Psychometric Review of the ADHD Spectrum Test

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

21 August 2026

I conducted a preliminary psychometric review of the IDRlabs ADHD Spectrum Test using three datasets comprising 35 items and 39,934 valid test sessions. The purpose was to evaluate data quality, score distribution, internal consistency, item behavior, and the extent to which the seven proposed facets function as related components of the broader measure.

1. Data quality

The dataset is unusually large for an initial psychometric evaluation. The original export contained 40,067 records. After duplicate sittings were identified and removed, 39,934 valid test sessions remained.

The supplied data include individual item responses, seven facet scores, overall scores, test language, and limited geographic and technical information where available. No personally identifying names or contact information are included.

A small number of respondents, 263 or approximately 0.66% of the sample, selected the same response option for every item. These cases were retained in the supplied dataset rather than removed retrospectively. This is a reasonable transparent approach, although future validation work may wish to specify an a priori careless-response criterion.

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

2. Score distribution

The overall score distribution is statistically well behaved:

- Mean = 63.15/100
- Median = 64
- SD = 17.75
- Skewness = −0.42
- Excess kurtosis = −0.01

The distribution is therefore moderately negatively skewed but otherwise close to a conventional continuous distribution. There is no evidence of severe floor or ceiling compression across the sample as a whole.

Several isolated score spikes are attributable to straight-line response patterns. Importantly, these involve less than 1% of participants and do not materially determine the overall distribution.

The seven facet means also span a reasonably broad range, from approximately 43 for Developmental Delays to approximately 72 for Emotional Dysregulation. This indicates that the measure is not simply producing the same score profile across every proposed dimension.

3. Internal consistency and item behaviour

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

Cronbach's α = .931

This remains essentially unchanged when straight-line responders are excluded (**α = .929**), indicating that the high internal consistency is not an artifact of these response patterns.

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

- Inattention: α = .733
- Forgetfulness: α = .820
- Hyperactivity: α = .780
- Developmental Delays: α = .731
- Inflexibility: α = .755
- Emotional Dysregulation: α = .689
- Impulsivity: α = .730

For brief five-item scales, these values are generally acceptable to good. Emotional Dysregulation is slightly below the conventional .70 benchmark, but the difference is small and should not be interpreted mechanically.

Corrected item-total correlations range from **r = .38 to .60** across the 35 items. All items therefore contribute positively to the overall measure, and there is no obvious item with extremely poor discrimination in this dataset.

These findings support the internal coherence of the instrument. They do not, by themselves, establish that every facet represents a distinct psychological construct or that the overall score specifically measures ADHD.

4. Facet structure

The seven facet scores are moderately correlated with one another. Pairwise correlations range approximately from **r = .41 to .66**.

This pattern is compatible with the intended structure of related but non-identical domains: the facets share substantial variance while still retaining meaningful differentiation.

However, correlations alone are not sufficient to establish the proposed seven-factor structure. A confirmatory factor analysis or comparable latent-variable analysis would be required to determine whether the seven-facet model provides an adequate representation of the item-level data and whether an overarching general factor should also be modeled.

5. Clinical and construct interpretation

The principal limitation at this stage concerns external validity rather than the basic statistical performance of the test.

The available dataset demonstrates that the instrument has strong overall internal consistency, generally adequate facet reliability, useful item discrimination, and a well-behaved score distribution. It does not establish how scores correspond to independently diagnosed ADHD or to established ADHD instruments.

Consequently, these data cannot determine diagnostic sensitivity, specificity, positive predictive value, or clinically meaningful cutoff scores.

Further validation should ideally examine:

- convergent validity against established ADHD measures;
- criterion validity using independently diagnosed clinical groups;
- discriminant validity against related conditions and traits;
- confirmatory factor analysis of the proposed seven-facet structure;
- test-retest reliability;
- measurement invariance across language, sex, age, and other relevant groups.

6. Overall conclusion

The initial psychometric evidence for the ADHD Spectrum Test is encouraging.

The 35-item instrument demonstrates high overall internal consistency, generally acceptable-to-good reliability across its seven facets, positive item discrimination throughout the scale, and an overall score distribution suitable for further psychometric work.

The large sample also provides a strong empirical foundation for subsequent validation analyses.

At present, however, the findings support interpretation of the instrument as a self-report measure of ADHD-related traits and experiences, rather than as a diagnostic test. Establishing diagnostic or clinical interpretation will require comparison with validated instruments and independently assessed clinical samples.

Subject to those limitations, the available evidence supports continued development and formal validation of the measure.

Signature:

Jennifer Schulz, Ph.D., MSW, LSW

21 August 2026