Title Page

Measuring Medical Students' Empathy: Exploring the Underlying Constructs of and Associations Between, Two Widely Used Self-Report Instruments in Five Countries.

Patrício Costa, PhD , Marco Carvalho-Filho, MD, PhD, Marcelo Schweller, MD, PhD, Pia Thiemann, PsyD, Ana Salgueira, PsyD, John Benson, MD, MRCP, FRCGP, Manuel João Costa, PhD, Thelma Quince, PhD.

Patrício Costa is Assistant Professor at the Health Sciences School of the University of Minho and researcher in medical education at the Life and Health Sciences Research Institute of the University of Minho, Portugal.

Marco Antonio de Carvalho-Filho is Professor of Medicine, Department of Emergency Medicine, Faculty of Medical Sciences, State University of Campinas (Unicamp), Campinas, Brazil.

Marcelo Schweller is an Emergency Medicine Physician and Preceptor, in Faculty of Medical Sciences, University of Campinas (Unicamp), Campinas, Brazil.

Pia Thiemann is Research Assistant in the Palliative Care Group, Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, UK.

Ana Salgueira is a PhD student at the Health Sciences School of the University of Minho and researcher in medical education at the Life and Health Sciences Research Institute of the University of Minho, Portugal.

John Benson is a Senior Lecturer in General Practice and Director of the General Practice Education Group, Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, UK.

Manuel Costa is Associate Professor at the Health Sciences School of the University of Minho and researcher in medical education at the Life and Health Sciences Research Institute of the University of Minho, Portugal.

Thelma Quince is Medical Educational Research Associate in the Primary Care Unit, Department of Public Health and Primary Care and the School of Clinical Medicine, University of Cambridge, UK.

Corresponding author: Dr. Thelma Quince,

The Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, IPH, Forvie Site, Robinson Way, Cambridge CB2 0SR. UK

taq1000@medschl.cam.ac.uk

Tele: 0044(0)1223 330364 Fax: 0044(0) 1223 762512

Word Count 3047

Abstract

Purpose: The definition of clinical empathy is unclear and evidence about its development among undergraduate medical students conflicting. These problems may stem from the instruments used to measure medical students' empathy. The authors sought to enhance understanding of the underlying constructs of two of the most widely used self-report instruments: Davis's Interpersonal Reactivity Index (IRI) and Jefferson Scale of Empathy (Student version) JSE (S), and of the distinctions and associations between these instruments.

Methods: IRI and JSE-S were administered in three separate studies between 2007 and 2014, in 5 countries (Brazil, Ireland, New Zealand, Portugal, United Kingdom). Data from 3069 undergraduate medical students were collected. Exploratory factor analyses, correlation analyses and multiple linear regression analyses were performed.

Results: Exploratory Factor Analysis yielded identical results in each country, confirming the subscale structures of each instrument. Results of correlation analyses indicated significant but weak correlations (r=.313) between the total IRI and JSE-S scores. All inter-correlations of IRI and JSE-S subscale scores were statistically significant but also weak (range r=-.040 - r=.306). Multiple linear regression models revealed that the IRI subscales were weak predictors of all JSE-S subscale and total scores. The IRI subscales explained between 8.9% and 15.3% of variance for JSE-S subscales and 19.4% for JSE total score.

Conclusions: The IRI and JSE-S are only weakly related, suggesting that they measure different constructs. Research into, and interventions addressing medical student empathy need clearer understanding and definition of the construct under consideration as results from the two scales are not comparable.

Introduction

Empathy is a core element in patient care. It may enhance patients' satisfaction and trust, so facilitating compliance and adherence to therapy. [1-4] Receipt of empathy may be therapeutic in its own right. [5-7] Greater trust by the patient may encourage better exchange of information in consultations, enabling better diagnosis and shared decision making. [1,4,8] From the doctor's perspective, empathy may lead to better clinical decisions, [9-11] greater job satisfaction, and enhanced psychological well-being. [10,12-15] The development of empathy among medical students would seem crucial to future patient care.

However "clinical empathy" is poorly defined and measured.[9,16-18] It has been seen as the ability to:

1. understand the patient's situation, perspective and feelings (and their attached meanings)
2. communicate that understanding and check its accuracy
3. act on that understanding with the patient in a helpful (therapeutic) way. [19]

This definition implies a multi-dimensional construct incorporating affective, cognitive, behavioural and moral components.

For patients it is the empathetic behaviour they receive which is important. However asking patients to assess medical students' empathy is problematic and studies using simulated or standardised patients have produced mixed results. [20-22] Most studies of medical student empathy rely on self-report measures, rather than direct observations.[18] The most widely used instruments are Davis's Interpersonal Reactivity Index (IRI) and the Jefferson Scale of Empathy (Student version) JSE- S. [23,24]

Davis considered empathy to be a set of related constructs, concerning responsivity to others, but each discriminable from each other. [23] The IRI comprises 28 items (9 negative) forming four, 7 item, subscales: Perspective Taking (IRI-PT) assessing consideration for the psychological point of view of the other person; Empathetic Concern (IRI-EC) assesses consideration for their feelings and concerns, Personal Distress (IRI-PD) assessing personal anxiety in tense interpersonal settings and Fantasy Scale (IRI-FS) assessing tendencies to transpose oneself imaginatively into the feelings and actions of fictional characters. [23] IRI-EC and IRI-PT have be seen as "other-oriented" and IRI-PD and IRI-FS as "self-oriented".[25,26] IRI-EC and IRI-PD relate to affective aspects while IRI-PT and IRI-FS to cognitive aspects. Respondents rate the extent to which statements apply, from "Does not describe me very well" to "Describes me very well" on a 5-point Likert scale. [23]

The IRI has been used in a wide variety of contexts including neurological studies,[27] clinical conditions [28,29] and criminology. [30] It has been found to have good psychometric properties and is regarded as a valid, and reliable instrument for measuring empathy. [31] Although less well used with medical students the factorial structure proposed by Davis has been supported in studies among college students.[32]

The JSE was developed as measure of empathy applicable to patient care.[24] This 20 item scale comprises three underlying factors: Perspective Taking (10 positively worded), Compassionate Care (8 negatively worded) and Standing in the Patient's Shoes (2 negatively worded). Most studies of medical student empathy report only the total JSE-S score. [33] Respondents rate their level of agreement with each statement on an ascending 7-point Likert scale (1 to 7). Used in a variety of cultural settings for assessing the empathy of medical students, nurses and other healthcare students its validity and reliability have been well supported [18,34-38]

Among healthcare students and practitioners the IRI subscales commonly used are the "other oriented" scales of IRI-EC and IRI-PT.[39,40] The distinction between cognitive and affective components is less clear in the JSE with both "Standing in the patient's shoes" and "Perspective Taking" appearing to reflect the cognitive component of empathy. [41]

The IRI and JSE-S were conceived with different populations in mind. The IRI is applicable to the general population and seen to reflect generic or dispositional empathy.[23] The JSE is applicable to those engaged in healthcare and hence seen to measure empathy specific to that context.[24] Studies of undergraduate medical students in different countries using the IRI have shown that they fall within the norms for IRI-EC and IRI-PT.[42] It would seem reasonable to expect at least a moderate associations between some of the IRI and JSE-S subscales. Further, a study of medical students found a moderate correlation between the total scores of the JSE-S and IRI ($r = .45$, $p < .01$). [44] However, unlike the JSE-S, the IRI subscales are not normally summed to a total score.[18]

To consider the underlying structural and conceptual differences of the IRI and JSE this study asked:

1] Whether the underlying factorial structures of the IRI and JSE-S reflected the dimensional constructs of empathy indicated by their respective subscales:

2] How the scales related to each other in terms of their total and subscale scores:

3] Whether scores on the IRI as a generic measure of empathy predicted scores on the JSE-S as a measure of empathy specific for healthcare.

Method

Data were obtained from three separate studies.

Study one included data from 16 UK medical schools, one in Ireland and one in New Zealand. All students beginning and all students approaching the end of, their undergraduate medical education were invited by email to participate in an international comparison. An online questionnaire survey took place between September 2013 and July 2014, and examined empathy, (IRI and JSE-S), psychological wellbeing, death anxiety and attitudes towards end of life care. Overall ethical approval was granted by the Psychology Research Ethics Committee of the University of Cambridge and by the relevant bodies in each participating school.

Study two was based in one Portuguese University. For each year between 2007 and 2014 students beginning their undergraduate medical education were invited to complete a paper questionnaire covering the JSE-S. In January and February 2013 students in all years were invited to complete a paper questionnaire covering the IRI. Data collection and storage were authorized by the Portuguese Commission for Data Protection (CNDP: 10432/2011). Retrospective approval was obtained: - Subcomissão de ética para as Ciências da Vida, process SECVS - 071/2013.

Study three was undertaken in one university in Brazil. In 2011 and 2012, all undergraduate medical students in years 4 and 6 were invited to complete paper questionnaires covering both the IRI and JSE-S. Ethical approval was granted by the Research Ethics Committee in Human Beings at the Faculty of Medical Sciences of Unicamp.

In all studies participants gave prior consent either in writing or online, and participation was voluntary and anonymous with no incentives offered.

Participants:

Table 1. Sample characterization (Sex, Country and Entry Scheme)

Medical schools in the studies offered "standard" courses lasting 5/6 years, with students typically aged 18 or 19 on entry. Some schools also offered 4 year accelerated "graduate entry" courses for students typically aged 21 or over on entry who had obtained a first degree.

The timing and balance of biomedical science and clinical course components in the participating schools varied. Some schools devoted the early years largely to biomedical sciences, others adopted a more integrated approach. This study did not set out examine in detail the nature of the courses offered and simple labels such as "integrated" may not fully represent course content and structure.

The sample comprised 3,069 medical students (Table 1) of whom 2059 (67.1%) were from the UK and 1887 (61.5%) were female. The majority of students (2619, or 85.3%), had entered standard courses. A statistically significant, but small in terms of effect size, difference in gender composition of samples in each country was found, with proportionately fewer males among the Portuguese sample and proportionately more males among the Brazilian sample. ($\chi^2_{(4,\ n=\ 3069)} = 9.6$, p =.047, Cramer's V=.056).

Instruments:

We used the JSE-S (student version) in all countries. The IRI Portuguese version is 24 items as opposed to 28 and was the result of a validation study which demonstrated

factor loadings <.35 for items numbered 1, 15, 18 and high standardized residual for item 10. [44] We adjusted the item numbers of the 28 item IRI used in other countries to those of the Portuguese version to allow IRI data to be merged.

Data analysis and modelling strategy:

We merged the JSE and IRI items and converted them into in the same scale using z scores. We used exploratory factor analysis (EFA) to explore the factorial structure of all IRI and JSE items, using the scree plot, the Kaiser's eigenvalue > 1 method and Parallel Analysis (PA) to explore the optimal number of factors and principal axis factoring (PAF) with oblimin rotation. We used Cronbach's Alpha to measure internal consistency and Pearson correlation coefficient to examine associations between subscale and total scores of each scale. We examined the effects of country and sex on subscales scores using MANOVA and on the total JSE-S scores using ANOVA. (The results of these are presented in the appendix.) We used multiple linear regression to examine the extent to which IRI scores predicted JSE scores (total and subscale) with IRI subscales, country, sex and entry scheme being independent variables. Data analyses were performed using IBM SPSS Statistics v22 and the R.Commander [45] and the *psych* package. [46] We considered P values of 5% as significant and interpreted effect sizes according to values given by Cohen(1988). [47]

Results

Latent dimensions of the IRI and JSE: (Table 2)

Table 2. Exploratory Factor Analysis (communalities and loadings) for IRI and JSE items and Cronbach's Alpha scores

An EFA performed on the combined IRI and JSE-S datasets z-scores, resulted a **nine** factor solution according to the Kaiser's eigenvalue >1 method and an **eight** factor

solution according to the PA and scree plot analysis (the line straightens after the eighth factor). Both solutions produced some dimensions with critical internal consistency values (Cronbach's alphas lower than 0.60). Therefore, the theoretically anticipated solution of **seven** factors accounting for 44.6% of variance was tested and led to higher and more acceptable internal consistency values. For the final EFA **seven** factors solution, measures of appropriateness of factor analysis were checked including KMO = .873 and Bartlett's test ($\chi^2_{(946)}$ = 33016, p <.001).

Considering the theoretical structure, a practical significance of 5% and an acceptable factor loading of >/=0.224 were found for all item. All items clustered as expected and recorded the highest loading on their original dimension with the exception of JSE-S item 14, which loaded higher onto JSE-PT than onto its original JSE-CC dimension. Nine items showed significant double loadings, but none crossed the two scales. In each of the five countries the seven factor structure revealed a satisfactory fit, (Table 2) with the exception for JSE-SPS dimension in Ireland (Cronbach's alpha=0.472).

Pearson correlations for all IRI and JSE subscales: (Table 3)

Table 3. Pearson correlations for IRI and JSE subscales and total scores.

Within scale associations: For both the IRI and JSE-S correlations between each subscale score and the total score were statistically significant: for the IRI generally strong (r= .431 to r=.712), for the JSE-S, moderate (r=.377) to very strong (r=.854). Correlations between the subscales within each scale were significant but less strong. For the IRI these ranged from r=.061 between IRI-EC and IRI-PD to r=.403 between IRI-EC and IRI-PT. A negative association was found between IRI-PD and IRI-PT. For the JSE the range was r=.114 between JSE-SPS and JSE-PT and r=.467 between JSE-PT and JSE-CC.

Between scale associations: The correlation between total scores of JSE-S and IRI was positive and significant, but weak r=.313. All inter-correlations of JSE-S and IRI subscale scores were statistically significant but weak, ranging from r=-.040 (JSE-PT with IRI-PD) to r=.306 (JSE-PT with IRI-EC). The only exception was the non-significant, negative correlation between IRI-PD and JSE-CC (r=-.016).The correlation between the subscales scores of one scale and the total score of the other scale were also all statistically significant but weak. IRI-PD was negatively associated with all JSE subscales scores.

Multiple linear regression models:

Table 4. Multiple linear regression models for JSE dimensions.

The multiple linear regression analyses tested whether the IRI subscales, gender, country, and entry scheme significantly predicted JSE subscale and total scores. The reference categories were female, UK and standard entry (Table 4.) All regression models were significant, with a relatively low adjusted R squared, varying between 8.9% and 15.3% of explained variance for JSE-S subscales and 19.4% for JSE total score.

With the exception of IRI-PD, all IRI subscales were significant, positive, predictors of each JSE subscale. Sex, was significant in all regression models except for JSE-PT. The extent to which students in countries differed from those in the reference country (UK) varied between instruments and between subscales of each instrument. Overall students in Brazil differed most from those in the UK whereas students in Ireland differed least. Entry scheme was not significant in any of the four tested models. The most pronounced predictor of total JSE-S score was IRI-EC.

Discussion

This study found that the dimensional structure of each instrument reflected its composite subscales with strong internal consistencies. The EFA results supported the cross-cultural construct validity and stability of both scales. For the IRI, our study confirmed Davis's 4 factor structure in 5 countries. To the authors' knowledge this factorial structure has been confirmed in studies of college students albeit with minor variations [32] but never before among medical students .

For the JSE, our results broadly accord with Hojat's original 3 factor structure and within that, the prominence of Perspective Taking (JSE-PT).[24] The only exception to this was the result for JSE-SPS in Ireland, possibly resulting from a combination of small sample size and small number of contributing items (n=2).

Our findings accord well with international JSE-S studies of medical students which, broadly support the 3 factor structure and their respective relative importance[37, 38] but with minor variations. For example studies of German and Japanese medical students support the JSE-PT construct but report variations in JSE-CC, possibly attributable to cultural differences. [37,49] A recent US study found the factorial structure of the JSE-S varied between preclinical and clinical medical students. Such analysis was beyond the scope of our study. [41]

The shared variance between the scales and subscales found in this study support the view that the scales measure different but related constructs. This view is further supported the correlation results which revealed only weak correlations despite an expectation of moderate correlations particularly in respect of subscale scores of IRI-PT and JSE-S-PT and IRI-EC and JSE-S-CC. Multiple linear regression models similarly suggested that all IRI subscales were weak predictors of the JSE-S subscale

scores and total score, with the strongest predictor of the JSE-S total score being the IRI-EC.

The study supports the view of gender differences in respect of empathy with women recording higher scores on self-report measures.

The suggestion that the two scales measure different but related constructs has implications for medical education, and medical education research. Care is needed in comparing studies using different scales. Conflicting results of studies of the trajectory of empathy during undergraduate medical education may, in part, be attributable to the use of instruments which are not comparable.[33,39] Similar implications may apply to intervention studies.

The suggested difference between to the two scales points to the need to clarify the constructs being measured. Whereas the IRI measures generic empathy the JSE-S may measure some idealized view of an empathic doctor-patient relationship. This distinction is reflected in differences in the wording of the scales. The IRI asks respondents the extent to which each statement "describes" his or herself, with all items containing the words "I" or "me". The JSE-S asks respondents for their level of agreement with statements about either how "doctors" should behave or the doctor-patient relationship, with only 4 items relating to the individual.

The IRI and JSE-S were conceived with different populations in mind. Generic empathy may be shaped by personality, certain life experiences and possibly culture. Studies in various cultures suggest that psychological conditions exert the largest influence. [42] As an idealized view of an empathetic doctor-patient relationship JSE-S scores may be shaped by cultural influences affecting both medical education and patient expectations. [49] These may be more amenable to training and education than IRI scores.[50] Studies examining the impact of educational interventions aimed at enhancing

empathy have found a larger increase in JSE-S scores than in IRI scores. [22] However idealized views may also be more vulnerable to the hidden curriculum. [16]

To characterize and clarify how the IRI and JSE-S constructs relate to each other, and how they change during medical education there is a need for more studies using both instruments, for more qualitative and mixed methods work and for more longitudinal work. If, as suggested the JSE-S measures context specific empathy then greater attention needs to be paid to that context including perhaps critical incidents and medical course content and structure. Our study only included undergraduate students. Comparable studies of post graduate medical students and/or physicians are needed.

This is one of the few studies of medical students using both the IRI and JSE-S and to the authors' knowledge the only study to include European, Brazilian and New Zealand data. One of its strengths is the large number of participants drawn from 5 countries. Whilst sample size in each country differed this was not a major limitation since one a main goal of the study was to explore the latent structure of IRI and JSE-S. Another limitation is that the analyses were run on the 24-item version of the IRI and did not include age per se. Our study drew data from countries with essentially "European" values which may explain the absence of marked cultural differences. Studies comparing the IRI and JSE-S among medical students in countries with very different cultural backgrounds, particularly those in which extreme scores have been recorded would be valuable in identifying differences between generic empathy and what is perceived to be an appropriate empathetic doctor/patient relationship. [42]

Conclusions

The factor analysis undertaken in this study supports the accepted factorial structure of the IRI and JSE-S and reaffirms the relationship of their respective subscales to the underlying dimensions of empathy: affective and cognitive, and for the IRI self-oriented versus other-oriented. These results are enhanced by being confirmed in 5 countries. However, this study suggests that the IRI and JSE are structurally different, weakly related concepts: the former generic or dispositional empathy, the latter context specific empathy. Consideration of this distinction may give rise to implications for medical education and may have implications for patient care. There is a need for more studies using both instruments, involving those at different stages in medical training, and for more longitudinal and qualitative studies in order to understand the practical implications of this distinction.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

Patrício Costa had full access to of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

The authors are indebted to the following for their help and support in data collection: James Brimicombe, BSc., University of Cambridge, David Bristow, PhD., University of Plymouth, Ana Da Silva, PhD., University of Swansea, Jonathan Hales, MBChB MRCGP., University of Leicester, Joanne Hart, PhD. C Psychol., University of Manchester, Peter Hayes, MBChB BAO MICGP, University of Galway, Republic of Ireland, Kathy Jackson, MBChB MRCGP MD FHEA., University of St Andrews, Susan Jamieson, PhD., University of Glasgow, Margaret Jones, MB ChB MRCGP., University of Nottingham, Paul Kinnersley, MBChB., University of Cardiff, Paula McDonald, MB ChB MRCGP MPH MFPHM Brighton & Sussex Medical School, Jean

McKendree, PhD., Hull/York Medical School, Helen Moriarty, MBChB MGP DPH., University of Otago, New Zealand, Anne Stephenson, MBChB MRCGP PhD FHEA., King's College University of London, Anna Stickland, MB ChB MRCGP., University of Liverpool, Kate Thomas, MBChB FRCGP., University of Birmingham, Trevor Thompson, MBBS MRCGP, PhD., University of Bristol, Helen Watson, PhD., University of Exeter.

References

1. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. Eval Health Prof. 2004;27:237-51.
2. Vermeire E, Hearnshaw H, Van Royen P, Denekens J. Patient adherence to treatment: Three decades of research. A comprehensive review. J Clin Pharm Ther. 2001;26:331-42.
3. Derksen F, Bensing J, Lagro-Janssen A. Effectiveness of empathy in general practice: A systematic review. Br J Gen Pract. 2013;63:e76-e84.
4. Parkin T, de Looy A, Farrand P. Greater professional empathy leads to higher agreement about decisions made in the consultation. Pat Educ Couns. 2014;96:144-150.
5. Hojat M, Louis DZ, Markham FW, Wender R, Rabinowitz C, Gonnella JS. Physicians' empathy and clinical outcomes for diabetic patients. Acad Med. 2011;86:359-64.
6. Rakel D, Barrett B, Zhang Z. et al. Perception of empathy in the therapeutic encounter: Effects on the common cold. Pat Educ Couns. 2011;85:390-7.
7. Canale SD, Louis DZ, Maio V. et al. The relationship between physician empathy and disease complications: An empirical study of primary care physicians and their diabetic patients in Parma, Italy. Acad Med. 2012;87:1243-9.
8. Steinhausen S, Ommen O, Thüm S. et al. Physician empathy and subjective evaluation of medical treatment outcome in trauma surgery patients. Pat Educ Couns. 2014; 95:53-60.
9. Larson EB, Yao X. Clinical empathy as emotional labor in the patient-physician relationship. JAMA. 2005;293:1100-6.

10. Coulehan JL, Platt FW, Egener B, et al. "Let me see if I have this right.. ": Words that help build empathy. Ann Intern Med. 2001;135:221-7.
11. Beckman HB, Frankel RM. Training practitioners to communicate effectively in cancer care: It is the relationship that counts. Pat Educ Couns. 2003;50:85-9.
12. Thomas MR, Dyrbye LN, Huntington JL.et al. How do distress and well-being relate to medical student empathy? A multicenter study. J Gen Intern Med. 2007;22:177-83.
13. DiLalla LF, Hull SK, Dorsey JK. Effect of gender, age, and relevant course work on attitudes toward empathy, patient spirituality, and physician wellness. Teach Learn Med. 2004;16:165-70.
14. Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. JAMA. 2002;287:3143-6.
15. West CP, Huschka MM, Novotny PJ, et al. Association of perceived medical errors with resident distress and empathy: A prospective longitudinal study. JAMA. 2006;296:1071-8.
16. Neumann M, Bensing J Mercer S, Ernstmann N, Ommen O, Pfaff H. Analysing the “nature” and “specific effectiveness” of clinical empathy: A theoretical overview and contribution towards a theory based research agenda. Pat Educ Couns. 2009; 74:339-346.
17. Pedersen R. Empirical research on empathy in medicine-A critical review. Pat Educ Couns. 2009;76:307-22.
18. Hemmerdinger J, Stoddart SD, Lilford RJ. A systematic review of tests of empathy in medicine. BMC Med Educ. 2007;7, doi:10.1186/1472-6920-7-24.
19. Mercer SW, Reynolds WJ. Empathy and the quality of care. Br J Gen Pract. 2002:52 (suppl):S9-13.

20. O'Connor K, King R, Malone KM, Guerandel A. Clinical examiners, simulated patients and student self-assessed empathy in medical students during a psychiatry Objective Structured Clinical Examination. Acad Psychiatry. 2014; 38:451-457.

21. Berg K, Majden JF, Berg D.Veloski J, Hojat M. Medical students self-reported empathy and simulated patients' assessment of student empathy: An analysis by gender and ethnicity. Acad Med. 2011;86:984-988.

22. Schweller M, Costa FO, Antonio MARGM, Amaral EM, De Carvalho-Filho MA. The impact of simulated medical consultations on empathy levels of students at one medical school. Acad Med. 2014 89:632-637.

23. Davis MH. Measuring Individual differences in Empathy: Evidence for a Multidimensional Approach. J Pers Soc Psychol. 1983 44:113-126.

24. Hojat M. Empathy in Patient Care: Antecedents, Development, Measurement, and Outcomes. Philadelphia, PA: Springer; 2007.

25. Banissy MJ, Kanai R, Walsh V, Rees G. Inter-individual differences in empathy are reflected in human brain structure. Neuroimage. 2012;62:2034-2039.

26. Juujärvi S, Myyry L, Pesso K. Does care reasoning make a difference? Relations between care, justice and dispositional empathy. J Moral Educ 2010;39:469-489.

27. Lamm C, Batson CD, Decety J. The neural substrate of Human Empathy: Effects of Perspective-taking and Cognitive Appraisal. J Cogn Neurosci. 2007;19:42-58.

28. Lehmann AQ, Bahcesular K, Brockmann TM, et al. Subjective experience of emotions and emotional empathy in paranoid schizophenia. Psychiatry Res. 2014;220:825-33.

29. Narme P, Mouras H, Duru C, Krystowiak KP, Godfrey O. Emotional and cognitive social processes are impaired in Parkinson's disease and are related to behavioural disorders. Neuropsychology. 2013;27:182-192.
30. Schaffer M, Clark S, Jeglic EL. The role of empathy and parenting style in the development of antisocial behaviors. Crime Delinq. 2009;55:586-599.
31. Baron-Cohen, S, Wheelwright, S. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism and normal sex differences. J Autism Dev Disord. 2004;34:163–175.
32. Fernández AM, Dufey M, Kramp U. Testing the psychometric properties of the Interpersonal Reactivity Index (IRI) in Chile. Eur J Psychol Assess. 2011;27:179-185.
33. Roff S. Reconsidering the "decline" of medical student empathy as reported in studies using the Jefferson Scale of Physician Empathy-Student version (JSPE-S). Med Teach. 2015;37:783-786.
34. Loureiro J, Goncalves-Pereira M, Trancas B, Caldas-de-Almeida JM, Castro-Caldas A. Empathy in the doctor patient relationships as viewed by first year medical students: data on validity and sensibility to change of the Jefferson Measure in Portugal. Acta Med Port. 2011, 24(Suppl 2);431-42.
35. Fields SK, Mahan P, Tillman P, Harris J, Maxwell K, Hojat M. Measuring empathy in healthcare profession students using the Jefferson Scale of Physician Empathy: healthcare provider- student version. J Interprof Care. 2011;25:287-93.
36. Ward J, Schaal M, Sullivan J, et al. Reliability and validity of the Jefferson Scales of Physician Empathy in undergraduate nursing students. J Nurs Meas. 2009;17:73-88.

457 37. Preusche I. Wagner-Menghin M. Rising to the challenge: cross-cultural adaption
458 and psychometric evaluation of the adapted German version of the Jefferson
459 Scale of Physician Empathy for Students –JSPE-S. Adv Health Sci Educ Theory
460 Pract. 2013;18:573-587.
461 38. Myoung-Sun R, Bong-Jin H, Dong Hun L, et al. Evaluation of empathy among
462 Korean medical students: a cross-sectional study using the Korean version of the
463 Jefferson Scale of Physician Empathy. Teach Learn Med. 2010;22:167-171.
464 39. Quince T A, Parker RA, Wood DF, Benson JA. Stability of empathy among
465 undergraduate medical students: a longitudinal study at one UK medical school.
466 BMC Med Educ. 2011;11: 90. doi 10.1186/1472-6920-11-90.
467 40. Paro HBMS, Silveira PSP, Perotta B. et al. Empathy among medical students: Is
468 there a relation with quality of life and burnout? PLoS ONE. 2014;9:e94133.
469 41. Stansfield RB, Schwartz A, O'Brien CL, et al. Development of a metacognitive
470 effort construct of empathy during clinical training: a longitudinal study of the
471 factor structure of the Jefferson Scale of Empathy. Adv Health Sci EducTheory
472 Pract. 2015; doi 10.1007/s10459-015-9605-1.
473 42. Quince T, Thiemann P, Benson J, Hyde S. Undergraduate Medical Student
474 Empathy: Current Perspectives. Advances in Medical Education and Practice.
475 2016 Forthcoming.
476 43. Hojat M, Mangione S, Kane GC, Gonnella JS. Relationship between scores on
477 the Jefferson Scale of Physician Empathy JSPE and the Interpersonal Reactivity
478 Index (IRI). Med Teach. 2005;27:625-8.
479 44. Limpo T, Alves RA, Castro SL. Medir a empatia: Adaptação portuguesa do
480 Índice de Reactividade Interpessoal. Laboratório de Psicologia. 2013;8:171-184.

45. Fox J. The R Commander: A Basic Statistics Graphical User Interface to R. Journal of Statistical Software 2005;14:1-42.

46. Revelle W. Psych: Procedures for Personality and Psychological Research. Northwestern University, Evanston, Illinois, USE, http:/CRAN.R-project.org/package=psych Version=1.5.8.

47. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Earlbaum Associates. 1988

48. Roh MS, Hahm BJ, Lee DH, Suh DH. Evaluation of empathy among medical students: A cross-sectional study using the Korean version of the Jefferson Scale of Physician Empathy. Teach Learn Med. 2010; 22:167-171.

49. Kataoka HU, Koide N, Ochi K, Hojat M, Gonnella JS. Measurement of empathy among Japanese medical students: psychometrics and score differences by gender and level of medical education. Acad Med. 2009;84:1192-7.

50. Hegazi I, Wilson I. Maintaining empathy in medical schools: it is possible. Med Teach. 2013;35:1002-8.