

Investigating the Reliability and Validity of Self and Peer Assessment to Measure Medical Students' Professional Competencies

Tyrone Donnon¹, Joann McIlwrick², Wayne Woloschuk²

¹Office of Health Medical Education Scholarship, Faculty of Medicine, University of Calgary, Calgary, Canada

²Office of Undergraduate Medical Education, Faculty of Medicine, University of Calgary, Calgary, Canada
Email: tldonnon@ucalgary.ca

Received April 10th, 2013; revised May 12th, 2013; accepted May 20th, 2013

Copyright © 2013 Tyrone Donnon 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 use of peer assessment through a multisource feedback process has gained recognition as a reliable and valid method to assess the characteristics of professionals and trainees. A total of 168 first-year medical students completed a 15-item questionnaire to self-assess their professional work habits and interpersonal abilities. Each student was expected to identify 8 first-year classmates to complete a corresponding 15-item peer assessment. Although the self and peer assessment questionnaires had strong reliability (Cronbach's $\alpha = 0.85$ and 0.91 , respectively), an exploratory factor analysis resulted in a 3- and 2-factor solution, respectively. The third factor was associated with items related to students' *personal attributes*. Significantly lower mean score differences for the self-report assessment were found for all 15 items (Cohen's $d = 0.27$ to 1.39 , $p < 0.001$). A decision study analysis found that 7 peer assessors were needed to achieve a generalizability coefficient of 0.70. The findings suggest some inconsistencies in regards to the construct validity and stability of measures between self and peer assessment measures. The need for self-awareness of students' strengths and limitations, however, is recommended as part of their development in a profession that emphasizes self-regulation.

Keywords: Self Assessment; Peer Assessment; Multisource Feedback; Medical Students

Introduction

Professionalism is a cornerstone of medical practice, reflecting the qualities we demand of our practitioners and the expectations we have for the medical students we accept into the profession (Papadakis et al., 2005). Although there are various approaches that can be used to assess professionalism, the use of direct observation compiled through a 360 degree evaluation or multi-source feedback (MSF) format has been recognized as one of the most effective methods for the assessment of professionalism (Arnold, 2002; Bandiera et al., 2006; ACGME, 2013). Although professionalism is a multi-faceted construct that includes a range of measures from the personal (e.g., ability to self-reflect and regulate), attitudinal (e.g., altruistic, honesty, integrity) and behavioral (e.g., dutifulness, collegiality), the use of MSF has the potential to be used as a formative or summative feedback process for evaluating specific components of professionalism. For example, within the context of medical practice MSF has been shown to be successful in evaluating physicians' professional attitudes and behaviors from a range of stakeholder perspectives that include other physicians, coworkers and the patients' themselves (Violato & Lockyer, 2006; Allerup et al., 2007; Brinkman et al., 2007; Lockyer & Clyman, 2008).

The introduction of MSF at the medical school level to assess professional attitudes and behaviours reflects a progressive

move towards acknowledging the importance the profession places on this role or core set of competencies (Epstein et al., 2004; ACGME, 2013). There are constraints, however, to introducing a MSF process at medical schools in that as students the assessment is restricted to other peers as they are the only other persons that will have observed and know each other well enough. In addition, the peer assessment is constrained by what other students can realistically observe and assess about each other as it relates to the professional attributes displayed through education activities shared mutually. In particular, the peer assessment protocol has shown promise as a reliable and valid method to assess the professional competence of medical students (Epstein et al., 2004; Dannefer et al., 2005). Used in part to assess the domains of competence such as interpersonal, humanistic, and teamwork skills, the 15 item peer assessment form was shown to distinguish between two domains of students' professional competence: consistency in their work related habits and interpersonal habits.

The subsequent research on the use of the peer assessment protocol with undergraduate medical students has expanded to explore the effects of rater selection (Lurie et al., 2006a), changes in self-perceived abilities among men and women (Lurie et al., 2007a), the relationship between peer assessments and other measures (e.g. Dean's letter rankings and ratings by internship directors) (Lurie et al., 2007b), and temporal and

group-related trends (Lurie et al., 2006b). In each case, the structure of the peer assessment form is premised on the measure of two identified dimensions referred to as *work habits* and *interpersonal abilities*. In a qualitative study of the impact that the peer assessment protocol had at the University of Rochester School of Medicine and Dentistry, medical students reported that they found the process transformative and a useful source of feedback that contributed to their own professional development (Nofziger et al., 2010).

In this study, we describe the implementation and results of a formative self and peer assessment protocol as a measure of medical students' professional competence in their first and second years at a medical school in Canada. The responses were analyzed to determine the reliability (i.e., internal consistency, test-retest, and generalizability coefficients) and construct validity (i.e., factor analysis) of the self and peer questionnaires and to explore relationships between self and peer assessment ratings.

Method

Participants

As a component of the student feedback process in the undergraduate medical education program, a total of 168, first year medical students from the University of Calgary completed an online 15 item, self assessment form during the middle of their first year and, then again, at the middle of their second year. All of the participants were asked to identify 8 classmates to complete a peer assessment version of the questionnaire in their first year using the same 15 items. This study was approved by the Conjoint Health Research Ethics Board of the University of Calgary and signed consent was obtained by all participants.

Questionnaire

The 15 item Likert form was developed through initial research by Dannefer et al. (2005) on a peer assessment protocol to measure medical students' professional competencies as a function of observable behaviors (**Appendix**). In particular, their findings supported a two factor or subscale measure with high internal reliability (Cronbach's alpha greater than 0.80 for each subscale) that assesses students' professional *work habits* (WH) and *interpersonal abilities* (IA). The 15th item is not considered to be connected to either factor or subscale and is treated as an overall assessment of the individual's potential for professional competency; reflecting on whether or not the assessor is concerned about this person's future patients. The scoring of each item is based on a 5-point scale anchored by descriptors at each end and the option to circle UA for "unable to assess". For example a score of 1 or 2 for item number 13 states that the person's "behavior is frequently inappropriate" to options 4 or 5 for "behavior is always appropriate".

Statistical Analysis

The construct validity of the self and peer assessment questionnaires were investigated with exploratory factor analysis using principal components and varimax rotated solutions. The internal reliability (Cronbach's alpha), and generalizability coefficients were calculated in a decision study analysis to determine the optimal number of peer assessors required to obtain

a generalizability coefficient of greater than 0.70 (Brennan, 2001).

A comparison of mean differences between self and peer assessment ratings on each of the 15 items was conducted using independent samples *t*-tests with an effect size difference (Cohen's *d*) analysis. The interpretation of the magnitude of the effect size for mean differences is based on Cohen's (1988) suggestions of *d* of 0.30 as "small", *d* of 0.50 as "medium", and *d* of 0.80 as "large". In addition, for the independent variable sex (men vs. women) mean differences in the self and peer assessments were also investigated using independent samples *t*-tests and effect size difference analyses.

As there have been found to be discrepancies between self and peer assessment on other measures of professional competencies using multisource feedback, medical students' self assessments were classified into quartile categories (i.e., $\leq 25^{\text{th}}$ percentile, 26th to 50th percentile, 51st to 75th percentile, and $\geq 76^{\text{th}}$ percentile) based on their mean total scores across all 15 items on the questionnaire and compared with their corresponding peer assessments.

Results

The Cronbach's alpha coefficients for the self and peer assessments questionnaires were $\alpha = 0.85$ and 0.91 , respectively. An average of 7.5 peer assessment questionnaires were completed for each medical student self assessment. On all 15 of the items shown in **Table 1**, the mean ratings of the medical students on the self assessment questionnaire were significantly lower than those on the students' corresponding peer assessment ratings ($p < 0.001$). Overall, the medical students consistently score themselves lower than their peers with a mean effect size difference of $d = 0.69$ across all 15 items (range from $d = 0.27$ to 1.39). A large effect size difference ($d = 1.17$, $p < 0.001$) was found across a total score comparison between the self and peer assessment groups.

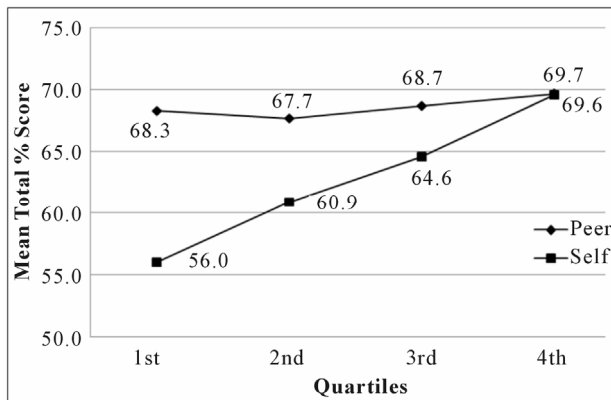
On the self assessment questionnaire items women ($n = 88$, 52%) were significantly more likely to report themselves as having higher "respect, compassion and empathy" [Item #4: $M = 4.64$ ($SD = 0.50$), $p < 0.01$, $d = 0.43$], "seeks to understand others/sensitivity" [Item #5: $M = 4.50$ ($SD = 0.60$), $p < 0.05$, $d = 0.36$], "seeks responsibility" [Item #8: $M = 4.13$ (0.57), $p < 0.001$, $d = 0.53$], and "behavior is always appropriate" [Item #13: $M = 4.63$ ($.49$), $p < 0.01$, $d = 0.42$] than did men ($n = 80$, 48%). Peer assessors also rated women significantly higher than men on three of these four items but the effect size differences were minimal in comparison: [Item #4: $M = 4.73$ ($SD = 0.52$), $p < 0.001$, $d = 0.22$], [Item #8: $M = 4.60$ ($.54$), $p < 0.001$, $d = 0.18$], and [Item #13: $M = 4.79$ (0.44), $p < 0.001$, $d = 0.21$]. In addition, women were rated significantly higher than men by their peer assessors on being "consistently well prepared" [Item #1: $M = 4.48$ ($SD = 0.58$), $p < 0.01$, $d = 0.16$] and that their "dress and appearance was always appropriate for the situation" [Item #12: $M = 4.83$ ($SD = 0.44$), $p < 0.01$, $d = 0.15$]. On Item #3, however, men students were rated by their peer assessors to be significantly higher on their ability to "explain clearly their reasoning process" [$M = 4.58$ ($SD = 0.58$), $p < 0.01$, $d = 0.17$] than women.

As shown in **Figure 1**, when students' self assessment scores were compared with their corresponding peer assessments by quartile groupings, medical students' in the lowest three quartiles score themselves significantly lower than their peer assessors

Table 1.

Means (standard deviations) and effect size differences between items on the self and peer assessment instruments.

	Self Assessment ^a (<i>n</i> = 168) Mean (SD)	Peer Assessment (<i>n</i> = 1267) Mean (SD)	Effect size Difference Mean (SD)
1) Consistently well prepared	3.55 (0.60)	4.42 (0.65)	1.39
2) Identifies and solves problems	3.98 (0.48)	4.54 (0.57)	1.06
3) Clearly explains reasoning processes	4.03 (0.59)	4.54 (0.59)	0.86
4) Demonstrates respect, compassion	4.52 (0.59)	4.68 (0.58)	0.27
5) Seeks to understand others/sensitivity	4.39 (0.62)	4.59 (0.60)	0.33
6) Takes initiative	3.88 (0.75)	4.37 (0.67)	0.69
7) Shares information with others	4.25 (0.64)	4.52 (0.64)	0.42
8) Seeks responsibility	3.96 (0.65)	4.56 (0.59)	0.98
9) Asks for and implements feedback	3.72 (0.73)	4.44 (0.62)	1.06
10) Trustworthy	4.40 (0.61)	4.74 (0.50)	0.61
11) Admits mistakes	4.41 (0.61)	4.71 (0.52)	0.53
12) Dress and appearance appropriate	4.55 (0.60)	4.80 (0.45)	0.48
13) Behaves appropriately	4.52 (0.55)	4.74 (0.51)	0.42
14) Thinks and works independently	4.37 (0.66)	4.71 (0.50)	0.59
	4.25 (0.61)	4.64 (0.56)	0.67
Total Score	62.38 (5.53)	68.97 (5.76)	1.17

Note: ^aIn independent samples *t*-tests between Self and Peer Assessment items, $p < 0.001$.**Figure 1.**

Mean percentile on four quartile groups comparing self to peer assessment total scores.

($p < 0.001$). In particular, those in the lower self assessment quartile underestimated their performance competency in comparison with their peer assessors by 12.3% of the total score. There was, however, no significant differences found between self and peer assessment total means scores in the top 4th quartile. Regardless of the discrepancies found in the self assessment percentile rankings, peer assessments were found to be not significantly difference and consistent across all quartiles (67.7% to 69.7%).

As shown in **Table 2**, the exploratory factor analyses resulted in 3 and 2 factor solutions for the self and peer assessment instruments, respectively. In both cases two of the factors were identified based on a previous factor analysis study as WH and IA (Dannefer et al., 2005). The third self assessment factor derived was labeled *personal attributes* (PA) as they reflect items that describe individual's attributes in asking for and implementing feedback (item loading = 0.751), admitting mistakes or being truthful (0.634), and collaborative through the sharing of information or resources (0.626). Although there are common items between the self and peer assessment questionnaires that are associated with only the WH (items 1, 2, 3 and 14) or IA (items 4, 5, 12 and 13) subscales, the remaining items

tend to cross load between factors (items 6, 7, 8, 9, and 10). The internal reliability coefficients for the peer assessment questionnaire were $\alpha = 0.87$ for the WH and $\alpha = 0.86$ for the IA subscales. For the three factor solution derived for the self assessment questionnaire the internal reliability coefficients were lower at $\alpha = 0.60$ (WH), $\alpha = 0.80$ (IA), and $\alpha = 0.77$ (PA); reflecting in part the reduced number of items associated with each factor or subscale. The total percentage of the variance accounted for by the self and peer assessment questionnaires were found to be 53.6% and 57.3%, respectively.

The generalizability (G) coefficient for the peer assessment with 8 raters for the 15 item checklist was $Ep^2 = 0.73$. A decision study analysis was used for this single-facet nested design (i.e., peer assessors nested within individual medical students) to derive mean G coefficients of 0.58 to 0.80 for 4 to 12 peer assessors, respectively (**Figure 2**). The proportion of variance accounted for by persons (medical students) in the analysis was 25% and the remaining 75% of the variance was contributed through the peer assessors nested within student interaction effect.

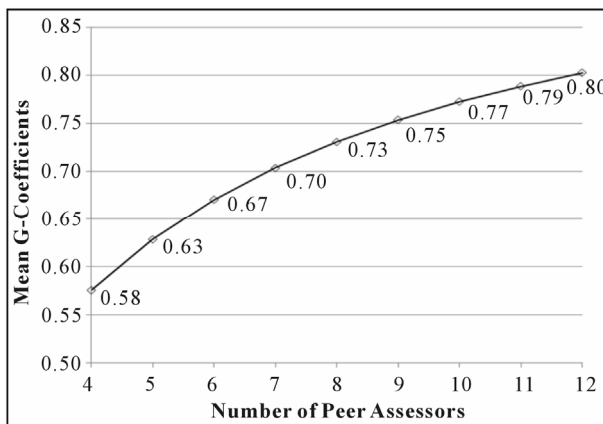
A subsequent administration of the 15 item questionnaire was completed approximately 12 months later in the middle of the students' second year. A test-retest reliability analysis resulted in a correlation coefficient of $r = 0.44$ and an overall significant increase in students' total scores from year one ($M = 62.5$, $SD = 5.47$) to year two ($M = 65.6$, $SD = 4.54$, $p < 0.001$; $d = 0.62$). A subsequent exploratory factor analysis on the completed self assessment questionnaire for year two also confirmed a three factor solution (accounting for 52% of the variance). In paired sample *t*-tests between the two administrations, students were found to rate themselves significantly higher on each of the three subscales identified: WH (from $M = 19.9$, $SD = 1.94$ to $M = 20.8$, $SD = 1.75$; $p < 0.001$; $d = 0.48$), IA (from $M = 17.9$, $SD = 1.87$ to $M = 18.4$, $SD = 1.63$; $p < 0.01$; $d = 0.28$), and PA (from $M = 20.6$, $SD = 2.30$ to $M = 21.8$, $SD = 1.90$; $p < 0.001$; $d = 0.57$). In addition, students' rated themselves significantly higher in year two on question 15 regarding the effectiveness of their healthcare practice with patients in the future (from $M = 4.25$, $SD = 0.60$ to $M = 4.53$, $SD = 0.50$; $p < 0.001$; $d = 0.51$).

Table 2.

Two and three factor solutions for self and peer assessment instruments, respectively.

Item	Self Assessment (n = 168)			Peer Assessment (n = 1267)	
	Work Habits	Interpersonal Abilities	Personal	Work Habits	Interpersonal Abilities
1) Consistently well prepared	0.622			0.776	
2) Identifies and solves problems	0.666			0.811	
3) Clearly explains reasoning processes	0.679			0.759	
4) Demonstrates respect, compassion		0.792			0.841
5) Seeks to understand others/sensitivity		0.671	0.371		0.793
6) Takes initiative	0.343		0.534	0.654	
7) Shares information with others			0.626	0.539	0.486
8) Seeks responsibility	0.485		0.439	0.673	0.425
9) Asks for and implements feedback			0.751	0.443	0.515
10) Trustworthy		0.499	0.561		0.705
11) Admits mistakes			0.634		0.687
12) Dress and appearance appropriate		0.682			0.525
13) Behaves appropriately		0.840			0.740
14) Thinks and works independently	0.619			0.686	
Number of items for each factor	5	4	5	7	7
Cronbach's α for each factor	0.60	0.80	0.77	0.87	0.86
Variance explained by each factor ^a	20.1%	17.8%	15.7%	28.7%	28.6%

Note: ^aSelf and peer Assessment total variance is equal to 53.6% and 57.3%, respectively. Note that only item loadings greater than 0.32 are reported (i.e., accounting for greater than 10% of the variance).

**Figure 2.**

Decision study Generalizability (G-) coefficients for 4 to 12 student peer assessors.

Discussion

The major findings of the present study are that: 1) self assessment of professional competence was significantly lower in comparison with peer assessors on all questionnaire items, 2) distinct from the 2 factor solution for the peer assessment questionnaire, the 3 factor solution for the self assessment includes an additional subscale associated with *personal attributes*, 3) women medical students were rated significantly higher either by themselves or their peers on more items than were the men, and 4) on a one year test-retest of the self assessment questionnaire, students' self-reported rating were significantly higher from year one to two on total and subscale scores.

The discrepancies found between self- and peer-reported assessment across each of the items rated reflect a concern as to why individual students tend to perceive that their professional competencies are lower than the mean ratings provided by a group of peers they interact with on a regular, if not daily, basis.

When investigated by quartile groups based on self assessment totals, there appears to be a majority of medical students that significantly underestimate themselves in comparison with their peers. This miscalibration effect found between the self and peer assessment ratings is typical of physicians' self and peer assessment as well (Violato & Lockyer, 2006). It would seem that the miscalibration of professional competencies that was shown to be evident in experienced physicians begins early in one's medical career, as we see in the present results, and likely reflects a general discrepancy found in human self assessment.

Although methods for the assessing of competencies such as *work habits*, *interpersonal abilities* and *personal attributes* are less well developed or tested, the peer assessment protocol and 15 item self and peer assessment forms provide a potentially reliable and valid method to introduce sources of feedback that can help medical students to reflect on and enhance their own professional development. The findings in this study, however, demonstrates the difficulty in developing MSF tools or questionnaires that can consistently measure similar constructs between different types of raters (e.g., self, peer, co-workers, patients) as a function of observable behaviors (Lockyer & Clyman, 2008). Therefore, one of the main limitations of the use of the results is that there is some uncertainty as to the specific factors being measured as the self assessment questionnaire appears to be more multidimensional in that an additional third factor was identified (i.e., *personal attributes*). Another limitation is that overall students are left with the impression from their peer groups that they are actually rated much higher (or at the same level) on each of the 15 items than the self-reported ratings provided on the corresponding self assessment questionnaire (Colthart et al., 2008).

There is an expectation that medical schools need to provide students with feedback related to their clinical and professional performance. The current assessment formats for medical students focus primarily on testing clinical knowledge and skills, without any means of formative feedback for professional development. The use of a self and peer assessment MSF process

in the initial years of medical school provides an opportunity for students to become engaged in understanding how they are performing on other aspects of their non-cognitive skills development (i.e., ability to collaborate with others, communication effectiveness, managing their time and resources). Nevertheless, efforts to measure relevant constructs that have application to their roles and responsibilities as future physicians are still not well defined. With the move towards competency-based frameworks in residency programs, MSF assessments will need to better reflect measures associated with the key competencies identified of practicing physicians (ACGME, 2013; Frank, 2005).

Acknowledgements

The authors acknowledge the students, staff and administration at the Office of Undergraduate Medical Education for their continuing support in our research efforts to improve the quality of assessment methods and approaches used at our medical school.

REFERENCES

- Accreditation Council for Graduate Medical Education (ACGME) (2013). ACGME Outcome Project. URL (last checked 15 January 2013). <http://www.acgme.org/acgmeweb/tabid/159/DataCollectionSystems/AccreditationDataSystem.aspx>
- Allerup, P., Aspegren, K., Ejlersen, E., et al. (2007). Use of 360-degree assessment of residents in internal medicine in a Danish setting: A feasibility study. *Medical Teacher*, 29, 166-170. [doi:10.1080/01421590701299256](https://doi.org/10.1080/01421590701299256)
- Arnold, L. (2002). Assessing professional behavior: Yesterday, today, and tomorrow. *Academic Medicine*, 77, 502-515. [doi:10.1097/00001888-200206000-00006](https://doi.org/10.1097/00001888-200206000-00006)
- Bandiera, G., Sherbino, J., & Frank, J. R. (2006). *The CanMEDS assessment tools handbook: An introductory guide to assessment methods for the CanMEDS competencies*. Ottawa, ON: The Royal College of Physicians and Surgeons of Canada.
- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer-Verlag.
- Brinkman, W. B., Geraghty, S. R., Lanpher, B. P., et al. (2007). Effect of multisource feedback on resident communication skills and professionalism. *Pediatric Adolescence Medicine*, 61, 44-49. [doi:10.1001/archpedi.161.1.44](https://doi.org/10.1001/archpedi.161.1.44)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Colthart, I., Bagnall, G., Evans, A., et al. (2008). The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice: BEME Guide No. 10. *Medical Teacher*, 30, 124-145. [doi:10.1080/01421590701881699](https://doi.org/10.1080/01421590701881699)
- Dannefer, E. F., Henson, L. C., Bierer, S. B., et al. (2005). Peer assessment of professional competence. *Medical Education*, 39, 713-722. [doi:10.1111/j.1365-2929.2005.02193.x](https://doi.org/10.1111/j.1365-2929.2005.02193.x)
- Epstein, R. M., Dannefer, E. F., Nofziger, A., et al. (2004). Comprehensive assessment of professional competence: The Rochester experiment. *Teaching & Learning in Medicine*, 16, 186-196. [doi:10.1207/s15328015tlm1602_12](https://doi.org/10.1207/s15328015tlm1602_12)
- Frank, J. R. (2005). *The CanMEDS 2005 physician competency framework. Better standards. Better physicians. Better care*. Ottawa, ON: The Royal College of Physicians and Surgeons of Canada.
- Lockyer, J. M., & Clyman, S. G. (2008). Multisource feedback (360-degree evaluation). In: E. S. Holmboe, & R. E. Hawkins (Eds.), *Practical guide to the evaluation of clinical competence* (pp. 75-83). Philadelphia, PA: Mosby Elsevier.
- Lurie, S. J., Nofziger, A. C., Meldrum, S., Mooney, C., & Epstein, R. M. (2006a). Effects of rater selection on peer assessment among medical students. *Medical Education*, 40, 1088-1097. [doi:10.1111/j.1365-2929.2006.02613.x](https://doi.org/10.1111/j.1365-2929.2006.02613.x)
- Lurie, S. J., Nofziger, A. C., Meldrum, S., Mooney, C., & Epstein, R. M. (2006b). Temporal and group-related trends in peer assessment amongst medical students. *Medical Education*, 40, 840-847. [doi:10.1111/j.1365-2929.2006.02540.x](https://doi.org/10.1111/j.1365-2929.2006.02540.x)
- Lurie, S. J., Meldrum, S., Nofziger, A. C., Sillin, L. F., Mooney, C. J., & Epstein, R. M. (2007a). Changes in self-perceived abilities among male and female medical students after the first year of clinical training. *Medical Teacher*, 29, 921-926. [doi:10.1080/01421590701753559](https://doi.org/10.1080/01421590701753559)
- Lurie, S. J., Lambert, D. R., Nofziger, A. C., Epstein, R. M., & Grady-Weliky, T. A. (2007b). Relationship between peer assessment during medical school, Dean's letter rankings, and ratings by internship directors. *Journal of General Internal Medicine*, 22, 13-16. [doi:10.1007/s11606-007-0117-4](https://doi.org/10.1007/s11606-007-0117-4)
- Nofziger, A. C., Naumburg, E. H., Davis, B. J., Mooney, C. J., & Epstein, R. M. (2010). Impact of peer assessment on the professional development of medical students: A qualitative study. *Academic Medicine*, 85, 140-147. [doi:10.1097/ACM.0b013e3181c47a5b](https://doi.org/10.1097/ACM.0b013e3181c47a5b)
- Papadakis, M. A., Teherani, A., Banach, M. A., et al. (2005). Disciplinary action by medical boards and prior behavior in medical schools. *New England Journal of Medicine*, 353, 2673-2682. [doi:10.1056/NEJMsa052596](https://doi.org/10.1056/NEJMsa052596)
- Violato, C., & Lockyer, J. (2006). Self and peer assessment of pediatricians, psychiatrists and medicine specialists: Implications for self directed learning. *Advances in Health Science Education*, 11, 235-244. [doi:10.1007/s10459-005-5639-0](https://doi.org/10.1007/s10459-005-5639-0)

Appendix.Peer assessment^a questionnaire (Dannefer et al., 2005).

Instructions to the students:								
Please rate this student based on your personal knowledge of the student and your own interactions with him/her. Note that 1 is the lowest rating and 5 is the highest rating for each characteristic. If you have insufficient contact with the student to evaluate him/her on a particular characteristic, circle UA (unable to assess).								
		Low/Unsatisfactory					High/Exceptional	
1	Consistently seems unprepared for sessions; presents minimal amount of material; seldom supports statements with appropriate references	1	2	3	4	5	UA	Consistently well prepared for sessions, presents extra material, supports statements with appropriate references
2	Overlooks important data and fails to identify or solve problems correctly	1	2	3	4	5	UA	Identifies and solves problems using intelligent interpretation of data
3	Unable to explain clearly his or her reasoning process with regard to solving a problem, basic mechanisms, concepts etc	1	2	3	4	5	UA	Able to explain clearly his or her reasoning process with regard to solving a problem, basic mechanisms, concepts etc.
4	Lacks appropriate respect, compassion and empathy	1	2	3	4	5	UA	Always demonstrates respect, compassion and empathy
5	Displays insensitivity and lack of understanding for others' views.	1	2	3	4	5	UA	Seeks to understand others' views
6	Lacks initiative or leadership qualities	1	2	3	4	5	UA	Takes initiative and provides leadership
7	Doesn't share information or resources; impatient when others are slow to learn; hinders group process; tends to dominate group	1	2	3	4	5	UA	Shares information or resources; truly helps others learn; contributes to the group process; able to deter to the group's needs.
8	Only assumes responsibility when forced to or stimulated for personal reasons; fails to follow through consistently	1	2	3	4	5	UA	Seeks appropriate responsibility; consistently identifies tasks and completes them efficiently and thoroughly
9	Does not seek feedback; defensive or fails to respond to feedback	1	2	3	4	5	UA	Asks classmates and professors for feedback and then puts suggestions to good use
10	Pleases superiors while undermining peers; untrustworthy	1	2	3	4	5	UA	Presents him/herself consistently to superiors and peers; trustworthy
11	Hides his or her own mistakes; deceptive	1	2	3	4	5	UA	Admits and corrects his or her own mistakes, truthful
12	Dress and appearance often inappropriate for the situation	1	2	3	4	5	UA	Dress and appearance always appropriate for the situation
13	Behavior is frequently inappropriate	1	2	3	4	5	UA	Behavior is always appropriate
14	Dependent upon others for direction with regard to his or her learning agenda.	1	2	3	4	5	UA	Directs own learning agenda; able to think and work independently
15	I have concerns for his or her future patients	1	2	3	4	5	UA	I would refer my own family or patients to this future physician or ask this person to be my physician

Note: ^aThe corresponding self assessment questionnaire is identical, but written in the first person.