The Dynamics and Validity of the Group Selection Interview

TIMOTHY TRAN
MELINDA C. BLACKMAN

Department of Psychology
California State University, Fullerton

ABSTRACT. The authors examined the validity of the group interview, in which several applicants are interviewed simultaneously, vs. the traditional 1-on-1 selection interview. The authors also investigated perceptions of fairness for each interview format and whether the order in which the applicant was asked to respond in the group interview affected the quality of his or her response. Participants were 91 undergraduates who participated in 1-on-1 or group interviews for a scholarship award. The present results supported the hypothesis that interviewers in the 1-on-1 format were significantly better at predicting the applicants’ academic potential. Also, in the group interviews, the order in which the applicants were asked to respond affected the quality of their responses significantly.

Key words: group selection interview, multiple-applicant interview, one-on-one interviews versus group interviews

THE INTERVIEW PROCESS—during it, every organization should make an accurate assessment of each candidate’s potential. Certainly there are long-ranging financial ramifications for organizations that make inaccurate and poor personnel selection decisions (Cascio, 1998; Ramiall, 2004). The ultimate outcome of these poor decisions typically is employee turnover (Cascio; Hellman, 2000). The tangible costs of employee turnover (e.g., those of recruitment, interviewing, lost productivity, replacement costs, and training) equal about 150% of an employee’s annual salary (Hellman; Ramiall). Organizations in the United States have an annual turnover rate of about 12% (Pinkovitz, Moskal, & Green, 1997). To put these costs into perspective, consider an organization with 1,000 employees in which the annual salary of each employee is $50,000 and with a 10% turnover rate. Based on these numbers, the annual cost of turnover for this organ-
ization would be $7.5 million (Cascio). These astounding rates and costs have heightened the need for researchers and organizations alike to more closely examine personnel selection techniques (Cascio; Hellman; Pinkovitz et al.).

Group interviews are growing dramatically in popularity because professionals see them as a very versatile means for meeting several objectives (Fischbach, 2004; Jacques & Hasselkus, 2004; Kyriadides & Kelly, 2004). In addition to using the group interview as a focus group or feedback session, many organizations (e.g., Disney, Abercrombie & Fitch, Starbucks, The Gap, and Banana Republic) choose to use it as a selection tool, in which one interviewer asks questions of two or more applicants simultaneously (Fischbach; Kyriadides & Kelly). However, it is disturbing that there is little or no literature on using the popular group interview as a selection tool (Dipboye, Gaugler, Hayes, & Parker, 2001). The majority of available research covers the validity of the panel interview format, in which one candidate is interviewed by multiple interviewers (Huffcutt & Woehr, 1999; McDaniel, Whetzel, Schmidt, & Maurer, 1994). Because so many organizations have recently begun implementing the group interview, it would be prudent—both financially and otherwise—for management in organizations to better understand the validity and other inherent components of the group selection interview. In the present study, our aim was to secure such understanding.

The objectives of the group selection interview regarding screening can range from selecting personnel, promoting personnel, assessing team skills and the individual’s ability to work with others to admitting potential students into a university and awarding scholarships (Campion, Palmer, & Campion, 1997; Ulrich & Trumbo, 1965). More specifically, an interviewer might be particularly interested in seeing how well a job applicant manages stress by surrounding the applicant with competing applicants in a group interview. Other employers may use the group interview to assess and compare job candidates and to find the most assertive or dominant one. Two additional objectives of employers in using the group interview are standardization and cost efficiency. The group interview technique would be cost efficient for those organizations that need to conduct mass hiring within a short period of time. The technique eliminates most of the expense, time, and training of employing multiple interviewers. Also, the group selection interview still allows the interviewer to have personal contact with the applicants (Dipboye et al., 2001). Given these inherent positive aspects of the group interview, we feel confident that this interview technique is probably here to stay. An ongoing investigation of this selection method is necessary because recent trends have indicated that organizations have moved toward more cost-efficient but not necessarily valid interview practices (Blackman, 2002a, 2002b; Campion et al., 1997; Campion, Pursell, & Brown, 1988; Dipboye et al., 2001; Huffcutt & Woehr; McDaniel et al., 1994; Moscoso, 2000; Parsons, Liden, & Bauer, 2001).

The principal investigators of the present study examined through an experimental design three aspects of the group selection interview that a literature
review indicated were potentially problematic (Blackman, 2002a, 2002b; Cambion et al., 1997). The first factor that we investigated was the validity and utility of the interviewer’s selection decision.

**Hypothesis 1:** Interviewers who implement the group selection interview method will be compromising the validity of their decisions because of several of the negative inherent aspects of the method.

One negative inherent aspect of the group interview format is the multitasking (attending to and prioritizing multiple stimuli simultaneously) that the interviewer must do (e.g., remembering each applicant’s response, name, and experience level). Researchers have consistently shown that multitasking leads to performance degradation (Latorella, 1997). Researchers hypothesize that the associated information overload and multitasking that the interviewers must do will lead to performance degradation and thus will compromise the quality and the validity of the interviewer’s selection decision (Bowers et al., 2000; Latorella).

Another issue that might affect the general validity of the group selection interview is the lack of accurate information obtained from the applicant. Researchers have documented that stress during an interview can lead to a job applicant’s lowered perceptions of self-efficacy and lower overall performance (Hotard, 1995; Knudstrup, Segrest, & Hurley, 2003). Because of the lack of privacy of the group interview with so many other applicants in the room, applicants may feel apprehensive and self-conscious, and thus they may self-disclose less and ultimately perform significantly more poorly (Knudstrup et al.; Purves & Erwin, 2004; Ruggiero et al., 2004). This lack of information from the applicant ultimately gives the interviewer a poorer quality and smaller quantity of information for making a selection decision. Researchers have shown that interviewers who receive poor quality or incomplete verbal and nonverbal information from the job applicant during the interview process produce inaccurate judgments (Blackman, 2002a, 2002b; Blackman & Funder, 1998; Funder, 1995). In the present study, we believed that the multitasking demands placed on the interviewer and the likelihood of poor quality responses from the applicants have the potential to compromise the validity of the selection decision made by interviewers who implement a group selection interview.

The second aspect of the group interview format that we investigated in the present study was the actual quality of the applicants’ responses and how these responses are affected by the order in which the applicants are asked to respond. We hypothesized the following, on the basis of research in other comparable domains:

**Hypothesis 2:** The order in which the applicant in the group interview format is asked to respond to the question will affect the quality and originality of each response (Friedrich & Smith, 1998; Haradlsson & Gissurarson, 1985; Kerstholt & Jackson, 1998; Newman & Fuqua, 1986).
Although research on this particular topic in the domain of selection interviews is limited, in the present research we believed that the applicant who responds last to the questions will give significantly more thorough, innovative, and original responses than will the other applicants. We believe that the last-response applicant will have had more time to think about the questions before responding and will be able to build off of the other applicants’ responses and the interviewer’s reactions. If the present findings support this hypothesis, its implications will be long ranging. This circumstance may cause the interviewer to inaccurately perceive the last-response applicant as the most qualified candidate and to offer him or her a job. Likewise, job applicants who sense that the job applicant who is last to answer the questions has a definite advantage may perceive the interview process as unfair. This perception of unfairness may in some way ultimately affect the quality of the applicant’s responses as well. If the present findings support our hypothesis, that the quality of an applicant’s response is affected by the order in which he or she is asked to respond, then interviewers who typically conduct group interview formats will need to be aware of this bias in the interview process and will need to try to remedy it. Interviewers could possibly remedy this bias by systematically rotating which applicant in the group interview process is asked to respond to the question first or last, using a corrective procedure that we incorporated into the present study’s method.

The last aspect that we investigated was the individuals’ perceptions of the fairness of the interview process. (See Table 1.) It is important to keep in mind that a selection interview is viewed as an interactive process through which organizations and individuals assess and select one another (Connerley & Rynes, 1997). Organizations typically use the interview as an opportunity to both recruit (sell the company or academic institution) and select (evaluate and screen) candidates (Connerley & Rynes). It is interesting that there has been a vast amount of research on the reliability and utility of selection procedures but relatively little empirical research on the applicants’ perceptions of such procedures until recently in the last decade (De-Jong & Visser, 2000; Gilliland, 1993; Kohn & Dipboye, 1998; Powell, 1991; Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993).

Applicants’ reactions to selection procedures, for example, in the domain of personnel selection, should be of concern for at least three reasons. First, Connerley and Rynes (1997) have well established that applicants form strong impressions of organizational recruiters. It is more important that researchers have shown that these impressions often affect how applicants approach subsequent steps of the recruitment process, including their eventual job choices (Connerley & Rynes; Rynes, Bretz, & Gerhart, 1991).

Second, applicants’ perception of fairness in the selection procedure may have both ethical and legal implications for the organization. Evidence indicates that selection procedures can also be viewed within a moral and ethical context.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Index</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfairness index</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Lack of scientific research</td>
<td>.37*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Not a logical approach</td>
<td>.65**</td>
<td>.54**</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Will not detect differences</td>
<td>.54**</td>
<td>.64**</td>
<td>.64**</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Impersonal and cold method</td>
<td>.58**</td>
<td>.15</td>
<td>.45**</td>
<td>.24</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. No right to obtain information in this manner</td>
<td>.50**</td>
<td>.02</td>
<td>.29</td>
<td>.10</td>
<td>.52**</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Invades personal privacy</td>
<td>.32*</td>
<td>.07</td>
<td>.21</td>
<td>.23</td>
<td>.40**</td>
<td>.43**</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>9. Not a widely used method</td>
<td>.62**</td>
<td>.51**</td>
<td>.68**</td>
<td>.46**</td>
<td>.25</td>
<td>.28</td>
<td>.38*</td>
<td>—</td>
</tr>
</tbody>
</table>

*Note. N = 40.*

*p = .05 (two-tailed). **p = .01 (two-tailed).*
Applicants are more likely to view those selection procedures that they perceive as lacking validity or as otherwise offensive or intrusive as unfair, unethical, or immoral, resulting in more complaints and court challenges (Smither et al., 1993).

Third, applicants’ perceptions of unfairness may indirectly affect both the validity and utility of a selection procedure. Perceived unfairness or low face validity may lower motivation to perform well during selection interviews or other procedures. This lower motivation in turn could result in biased or inaccurate scores and reduce the operational validity of the procedure (Smither et al., 1993). In short, selection procedures that decrease organizational attractiveness may also result in significant losses in utility as a result of qualified applicants’ either refusing job offers or admission into a university or being so “turned off” by the selection process that they withdraw before an offer can be made (Smither et al.).

When standing back and looking at the group selection interview from an applicant’s point of view, one may see many negative aspects, such as the perceived lack of individual attention from the interviewer, the lack of privacy within the interview, and the asking of the same question to each applicant possibly in the same order.

For the aforementioned reasons, we hypothesize the following:

Hypothesis 3: Individual applicants will perceive the group selection interview as negative and unfair in comparison to the one-on-one interview.

In the present study, we set out to test these potentially problematic aspects of the group selection interview. We compared the validity of the group interview format versus the traditional one-on-one interview format by determining which format best predicts the academic potential of scholarship candidates. Because very little research exists on the topic of group selection interviews, we felt the need to initially conduct an experimental investigation in a very controlled setting. We chose an academic scholarship as the screening objective to adequately motivate the applicants in this experimental setting. We also felt that all participants could knowledgeably partake in this particular type of selection process—for a scholarship award. We used the participants’ grade point average (GPA) and score on a test of mental ability as predictors of their academic potential.

Method

Participants

Participants were 91 undergraduate students (27 men and 64 women) from a large 4-year university. The age range was 18–44 years old (M age = 21.5 years old, SD = 3.41 years old). We recruited participants with a minimum cumulative GPA of 2.5 from two upper-division undergraduate psychology courses and gave each participant course credit for participating in the study.
Procedure

On the participants’ arrival, research assistants (RAs) confirmed that each participant was not an acquaintance of the interviewer or of any other participant (in both interview formats). RAs verbally solicited consent to participate and obtained a signed consent form giving the researcher access to the participant’s cumulative GPA from the University Registrar’s Office.

Next, we asked the participants to complete the Wonderlic Personnel Test (WPT; Wonderlic, 1966/2002) of mental ability. The WPT is a 50-item inventory for schools and businesses to measure general cognitive ability and the potential to learn in new environments. The administration time of the WPT is exactly 12 min, and one can administer it individually or in groups. The Cronbach’s alpha reliability of the WPT is generally .94 (Dodrill, 1983), and for the present sample it was .69. We acknowledge that this level is lower than that obtained by Dodrill, but we still feel justified in using this measure for the following reasons. The WPT has existed since 1937, and researchers have administered it to over 130 million people in both educational and workplace settings. Research results have extensively supported the validity of this test over the past 68 years (Dodrill & Warner, 1988). After conducting a thorough search of the literature, we could not find a measure that better fit our study’s goal of assessing the academic potential of participants within an educational or business setting than the WPT. In addition, after looking back at our results, we feel that the low alpha level was probably due to two outliers in our sample. For the aforementioned reasons, we feel very comfortable in choosing to leave the WPT in the present study as one of two measures of academic potential, with the other measure being GPA. Researchers have also shown the WPT to correlate strongly with intelligence, and it is a reliable measure of mental ability, whereas we chose GPA as a measure of academic potential because of its high validity for predicting future academic performance (Dodrill & Warner; Kulatunga, Chan, & Geoffrey, 2002).

After completing the WPT, we informed the participants that they would be participating in an interview for a cash scholarship award that one applicant would be chosen to receive. Each participant was provided with a detailed description of the scholarship award to review before the interview. The description stated that the award would recognize an individual who has demonstrated excellence in academic work, leadership, and community service. To test Hypothesis 1 and determine the predictability of the group interview versus the one-on-one interview, we then randomly assigned each participant to either a one-on-one interview format or the group format with two other candidates. In each interview format, the interviewers asked the same questions of each candidate with no follow-up questions. We based the interview questions for the present study on questions administered to the candidates of the President’s Scholars Program at the university. The program and an industrial and organizational psychologist had already carefully structured, standardized, and tested the predictive validity.
of the nine interview questions. The program and psychologist had designed the questions to tap into the facets of academic excellence, leadership, and community service.

We timed the one-on-one interviews to last approximately 10 min and the group interviews to last approximately 30 min. These time limits ensured that each participant was given approximately the same amount of time to respond during the interview. We counterbalanced the questions in the group interview format to control the effect of applicant interview position. Counterbalancing allowed each applicant to equally respond in all three positions: first, second, and third. We videotaped all interviews for later assessment and coding of observable behavior, including the quality of the applicant’s response.

After we debriefed the participants and thanked them for their participation, we provided the RAs or interviewers with a rating form on which they rated each applicant’s response to each question. The interviewers also made an overall rating (not ranking) of each applicant’s academic potential on a 10-point Likert scale, the criterion for the scholarship award.

We trained six female undergraduate RAs as the interviewers. The interviewers’ ages in years ranged from the early 20s to the late 20s. We required that the RAs be upper-division psychology majors with a high GPA in their major. We set these specifications to ensure the data collection was of high quality and that the interviewers were qualified to assess the criterion, the applicant’s academic potential.

All RAs were blind to the hypotheses. RAs (interviewers) received 1 week of training on how to conduct both the structured one-on-one interview and the structured group interview (three applicants being interviewed simultaneously by one interviewer). Specific examples and illustrations that were previously given during research-assistant training sessions anchored each rating scale.

The RAs conducted a minimum of three practice interviews for each interview condition, although they did not assess the target’s academic potential in these practice sessions. In addition, each RA watched a minimum of two mock videotaped interviews for the two interview formats. We then asked her to assess the target’s academic potential. We required each RA to achieve in their practice interviews an interrater agreement correlation Cohen’s kappa on their assessment of the target’s academic potential of at least .50 with two of the other RAs and at least .70 with the remaining three RAs. If there was a disagreement between the raters, we used the following policy. For every pair of raters whose interrater agreement correlation fell below the minimum, we asked both raters to independently watch the videotape again and reassess the target’s academic potential until they achieved the specified levels of agreement with one another and with the remaining raters. After the acceptable agreement level was achieved, the principal investigator brought the raters together to discuss their ratings and the reasons for the original disagreement. This meeting served as a learning session. The actual average interrater agreement correlation between all of the raters across all of
the training videos was .72. To ensure a realistic interview environment, we required the interviewers to dress in appropriate professional attire for each interview that they conducted. In addition, we required the interviewer to maintain eye contact with only the individual that she was currently speaking to. This technique controlled the interviewer’s time of attention to each candidate during the group interviews.

Video Coding

Our last measure was on a 5-point Likert scale that was made by the video coders for the group interview format regarding the quality of the applicant’s response. We used this coding to determine whether the order in which the questions were asked affected the quality and completeness of the applicants’ responses (see Hypothesis 2). We chose three additional RAs to receive training on how to assess and code the quality of each applicant’s response in the group interview format. These video coders watched a minimum of two mock videotaped group interviews and practiced coding the quality of each applicant’s response, to ensure the reliability and validity of those video coders’ judgments. The video coders independently viewed all of the applicants’ interviews. We asked them to assess the completeness of the applicants’ responses, to assess the innovativeness of their responses, and to determine whether their responses were verbally borrowed from another applicant. The video coders rated all responses on the scale. We aggregated the three raters’ ratings for each of the three categories, because the Spearman rank correlation coefficient for the three independent raters was significant at the .01 level (two-tailed). See Table 2. The video coders also watched the videotapes to ensure that the applicants did not in any way reveal their GPAs or give information related to their GPAs (e.g., information regarding dean’s list, academic probation).

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 2</td>
<td>.354**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td>.392**</td>
<td>.459**</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>4.69</td>
<td>4.52</td>
<td>4.24</td>
</tr>
<tr>
<td>SD</td>
<td>0.51</td>
<td>0.70</td>
<td>1.14</td>
</tr>
</tbody>
</table>

** $p = .01$ (two-tailed).
Perceptions of Fairness

To test Hypothesis 3 and determine whether participants in the group interview format perceived it as being less fair than those in the one-on-one format, we adapted questions from Steiner and Gilliland’s (1996) fairness perceptions of employment practices. We did not ask the original participants in the present study about their perceptions of the interview process because we believed that they would give socially desirable responses because a scholarship was at stake. Instead, 78 additional undergraduate students received course credit for completing the fairness survey. In a procedure similar to that in Steiner and Gilliland’s study, we told about one half of the participants to imagine that they were competing for a cash scholarship and would be interviewed by a single interviewer. At another time, we told the remaining participants that they would be participating in a group interview for a scholarship with two other participants. Because the questions would be counterbalanced, we told each participant to imagine that he or she would have an equal opportunity to answer two questions first, two second, and two third.

The survey consisted of nine questions on a 7-point Likert-type scale with appropriate anchors (ranging from 1 = least favorable to 7 = indicated most favorable). We designed the first two questions to assess interview selection method favorability (e.g., “If you did not get the scholarship based on the group interview selection format, what would you think of the fairness of this procedure?”). The Cronbach’s coefficient alpha for this scale in general was .73.

The remaining seven questions measured the bases for these reactions on seven procedural dimensions. More specifically, they measured the satisfaction and violation of these rules, which provide the basis for fairness reactions (e.g., “The group interview will detect the individuals’ important qualities differentiating them from others”). Of the participants, 40 answered questions that we targeted toward their perceptions of the group interview, whereas the remaining 38 participants answered questions that we targeted toward their perceptions of the one-on-one interview.

Testing of Hypotheses

We tested Hypothesis 1, in which we proposed that the one-on-one interview format will yield more valid judgments of the applicant’s academic potential than would the group interview format, by using average agreement correlations. More specifically, we compared the average agreement correlation between the interviewers’ ratings of the applicants’ academic potential and that of the applicants’ GPA and Wonderlic score to the agreement correlations in the one-on-one format. We hoped to find that the average agreement score in the one-on-one format was significantly higher than those agreement scores in the group format, thus supporting Hypothesis 1.
Next, we tested Hypothesis 2, by which we examined the effect that interview question order has on the quality of the applicant’s response, by using ratings made by coders who watched the videotapes of the group interviews. Optimally, the results would support Hypothesis 2 if the coders’ ratings of the quality of the applicants’ responses were systematically affected by the order in which applicants were asked to respond in the group interview. Last, we tested Hypothesis 3, by which we examined the applicants’ perceptions of the fairness of the one-on-one interview format versus the group interview format, by using ratings on a fairness scale from an additional group of participants who were not directly involved in the original portion of the present study to minimize social-desirability effects. The results would support Hypothesis 3 if the participants rated the one-on-one interview format as significantly more fair than they did the group interview format.

**Results**

*The Validity of the Interview Formats*

By the first set of analyses, we focused on the hypothesis that the one-on-one interview format would be significantly more valid in predicting the applicants’ academic potential than would the group interview format (see Hypothesis 1). We computed a Pearson product–moment correlation between the interviewer’s rating of each applicant’s academic potential and each applicant’s reported GPA in the one-on-one interview and then again in the group interview format. This process was repeated again, this time using each applicant’s score on the WPT instead of his or her GPA.

The Pearson product–moment correlation for the interviewer’s rating of academic potential with GPA in the one-on-one interview format was statistically significant at \( r = .82, p < .001, \) two-tailed, \( n = 40. \) As we hypothesized, the correlation between the interviewer’s rating of academic potential with GPA in the group interview format was not significant at \( r = .20, p = .091, \) two-tailed, \( n = 51. \) We then calculated a Fisher’s \( r \)-to-\( z \) transformation on the two correlations and indeed found a significant difference, \( z = 4.36, p < .001, \) one-tailed. A similar trend presented itself when the WPT was used instead of the applicant’s GPA, \( r = .40, p = .009, \) for the one-on-one format; \( r = .14, p = .222, \) for the group format. An \( r \)-to-\( z \) transformation of these two correlations revealed \( z = 1.29, p = .09. \) In general, the results indicated the possibility that the one-on-one interview had significantly better predictive validity with regard to the interviewer’s assessment of academic potential as opposed to the group interview format, thus supporting Hypothesis 1.

*The Effect of Response Order*

We directed the next set of analyses toward testing the hypothesis that the quality and completeness of the applicants’ responses would be significantly
affected by the order in which we asked applicants to respond to them in the group interview (see Hypothesis 2). We calculated analyses of variance (ANOVAs) to evaluate this relationship. The independent variable, the order of the applicants’ responses to the interview question, included three levels: first, second, and third levels. The dependent variable was the category of ratings: applicant’s completeness of response, innovativeness of response, or verbally borrowed responses. We aggregated the three raters’ ratings for each of the three categories because the Spearman rho correlation for the three independent raters was significant at the .01 level (two-tailed). See Table 2.

To help determine whether there was an order effect, we calculated for each of the six different interview questions or categories—each requiring three different ratings on completeness of response, innovativeness of response, and verbally borrowed response—18 one-way ANOVAs. The results of Questions 1, 4, 5, and 6 were significant for verbally borrowed responses. The ANOVAs revealed that there was a significant order effect for verbally borrowed responses for these questions, $F(2, 141) = 6.16, p = .003$; $F(2, 141) = 10.40, p < .001$; $F(2, 141) = 4.56, p = .012$; $F(2, 141) = 11.79, p < .001$, respectively, two-tailed. The effect sizes of these three analyses were .21, .26, .18, and .28, respectively. Also, the results of Question 5 indicated a significant order effect for the completeness of response at $F(2, 141) = 3.63, p = .029$, two-tailed, with an effect size of .16.

We conducted follow-up multiple comparison tests using Tukey’s honestly significant difference test because we assumed the variances between dependent variables to be equal. Tukey’s test provided the difference between each pair of means, standard deviations, yielding a matrix where asterisks indicated significantly different group means at an alpha level of .05, as Table 3 and Table 4 indicate.

For Questions 1, 4, 5, and 6, there were significant differences in the interviewers’ mean ratings of the applicants’ responses for the variable of verbally borrowed response, revealing an order effect. Also, we found a similar order effect for Question 5 for the variable of completeness of response. On the basis of these results, we inferred that the applicants who responded to interview questions in the second and third positions provided more complete and verbally borrowed responses than did applicants who responded to questions in earlier positions. These findings supported the hypothesis that in the group interview setting, the order in which the interviewer asks applicants to respond can significantly affect the quality and completeness of their responses (see Hypothesis 2).

Perceptions of Fairness

We executed the last set of analyses to test the hypothesis that participants in the group interview would have significantly more negative perceptions of the interview process than would those participants in the one-on-one interview (see Hypothesis 3). With regard to perceptions of fairness, we calculated a process favorability or fairness index by computing the mean response for each partici-
### TABLE 3. Differences Among Interview Positions in Verbally Borrowed Responses

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Position</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>First</td>
<td>Second</td>
<td>Third</td>
<td></td>
</tr>
<tr>
<td>Question 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td>First</td>
<td>4.94</td>
<td>0.24</td>
<td>0.33*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second</td>
<td>4.60</td>
<td>0.64</td>
<td>–0.33*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Third</td>
<td>4.46</td>
<td>0.97</td>
<td>–0.48**</td>
</tr>
<tr>
<td>Question 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td>First</td>
<td>5.00</td>
<td>0.00</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second</td>
<td>4.71</td>
<td>0.68</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Third</td>
<td>4.40</td>
<td>0.89</td>
<td>–0.60**</td>
</tr>
<tr>
<td>Question 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td>First</td>
<td>4.75</td>
<td>0.76</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second</td>
<td>4.85</td>
<td>0.50</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Third</td>
<td>4.38</td>
<td>1.08</td>
<td>ns</td>
</tr>
<tr>
<td>Question 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td>First</td>
<td>4.89</td>
<td>0.61</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second</td>
<td>4.67</td>
<td>0.93</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Third</td>
<td>3.89</td>
<td>1.39</td>
<td>–1.00**</td>
</tr>
</tbody>
</table>

Note. Difference was determined by Tukey’s honestly significant difference statistic. ns = nonsignificant differences between pairs of means.

* p = .05 (two-tailed), ** p = .01 (two-tailed).
In the present study, we examined the one-on-one interview and the group interview to determine the validity of the group interview format as a selection tool for predicting the academic potential of scholarship candidates. We found that interview format did affect the interviewer’s ability to select applicants with the most academic potential for the scholarship award. Specifically, the present findings support the hypothesis that the one-on-one interview format would produce significantly more accurate predictions of a candidate’s academic potential than would the group interview format (see Hypothesis 1).

The one-on-one interview format appeared to be more valid as a selection tool than did the group interview format because the former possibly placed less mental demands on the interviewer (i.e., remembering and rating each of the three applicants’ responses simultaneously; Bowers et al., 2000). On the other hand, the present authors believe that the one-on-one interview format afforded the interviewer with the opportunity to individually focus on and assess each applicant, which increased the validity of the interviewer’s ratings (Dipboye et al., 2001). In terms of the group interview format, the validity may have been less than that of the one-on-one interview format because of the greater demand on the interviewer to multitask during the interview (Latorella, 1997). The additional cognitive effort that multitasking behaviors require may have reduced the interviewer’s ability to accurately assess each applicant simultaneously.

Another possible reason for the present findings indicating that the interviewers in the group interview format produced less-valid academic predictions about the candidates is that the group format may have increased the applicants’
anxiety, lessened their openness, and consequently diminished the quality of their responses (Hotard, 1995; Purves & Erwin, 2004; Ruggiero et al., 2004). An applicant in a group interview is less likely to self-disclose information than one in the privacy of a one-on-one interview. Consequently, less accurate information may be available to the interviewer in evaluating the applicant (Blackman & Funder, 1998; Funder, 1995).

The present study’s Hypothesis 2 involved the order in which the participant was asked to respond to the interview question and how it affected the quality of his or her response. The results again supported our hypothesis that applicants who had to respond in the second and third positions gave more complete responses and verbally borrowed more responses than did the applicants who had to respond in earlier positions, thus yielding an order effect (see the present Hypothesis 2; Friedrich & Smith, 1998; Kerstholt & Jackson, 1998).

Verbally borrowed responses were more apparent with interview Questions 1, 4, 5, and 6. Complete responses were more apparent with Question 5. In general, we designed the interview questions to evaluate the applicant’s academic knowledge, skills, and abilities so as to assess his or her academic potential. The questions that yielded significant findings in the present study included questions that assessed the applicant’s successful transitioning into college life, academically related personality traits, future goals, and knowledge of current events. The two questions that yielded nonsignificant findings assessed the applicant’s “participation in community and university services” and his or her “leadership qualities.”

After considering the present findings, we find it plausible that the questions that produced verbally borrowed responses afforded applicants with universal responses, whereas the questions concerning their community service behavior and leadership qualities required more idiosyncratic or personally tailored responses. For example, many applicants responded to Question 5, the question regarding their future goals, with common responses indicating that they wanted to graduate or that they wanted to start their career. In contrast, questions on community service participation allowed only idiosyncratic responses from participants’ personal experiences and actual participation in community service.

Likewise, researchers can also interpret the present findings regarding the completeness of responses to Question 5 in terms of more idiosyncrasy in the responses. Question 5 assessed the applicant’s strongest and weakest personality traits. Idiosyncratic responses such as “independent” or “smart” could have been more difficult for the interviewer to assess for completeness, perhaps leading to the raters’ judging these responses as more complete than they were.

Last, we also solicited perceptions of fairness for each interview format to test Hypothesis 3. The present results revealed that the participants perceived the group interview format as significantly more unfair and inappropriate as a selection technique than they did the one-on-one interview, thus supporting Hypothesis 3. These findings indicate the possibility that participants have fairly strong perceptions about the fairness and suitability of each interview format. It is inter-
esting that such strong negative perceptions appeared for the group interview format although the respondents were told that the order of the interview questions would be counterbalanced during the group interview. The procedural justice correlates also enable insight into the bases of these strong negative perceptions of the group interview format. Among the prominent reasons about why the participants believed the group interview format was unfair was the interview’s lack of face validity, its inherently impersonal nature, and its tendency to invade the applicant’s personal privacy. These strong perceptions of unfairness would undoubtedly affect the candidate’s interview performance and his or her feelings toward the organization that uses the group interview as a selection tool.

Researchers should generalize the results of the present study to the preemployment interview process with caution because the present study was conducted in an academic environment with an academic criterion. Perhaps if the present study had been conducted in a professional workplace environment with a salary-paying job at stake, the motivation of the participants would have been greater. However, in the present study we took several steps to enhance its realism and generalizability. For instance, although we used mock interviews, a sizable cash scholarship adequately motivated the participants during the interview process. The training and selection of the interviewers and the carefully structured interview questions also enhanced the realism of the present study. However, a drawback of the present study might have been the age of the interviewers. They were of the same peer group and the same young age level as the participants. This circumstance might have caused the participants to not take the interviews as seriously as they would have otherwise.

After this controlled exploratory study, the next step would ideally be to test the present findings in an applied realistic setting, such as an actual workplace. Because of the large number of organizations that are turning to the group interview as a selection tool, it should not be too hard to find a test site. This type of environment would certainly provide all applicants with more than enough motivation for the interview. Two additional aspects of interview processes would be interesting in a realistic workplace environment: the quantity and quality of nonverbal behavior of the applicants in a one-on-one interview versus a group interview. These aspects are interesting because researchers have shown that accurately assessing nonverbal language during the interview process is a crucial part of making accurate judgments about the job candidate (Blackman, 2002a, 2002b). The question would then be the following: Are participants in the group format revealing less nonverbal information, thus compromising the accuracy of the interviewer’s decision? The present authors hope to test this question in a follow-up study.

In the present study, we hoped to educate interviewers with regard to selection interview practices, especially in the group interview. In general, interviewers should think twice before using the group interview format as a selection tool because it may not be as valid as the one-on-one interview format in predicting
academic potential or any other outcome potential. We believe that the group interview format requires more multitasking efforts of the interviewer, seeming to decrease the predictive validity of the format as a selection tool and degrading the interviewers’ ratings of the applicant. Also, interviewees seem to hold negative perceptions of the fairness and appropriateness of this selection technique. Interviewees may base these perceptions on the belief that those applicants answering the interview questions last have an unfair advantage. Although the group interview format is the less costly method in the short term, the present study indicates the possibility that it sacrifices accuracy and fairness for short-term efficiency.

REFERENCES


*Received October 26, 2004
Accepted July 26, 2005*