

RESEARCH

Improving Reliability of a Residency Interview Process

Michael J. Peeters, PharmD, Michelle L. Serres, PharmD, and Todd E. Gundrum, PharmD

University of Toledo College of Pharmacy & Pharmaceutical Sciences

Submitted December 2, 2012; accepted March 29, 2013; published October 14, 2013.

Objective. To improve the reliability and discrimination of a pharmacy resident interview evaluation form, and thereby improve the reliability of the interview process.

Methods. In phase 1 of the study, authors used a Many-Facet Rasch Measurement model to optimize an existing evaluation form for reliability and discrimination. In phase 2, interviewer pairs used the modified evaluation form within 4 separate interview stations. In phase 3, 8 interviewers individually-evaluated each candidate in one-on-one interviews.

Results. In phase 1, the evaluation form had a reliability of 0.98 with person separation of 6.56; reproducibly, the form separated applicants into 6 distinct groups. Using that form in phase 2 and 3, our largest variation source was candidates, while content specificity was the next largest variation source. The phase 2 g-coefficient was 0.787, while confirmatory phase 3 was 0.922. Process reliability improved with more stations despite fewer interviewers per station—impact of content specificity was greatly reduced with more interview stations.

Conclusion. A more reliable, discriminating evaluation form was developed to evaluate candidates during resident interviews, and a process was designed that reduced the impact from content specificity.

Keywords: psychometrics, interview, residency, reliability

INTRODUCTION

Postgraduate year 1 (PGY1) pharmacy residencies are increasingly prevalent in the United States; however, the pool of resident applicants has surpassed the number of available positions,¹ and professional organizations anticipate continued growth.^{2,3} Establishing a fair and objective process for selecting residents seems essential, and reliability of the process is a key characteristic in ensuring fairness in these candidate assessments.⁴ An essential element during a resident selection process is the interview. Lack of validity, objectivity, reliability, and structure of the interview process exists in medical residency,⁵⁻⁸ medical school⁹ and pharmacy school admissions.¹⁰

Selecting the best candidates for pharmacy residency training is a difficult task. The ideal interview and selection process for residency candidates would be one that was efficient, objective, and produced reliable feedback/information that could be used to make informed decisions. The survey questions asked and criteria used to

make decisions appear to be fairly consistent among residency programs.^{11,12} Psychometric developments over the past few decades may help in increasing reliability of tools used in the interview process. The objective structured clinical examination (OSCE) was first discussed in 1979 and its interview “offspring,” the multiple mini-interview (MMI), was described in 2003. Both showed promise for improving interview reliability. The overall concept and purpose of the MMI is to reduce the impact of content specificity, a concern in assessments, on the interview as compared to a traditional individual interview.¹³⁻¹⁵ In the MMI, candidates rotate through interview stations where different domains are assessed and, once completed, the candidate’s aggregate score is tabulated. This format has been increasingly used for admissions interviews for medical and pharmacy schools.^{16,17} It requires considerable planning prior to implementation and follows a structured format. These psychometric developments may also aid educators in making the pharmacy resident selection process more objective, which seems increasingly pertinent given the increasing number of candidates who apply for but are not offered a residency position.

Within this paper’s 3 study phases, the overall aim was to improve reliability of a resident interview process. Values included in the process were: (1) best evaluating

Corresponding Author: Michael J. Peeters, PharmD, Clinical Associate Lecturer, University of Toledo College of Pharmacy & Pharmaceutical Sciences, 3000 Arlington Avenue, MS1013, Toledo, OH 43614. Tel: 419-383-1946. Fax: 419-383-1950. E-mail: michael.peeters@utoledo.edu

qualities we desire in our residents, (2) time efficiency, and (3) maximizing fairness to resident applicants. The objective of the first phase was to psychometrically characterize the interview evaluation form, while the second and third phase objectives were to expand our investigation toward optimizing the reliability of the entire interviewing process.

METHODS

The University of Toledo IRB approved this study. This study consisted of 3 phases (Table 1). The first phase involved analysis and fine tuning of the evaluation form used by all interviewers for the college’s PGY1 pharmacy residency program. After gaining confidence from characterizing the reliability and validity of the evaluation form, the second phase was to analyze the reliability of the entire interview process. Building on the results from phase 2, the purpose of phase 3 was to confirm the reliability of the “optimal” process in a subsequent PGY2 residency interview cohort.

Phase One

There were 2 parts to this first phase. Because the instrument was modified in the first year of interviews (2008 residency recruiting cycle), a second year (2009 residency recruiting cycle) was also assessed to identify any unintended psychometric impact from revisions to the evaluation form. A formal structured process was not used to conduct the interviews in phase 1. Only 2 interviewers interviewed some candidates while 3 or 4 interviewers interviewed other candidates. Regardless of number, all interviewers independently rated each candidate using the same evaluation form.

The origin of the evaluation form used in this study was not known other than that other residency programs had used a prior version and that faculty members in this program had previously modified it. Interviewers’ scoring of evaluation form items was analyzed for rating scale functionality along with instrument reliability. Linacre’s rating scale guidelines were followed.¹⁸ Interviewers scored all items for each candidate. Prior to phase 2, the

rating scale was simplified further to a global 4-category scale and a “do not rank” checkbox was added (Appendix 1).

Facets software, version 3.67.1 (MESA, Chicago IL) was used for analysis because it allowed for a multiple variable model within a Many-Facets Rasch Measurement (MFRM). Within the MFRM, residency candidates’ scores were matched with their interviewers’ ratings to analyze the form’s construct validity and reliability.¹⁸⁻²² Important to our instrument’s use and reliability, the rating scale was analyzed for functionality among these interviewers by using rating scale step calibrations (ie, do interviewers’ rating scale responses align with a scaling of increased candidate ability among the evaluation form items?).

For reliability coefficients in general, thresholds of >0.7 for most low-stakes assessments and increased thresholds (>0.8 or >0.9) should be used as the stakes increase (ie, increasing from low-stakes [ie, a single classroom assessment], to high-stakes [ie, graduation is contingent upon the assessment or for a new job/position], to high-stakes [ie, professional licensing]).²³ For the residency interviews, a coefficient >0.80 was desired. As well, the MFRM will consider and adjust for rater severity/leniency among interviewers so that no single interviewer would disrupt the ranking of candidates by being too easy or too difficult in their ratings.²¹

Phase Two

In phase 2, the evaluation form was used in conducting PGY1 resident interviews from January through early February 2012, in anticipation of the college’s residency program expanding from 3 to 8 residency positions. Each invited candidate was scheduled for a half-day interview, where additional information about the program was discussed, candidates participated in patient care rounding, and 4 interview sessions were conducted. In this design, 2 interviewers were nested within each interview session (or station), for a total of 8 interviewers for each candidate. Interviewers did not move between stations.

Using the augmented evaluation form from phase 1 (Appendix 1), the process of PGY1 candidate interviews

Table 1. Flowchart of Study Phases

	Objective	Year	Participants	Interviewers	Interview Stations
Phase 1 (PGY1 MFRM)					
Part 1 (MFRM Explore)	Form reliability	2008	8 (576 form ratings)	8	Varied (often 2)
Part 2 (MFRM Confirm)	Form reliability	2009	4 (324 form ratings)	9	Varied (often 2)
Phase 2 (PGY1 G-theory)	Process reliability	2012	24	8	4
Phase 3 (PGY2 G-theory)	Process reliability	2012	8	8	8

MFRM= Many-Facets Rasch Measurement model, G-theory=Generalizability Theory model, PGY1=postgraduate year 1 pharmacy residency, PGY2=postgraduate year 2 pharmacy residency.

was analyzed. Based on each interviewer's ratings on the evaluation form items, a single overall, global rating was scored by each interviewer using the entire form. In general, a global rating scale has shown strong reliability²⁴ and has an advantage of quick, straightforward computation among interviewers for PGY1 candidate ranking afterward. Addition of median global score's (ie, 1, 2, 3, 4) from interviewers at each of the 4 stations was summed towards a total score (out of a possible 16;). Candidates were then sorted from highest to lowest total score for the program's rank list submitted for the ASHP Residency Match. Along with the global rating scale, a "do not rank" checkbox was added, but designed so that interviewers did not mistake it as being part of the global scale. Because unstructured interview panels can be just as reliable and possibly more valid than structured interview panels,²⁴ we kept questions unstructured so that after introductions, interviewers could ask whatever questions they deemed appropriate. Generally, the questions asked focused on clarifying curricula vitae statements, prior behaviors in specific situations, and other questions to assist the interviewer in scoring the evaluation form (Appendix 1). We did not train seasoned faculty interviewers specifically on how to use this form.

To go beyond form reliability and investigate the entire process's reliability, Generalizability Theory was used in this phase.²⁵⁻²⁹ The specific software program used herein was G_Strings IV, version 6.1.1 (Hamilton, ON, CA). The results output from a generalizability study will provide the relative contributions of different variance sources to variance in the entire data set, with a generalizability coefficient reported for reliability of the entire process (0.0-1.0 scale; similar to Cronbach's alpha).²⁶ Importantly, the generalizability study's analysis of variance among multiple variables further allows investigators to extrapolate results to a different numbers of variables (ie, stations, interviewers) and estimate the changes in reliability—to inform decision making.²⁹ In this generalizability study, the variables were candidates, interview stations, and interviewers, with interviewers nested within each station.

Phase Three

Based on phase 2's results, this third phase was conducted in late February 2012 and used interviews for our ASHP-accredited PGY2 critical care pharmacy residency at the same institution. Herein, candidates had one-on-one interviews with each of 8 interviewers, while program information was also discussed during each candidate's site visit.

All interviewers used a slightly modified version of the evaluation form from phase 2. Slight modifications

were that item 2 was revised from "hospital pharmacy experience" to "pharmacy practice experience" and item 3's "clinical experience" was revised to "critical care pharmacy experience."

G_String analysis was used again in phase 3. This data set would not allow any inter-rater variance to be calculated because multiple interviewers were not used on any of the interview panels, so variance could not be calculated between interviewers on the same panel. Similar to phase 2, this generalizability study's variables were the candidates and interview stations, with 1 interviewer nested in each station.

RESULTS

In phase 1, 8 interviewers interviewed 8 residency candidates and rated them on 9 items using the evaluation form. The resulting 576 items were analyzed within the MFRM. The original 5-point rating scale showed poor item functionality. Using Linacre's framework, rating scale categories 1 and 2 were collapsed, with the resulting 4-point rating scale functioning well. With the revised rating scale, the instrument was able to demonstrate 3 groups of candidates (separation=3.24) with a reliability index for this cohort of 0.91 (Appendix 1).

For confirmation, the following year, 9 interviewers each interviewed 4 candidates and rated them using the revised 9-item evaluation form. In that second cohort, the revised 4-point rating scale functioned well, with candidate separation of 6.56 while reliability was 0.98.

Phase Two

Twenty-four candidates participated in these PGY1 interviews. The generalizability coefficient for the entire process was 0.787. In decreasing order, the relative size of scoring variance was 74% from differences among candidates, 13.5% from candidate-station interaction (ie, content specificity), 3.5% from differences between interview stations, and only 2.5% from inter-rater variation. Extrapolating from the generalizability study, decision-studies were conducted to vary both number of interview stations and the number of interviewers within each interview station (Table 2).

Phase Three

Informed by phase 2, a third phase was conducted in which 3 interviews (1 at each station) used the revised form to rate 8 PGY2 candidates. The generalizability coefficient for the entire process was 0.922. In decreasing order, the relative size of scoring variance was 91% from differences among candidates, 7.7% from candidate-station interaction (ie, content specificity), and 1.3% from differences between interview stations. Extrapolating

Table 2. Decision Studies for our PGY1 Process (Phase 2)

Decision Studies	Interview Stations, No.	Interviewers in Each Station, No.	G-Coefficient (Reliability)
1 station, 2 interviewers	1	2	0.481
1 station, 4 interviewers	1	4	0.526
1 station, 8 interviewers	1	8	0.552
4 stations, 2 interviewers	4	2	0.787
8 stations, 2 interviewers	8	2	0.881
8 stations, 1 interviewer	8	1	0.847

from the generalizability study findings, d-studies were conducted to vary the number of interview stations (Table 3).

DISCUSSION

Content specificity as it related to student assessment and interviewing has been an underestimated yet important topic discussed throughout the health professions literature. Studies analyzing sources of variability in assessments, including the OSCE and MMI, have demonstrated the importance and perils of content specificity.^{9,13,14,17} For interviews, content specificity refers to the interaction between a particular candidate and the content of an interviewing station. The use of a reliable evaluation form during the interview process allows the interviewer to remain more objective and that a candidate is assessed based more abilities and not simply their personality or ability to present themselves well. This objectivity is *vital* to the overall reliability of assessments made in both interviewing and grading processes.¹³⁻¹⁷ Medical education analyses extended these findings to more general interviews and does not appear to be a feature of a single MMI structure.^{30,31} Increasing the number of interview sessions had a much larger effect on improving reliability

than adding evaluators. As demonstrated by this investigation, altering the evaluation form to produce positive changes in content specificity can have a significant impact on the overall reliability of a pharmacy resident’s assessment scores and ranking.

In phase 1, the evaluation form was revised to improve the reliability of residency interviews. Making the rating scale changes to the evaluation form improved the reliability of this instrument and resulted in a higher candidate separation, allowing for more confident ranking decisions to be made later. Unlike subsequent phases, phase 1 did not look specifically at content specificity.

Phase 2 was a more inclusive look at our interview rating process. Using a generalizability study allowed analysis of score variation from different sources (number of interviewers, number of interview stations, and content specificity) with a generalizability coefficient. This reliability estimate was broader in that it characterized the entire process, which included *both* internal consistency (ie, Cronbach alpha) and inter-rater reliability. With phase 2’s generalizability coefficient of 0.787 (with 4 stations of 2 interviewers), it should be improved further to a generalizability coefficient of at least 0.8, which is minimally desirable with high-stake decisions such as resident candidate rankings. Thus, for comparison, it is useful to look into the makeup of variability by the different modifiable variables (number of interviewers and stations). The majority of score variability was the result of inter-candidate differences. This was desirable as the program’s goal was to distinguish among candidates. Aside from variation in candidates, the rest of variation from other interactions represented “noise” in the data. The largest contributor to “noise” variability came from candidate-station interaction. Content specificity underpins this variability and accounts for more variation than the rest of the identified sources combined. This is a common finding in generalizability study analyses²⁷ and illustrates the concept that the perceptions (and subsequent ratings) of any one interviewer along with his/her specific questions greatly affect an assessment’s reliability; therefore, multiple interviewers should be sought. Using 8 interviewers within 8 separate interview stations was more reliable than having all 8 interviewers as 1 panel/station or even having 2 interviewers to 4 interview stations.

Phase 3 was designed to test the suggestion from decision studies in phase 2 of increasing the number of candidate-station interactions without increasing the overall number of interviewers or resources needed (Table 3). The most preferable generalization coefficient meeting this criterion was found when 8 interview stations were used and each station had a single interviewer. While results are not directly comparable given different candidates

Table 3. Decision Studies for the PGY2 Process (Phase 3)

Decision Study	Interview Stations, No.	G-Coefficient (Reliability)
1 station	1	0.597
2 stations	2	0.747
3 stations	3	0.816
4 stations	4	0.855
6 stations	6	0.899
8 stations	8	0.922

in both number and type of residency, the similarities of processes allow for moderate comparison between the results. The reliability with PGY2 critical-care candidate rankings in phase 3 was slightly higher than with PGY1 interviews. As well, candidate-station variability (or “noise” from content specificity) decreased with more stations. Thus, improvement in reliability probably can be at least partly attributed to reducing the impact from content specificity by having more interview stations.

In the past, researchers have spent a lot of time attempting to improve consistency focused only on inter-rater reliability. While this can be important, our series of study phases exemplifies a small possible change that attempts to augment inter-rater reliability could have on the overall process reliability compared with other manipulations in other variables such as candidate-station interaction (ie, content specificity). Phase 2 illustrated that only 2.5% of the variability was attributable to inter-rater variation among multiple interviewers without any rater training, while 3.4% was attributable to inter-station variation. Aside from our “signal” of candidate variability, only a quarter of “noise” data variation in phase 2 came from combined inter-rater and station variation. Meanwhile, over half of non-candidate “noise” variation was from content specificity. Variation from content specificity was 5 times larger than inter-rater variation in this interview process—implying that time and effort spent on improving inter-rater reliability and rater training may better be spent on improving content specificity.¹⁷

This study was primarily designed to determine the reliability of a program’s residency candidate interviewing process and improve its interview process quality. As such, the phases involved a relatively small number of candidates and different numbers of both PGY1 and PGY2 resident candidates; this study could be seen as a pilot study. The statistical analyses conducted were our attempts to work meaningfully with this relatively small amount of data from 2 different residency programs in 1 institution.³² Investigation with a larger and more diverse sample would improve confidence in the results. Confirmation of our findings by other programs could help to further the generalizability of this evaluation form and process.

Neither of the measurement models used in these studies are widely used in the pharmacy literature; however, both are complex, rigorous models, used in high-stakes testing such as in professional licensing examinations of physicians, nurses, chiropractors, osteopathic doctors, and pharmacists. Regardless of model used, reliability is a central characteristic of any assessment,

and legal challenges to decisions made using an evaluative instrument or process often focus on the poorly reliable assessment methods used.³³

CONCLUSION

This investigation demonstrated an improvement in the reliability of a residency candidate interview form and process. There was a clear preference for improvements that influenced content specificity. Beyond candidate differences, content specificity had the largest effect on the overall reliability for our entire interview process. In terms of candidate selection, having 1 or 2 interviewers in a single interview station had a reliability similar to flipping a coin, whereas having multiple interview stations improved reliability substantially, and was maximized with 1 interviewer in numerous interview stations. Increasing the number of separate candidate-interviewer interactions was a viable way to reduce the impact from content specificity, and possibly without even needing further resources. Continuous quality improvement in methods to optimize each pharmacy residency program’s interview reliability should be pursued.

ACKNOWLEDGEMENTS

We thank medical education’s Ralph Bloch, PhD, for his expert review of our G-Theory use. Additionally, we thank our prior PGY1 Residency Program Director, Chad Tuckerman, PharmD, for his assistance with study 1, and Julie Murphy, PharmD, for her review of the final manuscript. This study was presented as a poster abstract during the 2012 ASHP Annual Meeting in Las Vegas, NV

REFERENCES

1. National Matching Services, Inc. ASHP Residency Matching Program. Residency Match Statistics. <http://www.natmatch.com/ashprmp/aboutstats.html>. Accessed March 12, 2013.
2. American College of Clinical Pharmacy. American College of Clinical Pharmacy’s vision of the future: postgraduate pharmacy residency training as a prerequisite for direct patient care practice. *Pharmacotherapy*. 2006;26(5):722-733.
3. American Society of Health-System Pharmacists. ASHP Policy Positions 1982-2007. Residency training for pharmacists who provide direct patient care (policy position 0005). http://www.ashp.org/s_ashp/docs/files/About_policypositions.pdf. Accessed March 12, 2013.
4. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington DC: AERA; 1999.
5. Komives E, Weiss ST, Rosa RM. The applicant interview as a predictor of resident performance. *J Med Educ*. 1984;59(5):425-426.
6. Kovel AJ, Davis JK. Use of interviews in the selection of pediatric house officers. *J Med Educ*. 1988;63(12):928.

7. Dawkins K, Ekstrom RD, Maltbie A. The relationship between psychiatry residency applicant evaluations and subsequent residency performance. *Acad Psychiatry*. 2005;29(1):69-75.
8. DeSantis M, Marco CA. Emergency medicine residency selection: factors influencing candidate decisions. *Acad Emerg Med*. 2005;12(8):559-561.
9. Kreiter CD, Yin P, Solow C, Brennan RL. Investigating the reliability of the medical school admissions interview. *Adv Health Sci Educ Theory Pract*. 2004;9(2):147-159.
10. Latif DA. Using the structured interview for a more reliable assessment of pharmacy student applicants. *Am J Pharm Educ*. 2004;68(1):Article 21.
11. Mersfelder TL, Bickel RJ. Structure of postgraduate year 1 pharmacy residency interviews. *Am J Health-Syst Pharm*. 2004;66(12):1075-1076.
12. Mancusso CE, Paloucek FP. Understanding and preparing for pharmacy residency interviews. *Am J Health-Syst Pharm*. 2004;61(16):1686-1688.
13. Eva KW, Rosenfeld J, Reiter HI, Norman GR. An admissions OSCE: the multiple mini-interview. *Med Educ*. 2004;38(3):314-326.
14. Eva K. On the generality of specificity. *Med Educ*. 2003;37(7):587-588.
15. Peeters MJ, Cox CD. Using the OSCE strategy for APPEs? *Am J Pharm Educ*. 2011;75(1):Article 13.
16. Cameron AJ, MacKeigan LD. Development and pilot testing of a multiple mini-interview for admission to a pharmacy degree program. *Am J Pharm Educ*. 2012;76(1):Article 10.
17. Serres ML, Peeters MJ. Overcoming content specificity in admission interviews: the next generation? *Am J Pharm Educ*. 2012;75(10):Article 207.
18. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002;3(1):85-106.
19. Smith EV. Evidence for the reliability of measures and validity of measure interpretation: a Rasch measurement perspective. *J Appl Meas*. 2001; 2(3):281-311.
20. Bond TG. Validity and assessment: a Rasch measurement perspective. *Metodologia de las Ciencias del Comportamiento*. 2003;5(2):179-194.
21. Linacre JM. *Many-Facet Rasch Measurement, 2nd edition*. Chicago IL: MESA Press; 1994.
22. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences, 2nd edition*. Mahwah NJ: Lawrence Erlbaum Associates, Publishers; 2007.
23. Penny JA, Gordon B. *Assessing Performance*. New York NY: The Guilford Press; 2009.
24. Blouin D, Day AG, Pavlov A. Comparative reliability of structured versus unstructured interviews in the admission process of a residency program. *J Grad Med Educ*. 2011;3(4):517-523.
25. Streiner DL, Norman GR. *Health Measurement Scales, 4th edition*. New York NY: Oxford University Press; 2008.
26. Bloch R, Norman G. Generalizability theory for the perplexed: a practical introduction and guide: AMEE guide no.68. *Med Teach*. 2012;34(11):960-992.
27. Crossley J, Davies H, Humphris G, Jolly B. Generalisability: a key to unlock professional assessment. *Med Educ*. 2002;36(10):972-978.
28. Brennan RL. *Generalizability Theory*. New York NY: Springer; 2010.
29. Bloch R, Norman G. *G Strings IV User Manual*. Hamilton, ON, CA; 2011.
30. Axelson RD, Kreiter CD. Rater and occasion impacts on the reliability of pre-admission assessments. *Med Educ*. 2009;43(12):1198-1202.
31. Hanson MD, Kulasegaram KM, Woods NN, Fechtig L, Anderson G. Modified personal interviews: resurrecting reliable personal interviews for admissions? *Acad Med*. 2012;87(10):1330-1334.
32. Linacre JM. Sample Size and item calibration stability. *Rasch Meas Trans*. 1994;7(4):328.
33. Tweed M, Miola J. Legal vulnerability of assessment tools. *Med Teach*. 2001;23(3):312-314.

Appendix 1. Revised interviewer evaluation form.

Revised University of Toledo PGY1 Residency Candidate Interview Evaluation Form

Reviewer _____ Candidate _____ Date _____

COMMENTS ARE REQUIRED FOR EACH CATEGORY AND OVERALL IMPRESSION

Pharmacy Experience: Rate experience, knowledge and understanding of pharmacy.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Hospital Pharmacy Experience: Rate experience in hospital pharmacy practice, including unit dose, IV admixtures, pharmacy computer systems, and decentralized systems.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Clinical Experience: Rate experience in clinical activities based on clerkships, internship experiences and extracurricular activities.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Professional Goals: Rate the clarity of career goals, the compatibility between his/her goals and our residency program, reasons for interest in our program, and willingness to commit to an additional year of training.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Professional Maturity: Rate maturity, poise, confidence, appearance, personality and leadership potential.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Interpersonal Rapport: Rate abilities to communicate, foster and sustain a discussion, and answer questions.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
References: Rate candidate's acceptability based on available reference letters from work and faculty references. Consider not only what is said but what is not said.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Knowledge Base: Rate educational background based on candidate's content and quality of his/her PharmD didactic and experiential curriculum. You may also consider GPA, honors and/or awards received, and profession-related extracurricular activities.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
Fit to Program: Consider how well you think this candidate will both perform and fit within our residency program.	1 Weakness	2 Acceptable	3 Strength	4 Exceptional	Comments:
What impresses you most about this candidate?					What concerns you most about this candidate?
Overall (Revision for 2nd and 3rd study phases)					
<input type="radio"/> Exceptional	<input type="radio"/> Good	<input type="radio"/> Acceptable		<input type="radio"/> Concerns	
<input type="radio"/> Do Not Rank					