Ways to examine validity

First, when the latent variable being measured is a hypothetical construct, there is no one single empirical test for validity. Instead, there are several empirical things that can be examined which when put together build a case for the validity of a scale. Please steer away from using language like "validity was proven".

Second, Validity (as well as Reliability) are properties of an observable measurement (or scale) as it relates to a particular latent variable in a particular population. Thus it doesn't make sense to declare a scale valid and reliable without indicating what it is supposed to be measuring and what population it is used in.

Note, the actual names given to different assessments of validity are not used the same by all.

Ways to examine validity

- Qualitative validity (or content validity) - focus groups where face validity of questionnaire can be examined, errors in interpretation can be observed, reasons for question non-response, expert panel to rate relevance of each item, Example "A qualitative validation of the Minnesota Living with Heart Failure Questionnaire", *Quality of Life Research*, 13, 417-426. 2004.
- Predictive validity (or contrasting groups validity) - determine the ability of a measure of a latent variable to predict differences between groups who are expected to have different values of the latent variable. Example "Construct validity of the RAND-12 and Health Utilities Index Mark 2 and 3 in type 2 diabetes", *Quality of Life Research*, 13, 435-448. 2004.

Ways to examine validity

- Convergent validity (or concurrent validity or construct validity)- evaluate how closely a measure of a latent variable relates with other independent measures of conceptually similar constructs. For the best practice, make a priori assumptions about how highly correlated the constructs should be even if done loosely e.g. weak Pearson's < 0.30, moderate 0.30-0.50, and strong > 0.50. Example "Validation of the cancer needs questionnaire (CNQ) short-form version in an ambulatory cancer setting", 13, 1225-1233, 2004.

Note, correlations between measures with imperfect reliability will be weaker than the correlation between the underlying latent variables (what might be thought of as the true correlation), see slides on correlation attenuation.

Ways to examine validity

- Discriminant validity (or concurrent validity or construct validity)- investigate whether a measure of a latent variable is only weakly related to other independent measures which it should conceptually not be similar to.
  - Example "Attitudes of Japanese students toward people with intellectual disability" *Journal of Intellectual Disability Research*, 46(5), 365-78. Show that measures of empowerment beliefs and similarity beliefs are NOT correlated with social desirability measure.
  - Example "Using equity theory to examine the difference between burnout and depress", *Anxiety, Stress, and Coping*, 13, 247-268. Tests different factor structures underlying the set of items in order to conclude two distinct factors makes more sense than one.

Ways to examine validity - Not very good ways

- Sometimes authors will fit a confirmatory factor analysis model to a set of items and find that each of the factor loadings is significant and conclude this shows the scale (created from the items) is valid (or sometimes they'll use the term "internally valid" or sometimes use the general term "construct valid"). This may be evidence for the existence of some particular factor structure underlying the items, but it is not evidence that the factors are measures of what they are intended to be measures of. So it seems to be an improper use of the term valid.
- Sometimes authors will describe convergent and discriminant validity by examining the results of correlations between factors of a confirmatory factor analysis. This is correlations between factors in the same scale and thus is not comparing to anything external, hence is not real evidence for validity.

Impact of reliability on correlations

The correlation between two observed variables X and Y which are imperfect measures of the underlying latent variables f_x and f_y will be smaller than the true correlation between f_x and f_y. Let r_{xx} and r_{yy} be the reliability of X and Y. Then

$$\text{Correlation}(X, Y) = \text{Correlation}(f_x, f_y) \times \sqrt{r_{xx} r_{yy}}$$

Examples of impact of reliability less than 1 on correlation

true correlation of .5

	.2	.4	.6	.8	1.0
.2	.10	.14	.17	.20	.22
.4	.14	.20	.24	.28	.32
.6	.17	.24	.30	.35	.39
.8	.20	.28	.35	.40	.45
1.0	.22	.32	.39	.45	.50

true correlation of .7

	.2	.4	.6	.8	1.0
.2	.14	.20	.24	.28	.31
.4	.20	.28	.34	.40	.44
.6	.24	.34	.42	.48	.54
.8	.28	.40	.48	.46	.63
1.0	.31	.44	.54	.63	.70