RESEARCH

Evaluation of an Interview Process for Admission Into a School of Pharmacy

Michael P. Kelsch, PharmD, and Daniel L. Friesner, PhD

College of Pharmacy, Nursing, and Allied Sciences, North Dakota State University, Fargo

Submitted August 26, 2011; accepted October 31, 2011; published March 12, 2012.

Objective. To evaluate the doctor of pharmacy (PharmD) admissions interview process at North Dakota State University (NDSU).

Methods. Faculty pairs interviewed candidates using a standardized grading rubric to evaluate qualitative parameters or attributes such as ethics, relevant life and work experience, emotional maturity, commitment to patient care, leadership, and understanding of the pharmacy profession. Total interview scores, individual attribute domain scores, and the consistency and reliability of the interviewers were assessed.

Results. The total mean interview score for the candidate pool was 17.4 of 25 points. Mean scores for individual domains ranged from 2.3 to 3.0 on a Likert-scale of 0-4. Nine of the 11 faculty pairs showed no mean differences from their interview partner in total interview scores given. Evaluations by 8 of the 11 faculty pairs produced high interrater reliability.

Conclusions. The current interview process is generally consistent and reliable; however, future improvements such as additional interviewer training and adoption of a multiple mini-interview format could be made.

Keywords: interviews, admissions, interrater reliability, pharmacy students

INTRODUCTION

The Accreditation Council for Pharmacy Education (ACPE) requires that pharmacy programs use a variety of quantitative and qualitative measures when evaluating candidates for admission. Standard 17 of the 2007 self-assessment accreditation guidelines, version 4.0, states that candidates' written and verbal communication skills are assessed in a standardized manner, interviews are structured to consistently address key admission criteria for each applicant, interviewers have appropriate credentials and are trained in successful interview strategies and techniques, evaluation of professional attitudes and behaviors is a component of the student selection process, and the college or school develops and uses admission criteria that set performance expectations for admission tests, evaluations, and interviews used in selecting students who have the potential for success in the professional degree program and in the profession.¹ However, accreditation standards do not specify how the interviews must be performed, which allows flexibility in the

application of these standards. Prior to 2007, the PharmD program at NDSU offered admission to its program primarily based on quantitative measures such as grade point average (GPA), Pharmacy College Admission Test (PCAT) scores, and a written essay in response to a contemporary pharmacy or medical issue.

Evaluation of qualitative (and in some cases, non-cognitive) parameters such as students' verbal communication skills, maturity, integrity, compassion, and leadership qualities has proven valuable in selecting candidates for admission into pharmacy programs. These traits are typically assessed by interviewing potential students.^{2,3} Noncognitive admissions parameters significantly correlate with outcome measures such as PharmD program GPA and composite portfolio scores.⁴ Also, faculty interview scores significantly correlate with first-year GPA in the PharmD program.⁵

Structured interviews are preferred over unstructured interviews because they have better reliability and validity and they are more scientifically and ethically defensible than unstructured interviews. The structured interview process typically includes a set of standardized questions given to each candidate, a standardized scoring system, and a panel of at least 2 interviewers evaluating each candidate.⁶ Additionally, structuring the interview can decrease bias, which is more common in unstructured interviews. Bias refers to leniency, severity, or favoritism

Corresponding Author: Michael P. Kelsch, PharmD, Assistant Professor, North Dakota State University, College of Pharmacy, Nursing, and Allied Sciences, Dept of Pharmacy Practice (#2660), PO Box 6050, Fargo, ND 58108-6050. Tel: 701-231-6528. Fax: 701-231-6528. E-mail: michael.kelsch@ndsu.edu

shown to candidates by an interviewer when rating them. Ongoing interviewer training is essential to elicit interview information in a consistent and fair manner, minimize rating bias, and enhance overall interviewer performance.⁷

To our knowledge, no other report has been published describing a formalized assessment of a pharmacy admissions interview process. The researchers attempted to examine the general characteristics, consistency, and interrater reliability of the interview process used for the PharmD program at NDSU. The authors posited that the process was generally reliable and valid, as demonstrated by similar mean values and significant correlations found in interview evaluation scores among interview teams in the structured admissions interview process.

METHODS

In addition to the ACPE admissions interview requirements set forth in 2007, the admissions committee desired to implement a structured mechanism to evaluate 6 qualitative (and/or noncognitive) attributes of its candidates, such as ethics, relevant life and work experience, emotional maturity, commitment to patient care, leadership, and understanding of the pharmacy profession. Suggestions for structured interview development by Latif were reviewed prior to initiating the process at NDSU.⁶

In the fall semester, prior to the inaugural interview day in 2007, the aforementioned 6 attributes were selected by the college's admissions committee as professional qualities desired in pharmacists in training. After the attribute domains were selected, admissions committee members were assigned with designing interview questions that would evaluate the candidates' abilities in the respective domains. Questions were examined and debated by committee members at scheduled meetings to maximize clarity and application to the selected domain. Upon selection of the 6 questions, one per domain, the committee developed a grading rubric for each question with a 5-point Likert scale with an interval of 0 to 4. The committee identified potential answers that could be given by the candidates during the interview and placed them in the rubric scoring categories. The interview procedure, selected domains and questions, and grading rubric were presented to the full faculty at a retreat for further debate and modification. In each of the subsequent 4 years (2008-2011), the admissions committee made minor adjustments to the interview process based on feedback from the faculty. For example, in 2011 two questions were added to the interview process. One question (which was not evaluated) was added to the beginning of the interview to serve as an icebreaker, and 1 question was added to the end of the interview that asked the interviewee if

he or she had any questions for the interviewers. Although the intent of the questions was still to assess the originally selected domains, the majority of interview questions were changed each year.

Invitations to the interview day each year were determined by evaluating (through a rank-order process) candidates' grade point average in specific "core" courses (chemistry, cell biology, anatomy and physiology) and PCAT scores (including the writing component), among other factors. After all candidates either accepted or declined the invitation to the interview, the candidate pool was assembled. Matching the interviewers to the interviewees was, in part, determined by faculty members screening the candidate list for potential bias. For example, faculty members were not allowed to interview their own advisees. Additionally, other prior relationships (organizations, personal knowledge/interaction) precluded interviewers from interviewing a candidate. Upon completion of this initial screening, each candidate was randomly assigned to an interview team (composed of 2 members) by an administrative assistant. The pool of reviewers consisted of pharmaceutical sciences and pharmacy practice faculty members, and pharmacist practitioners who were affiliated with the school through teaching or committee appointments. Faculty members were expected to participate on the interview day as part of their job duties. Each interviewing team typically consisted of 1 member from the admissions committee and 1 non-committee member. Additionally, non-practicing pharmacist faculty members were typically paired with practicing pharmacist faculty members in an attempt to increase consistency of candidate ratings.

The interview day involved each faculty team interviewing a candidate for 15 minutes, with a 5-minute break between candidates for the team to complete the grading process for the previous candidate. Each interview team typically interviewed 10 to 12 candidates throughout the day, and there were typically 10 to 12 interview teams each year. The interviewers were not provided any specific background information about the candidates prior to, during, or after the interview. The interview teams were provided with a list of candidates who they would be interviewing, the times of each interview, a clock, grading rubric sheets, and a list of the interview questions in the event that a candidate desired visual assistance with any question. Candidates did not know which faculty members would be interviewing them prior to the interview. They were not allowed to bring any materials into the interview room except a notepad and a pen.

The interviewing teams were instructed not to share candidate grading information or impressions with each other to minimize grading bias. Each candidate's total

interview score was the sum of the 2 interviewer's scores. In addition to the rubric score, the interviewers also provided a subjective global assessment (definitely acceptable, acceptable, or unacceptable) of each candidate on the candidate's interview sheet. Unacceptable ratings required written justification on the scoring sheet. If a candidate received 2 unacceptable ratings, the admissions committee evaluated the comments to determine whether he/she should be eliminated from the candidate pool.

In 2011, the admissions committee invited 118 candidates to participate in the interview process. Eleven pairs of faculty members completed the evaluation process. Nine of the 11 faculty pairs interviewed 11 candidates, 1 faculty pair interviewed 10 candidates, and the remaining pair interviewed 9 candidates. Each interview lasted 15 minutes and included 6 questions to evaluate the qualitative (and/or noncognitive) attributes of the candidates. The questions were designed to evaluate verbal communication skills, ethics, relevant life and work experience, emotional maturity, commitment to patient care, leadership, and understanding of the pharmacy profession. The questions were conducted in English with no additional assistance provided to individuals for whom English was a second language. A written copy of the questions was available to any student who desired assistance during the interview.

Upon completion of the interview day, all data were entered into Microsoft Excel by an administrative assistant. Candidates' and interviewers' identification and interview scoring results were de-identified by the administrative assistant and then provided electronically to the researchers. The NDSU Institutional Review Board (IRB) granted approval for this project.

Given the relatively sparse literature assessing pharmacy admission interview processes, this analysis took/assumed a global, exploratory perspective. We gave primary emphasis to evaluating total interview scores across faculty pairs and applicants, rather than to a question-by-question analysis (question-specific analyses are available from the lead author upon request). In contrast to interview assessments in other disciplines (such as medicine),^{8,9} which focus primarily on correlations and measures of interrater reliability, this analysis focused first on mean differences across interview scores and subsequently examined correlations and Cronbach alpha scores across interview scores. This allowed for an examination of inconsistency or relative bias (one indicator of validity), as well as reliability.

We operated under 2 null hypotheses which were conservative in nature, yet inherently assumed that the interview process has a reasonable degree of usefulness in screening applicants. The first hypothesis was that no

mean differences in scores across each of the interview pairs would be found. That is, we had no prior information to suggest that any specific interviewer would rate candidates (at the mean) any differently than her/his interview partner. As such, there was no reason to expect any relative bias across reviewers. Analysis of mean values was conducted using one-way parametric (ANOVA) and non-parametric (Kruskal-Wallis) tests.

For all correlation analyses, we operated under the null hypothesis that no correlation would be found between reviewers. This null was conservative in that, without prior information, we had to assume that interviewer ratings would have little overlap. Rejection of this null would indicate a significant amount of reliability, and examination of the magnitudes of the Cronbach alphas and correlations could be used to assess the degree of overlap. As with the tests of means, we measured all correlations using parametric (Pearson) and nonparametric (Spearman) methods. As the pursuit of admission to a program (including the interview process) is often considered a high-stakes environment, Cronbach alpha values in excess of 0.80 were interpreted as exhibiting high interrater reliability.⁹ Lastly, all analyses were conducted with the SAS, Version 9.2 (SAS, Inc., Cary, NC) statistical package using a 5% significance level.

RESULTS

Table 1 contains the names, definitions, and descriptive statistics for each of the variables (one for each interview question) used in the analysis. Because each applicant was evaluated by 2 interviewers, the results in Table 1 are essentially a panel with 236 observations, with 2 observations per 118 applicants. As noted earlier, the interview consisted of 8 questions, the first of which was not scored, and thus omitted from the analysis. The eighth question asked the applicant whether they had any questions they would like to ask the interviewers. Applicants were given 1 point if they asked a question (without prompting) and otherwise received 0 points for this question. The remaining 6 questions (questions 2-7) were scored using the 0 to 4 Likert scale as mentioned earlier, and accounted for 24 of the 25 possible points.

The total interview score, when evaluated at the mean was 17.4 out of 25 possible points. Examining the quartiles of the total score distribution suggests that the majority of the data are symmetrically grouped between 15 and 20, with a median score of 17.3. The distributions of scores for each of the 6 evaluated interview questions show similar patterns. Mean scores ranged between 2.3 (question 6 – understanding of the pharmacy profession) and 3.0 (question 2 – perceptions of the attributes/skills of a “good” pharmacist). Standard deviations for each of the

Table 1. Descriptive Statistics for Interview Questions and Candidate's Total Interview Score

Variable	Description	Mean (SD)	Min	Range			
				25th Percentile	Med	75th Percentile	Max
Q2	Question on the candidate's perceptions of attributes/skills of a "good" pharmacist	3.0 (0.8)	1	2	3	4	4
Q3	Question on the candidate's perception of pharmacy ethics (medication errors)	2.8 (0.9)	0	2	3	3	4
Q4	Question on the candidate's commitment to patient care (medications for those unable to pay)	3.0 (0.8)	0	2	3	4	4
Q5	Question on the candidate's emotional maturity (dealing with working in groups and group dynamics)	2.8 (0.9)	1	2	3	3	4
Q6	Question on the candidate's understanding of the pharmacy profession (tasks beyond standard dispensing functions)	2.3 (0.9)	1	2	2	3	4
Q7	Question on the candidate's perception of pharmacy leadership (professional involvement and volunteer experience)	2.8 (0.9)	0	2	3	3	4
Q8	Indicator on whether the candidate has questions for the interviewers.	0.9 (0.3)	0	1	1	1	1
Score	Candidate's Total Interview Score (Sum of Q2-Q8: 25 Points Maximum)	17.4 (3.7)	9	15	17.3	20	25

Abbreviations: Min = minimum; Med = median; Max = maximum.

questions ranged between 0.8 and 0.9. The 6 evaluated questions (except question 6, which had a median of 2) had median values of 3, 25th percentiles equal to 2, and 75th percentiles equal to 3. Eighty-six percent of the applicants asked a question of the interview team.

Table 2 contains an analysis of mean scores across faculty pairs. In 9 of the 11 faculty interviewer pairs, the null hypothesis of no mean difference in total interview

scores could not be rejected. However, for 2 pairs of faculty members (pair 9 and pair 10) significant mean differences were found. For faculty pair 9, one interviewer gave the candidate a mean score of 16.5, while the other gave the candidate a mean score of 21.0. For faculty pair 10, the mean scores given were 20.4 and 17.6, respectively.

Table 3 contains the correlation analysis of the faculty interview pairs. Of the 11 faculty pairs, 8 had total

Table 2. Comparative Mean Analysis of Total Interview Scores

Faculty Pairs	No. of Applicants Interviewed	Interviewer 1, Mean (SD)	Interviewer 2, Mean (SD)	Analysis of Variance, <i>P</i>	Kruskal-Wallis, <i>P</i>
All	118	17.2 (3.7)	17.7 (3.6)	<0.001	<0.001
1	11	15.0 (3.7)	17.4 (4.4)	0.189	0.166
2	9	20.0 (3.5)	18.2 (3.2)	0.282	0.194
3	11	17.0 (3.7)	18.3 (2.7)	0.348	0.166
4	11	14.9 (3.2)	15.2 (3.6)	0.854	0.815
5	10	17.1 (3.4)	16.7 (2.6)	0.772	0.819
6	11	17.7 (2.3)	17.5 (4.3)	0.903	0.869
7	11	18.5 (3.9)	20.2 (2.7)	0.263	0.221
8	11	17.9 (3.9)	16.2 (2.4)	0.222	0.221
9	11	16.5 (2.3)	21.0 (2.8)	0.001	0.001
10	11	20.4 (3.0)	17.5 (2.7)	0.030	0.051
11	11	14.5 (3.8)	16.3 (4.5)	0.340	0.306

Table 3. Correlation Analysis of Total Interview Scores

Faculty Pair	No. of Applicants Interviewed	Raw Cronbach Alpha	Standardized Cronbach Alpha	Pearson Correlation	P	Spearman Correlation	P
1	11	0.135	0.136	0.073	0.831	0.274	0.415
2	9	0.814	0.816	0.689	0.040	0.708	0.033
3	11	0.879	0.905	0.826	0.002	0.799	0.003
4	11	0.995	0.998	0.996	<0.001	0.993	<0.001
5	10	0.839	0.855	0.747	0.013	0.783	0.007
6	11	0.255	0.302	0.178	0.601	0.235	0.486
7	11	0.889	0.919	0.850	0.001	0.820	0.002
8	11	0.802	0.860	0.755	0.007	0.800	0.003
9	11	0.817	0.826	0.703	0.016	0.647	0.031
10	11	0.558	0.559	0.388	0.239	0.299	0.372
11	11	0.871	0.877	0.782	0.005	0.833	0.001

interview scores that were significant (from 0) and whose Pearson correlations were 0.689 or higher. These same 8 pairs of faculty interviewers exhibited Cronbach alpha scores in excess of 0.80, indicating high interrater reliability. Faculty pairs 1, 6, and 10 had insignificant correlations and Cronbach alpha values below 0.80. Faculty pair 10 also has significantly different mean interview scores. As such, this faculty pair exhibited fundamentally different interpretations of the candidates and/or the use of the interview questions and rubric.

DISCUSSION

Given that the overall mean score for all interviews was just over 17, it is reasonable to conclude that one faculty interviewer in group 9 and one in group 10 gave applicants scores that may be interpreted as excessively high. Faculty pairs 1 and 6 gave similar scores on average, but the correlations from Table 3 suggest that these faculty members did not evaluate applicants consistently from one applicant to the next. Concomitantly, faculty pair 9 gave different total scores on average, but (given the significant correlations in Table 3) evaluated candidates in a relatively consistent or reliable fashion. The majority of the faculty pairs gave similar mean scores that were highly correlated, which suggests that the interview system used at NDSU is a generally consistent and reliable process (although not perfect). Also, several pairs of faculty members did not give the same scores (whether at the mean or in terms of correlations), which suggests that room for improvement exists. It is important for the admissions committee to develop and implement additional (or possibly refresher) training for interviewers to ensure greater consistency and reliability in interview scores. The training may need to focus on how to read and interpret the evaluation rubrics, as well as how to interpret the specific interview questions within the context of that

rubric. Discrepancies across reviewer pairs varied (ie, some had significant mean differences, some were not significantly correlated, and one faculty pair had different means and low correlations), which suggests that multiple interventions are necessary to address these discrepancies. In other words, different interviewers may need different types of training or retraining. Without a more general, exploratory analysis (and not just an examination of interrater reliability as is commonly done in the literature), not all of these issues would have been identified.

This study involved a single interview station for each candidate. There have been studies published in the medical school literature that show that multiple station mini-interviews increase reliability and validity of interview scoring compared to a single station with 1 interviewer.¹⁰ Multiple mini-interview stations are beneficial because they allow every interviewer to evaluate all interviewees. A limitation of this form of interviewing is the time required to complete the process.¹¹

Another related limitation of the analysis was the lack of available data for each faculty pair. In this analysis, each faculty pair only evaluated between 9 and 11 applicants; thus, most of the statistical analyses used less than 22 unique data points per calculation. In medical school (and other health science admission) interviews, each faculty member interviews the majority of applicants. Consequently, there are a larger number of observations per faculty pair available for analysis (usually in excess of 30 observations per faculty interviewer and/or per interviewee), which allows for more reliable (and powerful) statistical analyses and more reliable evaluations of applicants' skill sets. Both of these attributes facilitate better information when making admission decisions.^{12,13}

Although the interview process used in this study may apply to some colleges and schools of pharmacy, it

may not be applicable to all accredited pharmacy schools and colleges in the United States. While most colleges and schools use some form of interview process to assess potential applicants, the methods and structure of that process vary widely. Moreover, some colleges and schools use continuous admissions and/or interview processes, in which applicants are not assessed in the same timeframe, and possibly by faculty and staff members who are not consistently matched as interviewer pairs. In such cases, this study's findings may be of limited use.

A final limitation is that the current analysis represents a first step in a process of continuous quality improvement (CQI). While this study suggests that the interview process used at NDSU is largely successful, improvements to the process could be made. Once implemented, the improved interview process may be fundamentally different than the one evaluated in this study, making the latter obsolete. This, in turn, necessitates further research to evaluate whether (and if so, how) these changes have improved the quality, consistency, and reliability of applicant interviews. One notable opportunity for CQI is to add additional interviewer training each year to ensure that inconsistent reviews (such as conducted by faculty pair 10) are less likely to occur. This requires some additional time, planning, and research to identify best practices in the literature regarding interviews, and subsequently incorporate those practices into the interview process. It may also require more careful planning in selecting interviewer pairs, and require some prior assessment of the personalities of the interviewers to ensure that they are compatible. A second area of improvement would be to allow more interview time by revisiting and restructuring the entire interview day process. This would allow either more interview questions in each interview or for some analog of a multiple mini-interview process to be established.

CONCLUSIONS

The interview process can be an effective method to evaluate qualitative parameters or attributes of pharmacy school candidates such as ethics, relevant life and work experience, emotional maturity, commitment to patient care, leadership, and understanding of the pharmacy profession. The majority of faculty pairs in this study gave the candidates similar mean interview scores and subsequently achieved high interrater reliability. This indicates

that the interview process described is generally consistent and reliable; however, future improvements can be made through such interventions as continued interviewer training and using multiple mini-interview stations.

ACKNOWLEDGEMENTS

We are grateful to Lizbeth Frannea, administrative assistant in the Dean's office, for meticulous record keeping that provided tremendous support for this project.

REFERENCES

1. Accreditation Council for Pharmacy Education. Accreditation standards. <http://www.acpe-accredit.org/standards/default.asp>. Accessed January 24, 2012.
2. Latif DA. Including the assessment of nontraditional factors in pharmacy school admissions. *Ann Pharmacotherapy*. 2005;39(4):721-726.
3. Chestnut RJ, Philips CR. Current practices and anticipated changes in academic and nonacademic admission sources for entry-level PharmD programs. *Am J Pharm Educ*. 2000;64(3):251-259.
4. Stolte SK, Scheer SB, Robinson ET. The reliability of non-cognitive admission measures in predicting non-traditional doctor of pharmacy student performance outcomes. *Am J Pharm Educ*. 2003;67(1):Article 18.
5. Hardigan PC, Lai LL, Arneson D, Robeson A. Significance of academic merit, test scores, interviews and the admissions process: a case study. *Am J Pharm Educ*. 2001;65(1):40-43.
6. Latif DA. Using the structured interview for a more reliable assessment of pharmacy student applicants. *Am J Pharm Educ*. 2004;68(1):Article 21.
7. Joyner PU, Cox WC, White-Harris C, Blalock SJ. The structured interview and interviewer training in the admissions process. *Am J Pharm Educ*. 2007;71(5):Article 83.
8. Bandiera G, Regehr G. Evaluation of a structured application assessment instrument for assessing applications to Canadian postgraduate training programs in emergency medicine. *Acad Emerg Med*. 2003;10(6):594-598.
9. Bandiera G, Regehr G. Reliability of a structured interview scoring instrument for a Canadian postgraduate emergency medicine training program. *Acad Emerg Med*. 2004;11(1):27-32.
10. Lemay J, Lockyer JM, Collin VT, Brownell AK. Assessment of non-cognitive traits through the admissions multiple mini-interview. *Med Educ*. 2007;41(6):573-579.
11. Rosenfeld JM, Reiter HI, Trinh K, Eva KW. A cost-efficiency comparison between the multiple mini-interview and traditional admissions interviews. *Adv Health Sci Educ*. 2008;13(1):43-58.
12. Harris S, Owen C. Discerning quality: using the multiple mini-interview in student selection for the Australian National University Medical School. *Med Educ*. 2007;41(3):234-241.
13. Eva K, Reiter H, Trinh K, Wasi P, Rosenfeld J, Norman G. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ*. 2009;43(8):767-775.