

Exploring Reliability and Validity

Linnea S. Puckett

Liberty University

COSC 622: Assessment and Evaluation in School Counseling

Instructor: Joy Robertson

Date: Sept 6, 2025

Exploring Reliability and Validity

In this paper, I will analyze the psychometrics provided from the Values and Motives Questionnaire, addressing: 1) the type of reliability and validity used. 2) the areas of concern/strength for the Cronbach's alpha coefficients. 3) how the sample size and the nature of the population may influence the constructs that the test attempts to measure. 4) Present an opinion of the test using the information in this report.

Types of Reliability and Validity

The Values and Motives Questionnaire (VMQ) mentions one source of Reliability and one type of Validity. The Reliability mentioned is internal consistency. On the other hand, the kind of validity used is construct validity. Internal consistency reliability “evaluates the extent to which the items on a test measure the same ability or trait.”(Sheperis, et. al. 2020, pg. 140). Construct Validity is described as assessing “whether the characteristic which a test is actually measuring is psychologically meaningful and consistent with the test's definition.” (Psytech, n.d, pg. 16).

Reliability: Cronbach's Alpha Coefficients

With internal consistency, reliability coefficients are computed in this paper by using Cronbach's Alpha Coefficient Method. This method is used to measure test reliability when “items on a test are not scored dichotomously” (meaning there is no right or wrong answer, rather a scale from strongly agree to strongly disagree). (Sheperis, et. al. 2020, pg. 141). Upon looking at the scales, I laid them out in categories from low, acceptable, and high ranges below: In the low category(below 0.70) are: Infrequency (0.52), Achievement (0.53), S.D. (0.64), Independence (0.66), and Moral(0.68). In the acceptable category (0.70-0.79) are: Ethical (0.70),

Traditional (0.70), Affiliative (0.74), Altruistic (0.74), Affection (0.79), and Security (0.79). In the high ranges (0.80-0.89) are: Finance (0.83) and Aesthetics (0.83). According to these readings, there are 5 problematic internal consistencies out of 13 categories, which means the data is about 38% unreliable.

Sample Size and Nature of the Population

a. Validity: The VMQ sample size consisted of 159 Psychology and MBA students.

Another factor in the sample size and population used is gender. Describing that there were 68 males and 87 females who took this test. To break it down, 59 Psychology undergraduates volunteered to complete the VMI and MAPP as part of a validation study in exchange for feedback on their results. While 100 MBA students attending a prestigious London Business School completed both the VMI and the 16PF version 5 as part of an introduction to personnel assessment methods. (Psytech, n.d, pg. 20 & 22).

The VMI and MAPP show strong correlations between their readings, suggesting validity between the two tests. However, it is not the case with VMI and 16PF version 5; although there is correlation, it is not as strong, suggesting the validity might be weaker but not completely void between the two tests. (Psytech, n.d, pg. 20 & 22)

b. VMQ Norming Population: The VMQ represented students at the university level, both male and female, and only mentioned one cultural aspect (the MBA students were from London). The only things asked on the test itself are age and sex; no cultural aspect, such as race, is mentioned. Because of this lack of cultural representation, the test doesn't measure how certain values and/or motives align with certain diverse cultures, leaving a lack of multicultural understanding. Because of this, it is not broad enough to use with

other populations. They also did not sample enough diverse population to capture traits in the population.

My Opinion of the VMQ

In Summary, I would like to include my opinion on the reliability and validity of the VMQ. Although the VMQ does seem to be reliable when correlated with other exams working with the same population of students, I feel there is a lack of validity due to the loss of culturally diverse population measurements. The test does not mention any cultural aspects taken into account when conducting the exam, which means it lacks an understanding of multicultural competencies to consider when giving the exam to a culturally diverse population. As this test is on values and motives, which are key aspects in cultures, I feel the lack of a culturally diverse population means this test may not be valid to use in many situations and with many populations. Therefore, in my opinion, I would not use this exam until it had been tested for validity against a more diverse population of students.

References

Psytech. (n.d.). *The technical manual 5 VALUES and MOTIVES QUESTIONNAIRE measures of personality values & interests.*

<https://psytech.com/Content/TechnicalManuals/EN/VMIMan.pdf>

Robertson, J. (2025, September 1). *What Makes A Good Test* [Review of *What Makes A Good Test*]. Doc.google.com; doc.google.com.

<https://docs.google.com/document/d/1eF-X3hOPvQrn6VR1XoP8UNY2HvSvkSZO6iY5M25NAk/edit?tab=t.0>

Sheperis, C., Drummond, R. J., & Jones, K. D. (2020). *Assessment procedures for counselors and helping professionals* (9th ed.). Pearson.