1 CONSEQUENCES OF TEST INTERPRETATION AND USE: THE FUSION OF VALIDITY AND VALUES IN PSYCHOLOGICAL ASSESSMENT

Samuel Messick

This paper addresses the role of social values in educational and psychological measurement, with special attention to the consequences of testing as validity evidence, which is an inherently value-dependent enterprise. The primary measurement standards that must be met to legitimize a proposed test use are those of reliability, validity, and fairness, which are also value-laden concepts. Evidence of reliability signifies that something is being measured—the major concern is score consistency or stability. Evidence of validity circumscribes the nature of that something—the major concern is score meaning. Evidence of fairness indicates that score meaning does not differ consequentially across individuals, groups, or settings—the major concern is comparability.

The appropriate level of reliability depends on the meaning of the construct being measured because some constructs are conceived theoretically to be more consistent or stable than others. Hence, evidence of reliability consistent with the construct's meaning is simultaneously also evidence of construct validity. Within these limits, the measurement intent is to achieve sufficient score consistency or stability to warrant the use of the scores in decision making. Another way of putting it is that the uncertainty involved in determining score levels should be inconsequential for the proposed test use.

Validity is mainly concerned with the meaning and consequences of measurement. Accordingly, validation studies aim to accrue convergent evidence supportive of score meaning and its action implications as well as discriminant evidence discounting plausible rival interpretations. Validity is a unitary concept,


OVERVIEW

This chapter features Jackson’s and my (independent) work on the personality traits, values, and motives of productive scientists. It shows that the distribution of scientific achievement is J-shaped, implying that the underlying causes combine multiplicatively rather than additively. It reviews my own research using peer- and self-ratings on 29 traits, including those from Jackson’s Personality Research Form, in which impactful scientists were found to be ambitious, enduring, seeking definiteness, dominant, showing leadership, intelligent, aggressive, independent, not meek, and non-supportive. It reviews Jackson’s research finding that strong theoretical interests, curiosity, an autonomous personality, and achievement motivation are salient in successful scientists. It also reviews his suggestions that administrators need to take the unique attributes of scientists into account to create the conditions maximally conducive to productivity. These include freedom from bureaucratic interference, good financial rewards, and incentives for career advancement. It concludes by touching on the important role Doug Jackson played defending academic freedom in Canada.

A PERSONAL NOTE

Back in 1977 when I began my career at the University of Western Ontario, Doug Jackson (to whom I largely owed my appointment) was one of the most prolific and cited psychologists in North America. Year after year, he routinely published about 10 articles, book chapters, research bulletins, and psychology test manuals. I was keen to find out how he generated the quality and quantity of his output as I had developed a research program on productivity and impact (Rushton & Endler, 1977). I was also delighted to learn Doug shared my interest in the causes of high productivity.
Regular lunch partners, Doug and I have frequently discussed progress in science. By 1985 this mutual interest led us to co-host a conference on scientific excellence at the University of Western Ontario. Supported by grants from the Academic Development Fund at the university and the Social Sciences and Humanities Research Council of Canada, the conference resulted in an edited book, *Scientific Excellence: Origins and Assessment* (Jackson & Rushton, 1987).

**STUDIES OF SCIENTIFIC ACHIEVEMENT**

One of the most salient features of scientific achievement is its unequal distribution. Whereas personality and intelligence are normally distributed, scientific achievement is not. A very few scientists are responsible for the great majority of creative works. Across scientific disciplines, the most productive 10% of scientists typically account for 50% of the publications (Dennis, 1955; Shockley, 1957). Studies of academic psychologists show a similar distribution (Endler, Rushton & Roediger, 1978; Rushton & Endler, 1977; Rushton, 1989). Consider, for example, the citation and publication counts reported in Table 1. These cumulative percentage frequencies are based on 4,070 faculty members studied by Endler et al. (1978) in an analysis of the top 100 departments of psychology in the United States, Canada, and the United Kingdom. Over half (52%) of the sample did not publish an article in 1975 in any of the journals reviewed by the *Social Sciences Citation Index*. The picture is similar for citations, the great majority of academic psychologists having relatively few. For example, only about 25% of psychologists had more than 15 citations in 1975 and only 1% had more than 100 citations.

"Ageist," "sexist," and "elitist" factors contribute to the positive skew shown in these distributions. Productivity increases with age up to around 40-45 years, then gradually diminishes. Women are not only under-represented in science but, on a per capita basis, produce less than their male counterparts (Cole, 1981). Individuals who received doctorates from more prestigious institutions and/or who obtained their first academic positions at high prestige universities are more productive than those who graduated from or were appointed to less esteemed institutions (Gaston, 1978).

Why? As Walberg, Strykowski, Rovai, and Hung (1984) explain, the normal distribution does not apply to exceptional performance. Instead, J-shaped distributions such as those shown in Table 1, are characteristic. J-shaped distributions – monotonically decreasing at a decelerating rate – typically occur when the underlying causes combine multiplicatively rather than additively. (Additive causes typically produce normal distributions). Walberg et al. (1984) show that learning is a multiplicative, diminishing-returns function of student ability, time, motivation, and amount and quality of instruction. (Those instances