**Notes for question please no plag use references to cite**

wk 2 1. Briefly summary of the **comparison of the reliability and validity of responses on attitude scales**

Washtenaw Community College, Ann Arbor MI, Retrieved from <http://www4.wccnet.edu/departments/curriculum/assessment.php?levelone=tools>

**Strong words or moderate words: A comparison of the reliability and validity of responses on attitude scales**

A common assumption in attitude measurement is that items should be composed of strongly worded statements. The presumed benefit of strongly worded statements is that they produce more reliable and valid scores than statements with moderate or weak wording. This study tested this assumption using commonly accepted criteria for reliability and validity. Two forms of attitude scales were created—a strongly worded form and a moderately worded form—measuring two attitude objects—attitude towards animal experimentation and attitude towards going to the movies. Different formats were randomly administered to samples of graduate students. There was no superiority found for strongly worded statements over moderately worded statements. The only statistically significant difference was found between one pair of validity coefficients ( *r* = 0.69; *r* = 0.15; *Z* = 2.60, *p* ≤ 0.01) and that was in the direction opposite from expected, favoring moderately worded items over strongly worded items (total scores correlated with a general behavioral item). (PsycINFO Database Record (c) 2016 APA, all rights reserved) (Source: journal abstract)

wk 2 2. What are Effective ways to understand and organize data using descriptive statistics?

Organizing Quantitative Data

*Organizing quantitative data* [Video file]. (2005). Retrieved January 20, 2017, from <http://fod.infobase.com/PortalPlaylists.aspx?wID=18566&xtid=36200>

<http://fod.infobase.com/p_ViewVideo.aspx?xtid=36200>

Effective ways to understand and organize data using descriptive statistics. Analyzing data collected from studies of young music students, the video helps viewers sort through basic data-interpretation concepts: measures of central tendency, levels of measurement, measures of dispersion, and graphs. A wide range of organization principles are covered, including mode, median, and mean; discrete and continuous data; nominal, ordinal, interval, and ratio data; standard deviation; and normal distribution. Animation and graphics clarify and reinforce each concept. The video concludes with a quick quiz to assess understanding and focus on key areas. A viewable/printable instructor’s guide is available online. WE DISCUSSED HOW TO DESIGN AN EXPERIMENT AND CONTROL VARIABLES IN OUR FIRST VIDEO. AND NOW WE'RE GOING TO LOOK AT WHAT TO DO WITH ALL THE DATA THAT HAS BEEN COLLECTED. AN EXPERIMENT IS ONE OF THE MOST POWERFUL WAYS TO SHOW THE CAUSE OF AN EVENT AND ITS EFFECT ON OTHER THINGS. BUT REMEMBER THAT AN INVESTIGATION CAN ONLY BE A SCIENTIFIC EXPERIMENT IF IT HAS AN INDEPENDENT VARIABLE WHICH IS MANIPULATED WHILE ALL OTHER VARIABLES REMAIN CONSTANT. WE CAN THEN MEASURE THE EFFECT OF THE INDEPENDENT VARIABLE ON THE DEPENDENT VARIABLE. THIS IS CALLED "CAUSE AND EFFECT" BECAUSE WE CAN PRESUME THAT THE CHANGE OF ONE VARIABLE, THE IV, CAUSES THE EFFECT WE SEE ON THE DEPENDENT VARIABLE, KNOWN AS THE DV. SO, WE'VE CARRIED OUT OUR EXPERIMENT, AND WE ARE LEFT WITH ALL THE SCORES OR MEASUREMENTS. DON'T PANIC. WE'RE GOING TO SORT THEM OUT IN THIS PROGRAM USING DESCRIPTIVE STATISTICS. ANALYZING OUR DATA CAN BE AS EASY AS ONE, TWO, THREE. AND WE START BY LOOKING AT THINGS IN THREES. WE BEGIN WITH GENERALISABILITY, RELIABILITY, AND VALIDITY. THEN MODE, MEDIAN, AND MEAN AVERAGE. FOLLOWED BY NOMINAL, ORDINAL, AND INTERVAL DATA. THE NAMES MAY BE NEW TO YOU BUT YOU WILL ALREADY HAVE COME ACROSS SOME OF THEM UNDER ANOTHER NAME. RIGHT, HERE WE GO. WE NEED TO BE SURE THAT OUR RESULTS CAN BE APPLIED TO ALL PEOPLE OR EVENTS, NOT JUST OUR OWN PARTICIPANTS. IN SOCIAL SCIENCE, WE ARE INTERESTED IN FINDING OUT ABOUT HUMAN BEINGS IN GENERAL. IF WE HAVE CHOSEN OUR DESIGN AND SAMPLE CAREFULLY, WE MAY THEN BE ABLE TO GENERALIZE OUR FINDINGS TO ALL HUMANS, AND THEN THEY WILL BE REALLY USEFUL. BUT THIS IS ONLY POSSIBLE IF WE HAVE A RELIABLE AND VALID STUDY. RELIABLE MEANS THAT WE NEED TO BE FAIRLY SURE THAT ANYONE WHO REPLICATES OUR STUDY EXACTLY WILL GET THE SAME RESULTS. LIKE, SAY, MUSICAL SCALES. IF PLAYED ACCURATELY, THEY WILL GIVE THE SAME RESULT EVERY TIME WHOEVER PLAYS THEM AND WHEREVER THEY ARE.

wk 3 1 What are the differences in test construction between the Minnesota Multiphasic Personality Inventory (MMPI) and the NEO Personality Inventory?

wk 3 2. What happens when the psychometrics of testing development meets organization realities?

**NOTES:**

WHEN THE PSYCHOMETRICS OF TEST DEVELOPMENT MEETS ORGANIZATIONAL REALITIES: A CONCEPTUAL FRAMEWORK FOR ORGANIZATIONAL CHANGE, EXAMPLES, AND RECOMMENDATIONS.

Our standards for the construction of psychological tests give scant attention to the organizational context in which the tests are to be used. This paper describes 10 psychometric issues associated with the development of an electrical job knowledge test. The test was designed to replace seniority as a means of making promotional decisions within an organization. The 10 test-related issues are presented as a means to understand the underlying process of organizational change associated with the implementation of the new test. It is suggested that a closer link between the science and practice of our profession can be attained by achieving a greater understanding of issues associated with the practical implementation of theory-based interventions.

Psychological assessment is guided by published standards of professional principles for test construction (APA, AERA, NCME, 1999; SIOP, 2003). These published standards are revised periodically to reflect the latest professional developments in the field of psychological assessment. The revisions are often of a technical nature, as reflected in conceptions of validity (e.g., Jonson & Plake, 1998), measurement (e.g., Murphy & DeShon, 2000), and/or test score interpretation (e.g., Stricker, 2000). Research findings from such areas can become a basis for consideration of revision in our standards. Our profession is also subject to consideration of issues from legal sectors reflecting differing societal values of fairness in testing (e.g., Sackett & Wilk, 1994; Zedeck, Outtz, Cascio, & Goldstein, 1995). In fact, much of the debate within our profession about psychological testing pertains to the relationship between predictive accuracy and fairness in making assessments. However, there is another dimension to the conduct of psychological assessment that is rarely ad dressed in journal articles and books, and is simply not addressed at all in published standards on psychological testing.

Guion (1998a,b) has described how all personnel decisions are imbedded within an organizational context and how practice can depart from theory in making personnel decisions. Theorists explicate the way it should be done, but practitioners are entrusted with the responsibility for operationalizing the theory in a way that is acceptable to the particular organization. As Guion has noted, there can be considerable gaps between the stated theory and the actual practice. The gaps are caused (in part) by organizational pressures to achieve selected outcomes deemed beneficial to certain parties within the organization. In the best of cases, these organizational pressures complement the science and theory of personnel decision making. At other times, organizational pressures can be independent of scientific and theoretical principles. In such cases, our challenge as professionals is to fashion a system of personnel decision making that is scientifically and legally defensible within a larger organizational system that offers limited support for what we are trying to achieve. In the worst of cases, the organizational pressures are so severe as to confute scientific principles of psychological assessment. In such cases, we aren't even in a position to "finesse" the science through the quicksand of organizational realities. We may attempt to alleviate some of the inhibiting pressures by making predictions of possible dire consequences (often of a legal and/or financial nature) as a means of influencing the decision making process. However, it has been my experience that in a tussle between the science and theory of personnel decision making versus the pressures of organizational exigencies, the latter usually prevails. Time and money are usually far more salient organizational criteria in decision making than reliability and validity. Practitioners seemingly are engaged in the continuous art of compromise, navigating what they should be doing through the organizational seas of what they are allowed to do. The challenge is to guide organizations, which are typically staffed by influential decision makers with little or no training in psychological assessment, through the process of meeting two critical criteria. The first is acceptability of the finished product for use within the organization. Little is gained from a practical perspective if, for whatever reasons, the organization fails to implement the assessment system that was developed (i.e., relegates it to the file drawer), or finds maladaptive ways to contort the intended use of the test (i.e., expands the use of a test to make personnel decisions unrelated to the expressed purpose of the test). The second criterion is the psychological assessment must meet technical standards for quality. Our profession has developed a body of knowledge specifying scientifically based recommendations regarding test construction, use, and interpretation. There are also serious legal considerations associated with the misuse of tests that can adversely affect the organization. In short, there is a delicate balancing act that must be achieved between organizational reality and psychometric precepts. The recently published fourth edition of the Principles for the Validation and Use of Personnel Selection Procedures has recognized the importance of the organizational context in which assessment occurs. "Researchers face the challenge of ensuring high quality selection procedures in the context of the organization's history and current environment regarding employment-related strategies and practices as well as the cultural setting in which it operates." (SIOP, 2003, p. 36)

I believe that the implementation of a testing program in an organization can be viewed as a case of organizational change. Thus, the organizational change literature can be a resource to facilitate the practice of psychological assessment. As opposed to thinking of scientists and practitioners as two separate camps governed by two different sets of values (which can happen), techniques of organizational change can help to merge these two roles. That is, organizational change strategies can serve to implement in practice concepts and principles developed through scientific research. I have come to conclude that the science of I-O psychology has a wealth of knowledge to offer the workworld. In turn, the workworld has a great need for help in issues that are germane to I-O psychology. The major obstacle is the "translation" of the science into practice. As such, I have come to regard organizational change strategies as not only desirable, but imperative, for the merger of science and practice to occur. Whenever attempts are made to change an organization, resistance is likely to manifest itself in one or more forms. The skilled scientist-practitioner must be adept in identifying the source of the resistance, but also must possess an array of strategies to help mitigate the resistance. Among the strategies I have found useful are education, shared responsibility, serving as a negotiator/facilitator, and the overt manifestation of respect and recognition for the knowledge possessed by all parties.

This paper describes a real-life organizational context in which the principles of test construction were directed toward the development of a job knowledge test. Ten issues pertaining to test construction and use serve as the basis for a broader conceptual understanding of the organizational change process. However, the 10 issues also provide a context for understanding how the scientific principles of test construction can be challenged in developing a product that meets with organizational appeal. That is, the paper frames the process of test development within an organization into a larger theoretical context, that of organizational change. Specific recommendations are proposed to enhance the likelihood of attaining the twin criteria of organizational acceptance and psychometric integrity.

Background

The organization in question is a national leader (by size) within its industry and is in the manufacturing sector of the economy. The organization uses state-of-the-art manufacturing processes to produce a fairly narrow range of products. It is positioned early in the supply-chain process converting raw stock into mid-range products that are in turn converted by its customers into durable goods, that are in turn sold to retail vendors. The manufacturing process is heavily machine intensive. The maintenance of the machines is crucial for the operating efficiency of the company. Virtually all of the machines in the company use computer-aided manufacturing processes, requiring a high level of electrical knowledge among the employees responsible for maintaining the operating efficiency of the machines. Although the job titles of the employees responsible for machine maintenance differed slightly, the most accurate and representative job title would be electrical maintenance mechanic (D.O.T. Code 829.261-018; U.S.D.O.L., 1991).

The company operated eight production facilities within a 75-mile radius. Most of the older (30 years old) facilities were small (a few hundred employees) and utilized the least technically sophisticated manufacturing procedures. The newer production facilities employed approximately 1,000 employees at each location and utilized the most high-tech manufacturing procedures. The company was nonunion. Employees were hired into entry-level machine operator jobs and advanced as machine operators through a seniority system. The modal machine operator had a high school education but 20% did not. For the most part, there were relatively few problems using the seniority system of advancement within machine operator jobs. The major problem was the identification of machine operators who would become electrical maintenance mechanics. For many years the company had relied on the seniority system to select electrical maintenance mechanics. The selected maintenance mechanics were the most senior machine operators, who had the most "hands-on" work experience with the machines. However, over the past 5 years, the manufacturing process had greatly accelerated in technical sophistication. The electrical maintenance mechanics began to need increasingly higher levels of knowledge. The company had grown displeased with the seniority system of selecting electrical maintenance mechanics. The company expressed an interest in using an electrical knowledge test to select the electrical maintenance mechanics. The conversion from a seniority system of advancement to a knowledge test sys tem of advancement was alien to the company's culture. There were the stereotypical fears of potential litigation associated with the use of tests to make personnel decisions. However, the company felt it had little choice in the matter, as it was at a crossroads in its evolution. The seniority system of selection was not deemed successful, and the use of supervisor judgments of suitability to become an electrical maintenance mechanic was judged likely to produce charges of favoritism and bias in the selection process. Reichers, Wanous, and Austin (1997) noted the enthusiasm for change varies from person to person within an organization. I did not detect what Reichers et al. refer to as cynicism about the need for change, but rather skepticism as to whether a paper-and-pencil assessment used to make promotion decisions was operationally superior to seniority. Davis (1998) has stated the history of an organization with assessment can play a major role in its acceptance of a proposed assessment. This organization used the interview and background checks as a means of selecting machine operators. Although they were not predisposed against testing, they also had no previous positive experiences with it either. Thus, with considerable trepidation but with a conviction to advance, the company hired me to develop and implement a knowledge test to select electrical maintenance mechanics.

My primary contact in the company was the director of human resources, as the implementation of an employment test was deemed to be a "HR project." The culture in an organization has a major impact on what is seen to be "owned" by HR or the line, and whether HR is regarded as a business partner or a paper-pushing police force. In my opinion, one topic that is rarely discussed in our literature is the differential status and power of departments or units within an organization. This differential power not only affects what gets done in organizations, but how it gets done. The HR department did not have much political power within the company, as its primary function was record keeping. The most powerful department in the company was manufacturing, possessing more organizational clout than sales, marketing, or finance. Although the implementation of an electrical knowledge test was a "HR project," there was never any doubt that whatever I could accomplish would have to have the full support and blessing of manufacturing. The HR director convened a 4-person committee of top manufacturing personnel who would serve as my initial client within the organization. One of my first orders of business was to inform them that a number of commercially available paper-and-pencil electrical knowledge tests existed. I distributed copies of these tests to the committee for their reactions. It was quickly concluded by the committee that "the type of electrical knowledge needed to maintain the machines in this company is not the type of electrical knowledge measured by these tests." Simply put, they didn't like the questions, thus, they did not approve of the tests. Although my personal knowledge of electrical concepts is not sufficient that I could agree (or disagree) with the existence of "types" of electrical knowledge, I realized what I had offered the client was not acceptable. I then informed them that if the "buy" option was not available, the only other option would be to "make" our own test. I explained that I would serve as the project manager of the test construction process, but the company itself would have to be responsible for creating the very questions to be included on the test. In short, this company would also have to be in the test construction business as well as the manufacturing business. The members of the 4-person committee accepted the assignment and our joint responsibility for the overall project.

The next section of this paper describes 10 issues associated with the test construction process. They basically represent 10 phases of the design process, each grounded in established psychometric principles of test construction. Each of the 10 psychometric principles is presented, but they ultimately serve to illustrate critical issues associated with changing the organization and its culture. Each issues is: (a)

wk 3 3. Face validity is not technically a type of validity. However, it does have some merit in test construction. Why is this an important component of the test construction process?

Notes

Face validity relates to how it appears to others with regards to low or high. Cohen (2013) explains that test such as the Myers-Briggs Personality assessment has high face validity because it ask specific questions related to how the person acts in specific situations. For example, if a person is normally shy at parties, has high anxiety in public speaking, and typically does not like being in large groups. They would be considered an introvert, while a person who scores low on the same questions but high on question where they feel more comfortable in large groups, do well in public speaking, and can intermingle with others easily at parties would be considered Extroverted in personality. These types of questions are considered high face validity because they measure what they say they measure. However, if a person takes a Rorschach Inkblot test and they perceive specific pictures to mean something different than other populations, the face validity is considered to have low face validity because it does not always produce the same results. Face validity it used only by test takers and not test constructors. However, it is suggested that those who construct test measures what it is supposed to measure. The example given earlier about the Rorschach Inkblot, it was supposed to test personality by asking the participator what they saw, however, this was based on the assessors interpretation of the response from the participant and not the actual personality of the person participating in the assessment. Construction of test should allow both the constructor and participant to see the face validity as it is, that it measures what it purports to measure.

Cohen, R.J, Swerdlik, M.E, and Sturman, D.E. (2013). Psychological test and assessments: Introduction to test and measurements (8th ed). New York City, NY. McGraw Hill Inc.