

Straight Talk about 360-degree Feedback Validation

by Dennis E. Coates, Ph.D.

HR professionals researching 360-degree feedback often ask about the validation research that supports a given survey. The best answer to this question is not the one they expect.

The answer has three parts:

1. The process for validating a behavioral feedback survey is not the same as the process for validating a psychological instrument.

Since the technology of behavioral feedback surveys is an outgrowth of the tradition of psychological testing, it's natural to confuse the two kinds of assessments. However, even though they're both often referred to as "assessments," they're radically different kinds of tools, and the procedures for validating them are different.

Validity means one thing when evaluating a personality assessment or a psychological test, but it means something entirely different when evaluating a tool that gives direct feedback about behavior—a completely different kind of assessment.

Validity and reliability information has been important to determining the usefulness of tests that measure intelligence, values or psychological characteristics. These are important aspects of people that cannot be observed or measured directly, so questions are asked in order to draw inferences about these aspects. Therefore, it's important to determine whether the assessment really measures what it says it measures. In other words, are the constructs valid? Are the constructs that are calculated based on the answers highly correlated to phenomena in the real world? Do they make valid inferences based on the questions in the assessment? Will we get the same results, measurement after measurement?

Many of the early multi-source feedback instruments evolved from these traditions. Some even focused on traits, values, characteristics and other aspects that can't be measured through direct questioning or direct observation. Their purpose was to ask carefully researched questions from which one could draw inferences about these important qualities. The constructs created from the answers were the main purpose and product of these assessments. Therefore, research was needed to verify (validate) that the inferences were not just theoretical notions, but actually coincided with what was happening in the real world.

20/20 Insight GOLD is a wholly different technology. Quite simply, it's a computer-assisted method to gather and give feedback about specific observable behaviors. Unlike psychological instruments, the survey items (observable behaviors) themselves are the primary focus of the feedback—not constructs inferred from the item responses. In 360-degree feedback, many people are asked to report what they observed, and their aggregate response is reported. The responses themselves, not dimensions inferred from the responses, are the purpose and product of the assessment.

In fact, the 20/20 Insight surveys don't create or report constructs. The items are clustered into categories only to make it easier to relate and analyze the feedback. Therefore, there's no need to do research to verify the validity of inferences or

constructs, because no inferences or constructs are produced. Unlike personality or trait assessments, a "universal" validation that applies to all organizations is not appropriate.

This is not to say that validity isn't important. With 20/20 Insight, another kind of validity is essential. It's important to establish whether the feedback surveys used by the organization actually address the key workplace behaviors of that site. When surveys from the 20/20 Insight Survey Library are used, we encourage organizations to do internal competency research, then select, deselect, add or modify the items in the standard survey categories to align with local practices. These customized surveys need to be verified locally, because every organization has a different business, culture and priorities. Validation can be accomplished by using importance and frequency surveys, expert panel reviews and pilot assessment projects. To be valid, such a survey will accurately describe the key aspects of workplace behavior.

Reliability also doesn't have the same meaning with this kind of feedback tool. Reliability refers to the consistency of measurements. With a behavior-based feedback survey, you wouldn't expect or want to get the same results with each subsequent measurement, because the assessment focuses on observable skills, competencies and abilities, which are expected to improve over time.

Once again, with behavioral feedback surveys, the important validity issues are:

- (1) Did administrators ask the right questions?
- (2) Did people get the feedback they needed to begin self-improvement toward greater achievement?

2. "Executive Leadership" and other standard surveys contained in the 20/20 Insight Survey Library were researched and developed using standard behavioral feedback (competency) research and development practices.

The 20/20 Insight Survey Library features no psychological instruments. All the individual multi-source feedback surveys are behavioral feedback surveys. Optimally, standardized behavioral feedback surveys are the subject of not one, but two cycles of validation research, both of which are dramatically different from procedures used to validate psychological instruments: (1) author/publisher validation and (2) local validation.

The surveys featured in the Survey Library have been developed by experts who have extensive experience and learning in their specialized areas. Each used a standard method for researching competencies. The first step was a comprehensive review of the literature to determine what current research and authoritative writing has defined as highly desired competencies. Based on this research, the developer then created draft competency lists that model the target area of performance. These lists were then screened for redundancy, importance and validity, based on the survey authors' experience in organizations. The lists were then validated by the developers' client organizations, a process in which the models were examined by organization subject matter experts, used in training and development programs, checked for improvements in performance, and finally referenced in quantitative and qualitative feedback—an effort that often spanned several years. During that process, the survey items were also circulated to a variety of other experienced professionals for evaluation and suggestion. The final versions of the surveys were studied and revised as necessary by the publisher for acceptance in the Survey Library. The surveys were then offered to organizations during beta testing and later for general use in 360-degree feedback programs. The publisher reviewed the feedback from these groups and made appropriate revisions. In

the case of 20/20 Insight, the surveys included in the Survey Library have been systematically reviewed and used by hundreds of organizations.

The standard behavioral feedback surveys in the Survey Library were validated for developmental use, not for linkage to personnel or compensation decisions.

3. All behavioral feedback surveys must be customized and validated locally.

Different industries involve dramatically different business practices; what is understood as desired performance within organizations varies widely, as does the language that describes it. For example, in the U.S. Army, the Federal Reserve Bank and the Baptist church, leadership is defined and practiced differently. Yet, each desires to use 360-degree feedback to hold up a mirror to its leaders' performance. While there are several common behaviors, many are different.

Therefore, the second phase of validation must inevitably involve local validation. The author's research does not—and cannot—establish universal validity. While a one-size-fits-all instrument is possible and desirable in the field of psychological testing, it is neither possible nor desirable for feedback about specific workplace behavior. A published standard competency list is nothing more than an optimum start-point for organizational research, customization and validation locally.

This process requires people within the organization (or consultants) who are familiar with the area of competence to review the standard survey. Items inappropriate to the local workplace may be deleted and additional items added. Wording can be modified to use business-specific terms and align the items with the organizational culture. The new behavior list is then reviewed for size, frequency and importance by local subject matter experts and stakeholders; and additional revisions are made, if appropriate. A pilot assessment with a sampling of subjects tests effectiveness of the survey, and feedback from the process is studied for possible refinements. The final product is a locally validated list of observable behaviors.

For behavioral feedback surveys, validity equates to relevance to workplace performance. Does it focus on the most important workplace behaviors in this particular organization? In the end, the most valid survey process may not employ an entire competency list, but focus on portions of it that are deemed by executives to have the highest potential for individual and organizational improvement.

About the Author

Dennis E. Coates, Ph.D., is CEO of Performance Support Systems, based in Newport News, VA. He coordinates research and development and provides strategic direction for the company. He is the author of 20/20 Insight GOLD, an award-winning 360-degree feedback system (www.2020insight.net), and co-founder of the Train-to-Ingrain alliance (www.train-to-ingrain.com). Dr. Coates has over 40 years experience as a leader, manager and HR professional.

Copyright © 2006, Performance Support Systems, Inc. All rights reserved. This article may be reproduced for educational purposes only. Embodiment of this material in products or resale in any form is strictly prohibited.