Industrial and Organizational Psychology, **2** (2009), 514–516.

Babies, Bathwater, and Validity: Content Validity Is Useful in the Validation Process

JAMES A. TAN
St Cloud State University

Carrier, Dalessio, and Brown (1990) investigated the match between content and criterion-related validation strategies and found little support for their hypothesis. Carrier et al. noted that one may make a mistake in thinking that a content validation strategy may replace a criterion-related validation strategy. However, Carrier's study is limited in that they assumed there is a linear relation between content- and criterion-related validity, but this has yet to be shown (Moscoso & Salgado, 2001).

Although Ree, Earles, and Teachout (1994) argued that the use of specific ability tests may not yield better predictors of performance or even add incremental validity to general cognitive ability tests, there is some research that this may not be the case in certain instances. For example, Barrett, Polomsky, and McDaniel (1999) conducted a meta-analysis of the effectiveness of cognitive ability (general) and mechanical comprehension (specific) tests in predicting firefighter performance. Barrett et al. found that mechanical comprehension ($r = .26$; $\rho = .54$) was a better predictor of firefighter job performance than cognitive ability tests ($r = .20$; $\rho = .42$). In addition, composite cognitive ability and mechanical comprehension tests showed the highest validity coefficients ($r = .28$; $\rho = .56$).

Dye, Reck, and McDaniel (1993) investigated whether the validity of job knowledge tests was higher when the test and job contents matched. Dye et al. found that validity coefficients were highest for tests that had high job–test content similarity ($r = .31$; $\rho = .62$) compared with those with moderate ($r = .17$; $\rho = .35$) or low ($r = .16$; $\rho = .35$) job–test content similarity. In addition, Dye et al. investigated whether job complexity moderated this relation and found similar results. For jobs and tests that had high similarity, job knowledge tests were still better predictors of high-complexity jobs ($r = .33$; $\rho = .66$) than low complexity jobs ($r = .27$; $\rho = .55$). Dye et al. found a similar pattern when investigating situations where job and test content were moderately similar (high complexity $r = .26$; $\rho = .54$; low complexity $r = .14$; $\rho = .30$).

We can find similar results for other predictors. McDaniel, Whetzel, Schmidt, and Maurer (1994) showed that validities of interviews were highest for situational ($r = .27$; $\rho = .50$ corrected for range restriction) and job-related ($r = .21$; $\rho = .39$ corrected for range restriction) interviews compared with psychological interviews ($r = .15$; $\rho = .29$ corrected for range restriction), the purpose of which was to assess personal traits such as dependability. Other

Correspondence concerning this article should be addressed to James A. Tan.
E-mail: jatan@stcloudstate.edu
Address: Department of Management, G. R. Herberger College of Business, St Cloud State University, St Cloud, MN 56301–4498

https://doi.org/10.1111/j.1754-9434.2009.01185.x Published online by Cambridge University Press

researchers have also shown the effectiveness of matching job content with test content in improving interview validity (Latham & Sue-Chan, 1999; Taylor & Small, 2002). Finally, McDaniel, Morgeson, Finnegan, Campion, and Braverman (2001) showed that situational judgment tests based on job analysis had higher validity coefficients ($r = .29$; $\rho = .38$) than those that were not based on job analysis ($r = .22$; $\rho = .29$).

Additional benefits of content validation. In addition to the three benefits Murphy (2009) noted, I would like to add two more: (a) feasibility and (b) test security. Conducting criterion-related validation studies are often not feasible because of the time and expense associated with it. Other problems related to the use of criterion-related validation include: the missing persons problem, range restriction, differences between applicants and present employees, and the effects of job experience and training (Barrett, Phillips, & Alexander, 1981; Hough, 1998). It is often much quicker and less expensive to use a content validation strategy when developing a test.

Test security is a major concern in high-stakes testing. There have been allegations of cheating in public safety testing in recent years (Collins, 2007; Slack, 2008). To use a criterion-related strategy to validate a test would entail either administering the test to applicants or incumbents, either of which increases the risk that actual test questions would leak to future job applicants. Using content validation to validate the test in this case minimizes the probability that test items would be leaked to job applicants (Bellenger & Dean, 2008).

Throwing the baby out with the bath water? Both content and criterion-related validation evidence rely on expert judgment. Experts decide on the appropriate criteria in criterion-related validation studies. Furthermore, the criterion used is often judgmental in nature (i.e., supervisor ratings). As Sproule (2009) noted, ''If expert judgments are acceptable as part of the criterion-related validation process, why are they not acceptable to establish content validity evidence? All validation evidence relies on judgments. The best that we can do is to collect as much evidence as feasible using the scientific method, and analyze and present the evidence in a fair and objective manner.'' (p. 459).

Kane (2006) noted that it is difficult to obtain an adequate criterion, and once a criterion is selected, the question becomes how does one validate the criterion? Finding another criterion with which to validate the original criterion may not be feasible or practical, and therefore the original criterion has to be validated in another manner. The most efficient manner to validate the criterion is to establish a rational link between the procedures used to generate the criterion scores and the proposed interpretation or use of scores.

References

Barrett, G. V., Phillips, J. S., & Alexander, R. A. (1981). Concurrent and predictive validity designs: A critical reanalysis. *Journal of Applied Psychology*, *66*, 1–6.

Barrett, G. V., Polomsky, M. D., & McDaniel, M. A. (1999). Selection tests for firefighters: A comprehensive review and meta-analysis. *Journal of Business and Psychology*, *13*, 507–513.

Bellenger, B. L., & Dean, C. R. (2008, June). *Cheating on promotional public safety examinations: Strategies for responding to a widespread test security compromise*. Paper presented at the 32nd Annual IPMAAC Conference on Personnel Assessment, Oakland, CA.

Carrier, M. R., Dalessio, A. T., & Brown, S. H. (1990). Correspondence between estimates of content and criterion-related validity values. *Personnel Psychology*, *43*, 85–100.

Collins, D. (2007). Mayor: ''Strong evidence'' of fire department exam cheating. Retrieved November 30, 2007, from www.wbaltv.com/news/14735919/detail.html

Dye, D. A., Reck, M., & McDaniel, M. A. (1993). The validity of job knowledge measures. *International Journal of Selection and Assessment*, *1*, 153–157.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, *11*, 209–244.

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: Praeger.

Latham, G. P., & Sue-Chan, C. (1999). A meta-analysis of the situational interview: An enumerative review of reasons for its validity. *Canadian Psychology*, *40*, 56–67.

https://doi.org/10.1111/j.1754-9434.2009.01185.x Published online by Cambridge University Press

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, *86*, 730–740.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, *79*, 599–616.

Moscoso, S., & Salgado, J. F. (2001). Psychometric properties of a structured behavioral interview to hire private security personnel. *Journal of Business and Psychology*, *16*, 51–59.

Murphy, K. R. (2009). Content validation is useful for many things, but validity isn't one of them. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *2*, 453–464.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology*, *79*, 518–524.

Slack, D. (2008). Boston fire test will be voided. Retrieved February 29, 2008, from www.philly.com/inquirer/front_page/20090201_N_J_probes_allegations_on_fire_test.html

Sproule, C. F. (2009). *Rationale and research evidence supporting the use of content validation in personnel assessment*. Washington, DC: International Personnel Assessment Council.

Taylor, P. J., & Small, B. (2002). Asking applicants what they *would do* versus what they *did do*: A meta-analytic comparison of situational and past behavior employment interview questions. *Journal of Occupational and Organizational Psychology*, *75*, 277–294.

https://doi.org/10.1111/j.1754-9434.2009.01185.x Published online by Cambridge University Press