

The Dogmetry Dog Personality Inventory (DDPI)

Development and Validation — Technical Report, version 1.1

Nazar Akrami, professor of psychology (Uppsala University) · Universal Measures, Sweden · August 2026

This report describes how the DDPI was built and how it performed in validation. It is written for interested dog owners as well as researchers; the technical detail is real, but no statistics background beyond curiosity is required for the main text. Version 1.1 (August 2026) extends version 1.0 with a second, independent validation wave — combined N = 528 — a facet-level reliability table, and the confirmed short forms. Questions and requests (including researcher access to the full item bank and de-identified data): research@dogmetry.com.

1. What the DDPI measures — and what it does not

The DDPI is a questionnaire an owner completes about their dog. It measures personality: stable, general tendencies in how a dog typically behaves and reacts — how social, how excitable, how bold, how quick to learn. It does not measure, diagnose, or screen for behavior problems, and it is never used that way on the Dogmetry platform. Two dogs with the same personality profile can behave very differently, because behavior also reflects history, training, and context. The DDPI describes who a dog is; it does not judge how a dog was raised.

The DDPI is an informant-report instrument: the person who knows the dog best describes the dog. A meta-analysis of 31 studies shows that such owner ratings are about as consistent and valid as trained-observer codings of dog behavior (Fratkin et al., 2013; see also Gosling, Kwan, & John, 2003).

2. Structure and theoretical background

The DDPI measures five broad factors, each composed of three narrower facets, each measured by five items — 75 items in total. The five-factor structure follows the strongest convergence in the dog-personality literature (Jones, 2008; Ley et al., 2008; Svartberg & Forkman, 2002), aligned with, but not copied from, the human Big Five.

Factor	≈ Human analogue	Facets
Sociability	Agreeableness	Love for their people · Friendliness · Gentleness
Emotional Reactivity	Neuroticism	Shy with people · Scared of new things · Nervous and jumpy
Energy	Extraversion	Activity level · Excitability · Playfulness
Trainability	Openness / focus	Quick to learn · Listens and waits · Clever
Boldness	Self-assuredness	Confidence · Curiosity · Determination

All 75 items are original, written for the DDPI in plain, warm, phone-friendly language, and shown with the dog's name ("Bella loves being close to me" rather than "My dog..."). No item reproduces wording from existing instruments (C-BARQ, DPQ, MCPQ-R), which are referenced at the level of constructs only. Seventeen items are reverse-keyed as a guard against acquiescent responding. Items are rated on a 7-point scale (1 = Not at all like my dog, 4 = In between, 7 = Exactly like my dog), with a "Not sure / haven't seen this" option that is treated as missing data, never as a midpoint. Seven points were chosen over five to preserve variance at the top of the scale, where owner ratings of their own dogs concentrate (cf. Simms et al., 2019).

3. Forms and scoring

Three forms are drawn from the same 75-item pool: Snapshot (25 items; factor scores only; the free tier), Standard (45 items, 3 per facet; factor and facet scores; the default full assessment and co-rater form), and Complete (75 items; highest reliability; research and repeat testing). The forms nest: every Snapshot item is in the Standard form, and every Standard item in the Complete form.

Scoring is transparent: reverse-keyed items are reflected ($8 - x$), facet scores are the mean of their items, factor scores the mean of their three facets, all on the 1–7 metric. A facet score requires at least 3 of 5 items answered (2 of 3 on the Standard form); a factor score requires at least two valid facets. The platform recommends three assessment sittings at least a week apart and reports their mean as the dog’s stable profile — because single-occasion measurement captures the day as well as the dog.

Comparisons between dogs use verbal bands rather than statistical jargon: raw factor-score differences under 0.75 are described as “very similar”, 0.75–1.5 “somewhat different”, above 1.5 “clearly different”. Population comparisons (percentiles and verbal bands) are computed against the platform’s own frozen reference population — version 1 comprises the 528 full-form profiles from the two validation waves — which changes only at announced milestone updates, so a dog’s standing never drifts from day to day. The DDPI borrows no norms from other instruments, and no breed-specific norms are offered (breed explains only a modest share of individual behavioral variation; Morrill et al., 2022).

4. The validation studies

4.1 Participants and procedure

Two validation waves were run in August 2026 with participants recruited through CloudResearch Connect, a research panel, and compensated at a fair hourly rate (participant ratings: 4.4/5 in wave one, 4.9/5 in wave two). Eligibility: current dog owners, personally responsible for a dog that had lived in the household at least 3 months, residing in the US, UK, Canada, Australia, Ireland, or New Zealand. Wave-one participants were excluded from wave two, so every row is a different owner–dog pair. Each participant completed the full 75-item Complete form about one dog, plus screening questions, basic dog facts, and two embedded attention checks. Consent identified the responsible researcher and data controller; responses were anonymous and used for research and instrument development only.

538 questionnaires were completed across the two waves. Excluded before analysis, by fixed rules applied identically to both waves: 3 internal test runs, 3 duplicate sessions (first completion kept), 3 failed attention checks, and 1 respondent answering implausibly fast — leaving $N = 528$: 172 from wave one (one more than reported in version 1.0, from a session completed after that count was made) and 356 from wave two. Every retained respondent passed both attention checks. “Not sure” was used for just 0.14% of item responses — owners could answer essentially every item about their own dog.

4.2 The dogs

Characteristic	Value
Sex	54.4% male, 45.6% female
Neuter status	74.8% neutered, 24.1% intact
Age	$M = 6.9$ years (SD 4.0), range 0.4–22.6
Breed	40.0% mixed breed; 65 distinct pure breeds among the rest
Time in household	74.4% ≥ 3 years; 19.1% 1–3 years; 6.4% 3–12 months
Median completion time	6.7 minutes (IQR 4.9–8.7)

5. Results

5.1 Reliability

Internal consistency (Cronbach's α) for the five factor scales, by form:

Factor	Snapshot-25	Standard-45	Complete-75
Sociability	.66	.75	.84
Emotional Reactivity	.87	.92	.92
Energy	.80	.87	.87
Trainability	.86	.90	.90
Boldness	.77	.84	.87

All fifteen facets of the Complete form reach $\alpha \geq .70$ (range .71–.86); on the 3-item Standard form, thirteen of fifteen do (range .66–.90) — the facet-level table follows below. Factor reliabilities replicated across the two independent waves essentially to the second decimal (Emotional Reactivity $\alpha = .92$ in both). These values compare well with established instruments — the DPQ long form, for example, reports $\alpha .77$ –.88 at factor level (Jones, 2008). Sociability is the most modest scale throughout, largely because some of its items (“loves being close to me”) are true of nearly every dog: near-universal truths make poor differentiators, however valid.

Facet	α Complete (5 items)	α Standard (3 items)
1.1 Love for their people	.72	.76
1.2 Friendliness	.84	.77
1.3 Gentleness	.74	.66
2.1 Shy with people	.85	.88
2.2 Scared of new things	.81	.83
2.3 Nervous and jumpy	.82	.85
3.1 Activity level	.82	.84
3.2 Excitability	.71	.66
3.3 Playfulness	.86	.86
4.1 Quick to learn	.86	.90
4.2 Listens and waits	.77	.74
4.3 Clever	.85	.85
5.1 Confidence	.79	.79
5.2 Curiosity	.78	.75
5.3 Determination	.77	.82

5.2 Item performance

67 of 75 items performed well in wave one, and the same items led the table in the combined sample (median corrected item–total correlation .61). A systematic finding, stable across both waves: reverse-keyed items were weaker on average (.48 vs .63) — a well-known method effect in questionnaire research, not unique to dogs or to this instrument. Consequently, the short forms favor the strongest items regardless of keying, while the Complete form retains all 17 reverse-keyed items as a response-style check.

5.3 Structure

The five factors are distinct but not independent: notably, dogs high in Emotional Reactivity tend to be rated lower in Sociability ($r = -.69$ in the combined sample), and factor analysis at the facet level shows Friendliness and (low) fearfulness blending into the broad shyness–boldness dimension long documented in dogs (Svartberg & Forkman, 2002), while warmth toward the owner separates from it. This picture replicated in the second, independent wave at three times the original sample size — a stable property of owner-rated dog personality, not sampling noise. For users, the practical meaning is simple: some traits travel together, and Dogmetry's profiles report them separately while being honest about their relatedness.

5.4 Short-form selection

The initial Snapshot and Standard item selections had been made by expert judgment. Using the wave-one data, both were re-selected empirically — the items that best represent each facet and factor, with coverage rules so no facet is crowded out. The improvement was substantial: the Snapshot's weakest factor scale rose from $\alpha = .54$ to $.64$ and its strongest to $.89$; on the Standard form the number of facets below $\alpha = .70$ fell from seven to two. Wave two then served as the planned independent confirmation: on 356 new dogs, twelve of the fifteen Standard facet selections proved empirically optimal exactly as chosen; one item was exchanged (in Listens and waits), one near-tie was left in place, and one candidate exchange was rejected after an explicit test showed it would weaken the Snapshot's Boldness scale. With that, the short forms are no longer provisional: Snapshot-25 is unchanged and Standard-45 is confirmed with a single one-item revision, both frozen as of August 2026.

5.5 Reference values (validation sample)

The table below gives the score distributions observed across both waves ($N = 528$), on the 1–7 metric after reverse-keying. These values constitute the platform's frozen reference population, version 1: every percentile and verbal band shown on Dogmetry is computed against this reference, which changes only at announced milestone updates. The DDPI borrows no norms from other instruments.

Factor	M	SD	Facet	M	SD
Sociability	5.28	0.91	1.1 Love for their people	6.20	0.82
Emotional Reactivity	3.55	1.28	1.2 Friendliness	4.63	1.46
Energy	5.00	1.02	1.3 Gentleness	5.00	1.27
Trainability	5.13	1.01	2.1 Shy with people	3.54	1.59
Boldness	5.26	0.88	2.2 Scared of new things	3.76	1.40
			2.3 Nervous and jumpy	3.36	1.35
			3.1 Activity level	4.70	1.29
			3.2 Excitability	4.93	1.12
			3.3 Playfulness	5.37	1.40
			4.1 Quick to learn	5.26	1.25
			4.2 Listens and waits	4.75	1.17
			4.3 Clever	5.38	1.17
			5.1 Confidence	4.76	1.27
			5.2 Curiosity	5.54	0.98
			5.3 Determination	5.48	0.98

Owner ratings of their own dogs lean positive — most means sit above the scale midpoint — which is exactly why the DDPI uses a 7-point scale (variance is preserved at the top) and why comparisons use observed distributions rather than assuming the midpoint is “average”.

6. Limitations

Two validation waves totalling $N = 528$, recruited from an English-language research panel — a solid foundation, not a final word. The wave-two confirmation removed the main statistical caveat of version 1.0 (item selection evaluated on the same sample that chose it). What remains unestimated: test–retest stability, inter-rater agreement (owner vs. co-rater), cross-language measurement invariance, and links to outcomes — each has a planned study. The reference population consists of paid English-speaking panelists describing their dogs once; platform users may differ, and a panel-vs-platform equivalence analysis is planned as platform data accumulates. And as always: the DDPI describes personality; it is not a diagnostic, welfare, or training-outcome instrument.

7. Ongoing and planned research

(1) Platform cohort: dogs on Dogmetry complete the Standard or Complete form repeatedly, providing test–retest data and a growing pool for future reference updates. (2) Co-rater study:

household members and friends rate the same dogs, estimating inter-rater agreement. (3) Owner-personality study: participants from the validation waves will be re-invited to complete a human Big Five measure, quantifying how the rater's own personality colors dog ratings — and, more delightfully, how similar dogs and their humans actually are. (4) Translation and invariance: the DDPI exists in seven languages; formal invariance testing precedes each non-English launch.

References

- Fratkin, J. L., Sinn, D. L., Patall, E. A., & Gosling, S. D. (2013). Personality consistency in dogs: A meta-analysis. *PLoS ONE*, 8(1), e54907.
- Gosling, S. D., Kwan, V. S. Y., & John, O. P. (2003). A dog's got personality: A cross-species comparative approach to personality judgments in dogs and humans. *Journal of Personality and Social Psychology*, 85(6), 1161–1169.
- Hsu, Y., & Serpell, J. A. (2003). Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *Journal of the American Veterinary Medical Association*, 223(9), 1293–1300.
- Jones, A. C. (2008). Development and validation of a Dog Personality Questionnaire (Doctoral dissertation, University of Texas at Austin).
- Ley, J. M., Bennett, P. C., & Coleman, G. J. (2008). Refinement and validation of the Monash Canine Personality Questionnaire–Revised (MCPQ-R). *Applied Animal Behaviour Science*, 116(2–4), 220–227.
- Morrill, K., et al. (2022). Ancestry-inclusive dog genomics challenges popular breed stereotypes. *Science*, 376(6592), eabk0639.
- Rayment, D. J., De Groef, B., Peters, R. A., & Marston, L. C. (2015). Applied personality assessment in domestic dogs: Limitations and caveats. *Applied Animal Behaviour Science*, 163, 1–18.
- Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychological Assessment*, 31(4), 557–566.
- Svartberg, K., & Forkman, B. (2002). Personality traits in the domestic dog (*Canis familiaris*). *Applied Animal Behaviour Science*, 79(2), 133–155.
- Turcsán, B., Kubinyi, E., & Miklósi, Á. (2011). Trainability and boldness traits differ between dog breed clusters. *Applied Animal Behaviour Science*, 132(1–2), 61–70.