Evidence-Based Assessment Criteria

Directions: You will be asked to provide a likert-type rating for 5 separate psychometric properties. You will be providing your final ratings on the Online Qualtrics survey (https://iupsych.qualtrics.com/SE/?SID=SV_3ecXwED4ZQrDTox). This sheet is to be used for note-taking.

Please rate these psychometric properties on a scale of **0-4**:

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

1. **Internal Consistency:** *internal consistency measures correlations between different items on the same test and whether several items that propose to measure the same general construct produce similar scores.*

Note: (Only include internal consistency values pertaining to the full scale. We are NOT interested in consistency scores for each individual subscale.) Internal consistency measures correlations between different items on the same test and whether several items that propose to measure the same general construct produce similar scores.

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

0: α values are not yet available OR are only available for subscales

1: α values of < .60

2: α values of .60 - .69

3: α values of .70 - .79

4: α values of ≥ .80

NA: Internal Consistency measures are not applicable for this measure; OR, classical test theory anchors are not appropriate, results reported using item response theory

Notes

2. Structural Validity: *refers to the degree to which the interrelationships of dimensions measured by the test correlate with the construct of interest and test scores.*

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

0: No exploratory or confirmatory analysis has yet been performed, nor have any Item Response Theory tests of (uni-) dimensionality have been conducted. OR, percent variance explained is not reported.

1: The sample consisted of less than 5 times the number of items AND no factor analysis performed.

2: The sample consisted of 5 times the number of items but is less than 100 in total AND an exploratory factor analysis explained less than 50% of the variance OR a confirmatory factor analysis revealed an RMSEA of = .08 to .05 or CFI = .90 to .95

3: The sample consisted of 5 times the number of items and is greater than or equal to 100 in total OR the sample consisted of 5-7 times the number of items but is less than 100 in total AND in either case an exploratory factor analysis explained less than 50% of the variance OR a confirmatory factor analysis revealed an RMSEA of = .05 to .03 or CFI = .95 to .97.

4: The sample consisted of 7 times the number of items and is greater than 100 in total AND an exploratory factor analysis explained greater than 50% of the variance OR a confirmatory factor analysis revealed an RMSEA of < .03 or CFI > .97.

Notes

3. Predictive Validity: *Predictive validity refers to the degree to which the instrument can predict (or correlate with) an outcome of interest measured at some time in the future.*

NOTE: It is important that the instrument under review serves as the independent variable, not the dependent variable. Additionally, the instrument being evaluated for predictive validity MUST have been administered prior to the administration of instrument it is meant to predict.

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

0: Predictive validity not yet tested or failed to be detected in evaluation.

1: Evidence of small correlation (range: 0.1 to 0.29) between instrument and scores on another test (measuring a distinct construct of interest or outcome) administered at some point in the future.

2: Evidence of medium correlation (α range: 0.3 to 0.49) between instrument and scores on another test (measuring a distinct construct of interest or outcome) administered at some point in the future.

3: Evidence of strong correlation (α range: 0.5 to 1.00) between instrument and scores on another test (measuring a distinct construct of interest or outcome) administered at some point in the future.

4: Evidence of medium-strong correlation (α: 0.3 or higher) between instrument and scores on at least two other tests (measuring a distinct construct of interest or outcome) administered at some point in the future.

Notes

4. Norms: *Norms typically refers to the aggregate responses of a standardized and representative group are established for a test, against which a subject is compared. For our purposes, norms refer to the mean and sample size of the sample tested.*

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

0: Norms are not yet available.

1: Measures of central tendency and distribution for the total score (and subscales if relevant) based only on a small (n < 30) sample are available

2: Measures of central tendency and distribution for the total score (and subscales if relevant) based on a moderate (n = 30-49) sample are available

3: Measures of central tendency and distribution for the total score (and subscales if relevant) based on a medium (n = 50-99) sample are available

4: Measures of central tendency and distribution for the total score (and subscales if relevant) based on a large (n > 100) sample are available

5. Responsiveness: *Defined as the ability of an instrument to detect clinically important changes over time.*

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

0: The instrument has either not been administered both pre- and post-implementation to evaluate sensitivity to change, OR it has been administered and it did not demonstrate responsiveness across an implementation process

1: The instrument demonstrated significant change over time based on a small (n < 50) sample.

2: The instrument demonstrated significant change over time based on a medium sample (n > 50 but < 100).

3: The instrument demonstrated significant change over time based on a large sample (n > 100).

4: The instrument demonstrated significant change over time based on at least three large (n >100) samples AND change is considered meaningful based on information provided by investigators regarding norms, for instance.

Notes

Notes

6. Instrument Length: *This has to do with the ease of administration, but to simplify things and improve chances of achieving reliability on this criterion we are only rating instruments based on the number of items.*

0	1	2	3	4
None	Minimal/Emerging	Adequate	Good	Excellent

0: The instrument is not in the public domain.

1: The instrument has greater than 100 items.

2: The instrument has greater than 50 items but fewer than 100.

3: The instrument has greater than 10 items but fewer than 50.

4: The instrument has fewer than 10 items.

Notes