

Beyond the Big Five: What Drawmetrics Adds to the Description of Personality at Work

Executive Summary

Larry R. Price, Ph.D. *Version 2 — Revised for alignment with Price (2026)*, Behavioral Sciences, 16(1), 135

Purpose of This Summary

This summary describes what current research on Drawmetrics (DM) has established, what it suggests, and what remains untested. It is written to give organizations an accurate basis for deciding how, and how far, to use the instrument at its present stage of development.

Background

Organizations have used the Big Five personality model (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism) for decades to inform hiring, leadership development, and employee growth. The model has an extensive validation record and a well-established structure.

Its descriptive level, however, is broad. Knowing that an employee scores high on Conscientiousness does not tell a manager whether that person is task-focused, service-oriented, detail-driven, or forward-looking. Organizations frequently want vocabulary closer to the behavior they actually observe.

What Drawmetrics Is

Drawmetrics is a nonverbal assessment in which respondents create doodles from a set of symbols and then supply descriptive terms, ten relating to professional context and ten to personal context. Those terms are analyzed using computational semantic methods, including sentence-embedding models and network analysis.

DM is best understood at this stage as an **exploratory descriptive instrument**. It generates personality-relevant language in an unusual way, and research is underway to determine what that language measures and what it predicts.

What the Published Research Established

The peer-reviewed study (Price, 2026) examined 176 respondents who completed both DM and a Big Five inventory. Three findings are on the record:

1. DM terms can be interpreted within Big Five facet language. When DM descriptors were mapped onto IPIP-NEO facet vocabulary, the resulting semantic configuration aligned closely with Big Five structure ($RV = 0.937$; Procrustes = 0.94). This shows that DM output can be *translated* into a familiar framework.

2. That alignment was expected by design. Because DM terms and Big Five facet anchors were represented in the same language model, close configural correspondence was anticipated. The result demonstrates interpretable translation, not that the two instruments measure the same thing.

3. At the individual level, DM and Big Five scores did not converge. Correlations between matched DM and Big Five domains were near zero (mean $|r| \approx 0.067$, range -0.083 to 0.122), and correlations between non-matching domains were of similar size. **DM scores and Big Five scores are not interchangeable, and DM should not be described as a Big Five measure.**

What Preliminary Analyses Suggest

Additional analyses, **exploratory, unpublished, and not yet peer reviewed**, indicate that a substantial portion of the vocabulary respondents produce does not correspond to Big Five facet language. Preliminary clustering suggests groupings resembling work-relevant categories such as task orientation, service focus, and resourcefulness.

This is the most scientifically interesting observation in the DM research to date, and it is the focus of ongoing work. It should be understood as a **hypothesis under investigation**, not an established finding. Whether these clusters reflect meaningful individual differences, features of the doodle task itself, or measurement noise has not yet been determined.

What Has Not Been Established

Clear statement of current limitations, provided so that organizations can make informed decisions:

Question	Status
Do DM scores stay stable over time?	Not tested. No test-retest study has been conducted.
Do DM scores predict job performance, tenure, or training success?	Not tested. No criterion study has been conducted.
Do DM scores function equivalently across demographic groups?	Not tested. Demographic data have not been collected.
Does DM add predictive information beyond the Big Five?	Not tested. This requires criterion data.
Does DM function independently of language?	Not established. The current scoring pipeline is anchored in English-language models.

Question	Status
What are respondents doing when they produce DM terms?	Not studied. Cognitive interview research is recommended in future research.

Appropriate Use at the Present Stage

Based on the current evidence base:

Supported uses. Developmental conversations, coaching, team discussion, and self-reflection, contexts in which DM output serves as a prompt for dialogue rather than as a basis for consequential decisions.

Not currently supported. Hiring, promotion, placement, or any other selection decision. Under the *Uniform Guidelines on Employee Selection Procedures*, use of an assessment for selection requires validity evidence for that use. DM does not yet have it. Organizations using DM in a selection context assume legal and ethical risk that the current evidence base does not mitigate.

A note on interpretation. DM output describes the language a person produced in response to a specific task on a specific occasion. It should not be described as revealing subconscious content, bypassing conscious bias, or providing insight unavailable to the respondent. No research supports those characterizations.

Proposed Research Program

A staged validation program is in development, addressing in sequence:

1. **Structure** — recovering DM's dimensionality from a larger sample, independent of Big Five anchors
2. **Response processes** — cognitive interview research on what respondents are doing when they generate terms
3. **Stability** — test-retest study
4. **Prediction** — whether DM adds information beyond the Big Five in forecasting work outcomes

Each stage is designed so that its results can be reported accurately whether or not they favor the instrument.

Summary

Drawmetrics elicits work-relevant descriptive language through a method unlike conventional inventories, and preliminary evidence suggests that much of that language falls outside the Big Five's descriptive range. **That is a promising and unusual finding, and it should be investigated systematically.**

At present, DM is a developmental and exploratory tool. It is not yet a validated selection instrument, and organizations should **not** use it as one until the evidence supporting that use exists.

About the Author

Larry R. Price, Ph.D., is Professor of Biostatistics and Psychometrics. He has worked in psychometrics and measurement for more than three decades and is the author of *Psychometric Methods: Theory into Practice* (Guilford Publications, 2016; second edition in preparation).

Disclosure. Dr. Price has been engaged as a paid scientific consultant by AttituX in connection with research on Drawmetrics. He retains full independence over research findings, interpretation, and publication decisions.

This summary reflects the evidence base as of [date]. It will be revised as the research program produces new findings.