

OPINION PIECE:

Challenging the notion of face validity

by Tim Newfields

Dennis Roberts offered some arguments in favor of face validity in the October 2000 issue of this publication. In the spirit of a balanced discussion, I will present some counter-arguments and suggest why face validity may be an inappropriate concept.

First, face validity is a contradictory term. Matters involving surface appearance involve cosmetic value rather than validity per se. Validity should involve deeper factors such as logical veracity, consistency, and congruence. Face validity is concerned with popularity or common acceptance rather than scientific truth.

Second, if we regard testing as a rigorous discipline, face validity has little place because it is both atheoretical and imprecise. Face validity basically amounts to what Buck (2001) refers to as "faith validity" - the belief that a test is okay without empirical evidence. Empirical evidence is a *sin qua non* of testing. Since face validity is based primarily on the judgments of novices, this concept has value in terms of business and marketing, but it is not a yardstick test developers should focus on.

It could be argued that face validity encourages a cosmetic approach to test construction which emphasizes surface appearance rather than the operationalization of testing concepts. Perhaps one of the reasons why there are so many poorly constructed language tests is because there is such an obsession with face validity without adequate consideration of deeper forms of validity and reliability. Language test developers should be concerned with criteria such as task validity, content validity, construct validity, and reliability rather than the largely cosmetic notion of face validity.

Roberts (2001) claims that face validity is "an essential part" of the assessment process. However, there are many voices of dissent. Hajipournezhad (2000) mentions how this term is widely detested among testing scholars and quotes Mosier (1947, p. 194):

The concept is the more dangerous because it is glib and comforting to those whose lack of time, resources, or competence prevent them from demonstrating validity (or invalidity) by any other method. . . . This notion is also gratifying to the ego of the unwary test constructor.

Trochim (2002) cautions face validity is "the weakest way to try to demonstrate construct validity." Lacity and Jansen (1994) describe face validity in terms of persuasive appeal and note that test items can seem persuasive even if they lack internal validity.

In conclusion, face validity is essentially a cosmetic affair that should concern test marketers more than test developers.

References

Buck, G. (11 Nov. 2001,19:00). "Validities." Message posted on L-TESL Online Forum. Retrieved April 14, 2002 from <http://f05n16.cac.psu.edu/archives/ltest-l.html>.

Hajipournezhad, G. (2000). An Approach to the Validation of Judgments in Language Testing. In T. Newfields, S. Yamashita, A. Howard, & C. Rinnert. (Eds). *Proceedings of the 2003 JALT Pan-SIG Conference* held at Tokyo Keizai University on May 10 - 11, 2003. (p. 80 - 84). Retrieved April 14, 2002 from <http://jalt.org/pansig/2003/HTML/HajiPourNezhad.htm>.

Lacity, M., & Jansen, M. A. (1994). Understanding qualitative data: A framework of text analysis methods. *Journal of Management Information Systems*, 11, 137-166.

Roberts, D. M. (Oct. 2000). Face Validity: Is There a Place for This in Measurement? *SHIKEN: The JALT Testing & Evaluation SIG Newsletter*, 4 (2), 5. Retrieved April 14, 2002 http://jalt.org/test/rob_1.htm.

Trochim, W. (2002). "Measurement Validity Types." [Online] <http://trochim.omni.cornell.edu/kb/>. [Expired Link].

HTML:http://jalt.org/test/new_2.htm / **PDF:**<http://jalt.org/test/PDF/Newfields2.pdf>