December 2020

By: Nika Kartvelishvili

Job Assessment Validation: Basic Overview

What are Job Assessments?

Job assessments are tests that measure a candidate's knowledge, skills, and abilities (KSAs) to determine their fit for a particular job position. They are a popular tool in modern organizations, helping employers decide who gets the job, who gets promoted, and who gets placed in specific roles within the company. Job assessments are used to predict objective (e.g. performance efficiency, level of productivity) as well as subjective (e.g. job satisfaction, work engagement) outcomes if an applicant were to be hired or placed into a specific role[1].

Because of their pivotal role, it is crucial that these tests are well validated, meaning that they actually measure what they intend to measure and predict what they are expected to predict. However, designing a validation study can be daunting. *But fear not!* In this post, we will guide you through the main steps of the validation process so that you can be well on your way to using effective, quality assessments.

A Quick Note on Validity

You might be surprised to learn that validity is not an inherent property of the test itself. Rather, it's the extent to which conclusions drawn from the test results are in line with available theory and evidence. As such, we often need to look at information outside of the test to demonstrate its validity. Also, validity is not an all-or-nothing

---

[1] https://www.gmac.com/~/media/files/gmac/reflect-training/hogan-assessments-white-paper

affair; it exists in degrees! We can never demonstrate perfect validity, but we *can* demonstrate a level of validity that is sufficient for the test to be considered useful and appropriate.

General Procedure

Step 1: Reliability

Before we can demonstrate that our test is valid, we need to determine its *reliability*. In other words, we need to show that our test produces consistent results when administered to similar subjects under similar circumstances. One way to show this is via test-retest reliability, whereby the same test is given on two different occasions. If the scores from the first test have a fairly strong correlation with scores on the re-test, we can consider our test to be reliable.

A potential pitfall of this approach is that some measures may be less stable over time, which would artificially lower reliability. For example, people might learn more about a topic or improve on a skill over time, which would raise their scores at retest relative to first administration and weaken the correlation between scores.

An alternative approach would be to calculate Cronbach's alpha, which is a measure of internal consistency. It tells us the degree to which different test items are related to one another. It involves correlating the items with one another in every pair

combination. If our Cronbach's alpha coefficient is close to 1, our test can be said to have high *inter-item reliability*.

Step 2: Content Validity

Our next step is demonstrating *content validity*. This means that the test items should reflect all the main facets of the construct being measured in representative proportions. We must make sure none of our test items are irrelevant to the construct of interest, unless we are including them on purpose to see if the test takers are paying attention. We must also be careful to make sure we are not over-representing a minor facet by including too many items that measure it, as well as not under-representing a major facet by not including enough items that tap into it. To ensure content validity, it is strongly recommended to have the items drafted by experts in the field.[2]

Step 3: Internal Structure

Next, we want to examine the *internal structure* of our test to determine how the items relate to one another. One way to do this is via the Exploratory Factor Analysis (EFA), which is a statistical procedure that reveals the number of "factors" or "clusters" that the test contains. Factors/clusters are basically groups of test items that tap into different facets of the construct of interest. Based on this data, we can confirm whether the test has the structure that we originally intended (i.e. whether it

---

[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184912/

measures one dimension of our construct, multiple unrelated dimensions, or multiple related dimensions).

Step 4: Convergent & Discriminant Validity

The subsequent step of our process involves demonstrating *convergent* and *discriminant validity*. Convergent validity means that scores on our test should correlate with scores on other tests that measure a theoretically related construct. Conversely, discriminant validity means that scores on our test should *not* correlate with scores on tests that measure unrelated constructs. Both convergent and discriminant validity are aspects of *construct validity,* which simply means that the test measures what it's intended to measure.

Moving Forward: Predictive Validity

This final step can take additional time to complete, but it is extremely important. It involves finding evidence that scores on our assessment correlate (or "predict") some future variable that they would be expected to predict. For example, a test that measures cognitive ability would be expected to predict various aspects of job performance. Demonstrating predictive validity is essential because at the end of the day, this is the main reason we are designing our assessment to begin with! After all, what would be the purpose of a job test if it didn't predict future performance one way or another?

Final Words

Hopefully we helped you get a more clear idea of how to validate your assessment, and have made the process seem a little less overwhelming. Please keep in mind, however, that validation is costly and time-consuming, so whenever possible, it is better to use a test that has already been validated than to design (and validate) your own from scratch. Finally, please remember that a lot of the assessments on the market have not been properly validated, so choose wisely!

We Can Help!

Even though we tried to keep this guide as clear as possible, we acknowledge that in reality, validation is a complicated and challenging affair. If you find yourself in need of assistance to validate your own assessment *or* select a properly-validated test on the market, don't hesitate to reach out! Our team of consultants will be more than happy to assist. We can be reached by email at performancesolutions@ucf.edu.

References

1. GMAC. (n.d.). Hogan Assessments Whitepaper: Summary Technical Information for Validity & Norm Groups. *The Science of Personality,* 1-11. Retrieved from: https://www.gmac.com/~/media/files/gmac/reflect-training/hogan-assessments-white-paper

2. Sullivan, G.M. (2011). A Primer on the Validity of Assessment Instruments. *Journal of Graduate Medical Education, 3*(2), 119-120. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184912/