

## Research Article

# Correlation between Global Rating Scale and Specific Checklist Scores for Professional Behaviour of Physical Therapy Students in Practical Examinations

Kaitlin Turner,<sup>1</sup> Maegan Bell,<sup>1</sup> Lindsay Bays,<sup>1</sup> Carmen Lau,<sup>1</sup> Clara Lai,<sup>1</sup>  
Tetyana Kendzerska,<sup>2</sup> Cathy Evans,<sup>3</sup> and Robyn Davies<sup>3,4</sup>

<sup>1</sup> Physical Therapy Program, University of Toronto, 500 University Avenue, Toronto, ON, Canada M5G 1V7

<sup>2</sup> Faculty of Medicine, University of Toronto, 500 University Avenue, Toronto, ON, Canada M5G 1V7

<sup>3</sup> Department of Physical Therapy, University of Toronto, 500 University Avenue, Toronto, ON, Canada M5G 1V7

<sup>4</sup> Sunnybrook Health Sciences Centre, 2057 Bayview Avenue, Toronto, ON, Canada M4N 3M5

Correspondence should be addressed to Kaitlin Turner; [kaitlin@foundationphysio.com](mailto:kaitlin@foundationphysio.com)

Received 6 April 2014; Accepted 28 August 2014; Published 17 September 2014

Academic Editor: Karl Kingsley

Copyright © 2014 Kaitlin Turner et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The purpose of this study was to determine whether or not the specific item checklist (checklist) and global rating scale (GRS) scores are correlated in practical skills examinations (PSEs). Professional behaviour was evaluated using both the checklist and GRS scores for 183 students in three PSEs. Mean, standard deviation, and correlation for checklist and GRS scores were calculated for each station, within each PSE. Pass rate for checklist and GRS was determined for each PSE, as well as for each individual checklist item within each PSE. Overall, pass rate was high for both checklist and GRS evaluations of professional behaviour in all PSEs. Generally, mean scores for the checklist and GRS were high, with low standard deviations, resulting in low data variability. Spearman correlation between total checklist and GRS scores was statistically significant for two out of five stations in PSE 1, five out of six stations in PSE 2, and three out of four stations in PSE 3. The GRS is comparable to the checklist for evaluation of professional behaviour in physical therapy (PT) students. The correlation between the checklist and GRS appears to become stronger in the assessment of more advanced students.

## 1. Introduction

The “Essential Competency Profile for Physiotherapists (*the terms “Physiotherapist” and “Physical Therapist” are synonymous in Canada*)<sup>†</sup> in Canada” defines a professional as one that “[is] committed to the best interest of clients and society through ethical practice, support[s] profession-led regulation, and [has] high personal standards of behaviour ([1, page 14]).” Professional behaviour is the outward display, through actions, of the inner attitudes and values a person possesses [2]. Based on a comprehensive professionalism literature review, Arnold [3] identified specific elements of professional behaviour as altruism, reliability and responsibility, honesty and integrity, respect, and communication and interpersonal skills. In an internationally developed consensus statement on professionalism Hodges et al. [4] concluded

that professional behaviour is complex and multidimensional encompassing three scopes: individual (attributes and behaviours), interpersonal (interactions with others and in contexts), and macrosocietal (social responsibility, political platforms, economics, and moral). It is essential for practicing healthcare clinicians to behave in a manner that embraces all elements and scopes of professionalism. Employers report that students appear to have more difficulty with affective skills (including professional behaviour) compared to psychomotor or cognitive skills (clinical skills and reasoning) [5]. Therefore, it is important to evaluate students’ professional behaviour in clinical scenarios, such as Objective Structured Clinical Examinations (OSCEs), prior to commencement of clinical internships.

The OSCE was originally developed to evaluate medical students’ clinical competence and is accepted as a valid and

reliable method for such evaluation [6–8]. OSCEs involve the assessment of clinical skills and behaviours throughout series timed stations requiring demonstration of practical skills, problem solving strategies, and behaviours. Typically, performance is evaluated using checklists, either alone or in combination with a GRS [7, 9, 10].

Checklists are commonly used in OSCEs as they are believed to minimize examiner subjectivity [9, 11]. Whole skills are broken down into smaller measurable parts such that the examiner can mark either “yes” or “no” based on whether the skill was performed or not [11]. Checklists evaluate the thoroughness of performance such that greater competence of a skill is evidenced by completion of a greater number of checklist items [9, 11]. The GRS, in contrast, is a nonbinary evaluative tool requiring assignment of a score along a predetermined scale of generally five or seven points [12]. Some evidence supports use of the GRS over checklists [9, 11–15]. Checklists encourage a stepwise approach to problem solving and evaluate knowledge and skills that are easily objectified, while potentially ignoring the subtle nuances, such as empathy and respect, which are part of a competently completed station [9, 13]. Also, checklists may favour novice learners who often adopt a stepwise approach to problem solving [9, 11]. The GRS, alternatively, does not limit evaluators to specific items or promote memorization of required behaviours and may allow for recognition of higher levels of clinical competence, as it permits evaluators to appraise multiple approaches to problem solving and reward students who excel in skill performance [9, 11]. However, the definition of professional behaviour is broad and lacks accord, allowing raters to place varied importance on different characteristics and behaviour when evaluating students [16]. Despite this, the GRS is thought to have stronger psychometric properties even though the checklist may appear less biased [9, 14, 15].

In evaluating affective domain behaviours in OSCEs there is evidence that the GRS may be equally if not more reliable and valid than checklists [9, 14, 17, 18]. The GRS has been demonstrated to be superior to checklists in terms of interstation reliability and construct and concurrent validity [8–12, 17]. In the OSCE, domains of professional behaviour, such as communication, are often broken down into discrete observable items that reflect professional competence, such as coherence and nonverbal expression. Van der Vleuten et al. [19] summarized medical education literature and concluded that “objectified” methods of evaluation, such as checklists, do not necessarily lead to greater reliability. Currently, limited research is available to support the use of the GRS to evaluate professional behaviour in a PT population as psychometric properties of checklists and the GRS have been investigated primarily in medical students and residents [8–11, 14]. In addition, considerable variation in methodology exists across available studies, including whether the same, or different, evaluators completed both the checklist and GRS, varying complexity of the skills performed, and the different domains assessed (professional behavior versus clinical skills) [10, 12]. Additionally, most studies recruited volunteers and used one examination for students of multiple education levels [10, 11]. One study used PT students; however only the checklist was evaluated, and results demonstrated poor interstation

reliability and poor predictability of performance during subsequent clinical internships [20]. More evidence is needed to support the growing belief that the GRS may be a suitable alternative to the checklist for evaluating the affective domain of OSCEs, given that several academic institutions have begun to adopt the GRS into their OSCEs [9, 15, 17].

The PT program at the University of Toronto (Canada) is one such program that evaluates students’ clinical skills and professional behaviors with a PSE, similar to an OSCE, using both checklists and GRSs across the two-year program. Currently, the same checklist and GRS are used to evaluate professional behaviour at each station, within each PSE, across the two-year program. A passing score on the professional behaviour section is not required to pass the station overall; however, it can contribute to decisions regarding the final outcome of a student’s PSE.

The purpose of this study was to determine the correlation between checklist and GRS methods of professional behaviour evaluation in University of Toronto PT students during their PSEs. Study findings may provide insight into the usefulness and appropriateness of checklist items and current pass rates on the professional behaviour section of the PSEs.

## 2. Methods

Data from only the professional behavior component, from each PSE that occurred over the study period, were extracted and examined at the individual level. Of the three PSEs used in this study, PSE 1 tested first year cardiorespiratory skills with five stations, PSE 2 tested first year musculoskeletal skills with six stations, and PSE 3 tested second year musculoskeletal skills with four stations. All identifying information was removed from the score sheets, and the data was aggregated. The protocol for the study was submitted to the University of Toronto, Board of Ethics, and approved in October 2011 (number 27021).

*2.1. Participants.* All students enrolled in the University of Toronto PT program who completed at least one of three PSEs, between November 2011 and April 2012, were included in the study. A total of thirty-four evaluators examined students in the three PSEs. All evaluators were experienced clinicians who received individual training, in scoring and evaluation, from the course faculty prior to PSEs, as is the standard practice of the program.

*2.2. Measures.* A checklist consisting of seven items was used at each station, within each PSE to evaluate students’ professional behaviour. Students were expected to complete the items when interacting with standardized patients at each station and were graded with either a “yes,” indicating they completed and passed the item, or a “no,” indicating they did not complete and failed the item (Table 1). A score greater than or equal to 5/7 items marked as “yes” was considered a pass.

A GRS consisting of a five-item scale was used to evaluate students’ professional behaviour as a whole throughout the entirety of each station, within each PSE. In the GRS,

TABLE 1: Professional behaviour checklist items for PSEs.

Professional conduct			
1	Greets patient and introduces him/herself	Yes	No
2	Gives “lay” explanation about what he/she is about to do (including rationale)	Yes	No
3	Obtains verbal permission to proceed	Yes	No
4	Demonstrates ongoing monitoring of patient response and comfort throughout interaction	Yes	No
5	Verbal commands are appropriate in type and timing	Yes	No
6	Closes the session appropriately	Yes	No
7	Treats others with positive regard, dignity, respect, and compassion	Yes	No

TABLE 2: Professional behaviour global rating scale for PSEs. Based on your observation of the student’s performance at this station, please complete the following global rating scales for professional conduct, safety, and effectiveness as well as the overall global rating that indicates pass, fail, or pass with distinction.

Professional conduct (including communication, respect, compassion, and ensuring patient dignity)				
1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
Not able to demonstrate professional interaction.		Able to demonstrate professional interaction.		Able to <i>thoroughly</i> demonstrate professional interaction.

professional behaviour was inclusive of students’ “communication, respect, compassion, and insurance of patient dignity” (Table 2). A score greater than or equal to 3/5 was considered a pass.

2.3. *Data Collection.* Data used in the study were collected as standard evaluation component of the curriculum. Examiners completed score sheets, including a professional behaviour section, to evaluate students at each station, within each PSE. Examiners completed evaluations on paper score sheets at four of the five stations in PSE 1, four of the six stations in PSE 2, and one of the four stations in PSE 3. The remaining stations in the PSEs piloted the use of iPads (Apple, Cupertino, CA) for student evaluations. Identical electronic versions of the score sheets were uploaded onto the iPads (Apple, Cupertino, CA) for scoring and paper copies were provided as a backup.

Upon completion of each PSE, a faculty member collected all score sheets, removed identifying information, and extracted the professional behaviour section from the score sheets for analysis. Data collection was completed in November 2011 for PSE 1, April 2012 for PSE 2, and March 2012 for PSE 3.

2.4. *Data Analysis.* Checklist and GRS scores from the professional behaviour section of each PSE were entered into Microsoft Excel (Microsoft Corp., Redmond, WA) and transferred to SPSS, Version 20.0 (SPSS Inc., Chicago, IL) for statistical analysis. The pass rate for each checklist item (1–7), within each PSE, was calculated by dividing the number of students that received a “yes” on a checklist item by

the total number of corresponding completed checklist items. The pass rate for the checklist and GRS, within each PSE, was calculated by dividing the number of students that passed each evaluation method by the total number of completed checklist and GRS scores, respectively. The mean and standard deviation of the raw checklist and GRS scores were computed to look at the distribution of the scores for each station, within each PSE. Spearman correlation was used to examine the association between the checklist and GRS scores, within each PSE. Spearman correlation was used due to low variability in the data.

### 3. Results

A total of 183 University of Toronto PT students who completed PSEs between November 2011 and April 2012 were included in this study. Evaluations were collected for 101 students in PSE 1 from five stations, resulting in 505 checklists and 505 GRS scores; 104 students in PSE 2 from six stations, resulting in 624 checklists and 624 GRS scores; and 79 students in PSE 3 from four stations, resulting in 316 checklists and 316 GRS scores. With all PSEs combined, the final data set consisted of 1445 checklists and 1445 GRS scores; however, secondary to incomplete evaluations and/or technical problems, sample sizes varied depending on the variables to be analyzed.

The pass rates for checklist items one through five were greater than 90% for all PSEs. The pass rates for item six were 100% (PSE 1), 94% (PSE 2), and 79% (PSE 3). The pass rates for item seven were 90% (PSE 1), 95% (PSE 2), and 100% (PSE 3). Additionally, the pass rates for both the checklist and GRS were 100% (PSE 1), greater than 99% (PSE 2), and greater than or equal to 98% (PSE 3).

Mean professional behaviour scores for each station, within each PSE, were greater than or equal to 6.7 for the checklist and 3.8 for the GRS in PSE 1, 6.6 for the checklist and 3.9 for the GRS in PSE 2, and 6.6 for the checklist and 3.7 for the GRS in PSE 3 (Table 3).

With all stations combined 46% of students received 100% on both the checklist and GRS in PSE 1, and 42% of students received 100% on both the checklist and GRS in PSE 2. In PSE 3, 19% of students received 100% on both the checklist and GRS, with all stations combined. Additionally, 49% received a perfect score on the checklist and scored four on the GRS. Spearman correlation coefficients show a higher correlation

TABLE 3: Mean and standard deviation for checklist (/7) and GRS (/5) scores per station for each PSE.

Station	Evaluation	PSE 1		PSE 2		PSE 3	
		Mean ( <i>n</i> )	SD	Mean ( <i>n</i> )	SD	Mean ( <i>n</i> )	SD
1	Checklist	6.84 (99)	±0.37	6.82 (103)	±0.44	6.58 (65)	±0.68
	GRS	4.80 (101)	±0.40	4.40 (100)	±0.65	4.17 (64)	±0.52
2	Checklist	6.92 (97)	±0.28	6.78 (103)	±0.42	6.55 (78)	±0.70
	GRS	4.37 (101)	±0.64	4.77 (103)	±0.43	4.01 (75)	±0.74
3	Checklist	6.70 (97)	±0.52	6.63 (103)	±0.61	6.76 (74)	±0.59
	GRS	3.76 (99)	±0.85	4.06 (99)	±0.74	4.11 (76)	±0.78
4	Checklist	6.80 (98)	±0.41	6.90 (103)	±0.41	6.65 (79)	±0.51
	GRS	4.46 (98)	±0.58	4.69 (100)	±0.55	3.70 (79)	±0.65
5	Checklist	6.93 (99)	±0.30	6.73 (103)	±0.60		
	GRS	4.45 (100)	±0.69	4.38 (100)	±0.55		
6	Checklist			6.66 (102)	±0.59		
	GRS			3.88 (102)	±0.57		

TABLE 4: Correlation between checklist and GRS scores within each station for each PSE.

Station	PSE 1		PSE 2		PSE 3	
	rho ( <i>n</i> )	<i>P</i> value	rho ( <i>n</i> )	<i>P</i> value	rho ( <i>n</i> )	<i>P</i> value
1	-0.15 (99)	0.13	0.27 <sup>*a</sup> (100)	0.007	-0.08 (64)	0.54
2	0.07 (97)	0.48	0.31 <sup>*a</sup> (103)	0.001	0.62 <sup>*b</sup> (75)	<0.001
3	-0.05 (96)	0.65	0.45 <sup>*a</sup> (99)	<0.001	0.57 <sup>*b</sup> (74)	<0.001
4	0.41 <sup>*a</sup> (95)	<0.001	0.09 (99)	0.40	0.37 <sup>*a</sup> (79)	0.001
5	0.33 <sup>*a</sup> (98)	0.001	0.38 <sup>*a</sup> (100)	<0.001		
6			0.39 <sup>*a</sup> (102)	<0.001		

\*Correlations significant at  $\leq 0.01$  level (2-tailed) [20].

<sup>a</sup>0.25–0.50 fair correlation [20].

<sup>b</sup>0.51–0.75 moderate to good correlation [20].

between checklist and GRS as the student progressed through the program (Table 4).

#### 4. Discussion

This is the first known study to examine the evaluation of professional behaviours using checklist and GRS methods within the PT student population during OSCE-type examinations. In this study, mean scores for the checklist and GRS evaluation methods were typically high and showed a small degree of variation for all stations, within each PSE, demonstrating a large ceiling effect of both evaluation methods. In contrast, previous studies involving medical students reported lower mean scores and greater score variation for both checklist and GRS [10, 12]. Many factors may have contributed to the discrepancy of scores observed in this study, including differences in study methodology, variation in evaluated domains (e.g., professional behaviour versus clinical skills), station complexity, and participant recruitment [10, 12]. Also, participants in this study had the advantage of being familiar with the specific items on the checklist and the GRS prior to completing their first PSE. As students progress through the program the same measures of professional behaviour are used during PSEs allowing them to incorporate known markers for a professional performance into each station.

The pass rate for some individual checklist items of professional behaviour had varying patterns across each PSE. For example, percentage of students who passed item six (“Closes the session appropriately”) decreased as level of education increased. In PSE 3, completed by final year students, it may be possible that students ran out of time at more complex stations or placed a lower priority on completing this item. The trend seen in pass rate of this item may also reflect the systematic approach adopted by novice students in OSCEs [15, 17, 21]. Alternatively, the pass rate for item seven (“Treats others with positive regard, dignity, respect, and compassion”) increased as the level of education increased. Professional behaviour may be developed during clinical internships, thus becoming more natural and innate during any kind of patient interaction. This was consistent with the results of the study by Hochberg et al. [22], which demonstrated that checklist items of professional behaviour, such as “respect for patient values,” improved over time in surgical residents [22]. Additionally, professional behaviour is poorly operationalized [16], which may have contributed to the varied pass rates across checklist items within each PSE.

Generally, correlation between checklist and GRS professional behaviour scores increased as education level increased; however, overall pass rates for both checklist and GRS were the highest in PSE 1. Nearly half of the students that completed PSE 1 received the highest possible checklist and

GRS scores. Despite the perfect pass rate and high proportion of perfect checklist and GRS scores, the correlation between checklist and GRS scores for PSE 1 was only significant for stations four and five, and these correlations were only considered fair (Table 4) [23]. The correlation between the checklist and GRS scores for PSE 2 was significant and moderate for five of the six stations; correlation at station four was not significant (Table 4) [23]. In PSE 3, the checklist and GRS scores demonstrated significant and moderate to good correlation in stations two and three and significant and fair correlation in station four (Table 4) [23]. Multiple factors may affect the correlation. Those stations that did not correlate well may have involved more complex skills or problem solving strategies, potentially leading to completion of the clinical skill at the peril of professional behaviour items. Alternatively, examiner bias or error may have accounted for the differences as well.

In summary, given that performance on checklist items is typically consistent throughout all PSEs and that correlation between the checklist and GRS improves as students gain experience, the checklist may be a redundant tool in PSEs later in the program. Checklists may still be useful early in the program, as they incorporate the stepwise approach adopted by novice learners [9, 11]. The checklist also provides examples of foundational professional behaviour for students to bring into their first clinical internship. As students incorporate these behaviours into clinical practice, through multiple internships, the checklist items may become unsuitable for evaluations as professional interactions become more innate. Therefore, as students advance through the program and gain more experience, the GRS may be a more desirable tool to evaluate professional behaviour.

There were some limitations to this study, including the fact that data were analyzed cross-sectionally between two cohorts of students from only one university: first year students in PSE 1 and PSE 2 and second year students in PSE 3. The data was collected from one Canadian program, and, as such, programs using different methods and items for evaluating professional behaviour need to use caution when using results. Additionally, since this study looked at current practice for PSEs, we were unable to control evaluator setup and training. Each PSE utilized different clinicians to evaluate students, and one examiner scored both the checklist and GRS at each station. Since one examiner completed both the checklist and GRS, it is plausible that scoring the checklist prior to the GRS may have influenced GRS scores, such that students that performed better on the checklist may have scored higher on the GRS. However, Regehr et al. [11] demonstrated that reliability and validity of the GRS were not influenced by either the presence or absence of a checklist [11].

Although this study had some limitations, it can serve as a foundation for further research in evaluating the assessment tools used in OSCE-type examinations for PT students. The large ceiling effect of both the checklist and GRS evaluations of professional behaviour indicates the need to develop more discriminating methods of evaluating and defining professional behaviour. Incorporating clinician's feedback on professional behaviour during students' internships or conducting focus groups with clinical instructors may help

identify common issues and themes that could be integrated into the PSE professional behaviour evaluation. As well, understanding the examiners' rationale behind scoring the GRS, and how interpretation of professional behaviour differs among evaluators may aid in the operationalization of this construct. This may provide insight into variations seen among evaluators on GRS scores, which has been discussed in recent research [16]. The addition of a specific station requiring students to demonstrate professional behaviour, through a challenging interaction with a patient, may provide more insight into components of professional behaviour students commonly struggle with. Taking this proactive approach will assist in skill development prior to the start of the next clinical internship. Additionally, OSCE-type examinations only allow for evaluation of professional behavior at the individual level; future research should include evaluation of professional behaviour of PT students in clinical settings to incorporate the interpersonal and macrosocietal scopes as suggested by Hodges et al. [4].

## 5. Conclusion

The results of this study indicate that the GRS is a suitable summative measure and is comparable to the checklist, when assessing PT students on professional behaviour in PSEs. The pass rates and mean scores for professional behaviour were consistently high for both evaluation methods across the entire duration of the program. As students progressed through the program and continued to gain more knowledge and clinical experience, correlation between the checklist and GRS professional behaviour scores appeared stronger.

Given some of the challenges associated with checklists, such as memorization and their dichotomous nature, the GRS should be considered a strategy to augment evaluation of learners in practical skills exams.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

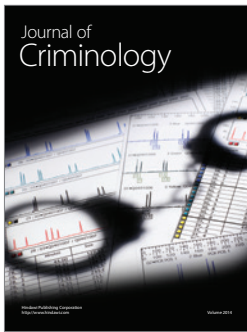
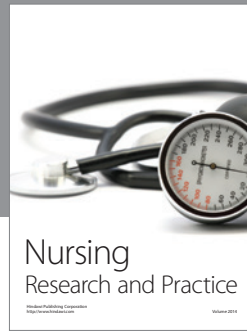
## Acknowledgments

The research team would like to acknowledge the support of Department of Physical Therapy at the University of Toronto, Toronto, Canada. This research was completed in partial fulfillment of the requirements for an MScPT degree at the University of Toronto, Toronto, Canada.

## References

- [1] Canadian Physiotherapy Association, "Essential Competency Profile for Physiotherapists in Canada," October 2009, <http://www.physiotherapyeducation.ca/Resources/Essential%20Comp%20PT%20Profile%202009.pdf>.
- [2] W. N. K. A. van Mook, S. J. van Luijk, H. O'Sullivan et al., "The concepts of professionalism and professional behaviour: conflicts in both definition and learning outcomes," *European Journal of Internal Medicine*, vol. 20, no. 4, pp. e85–e89, 2009.

- [3] L. Arnold, "Assessing professional behavior: yesterday, today, and tomorrow," *Academic Medicine*, vol. 77, no. 6, pp. 502–515, 2002.
- [4] B. D. Hodges, S. Ginsburg, R. Cruess et al., "Assessment of professionalism: recommendations from the Ottawa Conference," *Medical Teacher*, vol. 33, no. 5, pp. 354–363, 2011.
- [5] E. Mostrom, K. W. Hayes, G. Huber, J. Rogers, and B. Sanders, "Behaviors that cause clinical instructors to question the clinical competence of physical therapist students: invited commentary," *Physical Therapy*, vol. 79, no. 7, pp. 668–671, 1999.
- [6] S. S. Jain, S. Nadler, M. Eyles, S. Kirshblum, J. A. Delisa, and A. Smith, "Development of an objective structured clinical examination (OSCE) for physical medicine and rehabilitation residents," *The American Journal of Physical Medicine and Rehabilitation*, vol. 76, no. 2, pp. 102–106, 1997.
- [7] R. M. Harden and F. A. Gleeson, "Assessment of clinical competence using an objective structured clinical examination (OSCE)," *Medical Education*, vol. 13, no. 1, pp. 39–54, 1979.
- [8] G. Regehr, R. Freeman, B. Hodges, and L. Russell, "Assessing the generalizability of OSCE measures across content domains," *Academic Medicine*, vol. 74, no. 12, pp. 1320–1322, 1999.
- [9] J. P. W. Cunningham, A. J. Neville, and G. R. Norman, "The risks of thoroughness: reliability and validity of global ratings and checklists in an OSCE," *Advances in Health Sciences Education*, vol. 1, no. 3, pp. 227–233, 1996.
- [10] B. Hodges and J. H. McIlroy, "Analytic global OSCE ratings are sensitive to level of training," *Medical Education*, vol. 37, no. 11, pp. 1012–1016, 2003.
- [11] G. Regehr, H. MacRae, R. K. Reznick, and D. Szalay, "Comparing the psychometric properties of checklists and global ratings scales for assessing performance on an OSCE-format examination," *Academic Medicine*, vol. 73, no. 9, pp. 993–997, 1998.
- [12] R. K. Reznick, G. Regehr, G. Yee, A. Rothman, D. Blackmore, and D. Dauphinée, "Process-rating forms versus task-specific checklists in an OSCE for medical licensure," *Academic Medicine*, vol. 73, no. 10, pp. S97–S99, 1998.
- [13] K. Cox, "No Oscar for OSCA," *Medical Education*, vol. 24, no. 6, pp. 540–545, 1990.
- [14] S. van Luijk and C. van der Vleuten, "A comparison of checklists and rating scales in performance-based testing," in *Current Developments in Assessing Clinical Competence Montreal*, I. Hart, R. Harden, and J. des Marchais, Eds., pp. 357–382, Can-Heal Publications, 1992.
- [15] R. Cohen, A. I. Rothman, P. Poldre, and J. Ross, "Validity and generalizability of global ratings in an objective structured clinical examination," *Academic Medicine*, vol. 66, no. 9, pp. 545–548, 1991.
- [16] B. Hodges, G. Regehr, M. Hanson, and N. McNaughton, "Validation of an objective structured clinical examination in psychiatry," *Academic Medicine*, vol. 73, no. 8, pp. 910–912, 1998.
- [17] M. Zanetti, L. Keller, K. Mazor et al., "Using standardized patients to assess professionalism: a generalizability study," *Teaching and Learning in Medicine*, vol. 22, no. 4, pp. 274–279, 2010.
- [18] G. Regehr, R. Freeman, A. Robb, N. Missiha, and R. Heisey, "OSCE performance evaluations made by standardized patients: comparing checklist and global rating scores," *Academic Medicine*, vol. 74, no. 10, pp. S135–S137, 1999.
- [19] C. P. van der Vleuten, G. R. Norman, and E. de Graaff, "Pitfalls in the pursuit of objectivity: issues of reliability," *Medical Education*, vol. 25, no. 2, pp. 110–118, 1991.
- [20] J. Wessel, R. Williams, E. Finch, and M. Gémus, "Reliability and validity of an objective structured clinical examination for physical therapy students," *Journal of Allied Health*, vol. 32, no. 4, pp. 266–269, 2003.
- [21] J. Marshall, "Assessment of problem-solving ability," *Medical Education*, vol. 11, no. 5, pp. 329–334, 1977.
- [22] M. S. Hochberg, A. Kalet, S. Zabar, E. Kachur, C. Gillespie, and R. S. Berman, "Can professionalism be taught? Encouraging evidence," *The American Journal of Surgery*, vol. 199, no. 1, pp. 86–93, 2010.
- [23] T. Colton, *Statistics in Medicine*, Little, Brown, Boston, Mass, USA, 1974.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

