The Validity of Student Assessment Centers for Predicting Future Job Performance

Deidra J. Schleicher  
University of Tulsa

Ronald E. Riggio  
Claremont McKenna College

Bronston T. Mayes  
Cal. State University-Fullerton

Although the use of assessment centers in higher education has been touted for developmental and diagnostic purposes, it is an unanswered question whether or not such assessment centers can actually predict future job performance. The present paper details a follow-up criterion-related validity study for a student assessment center at California State University at Fullerton. Assessment center dimension ratings significantly predicted several criteria – including supervisors’ ratings of job performance – collected four years after assessment center participation. Results are discussed in terms of their contribution to previous research on, and controversies surrounding, assessment center validity.

Although assessment centers have been used in industry since shortly after World War II, it is only very recently that this method has been applied to educational settings, either as a developmental tool (e.g., Kottke & Schultz, 1997) or as a means of assessing managerial skill acquisition (AACSB, 1989; Mayes, Belloli, Riggio, & Aguirre, 1997). Given the recent increased attention to student outcome assessment in higher education (Brown & Glasner, 1999; Messick, 1999), it is no surprise that assessment centers have been used in this capacity. That is, they can provide the institution with a broad, performance-based assessment of the knowledge, skills, abilities, and other characteristics (KSAOs) of its graduates, which is important for school accreditation and for determining “work-readiness.” Also, assessment center methodology can determine deficiencies in students’ performance that can serve as feedback for curriculum improvement.

Something heretofore unexamined, however, is whether assessment centers in higher education can actually predict performance on the job. Although there is impressive and consistent evidence for the predictive validity of assessment centers for use in employment organizations (i.e., validity coefficients in the .30 - .40 range; e.g., Gaugler, Rosenthal, Thornton, & Benton, 1987), this evidence is all based on individuals who are out in the workforce. It is an unanswered (and unexamined) question whether the predictive validity of assessment centers would generalize to college student populations. Moreover, given that (a) the mission of most educational institutions is to prepare their students for success in their respective careers, as well as to develop various KSAOs, and (b) the major driving force behind the increased interest in outcome assessment is to demonstrate the “career readiness” of college graduates, such an examination of a more “ultimate” criterion for student assessment center success seems warranted. The current paper will detail a follow-up predictive validity study conducted on an assessment center used for student outcome assessment for undergraduate business school students at California State University, Fullerton (CSUF).

The Assessment Center

The Student Assessment Center (SAC) project at CSUF’s School of Business was a pilot project to explore the use of assessment center methods for student outcome assessment. Funded by the Department of Education’s Title VIII Cooperative Education Program, the SAC project was intended to develop methods for assessing undergraduate student learning while specifically examining the contribution of cooperative education and internship experience in developing undergraduate business skills and career readiness.

Assessees

Approximately 500 students participated in the Student Assessment Center (SAC) project at CSUF in 1994. These students had participated either as students undergoing outcome assessment (to assess learning outcomes) as they neared graduation, as entry-level (Junior year) students undergoing “income” assessment (to assess pre-existing levels of skills) to provide a baseline comparison for their eventual outcome assessment, or as students involved in other assessment projects. Modeled on traditional managerial assessment centers, this SAC consisted of a 5-hour combination of exercises and pencil-and-paper tests of relevant personality and skills to assess
students' work-readiness. Assessees completed the assessment center in groups of five. All assessment center exercises were videotaped.

Assessment Center Exercises

Three exercises were included in the present criterion-related follow-up study: a traditional leaderless group discussion (LGD), a presentation exercise, and a mock hiring interview. These exercises represent some of those most frequently used in assessment centers (Spychalski, Quinones, Gaugler, & Pohley, 1997; Thornton & Byham, 1982). Each exercise was designed to provide an opportunity to rate three dimensions: (a) decision-making skills (problem solving, reasoning, and learning; analytical and decision-making ability; creativity; and related areas of mental performance); (b) communication skills (clarity of communication, expressiveness, listening skills, appropriate use of language, organization of oral communication, and other factors related to communication abilities); and (c) leadership (persuasiveness, influence on others, leadership image, human relations abilities). In the LGD, each group of three to six assessees was provided with a problem and told to generate solutions. The problems assigned were issues familiar to undergraduate business students. The discussions lasted approximately 20 minutes. For the brief presentation exercise, students were given a business-related topic, along with some summary information on the topic. They were allowed 20 minutes to prepare (or practice) a very brief, persuasive argument, and then they made their one minute presentations. The final exercise in the present assessment center was a mock hiring interview. Students took part in this brief, simulated hiring interview where a trained interviewer asked a series of questions that might be included in a typical employment interview (e.g., “What special work-related talents or abilities do you have to offer an employer?” “What have you done that shows initiative and willingness to work?”). These interviews lasted approximately 15 minutes.

Assessment Ratings

Rating scales. Behaviorally Anchored Rating Scales (BARS) were used for rating performance on each of three dimensions in each of the three exercises. Following Reilly, Henry, & Smither (1990), the first step in the development of these scales was to identify behaviors relevant to the three dimensions that are elicited by specific exercises. A group of five undergraduate research assistants trained on the meaning of the three dimensions viewed the videotapes of all 500 assessees and recorded behaviors that were indicative of extremely high or extremely low performance on a particular dimension for a particular exercise (i.e., “critical incidents,” Flanagan, 1954). Procedures outlined by Smith and Kendall (1963) were then followed for selecting the behavioral anchors for the BARS. First, these behaviors were retranslated into the three dimensions by the five research assistants and the primary investigator, and a criterion of 83% agreement (5 out of 6) was used to select behaviors for each dimension. The research team then assigned effectiveness ratings (1 - 7) to these behaviors. Based on the means and standard deviations of these effectiveness ratings, behavioral anchors for the BARS were selected. Each behavioral anchor had three components, representing a behavioral example from each of the three exercises (such a rating format is considered a hybrid “task-BARS” approach; see Butler & Harvey, 1986; Harvey, 1991). Those behaviors that retranslated well but were not included on the BARS were used in teaching the theory of performance during training.

Assessor training. Assessors (faculty and students from a second university) received Frame of Reference (FOR) training (Bernardin & Buckley, 1981; Pulakos, 1984) before providing the ratings of assessees. These assessors received a complete description of the three exercises in the assessment center, including a videotaped demonstration of the exercises. They were then given copies of the dimension definitions and BARS to examine while the trainer read aloud the dimension definitions and scale anchors, including the different ways in which the dimension can be manifested in different exercises. Behaviors pertaining to different effectiveness levels for each dimension were also discussed by the trainer. The overall goal of this training was to create a common performance theory (i.e., frame of reference) among assessors so that they would agree on the standards used to evaluate asseseee behaviors.

Assessors then had the opportunity to practice rating using their new frame of reference on a videotape of training assessees. This videotape showed four assessees performing the exercises in the assessment center; it included three different assessees performing one each of the three exercises, and then a fourth asseseee performing all three exercises. The assessors were instructed to evaluate each participant after each exercise on each performance dimension. After each exercise, assessors were asked to share their ratings with the rest of the group; the trainer lead a discussion about the assigned ratings, clarifying discrepancies among
raters, and providing feedback on the appropriate effectiveness level on each dimension portrayed by each assessees. Training lasted approximately three hours.

**Target ratings.** Following training, each group of assessors viewed the videotape of five or six target assesses, and after each exercise, independently provided ratings on each dimension using the BARS introduced in training. The interrater reliability for ratings of each assessees on each dimension was computed for all groups of assessors. Following Schmitt (1977) and Schneider and Schmitt (1992), assessors in each group were treated as items (and assessors were treated as cases), enabling the calculation of internal consistency (i.e., coefficient alpha) estimates for each dimension. The dimension ratings demonstrated good interrater reliability, with alpha coefficients of .83 (communication skills), .89 (decision-making skills), and .75 (leadership skills). These reliability estimates further support previous conclusions that assessment centers generally demonstrate good interrater reliability (e.g., Howard, 1974; Klimoski & Brickner, 1987; Klimoski & Strickland, 1977; Sackett & Dreher, 1982).

Follow-up Validation of the Assessment Center

**Contacting Assesseees**

In 1998-1999 (approximately four to five years after completing the SAC), a follow-up survey assessment was conducted of the original assesses during their early post-graduation work careers. If assesses failed to return the surveys, a follow-up phone call was made after getting their phone numbers from an electronic telephone directory (i.e., PhoneDisc). Seventy-five respondents eventually returned the survey (20% of total valid surveys sent; 105 were returned from the US Postal Service as undeliverable). The breakdown of the sample was as follows: 54% male; 50% White, 25% Asian, 13% Hispanic, 9% Filipino, 2% American Indian, and 2% Middle Eastern; mean age was 29 years old (sd = 7.7); average tenure in organization was 32 months (sd = 26.3); 42% worked in the business sector; 25% worked in the service industry; 9% were in the manufacturing industry; and another 9% were in transportation/communication/public utilities.

Given the rather low (although not atypical) return rate, potential response biases are of particular concern. As a check on such biases, differences were examined between respondents and non-respondents on all demographic and personality variables for which such a comparison was possible. There were no differences between respondents and non-respondents in terms of age, gender, ethnic group membership, or any of the personality variables (all ps > .05). Taken together, these comparisons suggest little reason to be concerned that response bias is a significant factor in the present study.

**Criterion Variables**

The survey mailed to assesses asked for information relating to several occupationally-relevant variables, including: the numbers of interviews, job offers, and promotions the assesses had received since graduation; and ratings of (a) the quality of their exchanges with their supervisor (i.e., leader-member exchange, assessed via the 7-item LMX scale; Graen & Uhl-Bien, 1995); and (b) their job satisfaction (as assessed via the 20-item short form of the Minnesota Satisfaction Questionnaire; Weiss, Dawis, England, & Lofquist, 1967).

The survey also included several measures that were to be given to the assesses’ immediate supervisors/managers to be filled out and mailed directly back to the researcher. Three of these measures were the same three 7-point BARS (for leadership, decision-making skills, and communication skills) used by assessors, with the behavioral anchors modified so that they were tied less to the assessment center setting. In addition, the supervisors also completed two other measures of performance: an overall evaluation of the employee and an assessment of their future potential (both scaled from 1 - 7). Because single-item measures of performance are typically psychometrically inadequate, these five ratings were combined to create an overall index of job performance (alpha = .89). Finally, supervisors completed a 7-item scale assessing the quality of their exchanges with their subordinates (LMX-7; Graen & Uhl-Bien, 1995).

**Results and Discussion**

The intercorrelations, means, standard deviations, and reliability estimates for the assessment center ratings and occupational variables are reported in Table 1. As expected, the assessment center ratings were strongly predictive of job performance several years later. The overall SAC ratings (averaging across the three dimensions) significantly predicted supervisors’ ratings of performance in terms of decision-making skills ($r = .31$, $p < .05$), leadership skills ($r = .35$, $p < .05$), and the overall measure of performance (averaging across the three dimensions ratings and the overall performance and overall potential items; $r = .31$, $p < .05$).
The individual SAC dimension ratings of decision-making skills also predicted supervisor's ratings of leadership performance ($r = .32, p \leq .05$). The individual SAC dimension ratings of leadership skills significantly predicted the supervisors' ratings of decision-making skills ($r = .32, p \leq .05$), leadership skills ($r = .42, p \leq .01$), the overall performance item ($r = .31, p \leq .05$), and the overall average of the five items ($r = .36, p \leq .05$). The SAC communication skills dimension rating significantly predicted both the supervisors' rating of decision-making skills ($r = .32, p \leq .05$), and the overall average of the five items ($r = .30, p \leq .05$). Clearly the assessment center ratings (especially the leadership dimension and the overall average rating) showed substantial predictive validity.

In addition to the significant prediction of job performance demonstrated by the above results, there is also evidence to suggest that the ratings received in the assessment center provide incremental validity above and beyond other available predictors. That is, data was also gathered on asseesses' college GPAs (at the time they completed the assessment center) and their scores on the AACSB (American Assembly of Collegiate Schools of Business) content knowledge exam. The latter is a one-hour standardized examination developed by the AACSB that provides an indication of student knowledge in the core business disciplines of marketing, management, information systems, economics, finance, accounting, and quantitative methods. Unlike the SAC ratings, none of these other predictors were significantly correlated with job performance ($r$ ranged from -.10 to .23). In order to examine the incremental validity of the SAC ratings over these predictors, hierarchical regression analyses were computed, wherein GPA and the AACSB scores were entered simultaneously as a single block and then the assessment center ratings were entered. The $R^2$ associated with the first block of predictors was .05; for the second block (the SAC ratings), the $R^2$ was .24. This change in $R^2$ (.19) was significant ($p \leq .05$), indicating that the SAC ratings provided incremental validity beyond GRE and the AACSB exam scores in predicting job performance.

Unlike with the supervisors' ratings of performance, however, the assessment center ratings were not predictive of more objective outcomes of occupational success, such as the number of job interviews, job offers, or promotions obtained since graduation. Although the SAC dimension ratings also were not predictive of supervisors' assessments of the quality of their relationship with their subordinates (the asseesses), they were predictive of subordinates' (asseesses) ratings of the quality of their relationship with their supervisors ($r = .31, p \leq .05$). Such findings indicate that those performing better in the SAC also felt they had higher quality exchanges with their supervisors. Moreover, these asseesses who performed better in the SAC also reported higher levels of job satisfaction in their current position ($r = .31, p \leq .05$). Taken together, the results from the current study indicate that assessment center performance while a student can predict not only future ratings of job performance, but also future levels of satisfaction with both the job and the supervisor exchange relationship.

The criterion-related validity coefficients obtained in the present study approach the magnitude of the meta-analytic coefficients found in the literature (.37, corrected, Gaugler et al., 1987; .43, Hunter & Hunter, 1984; and .41, Schmitt, Gooding, Noe, & Kirsch, 1984). As such, the present study provides further support for the predictive validity of assessment centers. These findings are particularly good news for this specific assessment center, given that there were many factors in the present study that should, if anything, lead to a conservative estimate of the criterion-related validity.

The first factor that could possibly mitigate the criterion-related validity of the present assessment center is the fact that the asseesses were students. Although it has been documented that assessment centers are valid regardless of educational level of asseesses (Huck, 1973) and prior assessment center experience (Struth, Frank, & Amato, 1980), there was a great deal of variability in terms of how seriously the students approached the assessment center experience. Although some of them took it very seriously, other students clearly appeared less motivated to perform well. Such issues are presumably not a concern in most real world assessment center situations.

The nature of the jobs and the organizations included in the present study also may serve to make the obtained criterion-related validity an underestimate of the true validity possible with such an assessment center. That is, although assessment centers have been found to be predictive across many different types of jobs (Gaugler et al., 1987), any one assessment center validity study (with the exception of meta-analyses) typically only assesses the validity of the assessment center for predicting performance in a single job in a single organization. Moreover, it has been stressed that assessment center exercises and dimensions should be tied to the job via job analysis in order to be valid (Byham, 1978). Neither of these conditions existed in the present study, and yet the assessment center significantly predicted performance. Basing rating dimensions on job
analysis and having asseeses from a single organization would likely enhance predictive validity.

The final factor that could lead to a conservative estimate of the criterion-related validity of this assessment center is that, by design, the present study could capitalize on neither direct nor indirect criterion contamination, which are often hypothesized to be responsible for inflating assessment center criterion-related validity. Direct criterion contamination is believed to occur when assessment center results are used for making promotion or advancement decisions (i.e., the criteria). Norton (1977) has observed that not many organizations are willing to follow a purely predictive validity model by administering an assessment center to prospective managers and locking up the results. Indeed, assessment center validity studies without criterion contamination are relatively rare (Sackett, 1982 found two; McEvoy, 1986 found 10). The current study, however, was completely predictive and uncontaminated in design, eliminating this as a potential alternative explanation for the validity of the assessment center.

A more subtle form of contamination is what has been referred to as indirect criterion contamination. This is the belief that the criterion-related validity of assessment centers is a result of the shared prototypes and stereotypes of a good manager that both assessors and managers eventually providing criterion data bring to the rating task (McEvoy & Beatty, 1989; Turnage & Muchinsky, 1984). Many believe this subtle criterion contamination hypothesis cannot be ruled out as a significant contributor to assessment center validity (Klimoski & Brackner, 1987). However, the present study, with its inclusion of employees across several organizations and jobs (many of them nonmanagerial) and the fact that the assessors were not from the same organization as the asseeses, can rule out this indirect criterion contamination hypothesis as a significant source of inflation to the validity of the assessment center.

Similarly, there have often been criticisms of overreliance on advancement indices in assessment center validity studies, with the resulting coefficients believed to reflect a special case of indirect criterion contamination, wherein assessment center ratings are merely a proxy measure for organizational decision making (Klimoski & Strickland, 1977). Such criticisms have lead to a call for an examination of performance as criteria, in addition to just advancement and potential, often with disappointing results. That is, assessment centers frequently are good at predicting promotions, but less good at predicting performance (e.g., Cohen, Moses, & Byham, 1977; McEvoy & Beatty, 1989; Turnage & Muchinsky, 1984). In the present study, however, performance, not advancement, was measured; and the assessment center ratings were shown to be effective at predicting this performance.

In addition to ruling out direct and indirect criterion contamination as possible alternative explanations for the criterion-related validity of assessment centers, the present study also sheds some light on the sign versus sample (Wernimont & Campbell, 1968) debate surrounding the assessment center method. Specifically, the criterion-related validity coefficients obtained in the current study could be interpreted as strongly supportive of the “sign” position. That is, dimension ratings in an educational assessment center predicted job performance four years later across several different types of jobs and several organizations. It is not reasonable to assume that these exercises met the requirements for being considered work samples of the jobs now held by these asseeses. Therefore, the basis behind the criterion-related validity of the present assessment center must be an accurate evaluation of asseeses on dimensions believed to be important for job performance (i.e., a sign), rather than merely a work-sample test.

In summary, assessment center methods have been widely accepted for both management selection and development/training (Riggio & Mayes, 1997). In the context of higher education, a student outcome assessment center offers valuable data on student learning, curriculum effectiveness, and work readiness of students. As greater attention is given to outcome assessment in higher education, and to the issue of school-to-work transition, assessment center methodology will continue to increase in popularity because it allows for detailed evaluation of the relevant “hands-on” skills necessary for many professions. We are extremely encouraged by the results of CSUF’s pilot project, and believe that it can serve as a model for outcome assessment in higher education.

References


Table 1
Descriptive Statistics and Intercorrelations for AC and Occupational Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Alpha</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. AC-DMS Rating</td>
<td>4.4</td>
<td>1.1</td>
<td>.89</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. AC-LS Rating</td>
<td>4.3</td>
<td>1.1</td>
<td>.75</td>
<td>.92**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. AC-CS Rating</td>
<td>4.6</td>
<td>1.0</td>
<td>.83</td>
<td>.91**</td>
<td>.93**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. AC-Overall Rating</td>
<td>4.4</td>
<td>1.0</td>
<td>.82</td>
<td>.97**</td>
<td>.98**</td>
<td>.97**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Supervisor DMS</td>
<td>5.6</td>
<td>1.0</td>
<td>--</td>
<td>.26</td>
<td>.32*</td>
<td>.32*</td>
<td>.31*</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Supervisor LS</td>
<td>5.1</td>
<td>1.4</td>
<td>--</td>
<td>.32*</td>
<td>.42**</td>
<td>.29</td>
<td>.35*</td>
<td>.58**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Supervisor CS</td>
<td>5.4</td>
<td>1.3</td>
<td>--</td>
<td>-.01</td>
<td>.11</td>
<td>.11</td>
<td>.07</td>
<td>.60**</td>
<td>.59**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Supervisor Performance</td>
<td>5.9</td>
<td>0.9</td>
<td>--</td>
<td>.21</td>
<td>.31*</td>
<td>.28</td>
<td>.27</td>
<td>.72**</td>
<td>.68**</td>
<td>.67**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Supervisor Overall</td>
<td>5.7</td>
<td>1.0</td>
<td>.89</td>
<td>.25</td>
<td>.36*</td>
<td>.30*</td>
<td>.31*</td>
<td>.82**</td>
<td>.81**</td>
<td>.85**</td>
<td>.90**</td>
<td>.77**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Leader LMX</td>
<td>4.1</td>
<td>0.5</td>
<td>.79</td>
<td>.02</td>
<td>.01</td>
<td>.12</td>
<td>.02</td>
<td>.52**</td>
<td>.36*</td>
<td>.55**</td>
<td>.63**</td>
<td>.45**</td>
<td>.63**</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. # Interviews</td>
<td>3.7</td>
<td>4.9</td>
<td>--</td>
<td>-.01</td>
<td>.03</td>
<td>.05</td>
<td>.03</td>
<td>-.07</td>
<td>-.12</td>
<td>-.27</td>
<td>-.37*</td>
<td>-.33*</td>
<td>-.37</td>
<td>-.04</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. # Job Offers</td>
<td>1.6</td>
<td>1.6</td>
<td>--</td>
<td>.28</td>
<td>.22</td>
<td>.27</td>
<td>.26</td>
<td>.14</td>
<td>.04</td>
<td>-.21</td>
<td>-.24</td>
<td>.11</td>
<td>-.24</td>
<td>-.03</td>
<td>.46**</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. # Promotions</td>
<td>1.5</td>
<td>3.0</td>
<td>--</td>
<td>.11</td>
<td>-.05</td>
<td>-.02</td>
<td>.01</td>
<td>-.04</td>
<td>.10</td>
<td>-.14</td>
<td>-.07</td>
<td>-.06</td>
<td>-.07</td>
<td>-.12</td>
<td>-.02</td>
<td>-.08</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>15. Member LMX</td>
<td>3.9</td>
<td>7.7</td>
<td>.88</td>
<td>.26</td>
<td>.31*</td>
<td>.32*</td>
<td>.31*</td>
<td>.33*</td>
<td>.16</td>
<td>.24</td>
<td>.31*</td>
<td>.36*</td>
<td>.33*</td>
<td>.41**</td>
<td>.06</td>
<td>.20</td>
<td>-.10</td>
<td>--</td>
</tr>
<tr>
<td>16. Job Satisfaction</td>
<td>75.8</td>
<td>12.1</td>
<td>.88</td>
<td>.28*</td>
<td>.31*</td>
<td>.33*</td>
<td>.31*</td>
<td>.31*</td>
<td>.29*</td>
<td>.21</td>
<td>.32*</td>
<td>.36*</td>
<td>.32*</td>
<td>.14</td>
<td>-.15</td>
<td>.10</td>
<td>-.05</td>
<td>.53**</td>
</tr>
</tbody>
</table>

Note: Ns range from 40 to 75; DMS = decision-making skills; LS = leadership skills; CS = communication skills.

*p ≤ .05; **p ≤ .01