REFERENCES


Using an Assessment Center as a Developmental Tool for Graduate Students: A Demonstration

Janet L. Kottke
Kenneth S. Shultz

Department of Psychology, California State University
5500 University Parkway, San Bernardino, CA 92407-2397

Assessment centers are widely used in industrial settings to select, promote, and provide developmental feedback for executive-level talent. This paper describes the development and implementation of an assessment center (AC) administered to graduate students in a two-year industrial/organizational (I/O) psychology program as part of a practicum course requirement. To develop the AC, graduates of a master’s program in I/O psychology and several of their supervisors were interviewed for the core process skills necessary for success in organizational settings. Interviews were content coded and six competencies identified: written communication; oral communication; problem solving; organizing; interpersonal; and organizational survival skills. Four assessment center exercises were developed to measure these skills: a leaderless group discussion; an oral presentation; an in-basket, and a role play. Assessors rate graduate students in the program on the six competencies. Written feedback on their performance in the assessment center is provided to the students for their use in career development planning. Benefits include rich developmental feedback for the graduate students and a fresh view for the I/O faculty of the requirements of the MS I/O program.

Since their first use in military applications (U.S. OSS, 1948), assessment centers have become widely used in organizational settings to select, promote, and provide developmental feedback to executive-level talent (Byham, 1971; Gatewood & Feld, 1994). An assessment center may generally be described as “a variety of testing techniques designed to allow candidates to demonstrate, under standardized conditions, the skills and abilities most essential for success in a given job” (Joiner, 1984, p. 437). To make these assessments, the traditional assessment center

Authors’ Notes. The first author gratefully acknowledges the California State University for funding to support the identification of the critical competencies. The authors thank Diane Pfahler and three anonymous reviewers for their helpful comments on an earlier version of this paper.

Correspondence should be addressed to the first author (e-mail: jkottke@wiley.csub.edu).


requires assesses to participate in several simulations or exercises. Common activities used are leaderless group discussions, role plays, in-basket exercises, and personality tests. These multiple assessments have been found to be sufficiently valid for selection decisions (cf. Gaugler, Rosenthal, Thornton, & Bensoni, 1987; Hunter & Hunter, 1984); the developmental feedback from these assessments can also be very useful to the individuals who undergo an assessment center experience (Cochran, Hincke, & Dusenberry, 1987; Engelbrecht & Fischer, 1995; Gatewood & Feild, 1994). The use of an assessment center for developmental feedback is the focus of this paper. We will describe the development and implementation of an assessment center (AC) for graduate students in an industrial/organizational (I/O) psychology masters of science (MS) program where the AC is used exclusively for developmental purposes. We focus here on an AC within an I/O graduate program because of our association with such a program. However, we believe the strategy of developing an AC we describe here would have broad applicability to other academic programs that teach applied material. Our method of identifying the competencies and skills could be generalized to many other educational programs.

The use of an AC to evaluate students is not unique. Others have developed and used ACs with college students (See Aguirre, Mayes, & Riggio, 1995; Hakel, 1993; Ren, Rea, & Moomaw, 1990; Wendel & Joekel, 1991). However, we appear to be the first to use an AC for I/O students within the context of a required class. Before describing our AC, we will provide some background on the use of ACs for development.

Use of the Assessment Center to Yield Developmental Feedback

Most research about ACs has focused on their use as selection devices (Thornton & Byham, 1982). The AC has been found to have satisfactory predictive validity (Gaugler et al., 1987) and excellent utility (Cascio, 1991) for use in selection of managers, supervisors, executives, and even school administrators. Despite anecdotal reports about the use of ACs for feedback and how AC participants have reacted to feedback or have altered self-assessments (Fletcher, 1991; Fletcher & Kerslake, 1992; Schmitt, Ford, & Stults, 1986), surprisingly little research has been conducted to address the developmental value of the AC (Robertson & Rout, 1989). One study (Engelbrecht & Fischer, 1995) has directly addressed the issue of whether developmental feedback could translate into performance improvements on the job. Using a quasi-experimental design, Engelbrecht and Fischer found that first-line supervisors who participated in an AC were rated superior to nonparticipants on the majority of behavioral dimensions used to evaluate job performance three months later. Despite the limitations of their study, the authors provided empirical evidence that developmental feedback can be useful for assesses' career development and that an AC may improve performance on the job. Engelbrecht and Fischer used job performance as the criterion by which to evaluate AC feedback. In most organizations, the concern will be whether the feedback from the AC leads to job-related outcomes. However, ratings derived from an AC used within an academic setting could be used for individual developmental feedback and as criteria to evaluate student success (Aguirre, Mayes, Bellini, & Kubiak, 1997).

Assessing Graduate Training in Applied Psychology

The challenge of any applied graduate program is to produce graduates who will be competent in the technical skills of the profession and who will be able to successfully transfer those skills to the applied setting (Kottek, 1994). Academics and employers alike have expressed concern about the transferability of skills from the academy and have called for schools to perform outcome assessments. Additionally, outcome assessment may be necessary for higher education to acquire operating funds in the future (California State University Institute for Teaching and Learning, 1992), but currently very little outcome assessment is being done in psychology graduate education. In this respect, the applied social sciences appear to lag behind business, whose accrediting body, the American Assembly of Collegiate Schools of Business (AACSB) has revised its standards to require business schools to demonstrate that their students have achieved certain levels of knowledge and skill (Aguirre et al., 1995).

Because I/O psychologists often study people's work through the use of job and task analysis, we might expect that the curriculum of graduate study in I/O would be based in part on empirical investigations of graduates' subsequent placements. That is, we would expect that applied graduate programs would have been founded upon, or at least train for, the necessary knowledge, skills, and abilities required to perform as a professional. Instead, very little has been attempted to systematically analyze the tasks performed by applied psychologists at either the PhD or master's level (but see Schippmann, Hawthorne, & Schmitt, 1992 for a notable exception). I/O psychology is not alone in that it has not linked curriculum to requisite job skills. Several authors (Benassi & Fernald, 1993; Bourg, Bent, McHolland, & Stricker, 1989; Lowe, 1993; Lumsden, Grosslight, Loveland, & Williams, 1988; Smalley, Vineburg, Schippmann, & Prien, 1990) have expressed concern about graduate student preparation in several areas of psychology.

1 The Society for Industrial/Organizational Psychology (SIOP) has developed guidelines for master's level programs that focus on the technical content recommended for the master's curriculum. The Council of Applied Masters Programs in Psychology (CAMPP) also meets regularly to address master's level education issues.
In this paper we outline critical competencies we deem important to the success of graduates from an I/O psychology program. Our work is not the first in its attempt to link graduate training to job responsibilities—Schippmann et al. (1992) conducted task analyses of both master’s and PhD-level I/O psychology graduates’ work. What differentiates our effort from Schippmann et al.’s work is that we sought common skill competencies that are process oriented and “cut across” multiple applied settings (i.e., are not strictly technical or content-domain specific).

**Identifying the Competencies**

**Interviews.** Graduates of our master’s I/O program were contacted; 11 former students were interviewed with regard to the process skills and abilities necessary to perform successfully in their jobs. Students interviewed held a wide range of jobs consistent with the range of jobs master’s level applied psychology graduates typically hold (e.g., personnel research analyst, trainer, research associate). Interviews were conducted with three of these former students’ employers to define and amplify the competencies described in the graduates’ interviews.

The average length of an interview was two hours and in total, approximately 50 pages of notes were generated. The competencies were synthesized from commonalities that emerged from several hundred adjectives, descriptors, and behaviors identified in the written interview notes (Kottek, 1993).

One possible limitation in the sample that may have affected the final list of competencies was that the average length of time that a graduate had been working in the field was about three years. Hence, the majority of graduates were serving in entry-level jobs with little supervisory responsibility and no leadership dimension emerged from the interviews. The lack of a leadership competency may not be a serious limitation, because we are most concerned with skills that are deemed important by employers at time of entry.2

**Competencies.** The critical competencies identified, and their primary components, are listed in Table 1. These skills, some with different labels, are frequently found in the literature describing requisite managerial skills (Fandt, 1994; Mullin, Shaffer, & Grelle, 1991).

**Integration of the Competencies into the Graduate Program**

To provide a context, a brief description of our MS I/O program may be helpful. Our program is a two-year 82-quarter-unit program that requires a research thesis. Students take five quantitative-research courses, four I/O content courses, four quarters of I/O seminar, and three courses from the core content psychology curriculum (e.g., cognitive). The courses within the program allow for several of these process skills to be trained and assessed (i.e., written and oral communication) (Shultz & Kottek, 1996), but most have not been systematically addressed. A course required for graduation and a prerequisite to the internship, practicum in applied psychology, was structured to train these competencies through discussion and practice simulations. The AC is a required component of the practicum class; we use the AC as the primary mechanism to evaluate a number of these process skills3 among MS I/O graduate students.

---

2 The other university of which we are aware that is using an AC for developmental purposes within an I/O (PhD) program is Bowling Green State University. Their program assesses for skills perceived to be important at job entry. They do not assess for leadership skills, See Hakel, 1993.

3 The competencies identified here are not meant to be inclusive and generalizable to all master’s level training programs in I/O psychology. Rather, they are intended to be illustrative. Many of the typical curricular activities (e.g., group projects, internship placements) already present in many MA/MS applied programs can readily accommodate the development of the critical competencies in graduate students (see Shultz & Kottek, 1996).
TABLE 2  Descriptions of Assessment Center Exercises

In-Basket
Assessees:
- Plays role of assistant human resource director
- Given a week's accumulation from the "in-basket" of previous assistant director
- Has two hours to respond to items

Leaderless Group Discussion
Assessees (4-6) participate as one group
- Each given same materials
- No pre-designated leader
- Asked to arrive at a group decision after discussion
- Task must be completed in 45 minutes

Role Play
- One assessor plays supervisor and asks assessees to work weekend to finish special project
- Assessees play subordinate role who is told to resist working weekend
- One additional assessor observes and records behaviors
- 5 minutes to prepare and 5-10 minutes to role play

Oral Presentation
Assessees are given:
- Current materials on changes in US demographics and global business trends
- 30 minutes to prepare a 10 minute presentation
- Assessors (2-3) are audience who play roles as human resource specialists

Exercise Development
Four exercises were developed to provide a vehicle to measure these competencies: a leaderless group discussion (LGD), role play, oral presentation, and an in-basket. Table 2 describes the nature of each of the exercises. Exercise content was based on a review of the situations reported by alumni and their employers during the interviews and modeled after other managerial assessment centers. For example, for the oral presentation, participants are provided information regarding the changing demographics of the workforce and are required to prepare a briefing for a company’s human resources department on how the changing demographics will affect the company’s training in the future. These exercises were intended to provide a realistic simulation of applied problems and are of the type typically used in executive ACs. Table 3 shows the link between exercises and competencies. Linkages in Table 3 are based on consensus ratings of three I/O faculty and 12 advanced graduate students. Note that with the exception of written communication, each competency is addressed by at least two exercises.

AC Administration
First AC. The inaugural AC was conducted on a Saturday in several classrooms on our university campus. Assessors were trained in the morning and rated assessees in three exercises run in the afternoon. The in-basket was administered at a separate time. Nine assessees, nine assessors, and one administrator participated.

Subsequent ACs. Two ACs have been conducted since the initial AC. For the second AC, eight assessors and five assessors participated. To reduce the fatigue associated with requiring assessees to remain a full afternoon on-site, and the difficulty of recruiting assessors for a full day, the exercises were spread out over four days and took place during the pricicium class time. Although fatigue was reduced, the assessees reported that anxiety levels remained high throughout the assessment period and they felt that the experience was “disjointed.” Therefore, for the third AC—including 13 assessees and 8 assessors—we returned to conducting the AC in one afternoon.

Assessors and ratings. The first AC was staffed by former graduate students from the I/O program. For subsequent ACs, second-year students serve as assessors, rating first-year graduate students who are completing their first year. Assessors observe assessees in the exercises, record assessees’ behavior, and rate the assessees on the competencies. For most exercises, two assessors observe and rate one assessees. Assessors do not confer, but complete their rating forms independently. (The assessees

TABLE 3  Competency—Assessment Center Exercise Linkage

<table>
<thead>
<tr>
<th>Competency</th>
<th>Leaderless Group Discussion</th>
<th>Role Play</th>
<th>Oral Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizing Skills</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Oral Communication</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Written Communication</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Interpersonal Skills</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Organizational Survival</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Note: An “X” indicates that the competency is assessed within the specified exercise.
also rate themselves on the competencies prior to the AC. These ratings are not available to the assessors.

Assessor training. Assessors receive approximately 2.5 hours of training before rating. The assessors are first familiarized with the rating scales, with particular emphasis on the meaning of each competency subscale. In addition, the raters are given opportunities to practice writing behavioral statements.

Feedback and interrater reliability. Feedback to the assesses consists of behaviorally-oriented written comments about their performance on the competencies, ratings on the competencies of each exercise, and ratings on exercise-specific scales for the role play and in-basket. Assessors receive the rating sheets with the numerical ratings, recorded behavioral observations, and comments of the assessors. The practicum instructor reviews the forms before they are returned to the assessors. Assessors use these data and their own self-assessments to design a developmental plan for improving their skills. Their plans focus on implementing specific activities during their second year of graduate study. The practicum instructor reviews the developmental plan with the student and suggests any additional strategies for skill building and addresses any discrepancies in the student's plan.

For the most recent AC, interrater reliabilities (Pearson correlations) of the ratings using 7-point scales for all exercises were .28 for the in-basket, .45 for the leaderless group discussion, .61 for the role play, and .84 for the oral presentation. Certainly, we would have liked higher reliability estimates and have planned changes to the rater training. In training, emphasis has been placed on preparing behavioral statements. To improve the current training, we intend to use frame-of-reference (FOR) training (i.e., Bernardin & Buckley, 1981) in the next AC. The purpose of FOR training is to establish a common standard for all raters to use in making their ratings, with the goal to improve accuracy of rating rather than the elimination of rater errors such as halo or leniency (Hedge & Kavanaugh, 1988). If raters use common reference points in making their ratings, interrater agreement should improve.

4 The lower interrater reliability for the in-basket is probably a function of the complexity of the task. Also, training for the raters has focused on behaviors that occur in the five performances of the in-basket exercise (leaderless group discussion, role play, oral presentation). These data demonstrate a clear need to specifically address, in training, the rating of the behaviors of the in-basket.

5 If one uses Rosenthal and Rosnow's (1991) recommendation to calculate aggregate internal consistency of judges—i.e., effective reliability, the following reliabilities are obtained .45 for the in-basket, .69 for the leaderless group discussion, .76 for the role play, and .91 for the oral presentation. We thank an anonymous reviewer for this suggestion.

Evaluation of the AC and Value of the Feedback

Assessors and assesses from the initial AC were sent surveys requesting feedback about the experience of the AC and soliciting recommendations for future AC administrations. Five assessors and five assesses responded to the survey.

A critical question was whether to continue the AC in the future. All respondents recommended retaining the AC for future graduate students. In fact, comments from both the assesses and assessors were very favorable. Assessors, for example, commented on how "real-to-life" the exercises were and how they would have welcomed the experience when they were students. Negative comments by both assessors and assesses were directed at the context of the AC (i.e., scheduling, fatigue of testing in one day) rather than the AC itself.

A recent, extensive review of the effects of feedback interventions on performance by Kluger and DeNisi (1996) generally supports the positive influence of feedback interventions on increased performance. They do, however, caution that the feedback incentives tend to become less effective as the focus of feedback shifts from the task itself to the individual. Hence, if the AC experience is to be useful for providing feedback, three assumptions need to be met (Boehm, 1985). First, assesses must be able to make changes in behaviors or improve their skill levels as measured by the AC. Second, assesses must have sufficient motivation to engage in activities that will develop the relevant skills. Finally, the assessor must have developmental opportunities to remediate skill deficits identified through the AC.

We cannot directly address the first assumption because we do not know at entry, any given student's ability to learn or make behavioral changes. Yet, we think that a student who is able to gain entrance to a graduate program will have sufficient ability to learn. Also, we believe that the skills measured for can be taught and that the associated behaviors are sufficiently malleable to demonstrate the resulting changes.

The second assumption appears to be satisfied. Students are required for the practicum class to prepare a developmental plan that incorporates the feedback from the AC. Further, when the AC is conducted at the beginning of the course term, efforts toward improvement are incorporated into the grading of subsequent exercises.

Opportunities to engage in developmental activities, the third assumption, are also available and within the context of the practicum class. As required. The practicum class places the students into many simulations in which they practice the competencies. Outside of the class, substantial opportunities also exist. For example, within our master's program we
have instituted a mentoring program in which each first-year student is paired with a second-year student mentor. The first-year students are encouraged to seek out their mentors for help on improving their skills, using the written feedback as a base. Moreover, the developmental plan that the student develops in the practicum class must include explicit goals for the student's second year that address the competencies.

Finally, during the students' second year, students have completed their internships and are beginning their job searches. What the students will find is that these process skills, not technical skills, are many of the skills that distinguish successful from unsuccessful entry-level practitioners (Shultz & Lin, 1995).

**Benefits of the AC**

Though the primary objective for the AC has been to provide students feedback about their level of competence in several process skills identified as important to success in their first jobs, an additional outcome has been the impact the competencies have had on the I/O curriculum. The identification of core competencies necessary for graduate student success in various organizational settings has helped to inform the faculty within the MS I/O program and has given us the opportunity to review class requirements with these competencies in mind. For example, if a given class year as a whole scores low on a specific competency, the faculty can develop appropriate interventions as part of assigned class work to help students develop that competency. The identification of these competencies, therefore, has helped the faculty to reconsider the MS I/O curriculum and how we train for skills, both content and process oriented, within the program.

The behavioral feedback returned to the students within the context of the required practicum course is very rich and allows a student to plan and practice within the relative safety of the university. Another benefit of the AC feedback is that it is separate from any feedback or coaching faculty may have provided a student. Converging data from multiple sources validates for the student how the student is seen by his or her peers.

As a prerequisite to the internship, the practicum is intended to prepare the student to make the transition from school to the work world. The AC helps students to prepare and to remedy potential difficulties before taking an internship. Many of these benefits transfer to the entry-level job as well. Students gain practice with an AC before they may be asked to participate within one as a job requirement. Their experience as assessors gives them a sense of the responsibility of evaluating others before they may serve in such a capacity in their first job.

**Limitations and Future Directions**

Considering the time to train, observe, and rate assessors, the AC is relatively expensive to conduct. The start-up and maintenance cost may not be attractive to those who do not have an existing course that would accommodate the AC as we do in our graduate I/O program. In addition, the quality of the data is somewhat dependent on the quality of the assessors. Not all students are equally capable of providing high-quality data—of course, this problem is not unique to this academic-based AC—and students can be trained to produce ratings comparable to faculty (cf. Riggio et al., 1997). In the future, we plan to train assessors using the frame-of-reference technique to improve the reliability of the ratings.

Another possible limitation is that the competencies have not yet been validated against job requirements. The development of the competencies was based on a qualitative analysis of interview protocols. It is possible that different people might have interpreted the "armchair factor analysis" differently. The fact that others have found similar dimensions (Mullin et al., 1991) mitigates against this criticism somewhat, but work is underway (e.g., Sanchez & Kottek, 1996) to verify the dimensions used as the base for the AC exercises. We anticipate that this work will also help us to develop lists of frequently observed behaviors that will reduce the cognitive demands on assessors.

Despite these limitations, we believe that other disciplines may want to consider using AC technology. For example, education majors may benefit before they begin their student teaching by having exercises that require class management skills; certainly, an ability to provide clear oral presentation would seem to be important to education majors as well. Nursing students who will face difficult patients, and perhaps even more difficult physicians, may be assessed on their ability to handle difficult people in role plays. These are but two examples of two relatively disparate fields of study whose students may be helped by AC exercises and feedback before they graduate.

**Recommendations for ACs in Academic Settings**

Based on our experience, we have made and are planning further modifications to the AC. For those who might wish to use an AC for developmental purposes within an educational setting, we have some recommendations:

- Base the AC dimensions or competencies on a careful review of desired outcomes (i.e., do some form of job or task analysis).
- Develop behaviorally-oriented rating forms and lists of frequent behaviors for raters.
- Train the raters thoroughly to understand the core dimensions being assessed in the AC and how the rating scales should be used.
- Provide a context in which the AC ratings will be used by the assessees (students) so the experience will be seen as an integral part of the educational process.
- Run the AC exercises in a contiguous time block.
- Provide incentives for raters to produce high quality ratings.
- Educate the students and faculty about the purpose of the AC and the resulting ratings.

SUMMARY

In this paper, we have described the identification of critical competencies needed for graduates of a master's program in I/O psychology and the development of an assessment center to measure those competencies. The faculty have integrated knowledge of the competencies into the program by rethinking course requirements in terms of the competencies and have assigned work to remedy deficiencies when necessary. Further, students have an important source of feedback at an early stage of their careers to make adjustments before they enter their first job.

For our program, the AC, which is introduced as a course requirement, appears to be a viable approach to measure I/O graduate student process-oriented skills that is worth the expense of rater and administration time. In addition, we have also demonstrated that it is possible to design an AC at a state-supported university with relatively modest resources.

REFERENCES


Assessment Centers for Course Evaluations: A Demonstration

Bronston T. Mayes
Christopher A. Belloli
Ronald E. Riggio
Monica Aguirre

Student Outcome Assessment Center
California State University, Fullerton, CA 92634

A quasi-experimental design, employing assessment center (AC) pre- and post-measures of skills and knowledge, as well as traditional pencil-and-paper measures, was used to assess the amount of learning in two fundamentally different management classes. One class, the control group, focused on the theories and models of organizational behavior that emphasized bottom-up learning associated with committing material to memory and recalling memorized concepts. In the second class, the experimental group, students learned a variety of hands-on management skills. They then used bottom-up cognitive processes to identify critical cues in the task environment in order to diagnose problems requiring the application of specific skills. As expected, the theory class performed better on traditional tests requiring the recall of data from memory. The skills class performed better on assessment center exercises requiring the recognition of situational cues and the application of appropriate managerial action. The most significant gain in managerial skills occurred in the areas of oral communication and self-presentation in a mock employment interview. Issues related to AC use in higher education are student motivation, AC cost control, reliability of performance ratings, and statistical significance/power.

Academicians have long recognized the importance of evaluating the effectiveness of curriculum offerings. They have noted that the best criterion of good teaching is student learning. However, finding useful measures of learning has been a difficult task (McKeachie, Lin, & Mann, 1971). Because of the difficulty in finding acceptable measures of learning, course evaluations have tended to employ measures of instructional quality, which is believed to contribute to learning. One of the most frequently used assessment tools to evaluate the quality of instruction has been the student opinion survey.


American Psychological Association, San Jose, CA.


American Psychological Association, San Jose, CA.


