

Psychometric Review of the Misandry 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 IDRI labs Misandry Test using the supplied results package covering 15 items, three proposed facets, and 9,732 valid test sessions. The purpose was to evaluate data quality, score distribution, internal consistency, item behaviour, and the extent to which the three proposed facets function as distinguishable components of the broader measure.

1. Data quality

The dataset is large for an initial psychometric evaluation. The original export contained 9,820 records. After same-visitor identical-answer duplicates were collapsed, 9,732 valid test sessions remained. The test was represented across 15 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 notable 1,062 respondents, approximately 10.9% of the valid sample, selected the same response option for every item. This is sufficiently common that it should be treated as a substantive data-quality issue rather than a minor irregularity. Future analyses should pre-specify a careless-response rule and report sensitivity analyses with these cases both included and excluded.

The sample size is more than adequate for psychometric analysis. However, participants are self-selected internet users rather than a probability sample, so these data should not be treated as representative population norms or as estimates of the prevalence of negative attitudes toward men.

2. Score distribution

The overall score distribution is concentrated toward the upper end of the scale:

- Mean = 72.30/100
- Median = 78.3
- SD = 23.97
- Skewness = -0.96
- Excess kurtosis = 0.34

The distribution is substantially negatively skewed. The 75th percentile is 92, the 90th percentile is 98, and the 95th percentile is 100. The marked accumulation of scores near the maximum indicates an upper-end ceiling effect that may reduce differentiation among respondents with high scores.

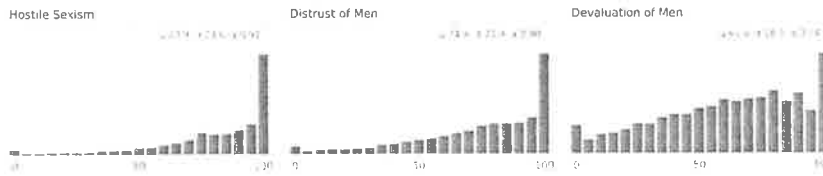


Figure 2. Facet-score distributions from the supplied results package.

Correlations alone cannot establish the intended three-factor structure. Confirmatory factor analysis, ideally including comparison with a one-factor or bifactor model, would be needed to determine whether reporting three separate facet scores adds reliable information beyond a general score.

5. Construct interpretation

The principal limitation at this stage concerns construct and external validity rather than basic internal consistency. The available evidence shows that the items behave coherently within this self-selected online sample, but it does not establish how the scores relate to established measures of attitudes toward men, observed interpersonal behaviour, or independent judgments.

The pronounced ceiling concentration and the relatively high rate of straight-line responding are particularly important when interpreting score levels. Current scores should not be treated as population norms, diagnostic classifications, or direct estimates of real-world behaviour.

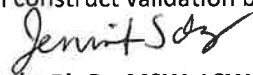
Further validation should ideally examine:

- convergent validity against established measures of negative or hostile attitudes toward men;
- discriminant validity against broader hostility, negative affect, social desirability, and related constructs;
- criterion validity using independent behavioural or informant-based indicators where feasible;
- confirmatory factor analysis comparing one-factor, three-factor, and bifactor representations;
- test-retest reliability;
- measurement invariance across language, sex, age, and other relevant groups;
- sensitivity analyses with pre-specified careless-response exclusions, particularly straight-line responding.

6. Overall conclusion

The initial psychometric evidence for the Misandry Test shows very strong internal consistency and strong item discrimination, but the score distribution has substantial upper-end compression, approximately one in nine valid sittings is a straight-line response pattern, and the three facets are very highly intercorrelated.

These features limit conclusions from the current package alone. The evidence supports continued development, with priority given to response-quality sensitivity analyses, factor-structure testing, and external construct validation before normative or high-stakes interpretations are considered.

Signature: 
 Jennifer Schulz, Ph.D., MSW, LSW
 25 August 2026