

Benchmark Exploring Reliability and Validity

Alden J. May

COSC 622: Assessment and Evaluation in School Counseling

Dr. Robertson

February 4, 2024

Types of Reliability and Validity

The first type of reliability is measured by looking at the stability of the scores on a scale over time. The kind of reliability that the authors of the VMQ manual used was internal consistency. (Values and Motives Questionnaire, n.d.) A test must have a 0.6 or greater score of the stability coefficients. If they are below 0.6, then it indicates that the behaviors being tested are too volatile. When looking at internal consistency, it is most common to use a measurement called Cronbach's Alpha. This helps the person giving the test see if the assessment is measuring something consistently. (Values and Motives Questionnaire, n.d.) The VMQ explains two different ways that Validity can be measured to assess reliability. One of these methods, called the normative scale, is when an individual's score is compared to a score from a typical score that is found in a homogeneous group score. This measure and the type of data that it provides is called normative data (Values and Motives Questionnaire, n.d.) The standardization of a test is important to make sure the scores are reliable. Also, it is important that the measurements from a test are reliable to make sure that the test is psychometrically sound. (Values and Motives Questionnaire, n.d.) Validity is explained by the Values and Motives questionnaire by describing it as the ability of a test to measure what it claims to be measuring. (Values and Motives Questionnaire, n.d.) Construct Validity is when the test assesses whether or not the test actually measures the psychological meaningfulness of the test. It is important to know if the test outcomes are relevant to what the test was supposed to be measuring. (Values and Motives Questionnaire, n.d.) The other type of Validity is called criterion validity. Criterion validity is when a score is interpreted in order to see what would happen if another variable is observed. Predictive Validity, which is one of two kinds of criterion validity, is used to test whether a person will be able to do something in the future. It could test if someone would be successful in

a future class they will take. The other kind of criterion validity is concurrent Validity. This is used to test things that are happening currently. (Values and Motives Questionnaire, n.d.)

Reliability: Cronbach Alpha Coefficients

The Cronbach alpha coefficients internal consistencies were gathered by looking at males and females. Eight of the items that were measured on the VMI scale were at 0.7 or above. Conversely, 5 of the items were below 0.7. (Values and Motives Questionnaire, n.d.) Affiliative and Altruistic had a reliability of 0.74, so they had an adequate interpretation of reliability. Affection and security had a reliability of 0.79, so they had an adequate interpretation of reliability. Achievement had a reliability score of 0.53, so it had limited applicability. Finance and aesthetics had a score of 0.83, so they had a good interpretation of reliability. Moral had a score of 0.68, so it has limited applicability. Traditional and ethical had a score of 0.70, so they had adequate interpretation of reliability. Independence had a score of 0.66, so it has limited applicability. S.D. had a score of 0.64, so it has limited applicability. Infrequency had the lowest score of 0.52, so it has limited applicability. (What makes a good test, 2021)

Sample Size and Nature of the Population

Validity

The population that was used for the VMQ was 159 psychology and MBA students. They completed this test as part of an introduction to a personal assessment course. The population of the students who took the MAPP validation study was 59 psychology students who voluntarily took the test. Since they took the test voluntarily, they received feedback on the results. (Values and Motives Questionnaire, n.d.) The population that took the 16PF was 100 MBA students from a prestigious school in London. These students also took the VMI test. The population of

students who took the OPP was 59 psychology students who took the test voluntarily in exchange for feedback. (Values and Motives Questionnaire, n.d.) The populations of some of the other tests are comparable when they have similar reliability constructs. An issue with some of the tests is that they voluntarily took the test in exchange for feedback. This could impact the way the students take the test.

VMQ Norming Population

I think it could be representative of the population, but it would be a good idea to sample a larger population of students. The VMQ only used 159 psychology and MBA students. It didn't give a big explanation of cultural representation or gender differences. I do not think the test was broad enough to extend to other populations of people. (Values and Motives Questionnaire, n.d.)

Opinion of the VMQ

The results of the VMQ were a good start, but I think further evaluation needs to be done. For example, I think the population size needs to be significantly increased. Also, there should be a greater explanation of the gender and cultural background of participants. The correlation was 0 - 0.57. This is a very modest score. This means the VMI scales are independent of each other. (Values and Motives Questionnaire, n.d.)

References

Values and Motives Questionnaire. psytech. (n.d.) Retrieved February 4, 2024

<https://psytech.com/Content/TechnicalManuals/EN/VMIMan.pdf> Cosc 622 (2021)What

Makes a Good Test? Liberty University.

[https://canvas.liberty.edu/courses/577293/assignments/9435825?](https://canvas.liberty.edu/courses/577293/assignments/9435825?module_item_id=63945906)

[module_item_id=63945906](https://canvas.liberty.edu/courses/577293/assignments/9435825?module_item_id=63945906)