in some cases. In many protocols, dogs were tested initially to determine whether they were motivated to attain the target at the time of testing, for example by providing freely accessible food and observing whether the dog eats it. These tests tended to focus on traits such as motivation, persistence, problem-solving, and various other cognitive abilities or styles. Historically, motivator-oriented stimuli were used most often for basic research purposes, but there has been

increasing use for applied purposes such as assessing working dog suitability (84).

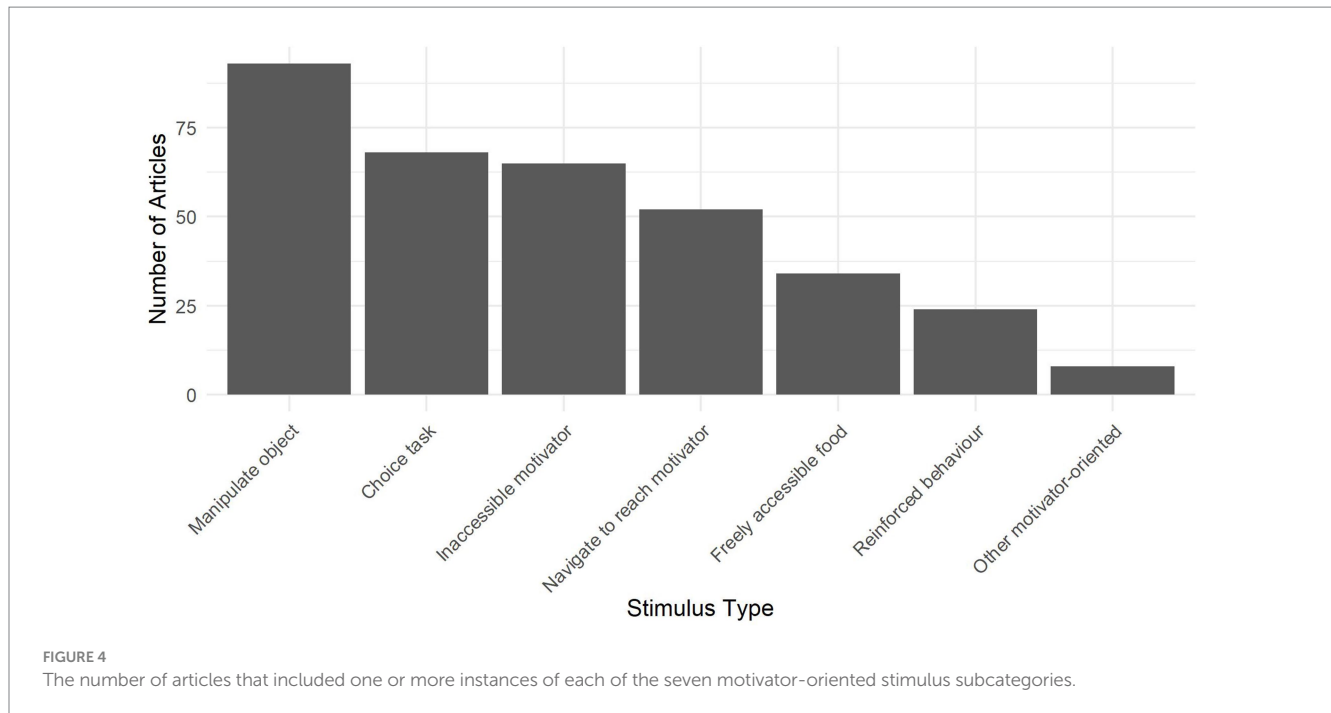
Table 4 sets out the subcategories and descriptions of motivator-oriented stimuli and their use in testing. Notably, Manipulating an Object to Reach a Motivator was used most frequently in this category, followed by other problem-solving tasks including Choice Tasks and Navigating to Reach a Motivator (see Figure 4). In these tasks, the dog could reach the motivator if they performed the correct behavior(s), which was expected to reflect a particular ability. For example, tests that involved detouring around a barrier required the dog to inhibit the impulse of moving directly towards a motivator and instead to first move away from it to reach it, which may reflect their level of inhibitory control or impulsivity (7). In choice tasks, understanding information about where a motivator was hidden would allow them to reach it, for example by following a human's pointing gesture, which may reflect their ability to understand human communicative cues (85). Also included were comparatively simple tests of manipulating an object to reach a motivator, such as extracting food from a tube or Kong™, to observe motor laterality or paw preference (86). Often, a test battery used many variations of the same paradigm, or stimulus type, to test different traits (87, 88).

In tests that used an inaccessible motivator, the dog was made aware that there was a resource out of reach, for example in a closed container. These were often referred to in the literature as an "unsolvable task" (89). This stimulus has been used most frequently to investigate interspecific social communication by observing whether dogs will make eye contact with a human when they are not able to reach the goal (90), as well as to reveal information about motivation, frustration, and persistence (91–93).

Tests involving reinforced behaviors trained the participant dog in a brief period to perform a behavior and then assessed traits related to learning, discrimination, and expectations. This was a diverse category that encompassed a variety of learned behaviors. An example of this stimulus type was "cognitive bias testing" that reinforced a dog for approaching an object when it was placed on one side of a testing set-up, but not when it was on the opposite side. They were then tested by placing the object in locations between these two sides, for which they had not yet learned the consequence. Latency to approach the object is taken as a proxy of optimism such that dogs approaching quickly are thought to be expecting reinforcement while those approaching slowly or not at all are not expecting reinforcement [e.g.,

TABLE 4 Subcategories of motivator-oriented test stimuli and their descriptions.

Subcategory	Description
Freely accessible food	Food is presented to the dog directly – it is freely accessible and they are able to consume it without interference or intentional distraction.
Navigate to reach motivator	A motivator is in the testing area and, in order to reach it, the participant dog needs to move towards it in a non-straightforward direction, for example around a barrier.
Manipulate object to reach motivator	A motivator is in the testing area and, in order to reach it, the participant dog needs to physically manipulate an object, for example by pushing open a container.
Inaccessible motivator	A motivator is presented but is physically inaccessible for the dog, for example within a container or out of reach.
Choice task	A motivator is placed in one of two or more discrete locations (e.g., buckets) and the dog may approach and check only one location. Information may be given about the location of the motivator, for example by a pointing gesture.
Reinforced behavior	A behavior is reinforced with a motivator so that the dog may learn to perform or inhibit a behavior.
Other motivator-oriented	A motivator-oriented stimulus which does not constitute one of the other categories.



(94)]. Similarly, in other reinforced behavior tests, the antecedents, behaviors, or consequences were manipulated to test various aspects of cognition.

In all subcategories, tests using motivator-oriented stimuli tended to facilitate the use of behavioral coding, which was advantageous for the intra-rater and inter-rater reliability of the measures. However, some tests may have issues with test–retest reliability due to the effect of learning. In cases in which learning is likely, researchers have reported high test–retest reliability for some tests (38) and low or mixed reliability for others (54, 95, 96). That said, overall, reporting of test–retest reliability was not common.

Since there are often no or very few established measures, such as questionnaires, for the cognitive traits that many motivator-oriented tests seek to measure, construct validity was often not established. However, some studies found good construct validity for traits with external measures, for example, impulsivity (95), and age-related cognitive decline (38, 97). In addition, criterion validity has been reported for motivator-oriented tests to identify suitable working dogs (84, 98–102).

3.2.4 Test batteries

In the literature we collated, test batteries were used much more frequently than standalone tests. These allow for observation of an individual dog in a series of contexts, to determine consistent behavioral patterns, from which aggregating scores, such as average ratings or factor scores derived from a principal component analysis, could be used to estimate a trait. Furthermore, test batteries readily facilitate measuring more than one trait at a time, which is often desirable.

4 Discussion

This review systematically assessed 392 peer-reviewed articles, published from 1948 to 2024, that used behavioral testing with the aim of measuring individual differences in psychological traits in dogs.

We sought to evaluate the extent of heterogeneity in methods and terminology in the field and then to find commonalities by categorizing and describing the stimuli used in testing protocols.

There has been a proliferation of canine behavioral testing literature since the last major review of the area (5), with over 60% of the articles included in the present review being published in the decade from 2015 to 2024. Many of the same issues, regarding heterogeneity in the literature, an overall lack of standardization in methods, reliability and validity reporting, and variable and often imprecise terminology, that were highlighted in previous reviews (4, 5, 29), persist. This may be unavoidable in such a broad research area, in which there has been considerable diversity in the purposes for employing tests, the sources and ages of participating dogs, and the measures used to collect data. These factors make it difficult, however, to compare and contrast tests such that they can be applied effectively and efficiently to address new research aims or applications.

Selecting the trait(s) to be measured is often the first consideration for researchers and practitioners when choosing appropriate behavioral tests. However, the terminology used to describe psychological and behavioral traits is inconsistent (4, 5). We found close to 400 terms that had been used to describe the traits or outcomes measured from behavioral testing. Often the same or similar testing protocols labelled the measured traits differently, sometimes due to researchers' preferences and often due to the outcomes of factor analyses. Although efforts should continue to be made to clarify or standardize the terminology that is used to describe psychological traits, given the current state of the literature, it may be difficult to select appropriate behavioral tests based solely on a particular trait descriptor.

Another aspect that makes test selection difficult is the vast number of extant behavioral testing protocols; in this review, we found over 1,000 unique tests. To synthesize these findings, we compared the stimuli that were presented in testing protocols and found three major categories. First, tests with human-oriented stimuli presented dogs with humans either behaving neutrally or interacting with them at

various levels of intensity (see Table 2). Second, tests with environmental stimuli presented non-human, sensory stimuli, including contexts, objects, odors, sounds, and physical sensations (see Table 3). Third, tests with motivator-oriented stimuli presented a tangible reward (food or play objects) to encourage the dog to engage in object-driven behavior (see Table 4). The subcategories within each of these three categories provide a reference for the types of canine behavioral tests reported and how they have been used. Knowledge of this structure may assist the selection, design, and use of behavioral tests in the future.

4.1 Practical considerations in testing

Our analysis of the literature, along with previous reviews of the field (4, 5, 29), highlights several important qualities that contribute to the practical use and accuracy of canine behavioral tests. We found that different categories of test stimuli had various advantages and disadvantages relating to these qualities.

4.1.1 Standardization

Standardization of the testing protocol(s) within a study is important to ensure that variation measured among individual dogs cannot be attributed to variations in methods or stimuli. This is a particular risk with human-oriented stimuli, especially those involving direct interaction. Some of the potential factors that impact social interactions with dogs and are difficult to control include eye contact, voice tonality, physical movement (103), and odor, which are all likely to vary with changes in a person's attributes, emotional state, or arousal level. Although some variation in human stimuli is unavoidable, this should be taken into account and particular care should be taken that the variation does not occur systematically according to the dog's attributes. For example, people may behave more enthusiastically or affably with dogs they personally find endearing.

In addition, certain environmental stimuli risk excessive stimulus variation. In particular, Environmental Walks that expose dogs to non-controlled, naturalistic locations may feature considerable variation among the stimuli that the dogs encounter. In addition, when presenting other dogs or animals, care must be taken as far as possible to encourage the stimulus animal(s) to show standardized behavior. Studies that do not report their efforts at standardizing testing stimuli may be difficult to replicate and should attract caution when interpreting the results.

4.1.2 Previous experience

In any behavioral test, it is relevant to consider the previous experiences that individual dogs may have had with a specific stimulus and how this might contribute to behavioral variation. For example, stationary and moving objects are often presented with the intention of testing dogs' response to novelty [e.g., (38, 67, 68)], as novelty is a typical way for researchers to assess fearfulness in animals (72). However, this is not always straightforward in canine testing because, except for laboratory dogs who have often had controlled exposure to stimuli, it can be difficult to control whether a dog has had experience with the same or similar stimuli. For example, a flashing toy car (77) might be familiar to dogs in a family with children, but entirely novel for those from a household without children. As such, previous experience with a test stimulus may amplify or mask actual differences in psychological traits.

4.1.3 Confounding effects

One of the drawbacks to behavioral testing is that it is susceptible to being influenced by state effects (transient variations in behavior) as well as other confounding factors, as opposed to purely measuring the trait or behavior of interest (5). One of the defining features of traits is that they are relatively stable over time, meaning that state and trait effects can be difficult to disentangle in data from a single snapshot in time. This can impact results for any test stimulus. For example, consider measuring the behavior of normally energetic and enthusiastic dogs on a day that they are unwell or fatigued. Repeated testing or a requirement to meet benchmark measures before proceeding with testing may ameliorate this issue, but it is difficult to avoid altogether in a behavioral testing context. Measuring and reporting test-retest reliability can indicate whether the test is sufficiently resilient against state effects, although this may be impossible for tests that require stimulus novelty.

Similarly, some of the variation in behavior observed in testing may be underpinned by characteristics other than what the test was intended to measure. For example, tests using motivator-oriented stimuli often seek to measure specific cognitive traits, but several other factors may affect dogs' performance in these tests. Over the course of a test battery, dogs are likely to have different attention, motivation, and energy characteristics and therefore may show poorer performance in later tests. Similarly, although attempts are usually made to ensure that dogs are interested in the motivator, dogs will generally have different degrees of motivation for the reward and this can fluctuate based on their affective state and arousal (104). Where possible, it would be desirable to account for some of this variation, for example by benchmarking an individual dog's test variables against their baseline, as well as designing protocols in a way that minimizes fatigue when resistance to fatigue is not a variable of interest for the test's purpose.

4.1.4 Multiple tests

As behavioral tests were most frequently administered in batteries of tests, it is necessary to consider individual tests in the context in which they were presented. The order in which behavioral tests are administered can impact dogs' perception of and responses towards each stimulus due to experiences in the preceding tests. For example, the cumulative effects of multiple stressors in succession can elicit a stronger stress response, known as "trigger stacking" (105). This is particularly relevant for test batteries where the dog is presented with several challenging stimuli to assess traits such as aggression or fearfulness [e.g., (55)]. Similarly, as discussed previously, a series of motivator-oriented tests could diminish motivation over time. These carry-over effects can sometimes be useful and intentional, for example, eliciting a stronger response to a stressor to assess the potential for aggression (52). When compiling test batteries, one should consider possible carry-over effects from each test to the next and, in particular, how such effects may influence the validity of the data for the intended purpose.

Due to these carry-over effects, the reliability and validity statistics reported in studies with multiple tests are applicable primarily for the test battery, as a whole, and not necessarily the individual tests. Although these metrics provide an indication that a test is likely to be useful and accurate, it is possible that a test is valid in the context of a battery but elicits different results if administered in isolation. Therefore, accuracy should be assessed whenever a new test battery is used, even if individual tests are replicated or adapted from extant validated batteries. When

appropriate, it is desirable to use an established test battery in its entirety to benefit from previous evidence of its validity.

4.1.5 Welfare considerations

Welfare should be a priority in all human-dog interactions and the responsible use of behavioral testing should ideally improve welfare outcomes for dogs and people. Behavioral testing may improve canine welfare by informing behavioral and training interventions as well as the recruitment of dogs that are psychologically well-suited for their roles. Such interventions could minimize unnecessary stress throughout a dog's lifetime.

Our review of the literature highlighted some important aspects relating to welfare in canine behavioral testing. Firstly, test stimuli that align with contemporary welfare standards should be selected by researchers and practitioners when designing and administering tests. Standards in welfare have changed over time and some of the test stimuli that were used in early canine behavioral tests [e.g., unpredictable electric shocks (12) and physically threatening or hitting dogs that have not been trained for protection (17, 106)] would be considered unethical today. Instead, behavioral testing should aim to be an enriching experience for participating dogs or, at a minimum, a neutral experience that does not cause any lasting psychological impact.

For some purposes, eliciting a degree of stress may be necessary to assess dogs' behavioral responses to stress. For example, in tests of aggression or to determine suitability for high-arousal working roles. The invasiveness or challenge of a test should be balanced against the necessity or importance of the information that is collected, and the least invasive or stressful option should be chosen. It may be possible, for example, to use stimuli that are likely to elicit only short-term effects. It might also be possible to offer emotional comfort and relief following the presentation of a stressful test stimulus.

Improving the methodology of current canine behavioral tests may produce less invasive tests while still being informative. For example, instruments that facilitate the precise measurement of subtle responses, such as HRV [e.g., (107, 108)] or automated movement tracking [e.g., (109, 110)], may reveal variations that can predict responses to stress without directly eliciting a high-level stress response. Continuing to critically evaluate behavioral testing practices and improve testing methods aligns with the ongoing goal to improve welfare outcomes.

4.2 Limitations

There were limitations to the current review process. Since we reviewed only peer-reviewed articles and did not include grey literature or protocols from industry sources, the review has a bias towards behavioral tests used for research purposes specifically. Information about behavioral tests that are routinely administered in practical contexts may be underrepresented since they are only infrequently reported in the scientific literature. Although there were articles included in the dataset that reported tests created and used for applied contexts, most articles reported tests used for research purposes. Future research that focuses on behavioral assessments used commonly in industry applications would help to reconcile scientific and practical knowledge.

Qualitative content analysis, which was conducted in this study to categorize the test stimuli extracted from test protocols, is a subjective and interpretive process. The aim in this case was to broadly discuss and evaluate a large number of protocols. As such, the stimulus categories that were discussed could not reflect all of the nuance and variation among testing methods and, instead, reduced them to their central stimulus to be useful for discussion and general overview. For example, the broad subcategory of "human interaction" could be further dissected and explored with greater specificity to determine the subcategories of interaction types and their ability to elicit certain responses. When selecting a test protocol to replicate or adapt, it is necessary to consider the fine distinctions between protocols that may elicit behaviors or responses specific to that test. Additionally, there were some limitations in specificity for our categories, as many studies did not provide precise details of the testing protocol, meaning that some potentially important details were unknown. There also appeared to be some overlap between categories or differing potential perceptions of the most salient stimulus within a behavioral test.

Finally, the scope of this review was necessarily limited. We sought to discuss test stimuli and how they are used in canine behavioral testing, with the perspective that this is a critical aspect of test selection and design. Beyond the use of test stimuli, there are many other important factors to consider for behavioral testing that will impact the quality of emergent data, such as the selection of the participant population, the measures that generate data, and the interpretation and analysis of results. Several reviews discuss these other aspects of canine behavioral testing (4, 5, 27, 29).

5 Conclusion

The body of scientific literature that uses canine behavioral testing is immense and increasing rapidly. Many researchers and practitioners rely upon behavioral testing as a tool to investigate research questions and make practical decisions around assessing dogs' suitability for roles. However, the field is vulnerable to issues relating to the standardization of methodology, terminology, quality reporting, and interpretation. This makes it difficult to critically evaluate, select, or design behavioral tests that are appropriate for an intended purpose. These difficulties could hamper the continued improvement of the methods used to assess canine behavior and cognition. The current review provides a comprehensive overview and practical reference of the methods and uses of published canine behavioral tests and offers a novel perspective by focusing on test stimulus categories. It is anticipated that this may help researchers and practitioners make informed decisions when choosing test protocols and interpreting responses from dogs.

Author contributions

AM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. MW: Formal analysis, Supervision, Writing – review & editing. WB: Conceptualization, Supervision, Writing – review & editing. PMG: Conceptualization, Supervision, Writing – review & editing. PB: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by an Australian Commonwealth Government Research Training Program (RTP) Scholarship.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Gosling SD. From mice to men: what can we learn about personality from animal research? *Psychol Bull.* (2001) 127:45–86. doi: 10.1037/0033-2909.127.1.45
- Isa d'R, Gerlai R. Designing animal-friendly behavioral tests for neuroscience research: the importance of an ethological approach. *Front Behav Neurosci.* (2022) 16:1090248. doi: 10.3389/fnbeh.2022.1090248
- Nielsen B. Asking animals: an introduction to animal behaviour testing CABI. CABI publishing (2020).
- Diederich C, Giffroy J-M. Behavioural testing in dogs: a review of methodology in search for standardisation. *Appl Anim Behav Sci.* (2006) 97:51–72. doi: 10.1016/j.applanim.2005.11.018
- Rayment DJ, De Groef B, Peters RA, Marston LC. Applied personality assessment in domestic dogs: limitations and caveats. *Appl Anim Behav Sci.* (2015) 163:1–18. doi: 10.1016/j.applanim.2014.11.020
- Barnard S, Marshall-Pescini S, Pelosi A, Passalacqua C, Prato-Previde E, Valsecchi P. Breed, sex, and litter effects in 2-month old puppies' behaviour in a standardised open-field test. *Sci Rep.* (2017) 7:1802. doi: 10.1038/s41598-017-01992-x
- Junttila S, Valros A, Mäki K, Väättäjä H, Reunanen E, Tiira K. Breed differences in social cognition, inhibitory control, and spatial problem-solving ability in the domestic dog (*Canis Familiaris*). *Sci Rep.* (2022) 12:22529. doi: 10.1038/s41598-022-26991-5
- MacLean EL, Snyder-Mackler N, vonHoldt BM, Serpell JA. Highly heritable and functionally relevant breed differences in dog behaviour. *Proc Biol Sci.* (2019) 286:20190716. doi: 10.1098/rspb.2019.0716
- Mehrkam LR, Wynne CDL. Behavioral differences among breeds of domestic dogs (*Canis Lupus Familiaris*): current status of the science. *Appl Anim Behav Sci.* (2014) 155:12–27. doi: 10.1016/j.applanim.2014.03.005
- Goodloe LP, Borchelt PL. Companion dog temperament traits. *J Appl Anim Welfare Sci.* (1998) 1:303–38. doi: 10.1207/s15327604jaws0104_1
- Morrill K, Hekman J, Li X, McClure J, Logan B, Goodman L, et al. Ancestry-inclusive dog genomics challenges popular breed stereotypes. *Science.* (2022) 376:eabk0639. doi: 10.1126/science.abk0639
- Royce JR. A factorial study of emotionality in the dog. *Psychol Monogr.* (1955) 69:1–27. doi: 10.1037/h0093736
- Svartberg K. The hierarchical structure of dog personality in a new behavioural assessment: a validation approach. *Appl Anim Behav Sci.* (2021) 238:105302. doi: 10.1016/j.applanim.2021.105302
- Kaiser MI, Müller C. What is an animal personality? *Biol Philos.* (2021) 36:1. doi: 10.1007/s10539-020-09776-w
- Svartberg K, Forkman B. Personality traits in the domestic dog (*Canis Familiaris*). *Appl Anim Behav Sci.* (2002) 79:133–55. doi: 10.1016/S0168-1591(02)00121-1
- Junttila S, Valros A, Mäki K, Tiira K. Do Cognitive Traits Associate with Everyday Behaviour in the Domestic Dog, *Canis Familiaris?* *Anim Behav.* (2024) 213:71–84. doi: 10.1016/j.anbehav.2024.04.012
- Fuller JL. Individual differences in the reactivity of dogs. *J Comp Physiol Psychol.* (1948) 41:339–47. doi: 10.1037/h0057680
- Vernouillet AAA, Stiles LR, Andrew McCausland J, Kelly DM. Individual performance across motoric self-regulation tasks are not correlated for pet dogs. *Learn Behav.* (2018) 46:522–36. doi: 10.3758/s13420-018-0354-x
- Cummings BJ, Head E, Afagh AJ, Milgram NW, Cotman CW. Beta-amyloid accumulation correlates with cognitive dysfunction in the aged canine. *Neurobiol Learn Mem.* (1996) 66:11–23. doi: 10.1006/nlme.1996.0039

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2024.1455574/full#supplementary-material>

- Kubinyi E, Bence M, Koller D, Wan M, Pergel E, Ronai Z, et al. Oxytocin and opioid receptor gene polymorphisms associated with greeting behavior in dogs. *Front Psychol.* (2017) 8:1520. doi: 10.3389/fpsyg.2017.01520
- Fox MW, Stelzner D. The effects of early experience on the development of inter and intraspecies social relationships in the dog. *Anim Behav.* (1967) 15:377–86. doi: 10.1016/0003-3472(67)90024-3
- Turcsán B, Tátrai K, Petró E, Topál J, Balogh L, Egyed B, et al. Comparison of behavior and genetic structure in populations of family and kennel beagles. *Front Vet Sci.* (2020) 7:183. doi: 10.3389/fvets.2020.00183
- Menaker T, Monteny J, Op de Beeck L, Zamansky A. Clustering for automated exploratory pattern discovery in animal behavioral data. *Front Vet Sci.* (2022) 9:884437. doi: 10.3389/fvets.2022.884437
- Wright JC. The effects of differential rearing on exploratory behavior in puppies. *Appl Anim Ethol.* (1983) 10:27–34. doi: 10.1016/0304-3762(83)90109-8
- Brady K, Cracknell N, Zulch H, Mills DS. A systematic review of the reliability and validity of behavioural tests used to assess behavioural characteristics important in working dogs. *Front Vet Sci.* (2018) 5:103. doi: 10.3389/fvets.2018.00103
- Bray EE, Otto CM, Udell MAR, Hall NJ, Johnston AM, MacLean EL. Enhancing the selection and performance of working dogs. *Front Vet Sci.* (2021) 8:644431. doi: 10.3389/fvets.2021.644431
- Mornement KM, Coleman GJ, Toukhsati S, Bennett PC. A review of behavioral assessment protocols used by Australian animal shelters to determine the adoption suitability of dogs. *J Appl Anim Welf Sci.* (2010) 13:314–29. doi: 10.1080/10888705.2010.483856
- Patronek GJ, Bradley J, Arps E. What is the evidence for reliability and validity of behavior evaluations for shelter dogs? A prequel to 'no better than flipping a coin'. *J Vet Behav.* (2019) 31:43–58. doi: 10.1016/j.jveb.2019.03.001
- Taylor KD, Mills DS. The development and assessment of temperament tests for adult companion dogs. *J Vet Behav.* (2006) 1:94–108. doi: 10.1016/j.jveb.2006.09.002
- Lindberg S, Strandberg E, Swenson L. Genetic analysis of hunting behaviour in Swedish flatcoated retrievers. *Appl Anim Behav Sci.* (2004) 88:289–98. doi: 10.1016/j.applanim.2004.03.007
- Ruefenacht S, Gebhardt-Henrich S, Miyake T, Gaillard C. A behaviour test on German shepherd dogs: heritability of seven different traits. *Appl Anim Behav Sci.* (2002) 79:113–32. doi: 10.1016/S0168-1591(02)00134-X
- Jones A, Gosling S. Temperament and Personality in Dogs (*Canis Familiaris*): A Review and Evaluation of Past Research. *Appl Anim Behav Sci.* (2005) 95:1–53. doi: 10.1016/j.applanim.2005.04.008
- McGarrity ME, Sinn DL, Gosling SD. Which personality dimensions do puppy tests measure? A systematic procedure for categorizing behavioral assays. *Behav Process.* (2015) 110:117–24. doi: 10.1016/j.beproc.2014.09.029
- Mitchell JT, Kimbrel NA, Hundt NE, Cobb AR, Nelson-Gray RO, Lootens CM. An analysis of reinforcement sensitivity theory and the five-factor model. *Eur J Personal.* (2007) 21:869–87. doi: 10.1002/per.644
- Cobb M, Branson N, McGreevy P, Lill A, Bennett P. The advent of canine performance science: offering a sustainable future for working dogs. *Behav Process.* (2015) 110:96–104. doi: 10.1016/j.beproc.2014.10.012
- Fratkin JL. Personality in Dogs In: J Vonk, A Weiss and SA Kuczaj, editors. Personality in nonhuman animals. Cham: Springer International Publishing (2017). 205–24.
- Gartner MC. Pet personality: a review. *Personal Individ Differ.* (2015) 75:102–13. doi: 10.1016/j.paid.2014.10.042

38. Piotti P, Piseddu A, Aguzzoli E, Sommese A, Kubinyi E. Two valid and reliable tests for monitoring age-related memory performance and neophobia differences in dogs. *Sci Rep.* (2022) 12:16175. doi: 10.1038/s41598-022-19918-7
39. McConnell I, Marker L, Rooney N. Preliminary investigation into personality and effectiveness of livestock guarding dogs in namibia. *J Vet Behav.* (2022) 48:11–9. doi: 10.1016/j.jveb.2021.10.006
40. Clay L, Paterson MBA, Bennett P, Perry G, Phillips CCJ. Comparison of canine behaviour scored using a shelter behaviour assessment and an owner completed questionnaire, C-BARQ. *Animals.* (2020) 10:1797. doi: 10.3390/ani10101797
41. Stolzlechner L, Bonorand A, Riemer S. Optimising Puppy socialisation—short-and long-term effects of a training programme during the early socialisation period. *Animals.* (2022) 12:67. doi: 10.3390/ani12223067
42. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* (2018) 169:467–73. doi: 10.7326/M18-0850
43. Covidence Systematic Review Software. (2024). Melbourne, Australia: Veritas Health Innovation. Available at: www.covidence.org
44. Vears DF, Gillam L. Inductive content analysis: a guide for beginning qualitative researchers. *Focus Health Prof Educ Multi-Prof J.* (2022) 23:111–27. doi: 10.11157/fohpe.v23i1.544
45. Bergamasco L, Osella MC, Savarino P, Larosa G, Ozella L, Manassero M, et al. Heart rate variability and saliva cortisol assessment in shelter dog: Human–animal interaction effects. *Appl Anim Behav Sci.* (2010) 125:56–68. doi: 10.1016/j.applanim.2010.03.002
46. Caddiell RMP, Cunningham RM, White PA, Duncan B, Lascelles X, Gruen ME. Pain sensitivity differs between dog breeds but not in the way veterinarians believe. *Front Pain Res.* (2023) 4:1165340. doi: 10.3389/fpain.2023.1165340
47. Schalke E, Ott SA, von Gaertner AM, Hackbarth H, Mittmann A. Is breed-specific legislation justified? Study of the results of the temperament test of lower saxony. *J Vet Behav.* (2008) 3:97–103. doi: 10.1016/j.jveb.2007.10.004
48. Dowling-Guyer S, Marder A, D'Arpino S. Behavioral traits detected in shelter dogs by a behavior evaluation. *Appl Anim Behav Sci.* (2011) 130:107–14. doi: 10.1016/j.applanim.2010.12.004
49. Lazarowski L, Rogers B, Krichbaum S, Haney P, Smith JG, Waggoner P. Validation of a behavior test for predicting puppies' suitability as detection dogs. *Animals.* (2021) 11:993. doi: 10.3390/ani11040993
50. Sinn DL, Gosling SD, Hilliard S. Personality and performance in military working dogs: reliability and predictive validity of behavioral tests. *Appl Anim Behav Sci.* (2010) 127:51–65. doi: 10.1016/j.applanim.2010.08.007
51. Clay L, Paterson M, Bennett P, Perry G, Phillips C. Early recognition of behaviour problems in shelter dogs by monitoring them in their kennels after admission to a shelter. *Animals.* (2019) 9:875. doi: 10.3390/ani9110875
52. Netto WJ, Planta DJU. Behavioural testing for aggression in the domestic dog. *Appl Anim Behav Sci.* (1997) 52:243–63. doi: 10.1016/S0168-1591(96)01126-4
53. Svartberg K, Tapper I, Temrin H, Radesäter T, Thorman S. Consistency of personality traits in dogs. *Anim Behav.* (2005) 69:283–91. doi: 10.1016/j.anbehav.2004.04.011
54. Turcsán B, Wallis L, Virányi Z, Range F, Müller CA, Huber L, et al. Personality traits in companion dogs—results from the VIDOPET. *PLoS One.* (2018) 13:e0195448. doi: 10.1371/journal.pone.0195448
55. Berg L, Schilder MBH, Knol BW. Behavior genetics of canine aggression: behavioral phenotyping of golden retrievers by means of an aggression test. *Behav Genet.* (2003) 33:469–83. doi: 10.1023/A:1025714431089
56. Arvelius P, Eken Asp H, Fikse WF, Strandberg E, Nilsson K. Genetic analysis of a temperament test as a tool to select against everyday life fearfulness in rough collie. *J Anim Sci.* (2014) 92:4843–55. doi: 10.2527/jas.2014-8169
57. Roth LSV, Jensen P. Assessing companion dog behavior in a social setting. *J Vet Behav.* (2015) 10:315–23. doi: 10.1016/j.jveb.2015.04.003
58. Kubinyi E, Gosling SD, Miklósi Á. A comparison of rating and coding behavioural traits in dogs. *Acta Biol Hung.* (2015) 66:27–40. doi: 10.1556/ABiol.66.2015.1.3
59. Borg JAM, Beerda B, Ooms M, de Souza AS, van Hagen M, Kemp B. Evaluation of behaviour testing for human directed aggression in dogs. *Appl Anim Behav Sci.* (2010) 128:78–90. doi: 10.1016/j.applanim.2010.09.016
60. Clay L, Paterson MBA, Bennett P, Perry G, Phillips CCJ. Do behaviour assessments in a shelter predict the behaviour of dogs post-adoption? *Animals.* (2020) 10:71225. doi: 10.3390/ani10071225
61. McGuire B, Jean-Baptiste K. Relationships between demographic characteristics of shelter dogs and performance on tests of a behavioral evaluation and between performance and adoption success. *J Vet Behav.* (2023) 66:11–9. doi: 10.1016/j.jveb.2023.06.007
62. McGuire B, Orantes D, Xue S, PARRY S. Abilities of canine shelter behavioral evaluations and owner surrender profiles to predict resource guarding in adoptive homes. *Animals.* (2020) 10:1702. doi: 10.3390/ani10091702
63. Mornement KM, Coleman GJ, Toukhsati SR, Bennett PC. Evaluation of the predictive validity of the behavioural assessment for re-homing K9's (B.A.R.K.) protocol and owner satisfaction with adopted dogs. *Appl Anim Behav Sci.* (2015) 167:35–42. doi: 10.1016/j.applanim.2015.03.013
64. Marcato M, Tedesco S, O'Mahony C, O'Flynn B, Galvin P. Prediction of working outcomes in trainee dogs using the novel assistance dog test battery (ADTB). *Appl Anim Behav Sci.* (2024) 272:106212. doi: 10.1016/j.applanim.2024.106212
65. Wilsson E, Sinn DL. Are there differences between behavioral measurement methods? A comparison of the predictive validity of two ratings methods in a working dog program. *Appl Anim Behav Sci.* (2012) 141:158–72. doi: 10.1016/j.applanim.2012.08.012
66. King T, Hemsworth PH, Coleman GJ. Fear of novel and startling stimuli in domestic dogs. *Appl Anim Behav Sci.* (2003) 82:45–64. doi: 10.1016/S0168-1591(03)00040-6
67. Ley J, Coleman GJ, Holmes R, Hemsworth PH. Assessing fear of novel and startling stimuli in domestic dogs. *Appl Anim Behav Sci.* (2007) 104:71–84. doi: 10.1016/j.applanim.2006.03.021
68. Stellato AC, Flint HE, Widowski TM, Serpell JA, Niel L. Assessment of fear-related behaviours displayed by companion dogs (*Canis Familiaris*) in response to social and non-social stimuli. *Appl Anim Behav Sci.* (2017) 188:84–90. doi: 10.1016/j.applanim.2016.12.007
69. Haverbeke A, De Smet A, Depiereux E, Giffroy J-M, Diederich C. Assessing undesired aggression in military working dogs. *Appl Anim Behav Sci.* (2009) 117:55–62. doi: 10.1016/j.applanim.2008.12.002
70. Gould TD, Dao DT, Kovacsics CE. The open field test In: TD Gould, editor. Mood and anxiety related phenotypes in mice: characterization using behavioral tests. Totowa, NJ: Humana Press (2009). 1–20.
71. Shin CW, Kim GA, Park WJ, Park KY, Jeon JM, Oh HJ, et al. Learning, memory and exploratory similarities in genetically identical cloned dogs. *J Vet Sci.* (2016) 17:563–7. doi: 10.4142/jvs.2016.17.4.563
72. Forkman B, Boissy A, Meunier-Salaün M-C, Canali E, Jones RB. A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiol Behav.* (2007) 92:340–74. doi: 10.1016/j.physbeh.2007.03.016
73. Wilsson E, Sundgren P-E. Behaviour test for eight-week old puppies—heritabilities of tested behaviour traits and its correspondence to later behaviour. *Appl Anim Behav Sci.* (1998) 58:151–62. doi: 10.1016/S0168-1591(97)00093-2
74. Völter CJ, Starić D, Huber L. Using machine learning to track dogs' exploratory behaviour in the presence and absence of their caregiver. *Anim Behav.* (2023) 197:97–111. doi: 10.1016/j.anbehav.2023.01.004
75. Gazzano A, Mariti C, Notari L, Sighieri C, McBride EA. Effects of early gentling and early environment on emotional development of puppies. *Appl Anim Behav Sci.* (2008) 110:294–304. doi: 10.1016/j.applanim.2007.05.007
76. Riemer S, Müller C, Virányi Z, Huber L, Range F. The Predictive value of early behavioural assessments in pet dogs – a longitudinal study from neonates to adults. *PLoS One.* (2014) 9:e101237. doi: 10.1371/journal.pone.0101237
77. Goddard ME, Beilharz RG. A factor analysis of fearfulness in potential guide dogs. *Appl Anim Behav Sci.* (1984) 12:253–65. doi: 10.1016/0168-1591(84)90118-7
78. Lensen RCMM, Moons CPH, Diederich C. Physiological stress reactivity and recovery related to behavioral traits in dogs (*Canis Familiaris*). *PLoS One.* (2019) 14:e0222581. doi: 10.1371/journal.pone.0222581
79. Menuge F, Marcet-Rius M, Jochem M, François O, Assali C, Chabaud C, et al. Early evaluation of fearfulness in future guide dogs for blind people. *Animals.* (2021) 11:412. doi: 10.3390/ani11020412
80. Hennessy MB, Voith VL, Mazzei SJ, Buttram J, Miller DD, Linden F. 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. *Appl Anim Behav Sci.* (2001) 73:217–33. doi: 10.1016/S0168-1591(01)00139-3
81. Bray EE, Levy KM, Kennedy BS, Duffy DL, Serpell JA, MacLean EL. Predictive models of assistance dog training outcomes using the canine behavioral assessment and research questionnaire and a standardized temperament evaluation. *Front Vet Sci.* (2019) 6:49. doi: 10.3389/fvets.2019.00049
82. Lazarowski L, Rogers B, Smith JG, Krichbaum S, Waggoner P. Longitudinal stability of detection dog behavioral assessment: a follow-up study of long-term working success. *Appl Anim Behav Sci.* (2023) 268:106082. doi: 10.1016/j.applanim.2023.106082
83. Asher L, Blythe S, Roberts R, Toothill L, Craigon PJ, Evans KM, et al. A standardized behavior test for potential guide dog puppies: methods and association with subsequent success in guide dog training. *J Vet Behav.* (2013) 8:431–8. doi: 10.1016/j.jveb.2013.08.004
84. MacLean EL, Hare B. Enhanced selection of assistance and explosive detection dogs using cognitive measures. *Front Vet Sci.* (2018) 5:236. doi: 10.3389/fvets.2018.00236
85. Lazarowski L, Thompkins A, Sarah Krichbaum L, Waggoner P, Deshpande G, Katz JS. Comparing pet and detection dogs (*Canis Familiaris*) on two aspects of social cognition. *Learn Behav.* (2020) 48:432–43. doi: 10.3758/s13420-020-00431-8
86. Wells DL, Hepper PG, Milligan ADS, Barnard S. Stability of motor bias in the domestic dog, *Canis Familiaris*. *Behav Process.* (2018) 149:1–7. doi: 10.1016/j.beproc.2018.01.012

87. MacLean EL, Herrmann E, Suchindran S, Hare B. Individual differences in cooperative communicative skills are more similar between dogs and humans than chimpanzees. *Anim Behav.* (2017) 126:41–51. doi: 10.1016/j.anbehav.2017.01.005
88. Stewart L, MacLean EL, Ivy D, Woods V, Cohen E, Rodriguez K, et al. Citizen science as a new tool in dog cognition research. *PLoS One.* (2015) 10:e0135176. doi: 10.1371/journal.pone.0135176
89. Marshall-Pescini S, Passalacqua C, Barnard S, Valsecchi P, Prato-Previde E. Agility and search and rescue training differently affects pet dogs' behaviour in socio-cognitive tasks. *Behav Process.* (2009) 81:416–22. doi: 10.1016/j.beproc.2009.03.015
90. Persson ME, Sundman A-S, Halldén L-L, Trottier AJ, Jensen P. Sociality genes are associated with human-directed social behaviour in golden and labrador retriever dogs. *PeerJ.* (2018) 6:e5889. doi: 10.7717/peerj.5889
91. Gould K, Iversen P, Sikkink S, Rem R, Templeton J. Persistence and gazing at humans during an unsolvable task in dogs: the influence of ownership duration, living situation, and prior experience with humans. *Behav Process.* (2022) 201:104710. doi: 10.1016/j.beproc.2022.104710
92. McPeake KJ, Collins LM, Zulch H, Mills DS. Behavioural and physiological correlates of the canine frustration questionnaire. *Animals.* (2021) 11:3346. doi: 10.3390/ani11123346
93. Turcsán B, Wallis L, Berczik J, Range F, Kubinyi E, Virányi Z. Individual and group level personality change across the lifespan in dogs. *Sci Rep.* (2020) 10:17276. doi: 10.1038/s41598-020-74310-7
94. Barnard S, Wells DL, Milligan ADS, Arnott G, Hepper PG. Personality traits affecting judgement bias task performance in dogs (*Canis Familiaris*). *Sci Rep.* (2018) 8:6660. doi: 10.1038/s41598-018-25224-y
95. Brady K, Hewison L, Wright H, Zulch H, Cracknell N, Mills D. A spatial discounting test to assess impulsivity in dogs. *Appl Anim Behav Sci.* (2018) 202:77–84. doi: 10.1016/j.applanim.2018.01.003
96. Olsen MR. An investigation of two ostensibly inhibitory control tasks used in canine cognition. *Appl Anim Behav Sci.* (2022) 256:105770. doi: 10.1016/j.applanim.2022.105770
97. Kubinyi E, Iotchev IB. A Preliminary Study toward a Rapid Assessment of Age-Related Behavioral Differences in Family Dogs. *Animals.* (2020) 10:1222. doi: 10.3390/ani10071222
98. Bray EE, Sammel MD, Seyfarth RM, Serpell JA, Cheney DL. Temperament and problem solving in a population of adolescent guide dogs. *Anim Cogn.* (2017) 20:923–39. doi: 10.1007/s10071-017-1112-8
99. Lazarowski L, Sarah Krichbaum L, Waggoner P, Katz JS. The development of problem-solving abilities in a population of candidate detection dogs (*Canis Familiaris*). *Anim Cogn.* (2020) 23:755–68. doi: 10.1007/s10071-020-01387-y
100. Lazarowski L, Bart Rogers L, Waggoner P, Katz JS. When the nose knows: ontogenetic changes in detection dogs' (*Canis Familiaris*) responsiveness to social and olfactory cues. *Anim Behav.* (2019) 153:61–8. doi: 10.1016/j.anbehav.2019.05.002
101. Tiira K, Tikkanen A, Vainio O. Inhibitory control – important trait for explosive detection performance in police dogs? *Appl Anim Behav Sci.* (2020) 224:104942. doi: 10.1016/j.applanim.2020.104942
102. Tomkins LM, Thomson PC, McGreevy PD. Associations between motor, sensory and structural lateralisation and guide dog success. *Vet J.* (2012) 192:359–67. doi: 10.1016/j.tvjl.2011.09.010
103. Payne E, Boot M, Starling M, Henshall C, McLean A, Bennett P, et al. Evidence of horsemanship and dogmanship and their application in veterinary contexts. *Vet J.* (2015) 204:247–54. doi: 10.1016/j.tvjl.2015.04.004
104. Starling MJ, Branson N, Cody D, Starling TR, McGreevy PD. Canine sense and sensibility: tipping points and response latency variability as an optimism index in a canine judgement bias assessment. *PLoS One.* (2014) 9:e107794. doi: 10.1371/journal.pone.0107794
105. Edwards PT, Smith BP, McArthur ML, Hazel SJ. Fearful fido: investigating dog experience in the veterinary context in an effort to reduce distress. *Appl Anim Behav Sci.* (2019) 213:14–25. doi: 10.1016/j.applanim.2019.02.009
106. Plutchik R. Individual and breed differences in approach and withdrawal in dogs. *Behaviour.* (1971) 40:302–11. doi: 10.1163/156853971X00447
107. Franzini de Souza CC, Maccariello CE, Dias DP, Almeida NA, Medeiros MA. Autonomic, endocrine and behavioural responses to thunder in laboratory and companion dogs. *Physiol Behav.* (2017) 169:208–215. doi: 10.1016/j.physbeh.2016.12.006
108. Katayama M, Kubo T, Mogi K, Ikeda K, Nagasawa M, Kikusui T. Heart rate variability predicts the emotional state in dogs. *Behav Process.* (2016) 128:108–12. doi: 10.1016/j.beproc.2016.04.015
109. Farhat N, Lazebnik T, Monteny J, Moons CPH, Wydooghe E, van der Linden D, et al. Digitally-enhanced dog behavioral testing. *Sci Rep.* (2023) 13:21252. doi: 10.1038/s41598-023-48423-8
110. Farhat N, van der Linden D, Zamansky A, Assif T. Automation in canine science: enhancing human capabilities and overcoming adoption barriers. *Front Vet Sci.* (2024) 11:1394620. doi: 10.3389/fvets.2024.1394620