Psychometric Review of the Big Tent Dark Traits Test

Jennifer Schulz, Ph.D., MSW, LSW
Associate Professor | BSW Program Director
Ashland University

25 August 2026

I conducted a preliminary psychometric review of the IDRlabs Big Tent Dark Traits Test using the supplied results package covering 35 items, seven proposed facets, and 52,739 valid test sessions. The purpose was to evaluate data quality, score distribution, internal consistency, item behaviour, and the extent to which the seven proposed facets function as related components of a broader measure.

1. Data quality

The dataset is unusually large for an initial psychometric evaluation. The original export contained 52,939 records. After same-visitor identical-answer duplicates were collapsed, 52,739 valid test sessions remained. The test was represented across 18 published locales.

The supplied package indicates that records are pseudonymous and do not contain respondent names or contact information. Incomplete or out-of-range sittings were excluded during preparation, while straight-line response patterns were retained.

A small number of respondents, 222 or approximately 0.42% of the valid sample, selected the same response option for every item. This is unlikely to determine the overall results, although future validation work should still specify an a priori response-quality criterion.

The sample size is more than adequate for psychometric analysis. However, participants are self-selected internet users rather than a probability sample, so the dataset should not be considered a representative normative sample of the general population.

2. Score distribution

The overall score distribution is statistically well behaved:

- Mean = 56.86/100
- Median = 57.1
- SD = 14.03
- Skewness = -0.15
- Excess kurtosis = 0.14

The distribution is close to symmetric, with only slight negative skew and little departure from normal kurtosis. The 5th and 95th percentiles are 34 and 79, respectively, indicating useful spread without severe floor or ceiling compression across the sample as a whole.

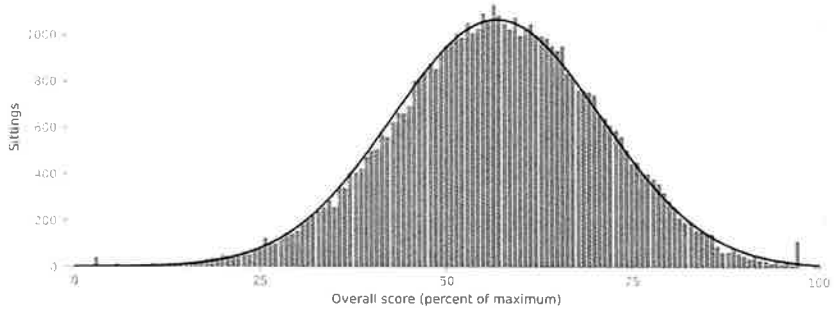


Figure 1. Overall-score distribution from the supplied results package.

Facet means range from 48.75 for Narcissism to 67.90 for Perfectionism. This spread indicates that the measure is not simply producing the same score profile across all seven proposed domains.

3. Internal consistency and item behaviour

Internal consistency is good at the level of the full 35-item measure.

Cronbach's α = .858

Internal consistency for the seven five-item facets is as follows:

- Machiavellianism: α = .691
- Narcissism: α = .771
- Psychopathy: α = .588
- Spitefulness: α = .742
- Greed: α = .520
- Perfectionism: α = .695
- Dependency: α = .718

For brief five-item scales, several of these values are acceptable for exploratory work, but reliability is clearly uneven. Greed and Psychopathy are the weakest facets in the supplied results, while Narcissism, Spitefulness, and Dependency are stronger.

Corrected item-total correlations with the overall score range from r = .04 to .56. Most items relate positively to the total, but several Dependency items have very weak overall item-total correlations. This is important because a low correlation with the total can reflect either a weak item or a genuinely distinct facet in a multidimensional instrument.

4. Facet structure

The seven facet scores show a heterogeneous correlation pattern, ranging from r = -.28 to .62. Machiavellianism, Psychopathy, Spitefulness, and Greed are moderately to strongly positively related, while Dependency has near-zero relationships with several of these facets and is negatively correlated with Narcissism (r = -.28).

The supplied first-component loadings reinforce this structural concern: the Dependency items have loadings close to zero or slightly negative on the first component. This suggests that Dependency may not participate in the same general dimension as the core dark-trait facets.

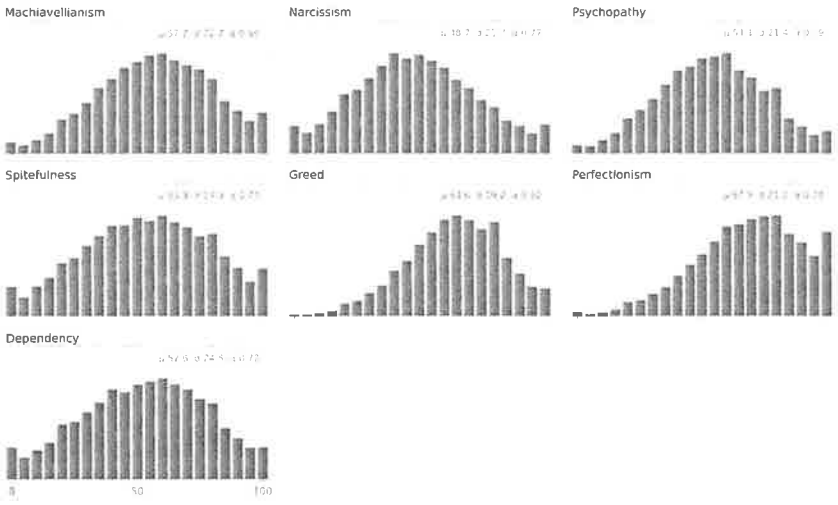


Figure 2. Facet-score distributions from the supplied results package.

Accordingly, the strong full-scale alpha should not be taken as sufficient evidence for a single overall dark-traits score. Confirmatory factor analysis or comparable latent-variable analysis is needed to determine whether a general factor is defensible, whether correlated facet scores are preferable, and whether Dependency belongs in the same higher-order construct.

5. Construct interpretation

The available results demonstrate useful overall score dispersion, good full-scale internal consistency, and meaningful differentiation among several facets. They do not establish that the seven domains are equally reliable or that a single total score has a clear unitary psychological interpretation.

The principal unresolved issue is structural validity. In particular, the behaviour of Dependency differs from the other facets, and the lower reliability of Greed and Psychopathy suggests that some subscales may benefit from item revision before strong individual-level interpretations are made.

Further validation should ideally examine:

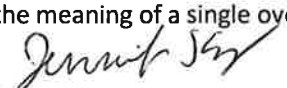
- convergent validity against established measures of each proposed dark trait;
- discriminant validity among the seven facets and against related personality constructs;
- confirmatory factor analysis comparing correlated-facet, higher-order, bifactor, and alternative models;
- item-level review and possible refinement of the lower-reliability Greed and Psychopathy facets;
- test-retest reliability;
- measurement invariance across language, sex, age, and other relevant groups.

6. Overall conclusion

The initial psychometric evidence for the Big Tent Dark Traits Test is mixed but useful. The 35-item instrument has good overall internal consistency and a well-behaved total-score distribution in a very large sample.

Facet reliability is uneven, and the correlation and loading patterns indicate that the seven domains do not all behave as manifestations of one simple common dimension. Dependency is especially distinct from the core dark-trait cluster.

The available evidence supports continued development and formal validation of the measure. Until its factor structure is tested directly, the seven facet scores should be interpreted on their own terms and the meaning of a single overall score should be treated cautiously.

Signature: 

Jennifer Schulz, Ph.D., MSW, LSW

25 August 2026