

Clinical Examiners, Simulated Patients, and Student Self-assessed Empathy in Medical Students During a Psychiatry Objective Structured Clinical Examination

Karen O'Connor · Romaine King · Kevin M. Malone ·
Allys Guerandel

Received: 14 December 2013 / Accepted: 3 April 2014 / Published online: 23 April 2014
© Academic Psychiatry 2014

Abstract

Objectives This study aims to assess and compare objective and subjective scores of empathy in final-year medical students by using firstly a validated student self-assessment just prior to the psychiatry objective structured clinical examination (OSCE), and then comparing this to clinical examiner's and simulated patient's (SP's) assessments of empathy of students using a Global Rating of Empathy scale (GRE) during a psychiatry OSCE.

Methods In 2011, all final-year medical students in the University College Dublin were invited to complete a subjective, self-assessed empathy questionnaire (The Jefferson scale of physician empathy—student version (JSPE-S)). They were also assessed for empathy in four OSCEs by the clinical examiner and the SP acting in that OSCE scenario.

Results Included in the analysis were 163 of 184 final-year students JSPE-S (88.6 %) questionnaires. The female students scores on the JSPE-S were significantly higher than those of their male peers ($t=3.34$, $p=0.001$). Concurrent validity was greater between the SPs' assessments of empathy in the OSCE and the JSPE-S score than between the clinical examiners assessments of empathy and the JSPE-S score ($r=0.23$, $p<0.005$; $r=0.14$, $p<0.08$). Inter-rater reliability of SP's and clinical examiner's using the GRE was found to be high ($F=0.868$ ($df=171, 171$), p value <0.001).

Conclusions SPs may be valid assessors of empathy in medical students during an OSCE.

Keywords Empathy · Simulated patients · Medical education

Empathy has been identified as a crucial element in the delivery of high-quality medical care and is seen as one of the personal qualities that define professionalism in medicine [1, 2]. Empathic communication skills are associated with improved health outcomes, increased diagnostic accuracy, better patient participation, better patient adherence, reduction in medical–legal risk, and improved patient satisfaction [3–7]. Empathy in the doctor–patient relationship may also benefit the doctor [8]. Displaying empathy may reduce the interpersonal challenges associated with practicing medicine and enhance job satisfaction [9]. Higher levels of burnout have been associated with diminished empathy, which in turn may be associated with increased likelihood of perceived medical error [10, 11].

Some gender differences have been found, which suggest that female gender is associated with higher levels of empathy [12]. In medical students, this finding has not been consistent [13]. Previous research has attributed gender differences to evolutionary and social learning factors [12–15].

However, while the evidence supporting the importance of empathy in good clinical practice continues to grow, not everyone agrees that empathy can and should be measured in medical students or doctors. Some authors have suggested that many of the most widely used definitions of empathy are overly reductionist and fail to recognize its true emotional and psychological complexity [16]. Such authors also suggest that true empathy derives from an experience of intersubjectivity, and this cannot be achieved in the doctor–patient relationship. Any mirroring of emotion by an “empathetic” doctor will always differ quantitatively and qualitatively from the patient's actual experience, and it is sympathy and not empathy that a caring doctor should aspire to feel for their patients [17].

K. O'Connor · K. M. Malone
University College Dublin, Dublin, Ireland

R. King · A. Guerandel
St. Vincent's University Hospital, Dublin, Ireland

K. O'Connor (✉)
Tallaght Hospital, Dublin, Ireland
e-mail: karenocconnor2@hotmail.com

Empathy is poorly defined in the medical literature [18]. However, for the purposes of this study, we have used Mercer and Reynolds' widely accepted definition [16]. This describes physician empathy as the ability of a physician to “(a) understand the patient's situation, perspective and feelings, (b) communicate that understanding and check its accuracy, and (c) act on that understanding with the patient in a helpful way” [19].

Empathy may be measured from three different perspectives [20], as follows: (1) Self-rating (first-person assessment)—the assessment of empathy using standardized questionnaires completed by those being assessed; (2) patient rating (second-person assessment)—the use of questionnaires given to patients to assess the empathy they experience from their health professional; and (3) observer rating (third-person assessment)—the use of standardized assessments by an observer to rate empathy in interactions between health professionals and patients, including the use of simulated patient (SP) encounters to control for observed differences secondary to differences between patients.

SPs are increasingly being used in undergraduate and post-graduate “high stakes” examinations in medicine. The benefits associated with their use have been well documented [21, 22]. These benefits include their ready availability, their facility for standardization, and the reduced risk of harm being done by an inexperienced student or trainee to an SP rather than to a real patient [20]. However, particular concerns have been raised about the use of SPs in the assessment of the interpersonal aspects of the doctor–patient encounter, such as the quality of empathy [23]. It has been proposed that the ability to understand, communicate, and act helpfully on empathy can be affected by the fact that both parties in an OSCE know that this is a simulated experience. Additionally, authors have postulated that factors beyond the scripted role may have an impact on the SP's affective experience, e.g., fatigue, acting ability, personal preoccupations, or personal experiences [23]. However, are some of these factors inherent to the difficulties associated with the assessment of “subjective” and complex interpersonal encounters? Would real patients really be any less or more prone to factors such as fatigue or distraction? Is the real difficulty in assessing empathy in an OSCE the complex and subjective nature of the item being assessed? It has been proposed that SPs may actually become authorities on physician communication and observe failings that a real patient might overlook [23].

However, in our medical school as is the case in many medical schools, it is neither patients nor SPs who assess empathy in “high stakes” OSCEs. Rather, it is clinical examiners. One published study has compared clinical examiners' assessments of doctor's interpersonal skills with those of “real” patients and found that there was a discrepancy between patient and examiner judgments of the more subjective elements of the examination [20]. Studies who have compared

SPs' assessment of empathy with that of clinical examiners' have to date been inconclusive in their findings [23, 24]. Some have suggested that there are discrepancies between the scores of SPs and clinical examiners [25]. While others found a moderate degree of agreement between examiner and SP scores for communication [26].

As such, though the weight of research underlies the importance of empathy in medical education and clinical practice, the best means of assessing empathy in the medical school setting is not agreed [20]. There is at least some theoretical and research evidence that in the context of a simulated clinical interaction such as an OSCE, the better assessor of empathy in the absence of the real patient may be their surrogate SP rather than the clinical examiner.

Aims of This Study

We had four aims for this study. Firstly, we sought to assess empathy in final-year medical students from an Irish university using a highly validated first person assessment: The Jefferson Scale of Physician Empathy—student version (JSPE-S). Secondly, we sought to assess whether there are gender differences in levels of empathy in a final-year medical student group from an Irish University. Thirdly, we aimed to assess empathy in final-year medical students using second and third person assessors (SPs and consultant psychiatrist examiners) and a Global Rating of Empathy scale (GRE) in an OSCE. Finally, we aimed to validate the use of SPs as assessors of empathy in a psychiatric OSCE by comparing the results of the first-, second-, and third-person assessments of empathy in this Irish final-year medical student sample.

We had two hypotheses. First, that empathy levels in Irish medical students, as assessed by the JSPE-S, would be comparable to their international peers with regards to absolute scores and gender differences. Our second hypothesis was that the assessment of empathy by SPs in an OSCE was a valid additional means of assessing empathy in medical students.

Subjects and Methods

In the University College Dublin in Ireland, psychiatry is taught in three stages in the medical curriculum: firstly, in the final preclinical year (year 3) as part of a 6-week introductory multisubject module (reproductive medicine/children/psychiatry); secondly, as a 6-week dedicated clinical module in final year; and finally, as part of a 6-week professional completion module just prior to graduation during which psychiatry knowledge is integrated into the overall patient assessment and management plan.

The final-year 6-week clinical psychiatry module includes 6 weeks clinical attachment program where students are

assigned to one of six centers, where they are allocated to a multidisciplinary clinical team. E-learning modules and clinical labs further support student learning. Development of empathy is a specific goal of this module and teaching and learning to this end is integrated in all aspects of the module. In particular the daily small group teaching sessions emphasize the patient perspective, facilitates student discussions on patient's experience, perspective, and feelings, and facilitates the development of empathetic communication skills by peer role playing. The final-year clinical psychiatry module is run four times each academic year. These four 6-week modules are April/May, May/June, September/October and October/November. All students completing the final-year clinical psychiatry module would previously have completed the pre-clinical module, which included some formal teaching on empathy. This preclinical module is assessed in part by an end of module OSCE during which the students ability to demonstrate empathy is assessed.

Following informed consent, all students completing the final-year clinical psychiatry module in 2011 were invited to complete a questionnaire which has been demonstrated to be both valid and reliable in the assessment of empathy in medical students; The JSPE-S. Students were invited to complete the questionnaire in the final week of the psychiatry module. The questionnaire was conducted quasi-anonymously. Students were invited to volunteer their student identification numbers. This identifying information was then coded prior to analysis. The University College Dublin Ethics Committee exempted this study from requiring ethical approval in April 2011.

The final-year module is assessed using a range of formative and summative methods and tools including continuous assessment and case presentation (20 %), a reflective essay (20 %), a multiple choice question paper (20 %), and an OSCE (40 %).

The OSCE is held on the last day of the final-year clinical psychiatry module. The OSCE is composed of four stations with SPs and two written stations where, students are expected to answer questions on a video recording/written clinical information. At each of the acted stations, one examiner assesses the students. Empathy is assessed as standard at each part of the four acted OSCEs. Empathy is awarded 20 % of the total marks at each acted OSCE station and students are fully aware of this from the commencement of the module.

The SPs are actors who receive additional training by the clinical tutors prior to the OSCE to ensure their performance is standardized and “life-like.” The acted OSCEs are examined by experienced consultant psychiatrists, one examiner at each station. They also receive additional training and guidance on the morning of the exam by the tutors to encourage standardized OSCE marking.

In 2011, all students completing the OSCE were assessed for empathy using a GRE by both the consultant psychiatrist

examiner and also the SP acting in that OSCE scenario, who were blind to each other's rating. The SP's and the examiner's ratings were completed “live,” i.e., at the time of the encounter. The GRE was completed in the one minute break between the OSCE stations by the examiner and the SP. Only the consultant psychiatrist's GRE mark was included in the student's summative assessment.

SPs and Examiners in OSCE Stations

The University College Dublin has an end of psychiatry module OSCE. This OSCE provides a summative assessment. A passing score is required to pass the clinical psychiatry module. Students have 5 min to perform a focused history on each of four SPs.

All SPs and examiners received written information on empathy, the assessment of empathy, the rating scale being used and the aims of the research study prior to the OSCE. On the morning of the OSCE this information was rehearsed orally with both the SPs and the examiners. Tutors were available to offer further clarification prior to and during the OSCE. The medical students were also made aware of the purpose and procedure of the study in anticipation of the OSCE. Students could choose not to participate in the study without prejudice.

Questionnaire

Jefferson Scale of Physician Empathy—Student Version

The JSPE-S includes 20 Likert-like items answered on a 7-point scale (1=strongly disagree, 7=strongly agree) [27, 28]. The JSPE-S has demonstrated to be both reliable and valid in the assessment of empathy in medical students [29–34].

Of the 20 items in the JSPE-S, 10 items are positively worded and linked to “perspective taking” and 10 items are negatively worded. Eight of the ten negatively worded items are concerned with “compassionate care,” and the remaining two items are concerned with “standing in the patient's shoes.” The minimum possible score on the JSPE-S is 20, and the maximum possible score is 140. The higher score indicates a more empathic behavioral orientation.

Global Rating of Empathy

SPs and examiners used a 5-point scale (5=excellent, 4=very good, 3=good, 2=fair, and 1=poor) to indicate global ratings of students' empathy. We developed this GRE on the basis of other GRE's, which have been used in previous studies to assess empathy in summative OSCE assessments [35].

We analyzed the data using the Statistical Package for the Social Sciences (SPSS) version 17. We replaced items of missing data with the mean. However, those students who did not provide a response to four or more items on the JSPE-S were not included in subsequent analysis. We assessed concurrent validity using the Pearson's correlation coefficient. We conducted inter-rater reliability testing using the intraclass correlation coefficient. We used conventional *t* tests and ANOVA to examine group differences.

Results

JSPE-S

One hundred and seventy-six of 184 final-year students returned JSPE-S questionnaires (94 %). However, 11 questionnaires were void due to high levels of missing data/spoiled responses.

Ten individual items on the 173 JSPE-S questionnaires were included in the analysis after making mean substitutions for missing data. One student failed to complete three items, two students failed to complete two items, and three students failed to complete one item on the JSPE-S. Sensitivity analyses were conducted including and excluding this data and demonstrated that these mean substitutions had no significant effect on the findings. Therefore, the six JSPE-S questionnaires, which had <4 missing items were included in the analyses. One hundred sixty-three JSPE-S (88.6 %) questionnaires were included in the analysis (74 males (45.3 %) and 89 females (54.7 %)).

The overall mean score on the JSPE-S was 113.6 (SD±12.3). For males (*n*=74), the mean score was 110.2 (SD±13.9). For females (*n*=89), the mean score was 116.5 (SD±10.1). Table 1 details the mean scores on the JSPE-S. The female scores on the JSPE-S were significantly higher than those of the male students (*t*=3.34, *p*=0.001).

Examiners Assessment of Empathy

The mean score of empathy at an OSCE station as assessed by an experienced consultant psychiatrist was 3.4±0.5. The mean total score of empathy in the OSCE (all four acted OSCE stations) was 13.94±1.9. Table 2 details the examiners assessments of empathy in the final-year students end of module

Table 1 Jefferson scale of physician empathy—student version

	Number	Mean score	SD
Final-year students	163	113.6	12.3
Males	74	110.2	13.9
Females	89	116.5	10.1

Table 2 Examiners assessment of empathy in OSCE

Total empathy score in OSCE (4 stations)	Number	Mean score	SD
Final-year students	174	13.9	1.9
Males	81	13.7	2.0
Females	93	14.2	1.8

OSCE. We found no significant gender difference in the consultant psychiatrist-examined OSCE stations.

SPs Assessment of Empathy

The average empathy score as assessed by an SP at an OSCE station was 3.8±0.6. The mean total score of empathy in the OSCE (all four acted OSCE stations) was 15.2±2.3.

Female students were more likely to achieve a higher empathy score than their male peers when assessed by an SP (*t*=2.01, *p*<0.05) (Table 3).

Concurrent Validity

A high level of correlation was found between the clinical examiners assessment of empathy and the SP's assessment of empathy in the OSCE (*r*=0.78, *p*<0.005). However, there was a higher level of correlation between the SP's assessments of empathy in the OSCE and the JSPE-S score than between the clinical examiners assessments of empathy and the JSPE-S score (*r*=0.23, *p*<0.005; *r*=0.14, *p*<0.08).

Inter-rater Reliability

Inter-rater reliability of SP's and clinical examiner's using the GRE was found to be high (*F*=0.868 (*df*=171, 171), *p* value <0.001). However, SPs marked students significantly higher on the GRE than the clinical examiners (*t*=5.76, *p*<0.0001).

Associations Between Assessments of Empathy and Other Exam Results

The significant associations found between JSPE-S, SP and examiner assessments with other exam results is described in Table 4. Gender has been identified as being associated with empathy as such, we completed a regression to assess the impact of gender on the above correlations. We only found

Table 3 Simulated patients assessment of empathy in OSCE

Total empathy score in OSCE (4 stations)	Number	Mean score	SD
Final-year students	175	15.2	2.3
Males	80	14.8	2.4
Females	95	15.5	2.2

Table 4 Significant associations between assessments of empathy and other exam results as assessed using Pearson's coefficient (r)

	OSCE	CAF	MCQ	Refective essay	Overall module score
JSPE-S score	$r=0.25^{***}$	NS	$r=0.25^{***}$	$r=0.21^{**}$	$r=0.23^{***}$
Simulated patient empathy score	$r=0.37^{***}$	NS	$r=NS$	$r=0.21^{**}$	NS
Clinical examiner empathy score	$r=0.51^{***}$	$r=0.23^{***}$	$r=0.29^{***}$	$r=0.32^{***}$	$r=0.22^{***}$

OSCE objective-structured clinical examination, CAF clinical assessment form, MCQ multiple choice examination, NS not statistically significant
 $*p<0.05$; $**p<0.01$; $***p<0.001$

gender and overall OSCE score to be significant ($r^2=0.21$, $p=0.05$).

Assessment of Empathy and Rotation/OSCE Order

We found no statistical difference between the JSPE-S scores ($F=1.35$, $p=0.10$) and rotation order. We found a statistical difference between SP's assessment of empathy ($F=2.55$, $p=0.001$) and examiners assessment of empathy ($F=2.59$, $p=0.05$) and rotation order. Students who completed the module in the April/ May 2011 performed significantly better in the SP assessments of empathy than students who completed the other modules. We found no association between rotation order and examiners assessments.

We found a significant relationship between overall OSCE score and rotation order ($F=4.34$, $p=0.005$). Students who completed the April/May 2011 module scored significantly higher than other students in the OSCE though not in other components of the examination and not overall (mean=63.49, SD=7.66). We found no significant relationship between other components of the exams and the rotation order.

We found no relationship between the order in which students completed the OSCE stations and the ratings they received from the SPs ($F=1.1$, $p=0.35$) or the examiners ($F=1.05$, $p=0.22$).

Discussion

In this study, we set out to assess empathy in final-year medical students from an Irish university using a highly validated first-person assessment (JSPE-S) and also using a GRE applied by second and third person assessors (SPs and consultant psychiatrist examiners). Then in an attempt to validate the use of SPs as assessors of empathy in an undergraduate psychiatric OSCE we set out to compare these assessments. Our hypothesis was that the assessment of empathy by SPs in an OSCE is a valid additional means of assessing empathy in medical students.

The Irish students' levels of empathy as assessed using the first person JSPE-S compare favorably with international samples. The mean score in the Irish final-year students (mean=113.6, SD=12.3) is significantly higher than that

found by Hojat et al. in an American sample of final-year medical students (JSPE-S mean=109.1, SD=14.1; $t=3.20$, $p=0.01$) and significantly higher than those reported by Kataoka et al. in a final-year sample in Japan (mean=107.8, SD=12.1; $t=4.56$, $p=0.001$) [33, 36].

Similar to the results with Mexican, Japanese, and the UK medical students our data demonstrated higher empathy scores among female students as measured by the JSPE-S [31, 33, 37]. Females were also more likely to score higher on the GRE when assessed by the SP rather than by the examiner. However, while the empirical evidence indicates that female gender is associated with higher levels of empathy, a number of studies conducted using the JSPE-S in medical student populations failed to demonstrate any gender difference [32, 34]. However, the consistency of these finding does suggest that medical educators need to consider whether recruitment processes, assessment processes, and/or the teaching of empathy need to explicitly consider gender issues.

This study found a high level of concurrent validity between SPs' assessment of empathy using the GRE in the OSCE and the highly validated JSPE-S. The SPs' assessment using the GRE was found to correlate more highly with the students' scores on the JSPE-S than the examiners' assessment using the GRE. As the JSPE-S has proven validity in the assessment of empathy in this population, this may indicate that SPs as second person assessors of empathy are more valid in their assessment than third person assessors who are observing rather than participating in the interaction.

One of the concerns raised in the psychiatric literature about the use of SPs in the assessment of empathy has been to what extent students and the actors are engaging in an exchange in which they both know the "rules" and so rather than engaging genuinely they are simply playing the game [35]. This may or may not be different from the situation with a real patient. The interaction with real patients may also have its own rules, as patients will have their own agenda and expectation of what they are hoping to get from the interaction, this agenda may distract from the 'realness' or genuineness of that interaction. Additionally, the reality of doctors day-to-day practice be it in hospital, in general practice or elsewhere is that it is rarely without external pressures. Time pressures, resourcing pressures, and other professional and personal issues attempt to impose themselves on the

interaction. As such, should students be able to 'suspend reality' and 'suspend the context' and engage empathetically with patients and SPs alike?

In this study, the association between students' self-reported empathy and SPs' perceptions of students' empathy, although statistically significant, was not large enough to merit that the two evaluations are redundant. This finding echoes of Berg et al. who in their study of 248 third-year American medical students demonstrated significant associations between students' self-reported scores on the JSPE and SPs' evaluations of students' empathy [31].

A high level of inter-rater reliability was found in the assessment of empathy by SPs and consultant psychiatrist examiners using the GRE as demonstrated by the high intra-class correlation coefficient. However, the SPs did tend to award students higher marks than the consultant psychiatrist examiners. As such, while SPs and examiners tended to agree on whether they thought a student demonstrated empathy or not, the SPs consistently marked students more leniently. This echoes the findings of previous studies, however a specific training program has been demonstrated to substantially improve inter-rater variability between SPs and clinical examiners [38].

The consultant psychiatrists' assessment of empathy in the acted OSCE stations was found to moderately but significantly correlate with all of the four summative assessments in this psychiatry module, i.e., overall OSCE score, continuous assessment, MCQ, and reflective essay. The SPs' assessment of empathy moderately but significantly correlated with only the overall OSCE and the reflective essay scores. One possible explanation for this is that consultant psychiatrists as clinical examiners are more influenced in their assessment of empathy by students knowledge of psychiatry than SPs. Students who have demonstrated good knowledge of the modules low-level learning objectives by scoring highly on the MCQ may also be marked more highly on empathy when assessed by the psychiatrist examiner but not by the SP.

Students who complete the April/May module in 2011 scored more highly on SP assessors rating of empathy and also higher in the total OSCE score. There has been some literature on the impact on OSCE performance of students' previous experience of OSCEs [39]. With different levels of experience in SPs being associated with differences in performance. However, this does not explain the finding in our sample, as this group was actually the first group of their academic year to complete the psychiatry module. Subsequent groups would have had more SP and OSCE experience. Anecdotally our faculty has noted an association between the module preceding the psychiatry module and students engagement and performance in the psychiatry module. However, we are not aware of any literature looking at this area.

Study Limitations

Limitations of our study include the fact that the study sample was from a single Irish medical school, which limits generalizability. In addition, the OSCE is intrinsically a standardized, checklist-driven appraisal with its own limitations that may result in unfair evaluations of students who may not 'play the game' as well as others.

The GRE used in this study has not been formally validated although a very similar GRE has been used in a previous study assessing empathy of medical students by SPs. The GRE was selected on the basis that it could be completed easily and quickly by both examiners and SPs in the one-minute interval between OSCE stations. If a tool is to be introduced to facilitate SPs assessment of students empathy during the OSCE its needs to be jargon free, require minimal training in its use and be easy and quick to complete accurately.

There could be a training bias, in that different SPs could be coached slightly differently and this could affect the integrity of their evaluations. This issue is minimized in this study because most cases were trained by the same module tutor. In addition, SPs capacity to accurately evaluate empathy could be affected by their own familiarity and ease with the case they are portraying. An SP who is more familiar and comfortable with the case may be able to put an anxious and tense student at ease thus facilitating more natural empathetic behavior.

Implications for Educators

- The female students' scores on the JSPE-S were significantly higher than those of their male peers. Medical educators need to consider whether gender issues need to be explicitly considered in the assessment processes and/or the teaching of empathy in medical schools.
 - The assessment of empathy by simulated patients (SPs) in an OSCE is a valid additional means of assessing empathy in medical students.
 - The use of empathy assessment tools by SPs in an OSCE may allow for educators to more objectively and validly demonstrate the need for improvement in empathy to individual learners.
-

Disclosures On behalf of all the authors, the corresponding authors states that there are no conflicts of interest.

References

1. Bleakley A, Bligh J. Students learning from patients: Let's get real in medical education. *Adv Health Sci Educ Theory Pract.* 2008;13:89–107.
2. Brownell AKW, Côté L. Senior residents' views on the meaning of professionalism and how they learn about it. *Acad Med.* 2001;76:734–7.
3. Stewart MA. Effective physician–patient communication and health outcomes: a review. *CMAJ.* 1995;152:1423–33.

4. Beckman HB, Frankel RM. Training practitioners to communicate effectively in cancer care: it is the relationship that counts. *Patient Educ Couns*. 2003;50:85–9.
5. Coulehan J, Platt F, Egner B, et al. "Let me see if I have this right": words that build empathy. *Ann Intern Med*. 2001;135:221–7.
6. Rees-Lewis JC. Patients views on quality of care in general practice: literature review. *Soc Sci Med*. 1994;39:655–71.
7. Levinson W, Gorawa-Bhat R, Lamb J. A study of patient cues and physician responses in primary care and surgical settings. *JAMA*. 2000;284:1021–7.
8. Larson EB, Yao X. Clinical empathy as emotional labour in the patient physician relationship. *JAMA*. 2005;293:1100–6.
9. Shapiro J. How do physicians teach empathy in the primary care setting? *Acad Med*. 2002;77:323–8.
10. Shanafelt TD, Bradley KA, Wipf JE, Back AL. Burnout and self-reported patient care in an internal medicine residency program. *Ann Intern Med*. 2002;136:358–67.
11. Shanafelt TD, West C, Zhao X, et al. Relationship between increased personal well-being and enhanced empathy among internal medical residents. *J Gen Intern Med*. 2005;20:559–64.
12. Hojat M, Gonnella JS, Nasca TJ, Mangione S, Vergare M, Magee M. Physician empathy: Definition, components, measurement and relationship to gender and specialty. *Am J Psychiatry*. 2002;159:1563–9.
13. Hojat M, Gonnella JS, Mangione S, et al. Empathy in medical students as related to academic performance, clinical competence, and gender. *Med Educ*. 2002;36:522–7.
14. Henry SG. Playing doctors. *JAMA*. 2005;294:2138–9.
15. Berg K, Majden J, Berg D, Veloski J, Hojat M. Medical students' self-reported empathy and simulated patients' assessments of student empathy: an analysis by gender and ethnicity. *Acad Med*. 2011;86:984–8.
16. Macnaughton J. The dangerous practice of empathy. *Lancet*. 2009;373:1940–1.
17. Stein and Stein. In: Stein E, Stein W, editors. *On the Problem of Empathy*. Washington DC: ICS Publications; 1989.
18. Pedersen R. Empirical research on empathy in medicine—a critical review. *Patient Educ Couns*. 2009;76:307–22.
19. Mercer SW, Reynolds WJ. Empathy and quality of care. *Br J Gen Pract*. 2002;52(suppl):S9–13.
20. Hemmerdinger JM, Stoddart SDR, Lilford RL. A systematic review of tests of empathy in medicine. *BMC Med Educ*. 2007;7:24.
21. Hodges B. Creating, monitoring, and improving a psychiatry OSCE: a guide for faculty. *Acad Psychiatry*. 2002;26:134–61.
22. Wallace J, Rao R, Haslam R. Simulated patients and objective structured clinical examinations: review of their use in medical education. *Adv Psychiatr Treat*. 2002;8:342–8.
23. Brenner A. Uses and limitations of simulated patients in psychiatric education. *Acad Psychiatry*. 2009;33:112–9.
24. Mellsop GW, Joanna MD, Selim El B, David M. Patients' appraisal of psychiatric trainee interview skills. *Academic Psychiatry*. 2012;36:374–79.
25. Whelan P, Church L, Kadry K. Using standardized patients' marks in scoring postgraduate psychiatry OSCEs. *Acad Psychiatry*. 2009;33:319–22.
26. Bennett AJ, Arnold LM, Welge JA. Use of standardized patients during a psychiatry clerkship. *Acad Psychiatry*. 2006;30:185–90.
27. Hojat M, Mangione S, Nasca TJ, et al. The Jefferson Scale of Empathy: development and preliminary psychometric data. *Educ Psychol Meas*. 2001;61:349–65.
28. Hojat M, Gonnella JS, Nasca TJ, et al. Empathy scores in medical school and ratings of empathic behavior in residency training 3 years later. *J Social Psychol*. 2005;145:663–72.
29. Hojat M, Mangione S, Nasca TJ, et al. An empirical study of decline in empathy in medical school. *Med Educ*. 2004;38:934–41.
30. Mangione S, Kane GC, Caruso JW, et al. Assessment of empathy in different years of internal medicine training. *Med Teach*. 2002;24:370–3.
31. Hojat M, Gonnella JS, Nasca TJ, et al. Empathy scores in medical school and ratings of empathic behaviour in residency training 3 years later. *J Soc Psychol*. 2005;145:663–72.
32. Di Lillo M, Cicchetti A, Scalzo AL, Taroni F, Hojat M. The Jefferson Scale of Physician Empathy: preliminary psychometrics and group comparisons in Italian physicians. *Acad Med*. 2009;84:1198–202.
33. Kataoka HU, Koide N, Ochi K, Hojat M. Measurement of empathy among Japanese medical students: psychometrics and score differences by gender and level of medical education. *Acad Med*. 2009;84:1192–7.
34. Rahimi-Madiseh M, Tavakol M, Dennick R, Nasiri J. Empathy in Iranian medical students: a preliminary psychometric analysis and differences by gender and year of medical school. *Med Teach*. 2010;32:e471–8.
35. Berg K, Majdan JF, Berg D, Veloski J, Hojat M. A comparison of medical students' self-reported empathy with simulated patients' assessments of the students' empathy. *Med Teach*. 2011;33:388–91.
36. Hojat M, Vergare MJ, Maxwell K, et al. The Devil is in the Third Year: A Longitudinal Study of Erosion of Empathy in Medical School. *Acad Med*. 2009;84:1182–91.
37. Tavakol S, Dennick R, Tavakol M. Psychometric properties and confirmatory factor analysis of the Jefferson Scale of Physician Empathy. *BMC Med Educ*. 2011;11:54.
38. Schwartzman E, Hsu DI, Law AV, Chung EP. Assessment of patient communication skills during OSCE: Examining effectiveness of a training program in minimizing inter-grader variability. *Patient Educ Couns*. 2011;83:472–7.
39. Colliver JA, Willis MS, Robbs RS, et al. Assessment of empathy in a standardized-patient examination. *Teach Learn Med*. 1998;10:8.