Interviewers' Experiences with Two Multiple Mini-Interview Scoring Methods Used for Admission to a Master of Physical Therapy Programme

Ina van der Spuy, PhD; Angela Busch, PhD; Julia Bidonde, PhD

ABSTRACT

Purpose: To describe participants' attitudes, beliefs, and experiences with the use of two methods of scoring the Multiple Mini-Interview (MMI) for admission to a Master of Physical Therapy program: a rank-based scoring system (RBS; used from 2007 to 2013) and a criterion-based scoring system (CBS; tested in 2014). The MMI uses short independent assessments to obtain an aggregate score of candidates' professionalism and interpersonal skills, based on behavioural questions within scenarios that assess one attribute at a time. **Method:** This qualitative descriptive inquiry sought to capture the experiences of 18 MMI interviewers primarily through semi-structured interviews. Interviews were transcribed verbatim, and the data were analyzed using thematic analysis. The results were validated by theoretical and investigator triangulation and member checking. **Results:** One major theme, scoring systems, and two sub-themes, CBS and RBS, emerged across all data. Participants unanimously agreed that CBS is a more fair and objective way to score candidates' interviews. **Conclusions:** CBS was well accepted by participants, and the majority preferred it over RBS. Participants felt that CBS presented a more accurate depiction of candidates.

Key Words: evaluation studies as topic; qualitative research; students.

RÉSUMÉ

Objectif : Décrire les attitudes, les croyances et les expériences des participants face à l'utilisation de deux méthodes d'attribution d'une note à la mini-entrevue multiple (MEM) pour l'admission à un programme de maîtrise en physiothérapie: un système de classement selon la note reçue (RBS, utilisé de 2007 à 2013) et un système de classement selon des critères (CBS, testé en 2014). La MEM utilise de brèves évaluations indépendantes pour produire, au sujet du professionnalisme et des techniques de relations interpersonnelles des candidats, une note agrégée basée sur des questions relatives au comportement dans les scénarios qui évaluent une qualité à la fois. **Méthode :** Cette recherche descriptive qualitative visait à saisir les expériences de 18 intervieweurs MEM principalement grâce à des entrevues semi-structurées. Les entrevues ont été transcrites textuellement et l'on a analysé les données en appliquant une analyse thématique. Les résultats ont été validés par triangulation entre la théorie et les chercheurs et par une vérification auprès des membres. **Résultats :** Un thème de premier plan, soit les systèmes de notation, et deux sous-thèmes, soit les systèmes RBS et CBS, ont émergé de toutes les données. Les participants ont convenu à l'unanimité que le classement selon des critères constitue une façon plus équitable et objective d'attribuer une note aux entrevues des candidats. **Conclusion :** Les participants ont bien accepté le classement basé sur des critères qu'ils préféraient en majorité au classement selon la note reçue. Les participants étaient d'avis que le classement basé sur des critères représentait plus fidèlement les candidats.

For many years, admissions for medical and physical therapy (PT) programmes have used the traditional interview.^{1,2} Recent studies have demonstrated, however, that traditional interviews may not be sufficient to select the best candidates.³ The main criticism of traditional interviews is that they may not be reliable or valid because they are context specific; that is, outcomes rely heavily on the interviewers, the questions being asked, and fac-

tors not relevant to candidates' suitability.^{1,2,4} Interviewer variability accounts for fully 56% of the variance in traditional interview scores.⁵

The Multiple Mini-Interview (MMI) addresses these concerns through its structure of multiple stations, interviewers, and questions.²⁻⁴ In this way, the MMI expands on the traditional interview and addresses context specificity.^{1,3} Both the MMI and the traditional interview

From the School of Physical Therapy, University of Saskatchewan, Saskatoon.

Correspondence to: Ina van der Spuy, School of Physical Therapy, College of Medicine, University of Saskatchewan, 1121 College Dr., Saskatoon, SK S7N 0W3; ina.vanderspuy@usask.ca

Disclaimers: The views expressed in this article are those of the authors and not the official position of the School of Physical Therapy, University of Saskatchewan.

Competing interests: None declared.

Acknowledgements: The authors thank Julie Penner, BScPT, MSc, for help with research on the background section of the article.

Physiotherapy Canada 2016; 68(2);179-185; doi:10.3138/ptc.2015-24E

aim to assess professionalism and interpersonal skills.^{1,2} Six characteristics are assessed in the MMI: communication and interpersonal skills; ethical and moral reasoning; critical thinking and logical reasoning; preparation and motivation to study in the designated field; leadership and teamwork; and integrity and honesty.³ Eva and colleagues² original 2004 study identified six benefits of the MMI over the traditional interview: The MMI allows multiple views into candidates' abilities; dilutes examiner bias and chance; is structured so that all candidates respond to the same scenarios and questions; allows for flexibility in what is assessed at particular stations; gives candidates an opportunity to start over with a new interviewer if they have difficulty with any particular station; and may require fewer resources. Several studies have confirmed that the MMI is reliable, valid, and suitable for assessing candidates' skills in non-fact-based knowledge categories such as communication.^{1,4,6}

Rank-based scoring (RBS) is used to score candidates' MMI for the medical programmes at McMaster University, where the MMI was first developed. The RBS approach determines candidates' scores relative to the performance of a pool of candidates.⁷

Scores on interviews can also be based on predetermined standards for specific performance of objectives or skills, also known as criterion-based scoring (CBS).^{8,9} CBS is reported to be better for testing the mastery of skills and to easily encompass items with a greater range of difficulty or ease. When scoring the performance of individuals in a homogeneous group, CBS has better curricular validity than norm-referenced tests, and this approach also performs well in heterogeneous groups.^{8,10,11}

Between 2007 and 2013, the Master of Physical Therapy (MPT) program at the University of Saskatchewan used the RBS system with the MMI; in 2014, we tested the CBS system. The purpose of this study, therefore, was to describe participants' attitudes, beliefs, and experiences with these two methods of scoring the MMI.

METHODS

Our study used a qualitative descriptive inquiry approach that was reflective and interpretive in nature to describe the participants' experiences with scoring the MMI using RBS and CBS. The appeal of a descriptive inquiry perspective is in the pedagogical reflections and the comprehensive summary of events that this approach elicits, which we considered necessary to understand participants' experiences with the two scoring methods.

Context

This study was conducted during the admissions process for the MPT program at the University of Saskatchewan. The MMI involves a circuit of multiple stations through which the candidates rotate. The MMI used for this study had three circuits consisting of the same eight stations (seven scenarios and one rest station); to

accommodate the 96 selected candidates, four cycles of MMIs were completed. Three stations per circuit in all four cycles in the morning and afternoon were designated as research stations, and the same three questions were used throughout the day. During their orientation on MMI day, candidates were informed about the research process and were instructed to ignore the second person in the room while being interviewed at the research stations.

Each station presented a scenario designed to evaluate factors such as the candidate's empathy, integrity, ethical judgment, and overall communication skills. Each candidate had 10 minutes at each station: 2 minutes to read the scenario and 8 minutes to respond. The interviewer then scored the candidate using CBS; at the research stations, observers simultaneously scored candidates using RBS.

Participants

We used an opportunistic sampling strategy, and all participants were drawn from the pool of 27 admissions interviewers (PT faculty, PT clinicians, and 2nd-year MPT students).¹² A letter of invitation was mailed to all potential participants.

Research ethics approval was obtained from the University of Saskatchewan, certifying that the study met national standards for research involving human participants, and informed consent was obtained from each participant. Each participant was paired and stationed with another participant; one served as the active interviewer for the candidate and the other served as an observer for each of the 96 candidates interviewed. Training for both interviewer and observer roles at the research stations was provided before MMI day. The training session included information on the admissions and MMI processes, detailed information on CBS, and practice in the use of both CBS and RBS.

Scoring of the Multiple Mini-Interview

Criterion-based scoring

Participants scored candidates' performance in three categories (critical thinking, communication, and professionalism) on a 10-point scale (ranging from 1–2 = *unsatisfactory/unacceptable* to 9–10 = *exceptional*; see detailed rubric in Appendix 1 online). A detailed rubric, rubric summary, and worksheet were provided to help participants with scoring.

Rank-based scoring

Participants assigned each candidate a single score on the basis of his or her overall performance on the main focus of the case relative to the seven other candidates in that circuit. Participants were encouraged to assign preliminary scores, which they revised later as they assessed the remainder of the candidates in the circuit.

Data collection

The primary data source was semi-structured one-on-one interviews conducted with the participants, by phone or in person, by two of the investigators (see Interview Guide in Appendix 2 online). The interviews were recorded and transcribed verbatim. Participants were given ample opportunity for input, allowing them to elaborate on their experiences and opinions related to the MMI scoring methods. Open-ended questions such as “Could you describe your experience scoring the MMI this year?” and “What do you think about the accuracy of the interview scores?” were used to increase the breadth of responses.¹³ The existing literature on MMI admissions experiences helped in the development of the interview guide.

All admissions interviewers completed assessor feedback forms at the end of the MMI. Data from the participants’ feedback forms were transcribed verbatim and used to further probe into questions addressed during the one-on-one interviews.¹⁴ The full assessor feedback form is available from the lead investigator.

To create an audit trail, detailed notes were recorded at every stage and included descriptions of the research steps taken, from the start of the project to the development and reporting of findings and preliminary thoughts about emerging experiences. These descriptions permitted a conceptual return to the setting during the data analysis.¹⁵

Data analysis

Each investigator completed an inductive line-by-line thematic analysis using QDA Miner 4 Lite (Provalis Research, Montreal). We read the transcripts and notes numerous times.¹⁴ Particularly revealing phrases were highlighted, coded, and assigned meaningful labels. In the event of discrepancies in codes or variance in the labels, a discussion ensued until we reached a common understanding. The data analysis continued by constantly comparing labels and phrases to determine whether they should be classified separately or whether they belonged to an existing code.¹⁶ We then determined the essential or invariant themes, that is, those that gave fundamental meaning to the experiences, as identified by their pattern regularities.¹⁴

Trustworthiness

The truth value or credibility of the data was enhanced via data saturation, that is, by repeating information and verifying previously collected data through subsequent interviews. The first two investigators established rapport with the participants during the MMI training and admissions sessions; they also knew the participants through their prior involvement with the program. As Guba^{17(p.85)} has stated, member checking is “the single most important action inquirers can take, for it goes to the heart of the credibility criterion”; therefore,

we also contacted participants by email and asked them to review their transcripts for accuracy, after which all participants signed a transcript release form. They were also asked to comment on whether they saw their experiences in the themes that emerged from the analysis. All participants responded favourably to the themes, and none suggested changes.

Finally, theoretical triangulation contributed to the study’s credibility because the research question was situated within the current literature on admissions to health professional programmes. Confirmability of our findings was supported by the use of multiple data sources (interview transcripts, assessor feedback data, and investigators’ detailed notes). To further reduce investigator bias, we used investigator triangulation.¹² Together, we possess knowledge of PT, admissions programmes, and qualitative inquiry; to bring dependability to our interpretation of the data, we coded the data independently and determined the emergent themes collaboratively. As noted, we maintained an audit trail (i.e., notes) recording methodological decisions, contextual notes, analytic comments, and information on the processes and progress of the study.¹⁸ The transferability of the findings to other contexts depends on similarities to the program and to the context and nature of the process described.¹² A detailed list of strategies used to ensure trustworthiness is available in Appendix 3 online.

RESULTS

Of the 27 potential participants, 18 confirmed their interest in participating in the research project by submitting signed consent forms, for a response rate of 67%. The 18 study participants (6 faculty members, 6 clinicians, and 6 second-year students) were aged between 23 and 62 years; 12 had prior experience and 6 had no experience with admissions and administering the MMI. All faculty and clinicians held at least a bachelor’s degree in PT; in addition, 2 held master’s degrees and 4 held doctoral degrees.

One major theme, scoring systems, and two sub-themes, CBS and RBS, were identified across all data. Within these themes, overlapping themes and category weighting were also discussed. Pseudonyms (element symbols) are used to identify participants instead of names and are indicated in parentheses at the end of the quote (i.e., *Li*).

Participants were unanimous in recognizing that both interpersonal and intellectual characteristics are important qualities for physical therapists to possess. They also agreed that the MMI consistently captured the profession’s expectations and emphasized that only through the interview can candidates’ personal characteristics be identified. The interview was regarded as a tool to capture essential characteristics of interest and weed out candidates not suitable for the profession:

I think interviews are good and important because I think the things that you measured are important ... they may have worked their tails off and had a good GPA but if they are not able to have good critical thinking, then they may not make a good physical therapist because there will be situations that come up that aren't in the textbook. (Li)

The printed materials (worksheet and rubric) provided were considered useful and appropriate for the tasks at hand. Some participants used these materials more than others; reasons for these variations were interviewer experience, reliance on memory (e.g., particular features such as t-shirt colour), and use of a personal system (e.g., sticky notes).

The participants had mixed views on whether sufficient time was allocated to each station. Although some participants considered an 8-minute interview with 2 minutes to finish off scoring as just right, others saw it as restricting their ability to make an appropriate judgment. Participants reflected on the nature of the station (i.e., some cases were more complex than others), which scoring system they were using, their own experience with interviewing, and the adoption of a tracking method as key to using the time wisely and effectively.

Most participants experienced some degree of physical or mental fatigue over the course of the day, depending on their role. Participants reported that being an observer or an interviewer required more or less concentration, and each role brought particular challenges. Multi-tasking was complex or even unfamiliar, and new participants reported that the process took more time and energy than they had expected it would. In addition, some candidates, or sets of candidates, were more demanding than others, requiring more concentration or effort (e.g., prompting, engaging) from the interviewer, which added to an already demanding job.

Scoring systems

Criterion-based scoring

Participants agreed that the three categories—critical thinking, communication, and professionalism—appropriately captured the necessary outcomes for the profession. For this reason, CBS was assessed as fair in terms of its capacity to evaluate candidates appropriately and accurately:

I thought it was much more accurate. I thought you could get a better feel for the person in the three areas. ... I feel that the CBS does give a more accurate representation of the candidates. (Na)

This scoring system ensured greater objectivity, structure, and consistency. Participants emphasized that CBS enabled them to link these properties with the desirable and sought-after qualities or characteristics. Many participants felt that CBS was both a fair and an effective tool to evaluate affective characteristics.

Furthermore, participants were confident that CBS not only portrayed candidates more accurately but had the ability to discriminate between excellent, good, and poor candidates:

Candidates who were exceptional would likely excel in all three and likewise on the other end if they were unacceptable, then they were likely low in all three. But in the middle is where it really clarified things. (Mg)

However, the accuracy of CBS in discriminating between the categories assessed was not always clear. Professionalism, in particular, was challenging: Although most participants defined professionalism as capturing what it means to be a professional and what is expected of them, the concept is complex and has several dimensions. The difficulty may be rooted in personal beliefs about what constitutes professionalism:

I think the professionalism needs to be perhaps fine-tuned. (Al)

I think that professionalism is extremely important and I think it is great that it has its own category. (Si)

Communication was considered key to the profession. Participants' underlying assumption was that candidates would be highly articulate, with strong communication skills and fluency in English, and have the ability to share and express ideas in individual or group situations, orally or in writing. Language and communication were considered bearers of culture and identity, a powerful instrument and a key ingredient in a candidate's success. One participant noted that communication is far more than a neutral medium for conveying ideas and that mastering communication will have an impact on how physical therapists are perceived by others. Communication, including interpersonal and relationship skills (i.e., professionalism), was without hesitation considered essential to the profession.

Overlapping areas

However, participants also saw communication as overlapping with professionalism. Collectively, professionalism and communication demonstrated candidates' credibility and suitability for the program. One interviewer suggested either amalgamating these categories or breaking them down to better tell them apart:

Professionalism I felt overlapped a lot with particularly the communication one, and a lot of times I judged professionalism based on the communication delivered, so if I am thinking of some of the really poor candidates, they seemed unprofessional because they didn't communicate well. Their body language was poor; they spoke in a really high, rapid tone, without any interaction with the interviewer to see how I was receiving their information. (Cl)

Critical thinking was also considered essential to the profession; participants described it as a process, a way

of thinking, understanding, and expressing oneself. Critical thinking shows that one can accurately and carefully analyze ideas, observations, and experiences and explore the evidence. Participants felt that the critical thinking category was well defined and a stand-alone area; at the same time, however, it was hard to evaluate because participants were looking for specific answers:

It was hard. I think the critical thinking component is really important but it was harder to judge. (Ar)

I think that is a very well-defined criterion in that tool and I think that it quite stands alone (Ca).

Category weighting

Suggesting that CBS categories be weighted generated interesting comments. Assigning weight would imply agreement on the categories to be evaluated and reflect the importance of each category. Some participants opposed attaching a weight to any category, arguing that all are equally important; some were uncertain about weighting the categories, arguing that they could not decide which was most important:

I think it would be very hard to pick which should be weighted more. I think they are all very important when you are looking at candidates trying to get into the School of PT . . . you are going to get a really good overall picture if you weight them the same. (Sc)

I probably would put the most weight on the communication and critical thinking. (Cr)

Rank-based scoring

Although RBS was intended to help participants discriminate among candidates and choose those who present themselves as right for the profession—a worthy objective—participants often felt that the cure was worse than the disease. Overall, they expressed concerns about a lack of fairness, difficulties with multitasking, and ranking a small group of candidates. Participants were generally discontent with the performance of RBS.

There was a strong perception that RBS was not fair to all candidates. Some experienced participants questioned the validity of the scoring system because of the need to rely on a gut feeling:

Needed to base the score on a “gut feeling,” which would be difficult to justify if one had to justify the score you gave. RBS ends up being quite subjective. We had a group of stronger candidates followed by a group of weaker candidates. The ranking of a mediocre student among a group of strong students would result in a low score as compared to ranking a mediocre student among a group of weaker candidates. (Fe)

Some participants firmly believed that RBS did not assess everyone according to their abilities. The number of candidates pooled—only eight per circuit—was brought up as a particular problem; eight is too small a number to do an accurate job.

A secondary problem was that most candidates were forced into the middle. Participants felt that RBS limited the interviewer’s flexibility and was more of an obstacle than a help in selecting appropriate candidates:

There were some candidates that obviously did very well and some that did poorly, but the majority were in the middle. (Ni)

An important requirement of a scoring system is that it be able to discriminate among poor, good, and excellent candidates. Participants considered RBS both inaccurate and unable to discriminate; in general, they disliked RBS:

I think it was good at either end, being the poor or kind of unsatisfactory versus the exceptional, and it got a little bit more difficult with the in-between kind of ones. (Cu)

RBS did not allow interviewers to discriminate among factors such as communication, critical thinking, and professionalism. Personal preferences swayed their answers: Whereas communication was more important for one interviewer, critical thinking was paramount for others:

It was definitely a lot tougher. . . . It was hard because some students may have been really good in one regard, like problem solving or critical thinking, but not as good in communication, so then they got chopped to compare them to students who may have had different strengths. I found it tough to make those comparisons, and then it was up to me which one I valued more out of the three. (Zn)

DISCUSSION

This study explored participants’ attitudes, beliefs, and experiences related to two methods of scoring the MMI, both in detail and in depth. The interviews were not restricted to specific questions, although they were guided by the investigator.

Participants unanimously agreed that CBS was a more suitable method for selecting candidates for admission to the MPT program. Their main concern about using RBS was the difficulty of accurately ranking the small pool of candidates (eight per circuit). RBS also hampered their ability to determine a performance-based individual score for comparison. CBS, however, provided an objective structure and allowed each candidate to be assessed individually. Participants considered CBS to perform better as a whole and to produce a more accurate representation of candidates.

The literature supports the use of non-fact-based categories in making admissions decisions.^{19–21} According to the literature, the categories selected in this case—professionalism, critical thinking, and communication—are appropriate for the PT profession.^{22–28} The findings of this study confirmed that the CBS categories are valid; however, they are also overlapping and inter-related, especially communication and professionalism, and this question merits further investigation.

In the MMI admissions process, the interviewer plays a crucial role. In our study, experience with interviewing influenced the overall admissions experience; it is possible, however, that because this is the first time the program has conducted a qualitative examination after admissions, the opinions of new and more experienced participants are now understood better than before. Participants' level of fatigue, their use of the materials provided, and overall logistics (e.g., interactions with candidates, perceptions of time) were influenced by their past experience. It seems worthwhile to further explore this issue.

Although different cases are usually used for the morning and afternoon cycles, for the purpose of this project the same cases were used all day at the research stations. Participants expressed concerns about this resulting in fatigue and an inability to maintain the same standard throughout the day. Although the candidates were allocated randomly to the different MMI cycles, several participants agreed that the stronger candidates were in the morning's circuits. Although this would create a problem in an admissions process relying on RBS, CBS solves this problem because candidates are assessed individually rather than relative to the rest of their assigned circuit—that is, whether they are in a strong or a weak circuit, CBS should produce their true score.

Professionalism, communication, and critical thinking are thought to underpin success in a PT career.^{22–28} These categories are increasingly considered to be as important as academic credentials in determining professional and academic outcomes.^{22,27,29} Using CBS, participants evaluated candidates according to several criteria in these three categories, which were chosen with the goal of finding the right candidates with the highest potential, not just those with the best test scores (i.e., grade-point average). It is important to investigate the predictive value of the new scoring system (CBS in the MMI) in determining the likelihood of a candidate's completing the program successfully. In addition, the selection profile should be compared from year to year to determine what, if any, adjustments may need to be made.

Our study has several limitations. First, some of the study participants assisted in developing one of the cases, which may have biased their assessment of candidates' responses. Second, our analysis did not consider participants' subjectivity or the difficulty of the various stations.

CONCLUSION

To our knowledge, this is the first study to investigate CBS in the MMI. CBS was well accepted by participants, the majority of whom preferred it over the already-known RBS method. Participants felt that CBS evaluated a valid set of non-fact-based characteristics and, in the

end, presented a more accurate and fair depiction of the candidates. Participants agreed that CBS has potential as a method to identify and evaluate unique non-fact-based characteristics of the PT profession not identified by RBS and that it would assist in the PT admissions process. Although this study only included the MMI for admission to the MPT program at the University of Saskatchewan, the findings may be generalized to other health-related admissions processes in Canada.

KEY MESSAGES

What is already known on this topic

The Multiple Mini-Interview (MMI) has been shown to be reliable, valid, and suitable for assessing candidates' skills in non-fact-based knowledge categories such as communication.^{1,4,6}

What this study adds

Participants considered the criterion-based system to be more accurate and fair than the rank-based system for scoring the MMI. Criterion-based scoring also identifies desirable characteristics for the physical therapy profession. The findings of this study suggest that the use of criterion-based scoring in Master of Physical Therapy admission MMIs should be expanded.

REFERENCES

1. Uijtdehaage S, Doyle L, Parker N. Enhancing the reliability of the Multiple Mini-Interview for selecting prospective health care leaders. *Acad Med*. 2011;86(8):1032–9. <http://dx.doi.org/10.1097/ACM.0b013e3182223ab7>. Medline:21694560
2. Eva KW, Rosenfeld J, Reiter HI, et al. An admissions OSCE: the Multiple Mini-Interview. *Med Educ*. 2004;38(3):314–26. <http://dx.doi.org/10.1046/j.1365-2923.2004.01776.x>. Medline:14996341
3. Till H, Myford C, Dowell J. Improving student selection using multiple mini-interviews with multifaceted Rasch modeling. *Acad Med*. 2013;88(2):216–23. <http://dx.doi.org/10.1097/ACM.0b013e31827c0c5d>. Medline:23269299
4. Roberts C, Walton M, Rothnie I, et al. Factors affecting the utility of the Multiple Mini-Interview in selecting candidates for graduate-entry medical school. *Med Educ*. 2008;42(4):396–404. <http://dx.doi.org/10.1111/j.1365-2923.2008.03018.x>. Medline:18338992
5. Harasym PH, Woloschuk W, Mandin H, et al. Reliability and validity of interviewers' judgments of medical school candidates. *Acad Med*. 1996;71(1 Suppl):S40–2. <http://dx.doi.org/10.1097/00001888-199601000-00038>. Medline:8546779
6. Lemay JF, Lockyer JM, Collin VT, et al. Assessment of non-cognitive traits through the admissions Multiple Mini-Interview. *Med Educ*. 2007;41(6):573–9. <http://dx.doi.org/10.1111/j.1365-2923.2007.02767.x>. Medline:17518837
7. Admissions, Undergraduate Medical Program, Michael G. DeGroote School of Medicine. Manual for interviewers [Internet]. Hamilton (ON): McMaster University; 2013 [cited 2013 Jan]. Available from: <http://fhs.mcmaster.ca/mdprog/interviews.html>.
8. Ornstein AC. Norm-referenced and criterion-based tests: An overview. *NASSP Bull*. 1993;77(555):28–39. <http://dx.doi.org/10.1177/019263659307755505>.
9. Montgomery PC, Connolly BH. Norm-referenced and criterion-referenced tests. Use in pediatrics and application to task analysis of motor skill. *Phys Ther*. 1987;67(12):1873–6. Medline:3685115

10. Popham JW. Educational evaluation. Englewood Cliffs (NJ): Prentice-Hall; 1975.
11. Leonard J. How group composition influenced the achievement of sixth-grade mathematics students. *Math Think Learn*. 2001;3(2-3):175-200. <http://dx.doi.org/10.1080/10986065.2001.9679972>.
12. Creswell JW. Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks (CA): Sage; 2006.
13. Fontana A, Frey JH. Interviewing: the art of science. In: Denzin NK, Lincoln YS, editors. *Handbook of qualitative research*. Thousand Oaks (CA): Sage; 1994. p. 361-75.
14. van Manen M. Researching lived experience: human science for an action sensitive pedagogy. London (ON): Althouse Press; 1997.
15. Bogdan RC, Biklen SK. Qualitative research for education: An introduction to theory and methods. 2nd ed. Boston: Allyn & Bacon; 2003.
16. Wolcott HF. Writing up qualitative research. 2nd ed. Thousand Oaks (CA): Sage; 2001. p. 12-48.
17. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. *Educ Tech Res*. 1981;29(2):75-91.
18. Meadows LM, Morse JM. Constructing evidence within the qualitative project. In: Morse JM, Swanson JM, Kuzel J, editors. *The nature of qualitative evidence*. Thousand Oaks (CA): Sage; 2001. p. 187-200.
19. Grice KO. Use of multiple mini-interviews for occupational therapy admissions. *J Allied Health*. 2014;43(1):57-61. Medline:24598901
20. 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):10. <http://dx.doi.org/10.5688/ajpe76110>. Medline:22412209
21. Oliver T, Hecker K, Hausdorf PA, et al. Validating MMI scores: are we measuring multiple attributes? *Adv Health Sci Educ Theory Pract*. 2014;19(3):379-92. <http://dx.doi.org/10.1007/s10459-013-9480-6>. Medline:24449121
22. Vendrely AM. An investigation of the relationships among academic performance, clinical performance, critical thinking, and success on the physical therapy licensure examination. *J Allied Health*. 2007;36(2):e108-23. Medline:19759986
23. Roberts LC, Whittle CT, Cleland J, et al. Measuring verbal communication in initial physical therapy encounters. *Phys Ther*. 2013;93(4):479-91. <http://dx.doi.org/10.2522/ptj.20120089>. Medline:23197846
24. Pinto RZ, Ferreira ML, Oliveira VC, et al. Patient-centred communication is associated with positive therapeutic alliance: a systematic review. *J Physiother*. 2012;58(2):77-87. [http://dx.doi.org/10.1016/S1836-9553\(12\)70087-5](http://dx.doi.org/10.1016/S1836-9553(12)70087-5). Medline:22613237
25. Black LL, Jensen GM, Mostrom E, et al. The first year of practice: an investigation of the professional learning and development of promising novice physical therapists. *Phys Ther*. 2010;90(12):1758-73. <http://dx.doi.org/10.2522/ptj.20100078>. Medline:20930050
26. Øien AM, Steihaug S, Iversen S, et al. Communication as negotiation processes in long-term physiotherapy: a qualitative study. *Scand J Caring Sci*. 2011;25(1):53-61. <http://dx.doi.org/10.1111/j.1471-6712.2010.00790.x>. Medline:20384974
27. Cahalin LP. The Linda Crane lecture professionalism & core values in physical therapy: lessons learned from Linda Crane. *Cardiopulm Phys Ther J*. 2012;23(2):30-9. Medline:22833707
28. Aguilar A, Stupans I, Scutter S, et al. Exploring the professional values of Australian physiotherapists. *Physiother Res Int*. 2013;18(1):27-36. <http://dx.doi.org/10.1002/pri.1525>. Medline:22585620
29. Kelland K, Hoe E, McGuire MJ, et al. Excelling in the role of advocate: a qualitative study exploring advocacy as an essential physiotherapy competency. *Physiother Can*. 2014;66(1):74-80. <http://dx.doi.org/10.3138/ptc.2013-05>. Medline:24719513