Development of Competency-Based Assessment Centers

Hsin-Chih Chen  
Sharon S. Naquin  
Louisiana State University

Construct validity has been an issue in assessment centers literature. This study developed a model for competency-based assessment centers by integrating competency modeling, training design, assessment centers, and multi-source feedback assessment. Several propositions in conjunction with the model were also developed to help resolve construct validity issues and ensure success in competency-based assessment center implementation.

Keywords: Assessment Centers, Competency Modeling, Performance Assessment

A review of related literature indicates that researchers have not reached a clear definition of competency. The term sometimes refers to outputs of competent performers and sometimes refers to underlying characteristics that enable an individual to achieve outstanding performance (Dubois & Rothwell, 2004; Hoffmann, 1999; McLagan, 1997). Most definitions, however, relate to exemplary performers or performance in a specific job or job level (Boyatzis, 1982), while a relevant term “core competency” is tied to strategic, future-oriented, collective functions in an organizational level (Moigneon & Edmondson, 1996; Parhalad & Hamel, 1990). Thus, we have adopted an overarching perspective that combines both the performance and strategic aspects associated with the various definitions found in the literature. We consider competency to refer to the underlying individual work-related characteristics (e.g., skills, knowledge, attitudes, beliefs, motives, and traits) that enable successful job performance where “successful” is understood to be in keeping with the organization’s strategic functions (e.g., vision, mission, uniqueness, future-orientation, success, or survival).

A similar construct, competency development or competency modeling refers to the process of identifying a set of competencies representative for job proficiency. With the generic term just defined, competency development can be applied for various purposes including personnel selection, job promotion, training and development, training needs analyses, performance appraisal, individual career planning, organizational development, human resource planning, placement, strategic planning, succession planning, compensation, and recruitment (Byham & Moyer, in press; Howard, 1997; Lucia & Lepsinger, 1999).

Understandably, assessment strategies and methodologies are closely related to competency, and long standing assessment strategy is the use of assessment centers. “Assessment centers” are not brick and mortar research centers or buildings. They are rather an abstract concept that exists in practice and refer to standardized procedures used for assessing behavior-based or performance-based dimensions whereby participants are assessed using multiple exercises and/or simulations (Thornton, 1992). Common simulation exercises used in an assessment center include oral presentation, leaderless group discussions, role playing, in basket, oral fact-finding, business game, and integrated simulations (Thornton, & Mueller-Hanson, 2004). Dimensions for assessment (equated to competencies) are usually identified through job analysis. It should be noted that though job analysis is often used as interchangeably with competency modeling, the two differ in terms of assessment of reliability, strategic focus, and expected outcome (Shippmann, et al., 2000).

Research Problems and Purposes

Assessment center research has advanced greatly over the past few decades as the literature has moved from focusing on an understanding of what it is and how it works to establishing some criterion-related validity and generalizability (Howard, 1997). However, a research challenge remains unmet as the issue of establishing construct validity of assessment centers is still unresolved (Robie, Osburn, Morris, Etchegaray, & Adams, 2000). Construct validity refers to the degree to which a theoretical concept is operationalized and the operational inference exhibits consistency to what a researcher intends to measure. In current assessment center literature, it largely refers to discrepancies between competencies and the measures in assessment center activities that are used to demonstrate such competencies. Woehr and Arthur (2003) assert that the lack of construct validity in assessment center literature is largely due to issues of design and development.

Copyright © 2005 Hsin-Chih Chen & Sharon S. Naquin
This challenge also clearly relates to an ongoing debate as to whether the design of assessment centers should be based on dimensions/competencies or tasks/exercises (Byham, in press; Lowry, 1995; Joyce, 1994; Howard, 1997). Since HRD is deeply rooted in the design and development of learning activities across various levels, integrating the human resource development perspective into assessment centers has the potential to open the opportunity for resolving the construct validity issue and helping to move the HRD field away from its primary evaluation practices—cognitive or reactive assessments—to a higher level, behavioral or performance assessment.

Changes of individual behavior can be readily observed through assessment centers. However, the ability to assess implicit behavior (e.g., motivation and emotion) through assessment centers is limited. In contrast, multi-source feedback such as dual-ratings assessment can potentially be more effective in assessing implicit behavioral competencies, but these methods are not able to provide the level of information regarding tangible outcomes that assessment centers can. This is due primarily to the fact that assessment centers typically involve observation of outcomes or performance—behaviors—while the multi-source feedback method relies on perceptions and/or memories of behavior. Therefore, the purpose of this article is to develop a competency-based assessment center model that integrates competency development, assessment center, and multi-source assessment strategies. This model attempts to serve multi-purposes. First, it introduces a systematic approach to linking competency development, assessment center and multi-source assessment. Second, it provides a design process that has the potential to enhance the construct validity of an assessment center. Third, it helps develop a set of propositions for future research.

**Conceptual Framework**

Best practices in competency-based development and the current version of the “Guidelines and Ethical Considerations for Assessment Center Operations” which is endorsed by the international congress on the assessment center method (Joiner, 2000) guided our research efforts to develop a model for a custom competency-based assessment center. According to Joiner (2000), an assessment should include ten key components: job analysis (equated as competency development in this context), behavior classification, assessment techniques, multiple assessments, simulators, assessors, assessor training, recording behavior, reports, and data integration. When developing such a model, comment errors of assessment centers (Caldwell, Thornton, and Gruys, 2003) were also considered. The errors suggested by Caldwell et al. (2003) include poor planning, inadequate job analysis, weakly defined dimensions, poor exercises, lack of pretest evaluation, unqualified assessors, inadequate assessor training, inadequate candidate preparation, sloppy behavior documentation and scoring, and misuse of results. The model can be found in Figure 1.

![Figure 1 Integrated Competency-Based Assessment Center Model](35-2)
Integrated Competency-Based Assessment Center Model

Building Hierarchical Competency System

The first step in developing a competency-based assessment center is to build a hierarchical competency system that breaks a whole into supporting parts. This step is critical because it layouts a framework to guide training design and assessment center measurement. The number of levels of competencies depends on the complexity of a system. The task of identifying competencies in current competency modeling practice takes various forms. Some identify competencies in specific ways such as in performance or behavioral indicators (e.g., respond to customer’s inquiry politely and consistently; adjust the equipment in terms of a mechanical manual, etc.). Others describe them in generic or abstract terms (e.g., communication, problem-solving, networking, team building, etc.). As Holton and Lynham (2000) point out, "competencies are less specific than tasks but more job related than learning objectives alone."

For discussion purpose, we divide the hierarchical competency system into three levels: competencies, sub-competencies, and procedures or steps. Competencies are described in collective, abstract form while its supporting sub-competencies are more measurable, specific but less collective than competencies. Sub-competencies are normally composed of a set of observable, specific, behavior-based steps. The three-level hierarchical competency systems appear to be effective in communicating with stakeholders and linking competency to training design and assessment center. The first level, which is in abstract form, can be easily communicated in discussing competency issues with stakeholders. The second level (the sub-competencies) is the action statement that supports competencies. The third level provides detailed descriptions how to achieve the action statement in level two.

In the hierarchical competency system, the sub-competencies serve as connectors to link to training and assessment center designs. In a training design, the sub-competencies can be desired job outcomes representing what training participants are expected to perform when they return to their job. In an assessment center design, the sub-competencies can be used as performance objectives that participants are expected to perform in assessment centers. The relationships among competency model, training program, and assessment center can be found in figure 2.

Proposition 1: linking sub-competency to training design and assessment center will improve construct (competency) validity of the competency-based assessment center.

1. Competency model triggers training and competency-based assessment center designs.
2. Competency-based assessment center includes a traditional assessment center and a multi-source feedback assessment.
3. Competencies are in collective, abstract form.
4. Sub-competencies are more measurable, specific but less collective than competencies.
5. Steps or procedures are observable, specific, and behavior-based. Steps or procedures are in very specific form and described in support of sub-competencies, which are in term to support competencies.
7. Learning objectives are in support of job outcomes in a training design.
8. Job outcomes are work-related outcomes while learning objectives are supported by training materials.
9. Performance objectives are general indicators that assessment center and multi-source feedback assessments are targeted to measure.
10. Performance indicators are specific indicators in support of performance objectives.

Figure 2 Relationship between competency model, Training Program, Assessment Center, and Multi-source Feedback Assessment
Much of the literature in competency development addresses validity issues relies on face or content validity. Specifically, determining whether or not the competencies are valid is most often based on subject matter experts or managers’ judgment. In other words, the validity is through a qualitative rather than quantitative lens. The hierarchical competency system can serve as a conceptual framework for quantitative research to enhance construct validity of competencies. For example, the competencies can be used as constructs to be assessed, while the sub-competencies are variables to represent the constructs. Researchers can use sub-competencies as questionnaire or survey items, distribute to a targeted sample, collect data from it, and conduct a factor analysis to examine the relationship between competencies and sub-competencies.

Caldwell, Thornton and Gruys (2003) point out a common error of assessment center practices – weakly defined dimensions or competencies. However, factor analysis can easily allow the researcher to determine how well the competencies were defined and can serve as a means to help refine the definitions of the competencies.

Proposition 2: Using a factor analysis, in addition to qualitative competency development, to examine construct validity of competencies will help refine definition of competency and enhance the validity of the competency-based assessment centers.

Developing Sub-Competency-Assessment Center Activity Matrix

Developing a competency-exercise matrix is a basic requirement for assessment center development (Joiner, 2000). Current practices for developing such a matrix are conducted at competency level, which is an abstract level (e.g., Halman and Fletcher, 2000). However, using an abstract competency to develop assessment center exercises can potentially jeopardize the validity of selected assessment center activities because such a matrix cannot identify the most appropriate activities to assess the competencies. For example, from a generic view, one may select role-play activities to assess an individual’s communication competency. However, the communication competency can encompass written and oral skills yet the role-play does not address all necessary communication skills.

Proposition 3: Using sub-competencies, which collectively represent competencies in a more observable way, to develop the assessment center activity matrix will enhance construct validity of competencies.

As previously mentioned, the traditional assessment center is a behavior-oriented approach that mainly relies on various exercises, activities, or simulations to observe individual behaviors. And the dimensions to be assessed in the assessment center are based upon a job analysis. However, this technique has limited ability in measuring implicit behavior such as motivation, values, and beliefs. Although one may argue that implicit behavior can be measured thorough transferring it to explicit format, it evidently cannot be effectively managed in assessment centers due to complexity of implicit behavior.

Consequently, assessing the implicit-behavioral competencies in traditional assessment center could have created more problems than it can be solved. It is very likely that the construct validity issue from an assessment center is due to lack of differentiation between explicit-behavioral and implicit-behavioral competencies. Therefore, multi-rater feedback assessments, a popular assessment mechanism, appear to be more effective in assessing implicit-behavioral competencies. This is because multi-rater assessment has the advantage of being able to in solicit implicit information. This supports the notion that multi-rater assessments and assessment centers can complement with each other (Howard, 1997). Integrating multi-rater feedback in the assessment center design also provides flexibility to reduce complexity and avoid common errors (See more discussions at later section).

Proposition 4: Differentiating between explicit-behavioral and implicit-behavioral competencies will improve construct validity of competencies where explicit-behavioral competencies are measured by traditional assessment center and implicit-behavioral competencies are assessed by multi-source feedback assessment.

Determining Appropriate Assessment Center Activities

Current research in developing the competency (or sub-competency)-activity matrix uses simple check marks to determine an exercise to be used for a particular competency. However, this approach provides no information on how well the competencies fit the exercises. We suggest using numeric ratings such as five-point Likert type scale to determine appropriateness of assessment center activities to the sub-competencies by treating the matrix as a questionnaire. This approach not only helps in alleviating subjective decisions but also provides more valid information on the degree to which sub-competencies fit assessment center activities.

This approach will require a group of participants to rate the questionnaire and then calculate aggregated scores on the collected data. Although it sounds impractical to involve a group of people to rate the matrix, if this approach...
can enhance the assessment center design validity, it should be considered. As a matter of fact, as long as each activity in a matrix development is clearly defined, any managers or trainers in an organization should be able to serve as raters for the questionnaire.

Moreover, according to the guidelines, to increase the likelihood of obtaining objective data, each dimension or competency should include more than one assessment exercise. The numeric scale has merit in assisting an assessment center designer to select the most appropriate activity to be used. The following strategies are designed to help determine the most appropriate activities to be used in an assessment center. They include:

- Select two top-ranked activities for each of the sub-competencies.
- If more than two activities are tied as top-ranked, take consider all of them.
- If no rating for a particular sub-competency is greater than 3.0, its applicability to any of the exercises is low; therefore, considering multi-source feedback questionnaire as a more appropriate approach to assess the sub-competency.

Proposition 5: Using numeric rating scale rather than a dichotomous scale will lead to appropriate assessment center activity selection; therefore, the numeric scale will indirectly influence validity of the competency-based assessment center.

Determining Performance Objectives for Activities

This step requires composing a list of appropriate sub-competencies/performance objectives (the two top-ranked activities) to each of the activities. These performance indicators will be aligned with the activity design. It is important to note that the assessment center designers should not over rely on quantitative data as presented here to design the assessment center activities since quantitative data are only meaningful as well interpreted. Activity designers should always review these sub-competencies and examine whether or not they are in a good fit. Our suggestion is to move less congruent sub-competencies in an activity to multi-rater feedback.

In addition, research suggests that one activity should not include too many measures; otherwise, the assessors’ rating could be biased by intuition due to the limitation of our cognitive abilities in differentiating complex situations in certain amount of time (Lievens & Klimoski, 2001). When the number of sub-competencies increases, an assessment center activity designer should use judgment to avoid the problem of measuring too many sub-competencies in a single exercise or activity. Thornton (1992) suggests five to ten dimensions (sub-competencies in this context) to be assessed for various assessment centers while Thornton & Mueller-Hanson (2004) state that, in practice, consultants only measure four or five dimensions in an exercise. Synthesizing findings and suggestions of these literatures, it is reasonable to assert that no more than 10 dimensions are practical for in an activity. On the other hand, from a cost-effective perspective, for an activity with fewer than five sub-competencies to be measured, it is also reasonable to eliminate the activity and move the sub-competencies classified in this activity to multi-rater feedback assessment.

Proposition 6: Measuring no more than 10 sub-competencies in an activity will enable assessors to accurately assess the sub-competencies that are supposed to be measured. Therefore, it increases the validity of the assessment center measurement.

Designing Competency-Based Assessment Center Materials

In developing assessment center materials, although two models, off-the-shelf and custom, can be applied, using a custom model adds more value to this systematic competency-based assessment center design. In addition, Thornton (1992) suggests that “fidelity” of assessment center design will help improve the validity of performance outcome. The notion of fidelity is essential to design activities or cases that close relate to participants’ day-in-the-life work situation.

Thornton and Muller-Hanson (2004) suggested several sets of exercise materials that must be designed for various persons involved in the exercises including participants, administrators, assessors, resource persons, and role players. In addition to determining the sub-competency or supporting performance indicators for exercise development, based on different exercises, an assessment center designer should also consider factors such as setting, technology, and level of difficulty of the indicators when designing exercise materials. In developing multi-rater feedback assessments, the supporting performance indicators should also be as action-oriented as possible (e.g., starting an action verb). An appropriate supporting performance indicator may include the performance to be measured, condition in which the performance occurs, and criterion to determine effectiveness or efficiency of the performance (Mager, 1997).
Proposition 7: Using custom assessment center materials which are designed with close relationship to participants’ work settings will lead to stronger predictive validity of competency-based assessment center.

Selecting and Developing Assessors

Selecting and developing qualified assessors usually go hand-in-hand. In selecting assessors, Spychalski, Quinones, Gaugler, and Pohley (1997) found that most practices use line or staff management as assessors and these assessors are generally two organizational levels higher than the individuals to be assessed. Only a small portion of assessment center practices used psychologists as assessors. In addition, research on the effect of assessor’s individual background shows a mixed result. For example, Gaugler, Rosenthal, Thornton, & Bentson, C. (1987) found that assessment centers that used psychologist assessors as assessors exhibited higher criterion validity than managerial assessors. However, Thomson (1970) found no significant differences between ratings of psychologist and manager-assessors. Although the assessment center guidelines suggested considering professional psychologists as assessors, from a practical standpoint, it is plausible to select assessors from the target organization. The more important aspect is perhaps to keep these selected assessors (e.g., managers) well trained on what and how to assess assesses’ performance before engaging in an assessment center activity.

In addition, according to the guidelines, the assessor training should clearly state training objectives and performance guidelines. The objectives of assessor training are to facilitate assessors gaining reliable and accurate judgments. Contents in the assessor training may include: knowledge and understanding on assessment dimensions, definitions of dimensions, relationship to job performance, examples of effective and ineffective performance, simulations on exercises to be assessed, ratings issues, data integration, feedback procedures, etc. Training length should be determined in connection with other considerations such as trainer and instructional design, assessor capability, and assessment program. It is also important to consider establishing a continually improving training system to help assessors keep up-to-date skills, knowledge, and attitudes. More detailed issues related to assessor training can be found in the assessment center guidelines (See Joiner, 2000).

Finally, a trainer of assessor training should be familiar with simulation exercises, have a deep understanding on issues related to assessor training, and continually communicate with assessment center designers and a program champion. This is because the assessment center designers master in functions of an assessment center design, while program champions have broader insights on how the program could work. Both, if a communication system is well established and utilized, can contribute to the success of assessor training.

Proposition 8: Well-trained assessors will contribute to criterion validity of a competency-based assessment center.

Communicating with stakeholders

Designing and implementing an assessment center is labor-intensive, time-consuming, expensive, and difficult to manage (Dulewicz, 1991). Perhaps these are the major reasons that assessment centers do not receive enough attention in HRD. Most assessment centers are often applied for selection and promotion but not tied to HRD practices. Competency-based assessment centers as presented in this article tie competency models to training program design, resulting in assessment center activities with meaningful implications to organizational success. This is because this design can be easily integrated with other human resource development functions through the common connector, the sub-competencies. On the other hand, current training effectiveness practices that focus on participant’s reaction and learning add little to how participants can apply what they learned to the jobs. A competency-based assessment center is a behaviorally oriented system that can help answer such a question. Moreover, a competency-based assessment center can serve as a diagnosis tool for trainee qualification because it can provide more useful information to determine the level of qualification than paper-and-pencil-based tests. Consequently, these facts seem obvious yet many organizational stakeholders do not naturally recognize them. Therefore, we propose several strategies that could be useful in communicating with organizational stakeholders the need to adopt competency-based assessment centers in their organization.

Educating organizational stakeholders. The major objective for this strategy is to advocate various advantages and applications of assessment center to organizational stakeholders. Specific information that can be provided may include: the reasons that assessment centers are important, how an assessment center can make changes, the differences between traditional assessment center and competency-based assessment centers, and benefits of competency-based assessment centers to individual development and organizational effectiveness as a whole.

Collecting and providing cost-effective information. Cost-effectiveness is always a concern for organizational decision makers. For a competency-based assessment center to be adopted by an organization, it is imperative to collect cost-effective information of existing assessment center practices from other organizations and provide cost-effective information of the competency-based assessment center program that will be implemented.
Providing comparative information. This strategy mainly deals with providing information on what other organizations in the same industry have done on assessment centers and how well the assessment centers have helped the organization improve performance. The information allows decision makers to justify how assessment centers can make change in their organization.

Articulating purposes of the competency-based assessment centers and data to be used for. Because assessment centers can be used for various purposes, it is important that assessment center designers and implementers fully articulate the purposes and describe how the data collected from the centers will be used for. When a new tool such as competency-based assessment centers to be used for an evaluation purpose, it is inevitable that resistance will be met. Anxiety and motivation to change are often related to resistance. Therefore, articulating purposes could be the key to reduce the anxiety and enhance the motivation for stakeholders to adopt the program.

Implementing a pilot test. The process of this competency-based assessment center design aims to facilitate a customized assessment centers. Implementing a pilot test can provide information of formative feedback for program design. It also helps examine practicability and other issues (e.g., culture fit) of assessment centers in an organization.

Communicating responsibilities. The implementation of a competency-based assessment center cannot solely rely on designers or implementers. Organizational stakeholders’ involvement and support will be the key to its success. Therefore, communicating responsibilities before a center is implemented is equally important to other strategies proposed here.

Proposition 9: Education statements about the needs, providing cost-effective and comparative information, articulating purposes, implementing a pilot test, and communicating responsibilities with stakeholders will lead to a successful competency-based assessment.

Conclusions

This model has several implications for research and practice. Five major distinctive components are included in this model. First, this model linking competency modeling to training design, and then to the assessment center provides a systemic view in designing assessment centers for human resource development practices. The systematically integrated design (e.g., sub-competencies, job outcomes, and performance objectives) can easily build in other human resource functions (e.g., instructional system design, performance appraisal).

Second, this model suggests that research goes beyond face validity by using a quantitative statistical technique—factor analysis—to verify constructs or competencies to be assessed and refine definitions of the competencies at the early stage of the competency-based assessment center development. Using such a scientific approach can primarily enhance the construct validity of the competencies. Therefore, the sub-sequential design and planning can accurately reflect to the competencies. We also acknowledge that this approach can be best utilized for a large organization that develops overarching competencies for corporate-wide use but not for small business units.

Third, applying numeric rating scale to the sub-competency-activity matrix allows practitioners to assess how well targeted sub-competencies fit the activities. This component has potential to enhance construct validity and provides congruent linkage from sub-competencies to activities design.

Fourth, building multi-rater feedback assessment into the design offers the ability to tackle implicit behaviors that traditional assessment center normally cannot do well. The built-in multi-rater feedback assessment also helps reduce the number of activities to be conducted for greater cost benefits. And finally, the model with embedded HRD practices provides a researchable framework and a set of propositions to be examined as future research direction.

References


