


RESEARCH NOTE

Open Access



Changes in character strengths after watching movies: when to use rasch analysis

Sirinut Siritikul¹, Sirikorn Chalanunt¹, Chitipat Utrapiromsook¹, Suchanard Mungara¹, Tinakon Wongpakaran^{2*} , Nahathai Wongpakaran², Pimolpun Kuntawong² and Danny Wedding³

Abstract

Objective: Professionalism is a critical part of a medical education, and various activities have been proposed to enhance professionalism among medical students. Watching films is an activity to promote character related to professionalism. Limitation of such is a single group pre-posttest design raising concerns about the errors of measurement. The study aimed to demonstrate a method to deal with this design using Rasch analysis.

Results: This study used a pre-posttest design with 40 first year medical students. All participated in a 3-day activity that involved watching four selected movies: *Twilight*, *Gandhi*, *The Shawshank Redemption* and *Amélie*. These films offer compelling illustrations of the themes of self-regulation, humility, prudence and gratitude, respectively. All participants completed a 10-item composite scale (PHuSeG) addressing these themes before and after watching the movies. When determining who benefitted from the intervention, paired t-tests on the results of a Rasch analysis were used to evaluate changes between pre- and posttest. Using Rasch analyses, we could document the stability of the items from pre- to posttest, and significant changes at both the individual and group levels, which is a useful and practical approach for pre- and posttest design. Moreover, it helps validate the psychometric property of the instrument used.

Keywords: Movies, Character strength, Composite scale, Prudence, Humility, Self-regulation, Gratitude, Medical students

Introduction

Medicine is a demanding profession. In addition to knowledge, medical students are expected to be skilled at relating to patients, the patients' caregivers and other healthcare professionals [1, 2]. The core attributes pertinent to human connection are found across all cultures. Such attributes include the ability to build a therapeutic relationship with patients, skills in providing

patient-centered care, effective communication and interpersonal skills [3, 4].

Character strengths and virtues are considered core characteristics valued by moral philosophers and religious thinkers [5]. The virtues associated with positive psychology are wisdom and knowledge, courage, humility, justice, temperance and transcendence [6]. Several character strengths related to the medical professionalism include medical ethics. Studies have shown that some clinical dilemmas require the character strengths of honesty, wisdom, prudence, kindness, courage, hope and wisdom to guide ethical decision-making [7, 8]. Culturally, the characters or virtues associated with being a good doctor include self-regulation, prudence,

*Correspondence: tinakon.w@cmu.ac.th

² Psychotherapy unit, Department of Psychiatry, Faculty of Medicine, Chiang Mai University, 110 Intawaroros Rd., T. Sriphum, A. Muang, Chiang Mai, Thailand

Full list of author information is available at the end of the article



humility and gratitude [1]. These characteristics are briefly described below.

Self-regulation includes behaviors such as calmness and patience. For medical students, self-regulation involves maintaining competence such as taking appropriate action to prevent conflicts of interest when dealing with a pharmaceutical company [9].

Prudence is a strength described as being careful about one's choices, such as not taking undue risks [10]. This value is commonly found among physicians who are noted for their clinical judgment. One topical and salient example of where empathy and prudence are especially needed is finding strategies and resources to improve quality of care and to ease the anxiety concerning nurses caring for patients with COVID-19 [11].

Humility is a character strength characterized by humility and freedom from arrogance. This value is related to work success [12], as well as psychological strength and effectiveness [13].

The last strength is *gratitude*, which is characterized by a general state of thankfulness; gratitude is defined as "the appreciation of what is valuable and meaningful to oneself" [14].

A number of strategies are available to promote character strengths such as mindfulness-based training programs and digital-free tourism [15, 16]; one activity involves watching movies. Cinema-education [17] and related interventions using films attempt to promote psychological health [18]. Research has shown that students' positive orientation and growth initiative can be promoted using a systematic movie-based teaching course [19, 20].

To measure the changes resulting from watching movies, we usually use a pre-post design. This design is used in both academic and in clinical settings. In a pedagogic setting, medical educators may be interested in how students change after a class or intervention. This pre-post evaluation can; however, be biased because of its ordinality. Rasch analysis is one way of addressing this limitation because data have been transformed onto an interval scale. In Rasch analysis, both item difficulty and person ability and parameters are considered and plotted on the same interval-level scale, to reduce errors of measurement [21]. Rather than only examining the change at the group level, the Rasch model analyzes change at an individual level. Rasch analysis allows us to have interval scores for each person, making the comparison between pre- and posttest more accurate.

The aim of the study was to demonstrate the advantage of using Rasch analysis in one group employing a pre-post design to allow (1) verification of the stability of the items, (2) assessment of who will or will not benefit from

this particular intervention and (3) assessment of the reliability of the measurement.

Main text

Methods

This research was approved by the ethics committee of the Faculty of Medicine, Chiang Mai University, Thailand.

Participants and procedure

Forty students joined the movie project as part of the extracurricular activity of the general education curriculum. The subjects comprised 40 medical students: 25 males and 15 females between 19 and 21 years old. Before watching the films, each participant completed a composite scale determining character strength. On the first day, the participants watched two movies, i.e., *The Shawshank Redemption* and *Twilight*. After viewing each movie, a group of five participants discussed and shared their opinions about the movies. Then all participants wrote down their own summary about their attitudes toward each movie. On the second day, the same process was repeated with the other two movies, i.e., *Gandhi* and *Amélie*. Then after viewing all four movies, all participants completed the same questionnaires (Additional file 1: Figure S1).

Measurement

The PHuSeG scale constitutes a composite scale measuring prudence, humility, self-regulation and gratitude [22]. The scale measures positive psychology character strengths associated with professionalism. The PHuSeG consists of ten items and five rating responses. In our previous study of construct validity using the Rasch measurement model, the scale was shown to be unidimensional, and all items had mean square fit statistics between 0.76 and 1.37, which fell within the recommended range of 0.5 to 1.5 [23], with good person and item reliability (0.80 and 0.91, respectively). Cronbach's alpha was 0.84. Three items assessed self-regulation, two assessed humility, four items were selected to measure prudence and one item measured gratitude. The summed raw scores of the PHuSeG scale range from 10 to 50. The higher the score, the greater the positive character strength is present. The study sample yielded a Cronbach's alpha of 0.81 for pretest and 0.84 for posttest.

Films

Four films were used to illustrate each positive attribute based on expert recommendations in the book entitled, "*Positive psychology at the movies: using films to build virtues and character strengths*" [24]. The authors selected the following films based on these criteria: *Gandhi* (Richard Attenborough, 1982), demonstrating humility;

Twilight (Catherine Hardwicke, 2008), demonstrating self-regulation; *The Shawshank Redemption* (Frank Darabont, 1994), demonstrating prudence and *Amelie* (Jean-Pierre Jeunet, 2001), demonstrating gratitude.

Data analysis

Descriptive analyses were conducted on sociodemographics. To compare differences of all dependent variables after intervention, paired t-tests were used. Both PHuSeG ordinal (raw) and interval (Rasch) scores were analyzed separately and compared. For all the analyses, levels of significance were set at $P < 0.05$ and IBM SPSS, Version 22 was used for all analyses.

Rasch models were adopted for this pre-post design analysis because they provided information at individual levels, allowing us to pinpoint who benefits from the intervention. For Rasch analyses, when the data fit the model, interval measures are collected from ordinal scores, yielding more accurate measures of change. Individuals or subjects can be measured within a common frame of reference covering different time points so that the measurement of change becomes a precise numerical representation on a shared linear scale in additive measurement units (logits). Rasch analysis also ensures the invariance (stability) of the instrument across time points. According to Wright [25], the Rasch model can provide answers for two different research questions: one focuses on changes in student performance; the other focuses on changes in item difficulty over time. Before any interpretations, examining the fit of the data to the model is required. Fit statistics along with a principal component analysis (PCA) can determine whether the assumption empirically supports unidimensionality [26].

For fit statistics, outlier-sensitive fit statistics mean square (OUTFIT.MnSq) and information-weighted fit statistics mean square (INFIT.MnSq) ranged from 0.5 to 1.5; these values are considered acceptable [23]. The separation and reliability for the persons and items were examined laying the support for the validity of interpretations. Acceptable values for separation cutoff, person reliability, and item reliability, provided by Wright and Stone [27], were ≥ 2 , ≥ 0.70 , and ≥ 0.8 , respectively. To test for item stability for pre-post comparisons, the differential item function (DIF) was evaluated. The significant DIF was considered when DIF contrast was ≥ 0.64 [28]. Winsteps, Version 4.7.0 was employed for Rasch analysis.

Results

Items with differential item functioning for pre-post comparisons were evaluated, and no significant DIF was observed (Additional file 2: Table S1).

Figure 1 compares the PHuSeG scale items being measured at pre- and posttest. Notably, almost all items maintain their location at the variable, denoting invariance of item calibration. The Wright maps indicate positively skewed person measures and large gaps between H and S items and P and G items.

Additional file 2: Table S1 presents item calibration for the PHuSeG scale showing similar calibration between pre- and posttest. S15 and P8 was rated more difficult after watching movies; while P1 was rated as easier after watching movies. All items showed fit statistics ranging from 0.55 to 1.49. No significant DIF was observed.

The person reliabilities of the scale for pre-posttest were 0.82 and 0.81, while item reliabilities of the scale for pre-posttest were 0.99 and 0.98. PCA confirmed that all items composing the construct of positive character strength did not violate the assumption of unidimensionality.

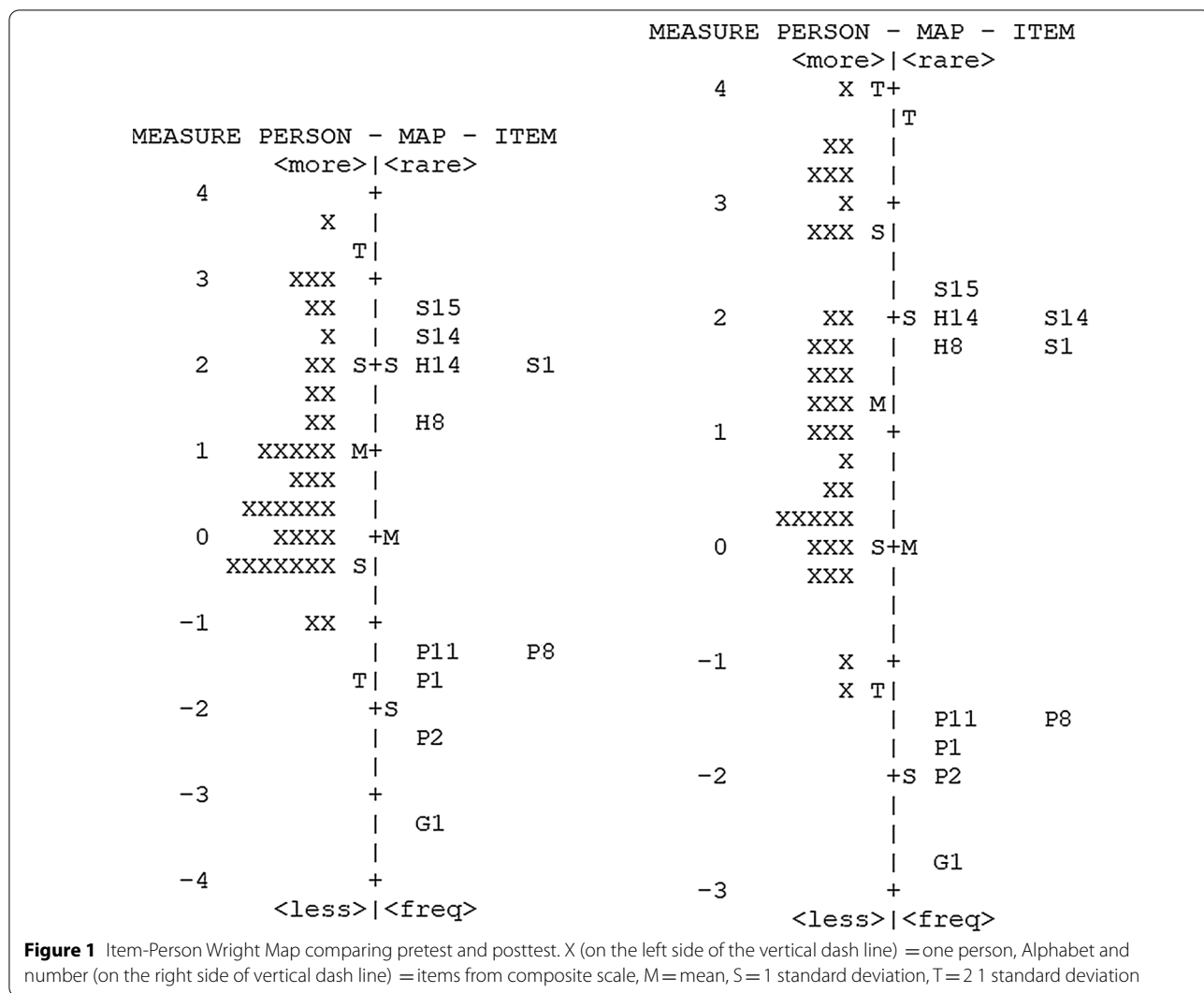
Figure 2 shows that all ten items functioned the same way at both times. G1, S15 and H8 fall on the border of the CI. However, no significant difference was observed in the calibration, and all items were invariant between the two times.

Figure 3 shows that seven persons did not change in PHuSeG score (dots along dashed line), 14 gained higher scores after watching the movies (dots on the left side of the centered dash line), while 19 tended to score lower at posttest (dots on the right side of the centered dash line). Only two scored significantly higher at posttest ($t = -4.45 = p < 0.0001$, while 1 scored significantly lower on posttest ($t = 2.76, p < 0.01$). All the rest showed non-significant change.

Discussion

The study aimed to demonstrate how using Rasch model analysis could be applied to a real-life situation using a pre-post design. The overall results showed that Rasch analysis benefitted this type of design. As previously found, interval measures, produced by Rasch analysis are defined by measurement units that are invariant over the entire domain so that the measurement of change is more accurate [29–31]. Because no DIF between pre- and posttest was evident, these changes, if significant, could be affected by the intervention provided.

Concerning the individual level, 37 individuals showed nonsignificant change; only one improved and one worsened significantly. Rasch models solve current shortcomings when assessing change. A measure is determined for each person so that the change can be measured at the individual level. The statistical significance of change is analyzed by means of the standard errors that define the measures.



As documented in related research, the Rasch model is well-developed and used in the field of educational sciences [32–34]. Our study has supported using Rasch analysis for educators to promote growth of character strengths. By that, investigating students who changed might help identify specific characteristics of students linked to positive responses. These factors can then be used to identify those students who are good candidates for the intervention in advance. Second, the intervention does not automatically have to be the same for all students, but can vary.

Rasch analysis showed that all items of the PHuSeG scale were fitted to the model, and were stable over time. However, the skewness of data generally reduced reliability, and the reliability was expected to increase a little in the sample with normal distribution. The large gap shown on the Wright Map also provided us

important information that new items of appropriate difficulty should be added.

However, item stability could be questioned due to small sample sizes, and further investigation is warranted. In addition, it seems that the PHuSeG is not sufficiently sensitive to detect the change. PHuSeG was notably derived from the 50 items of the four scales. More items that have responsiveness to change should be identified and included in the new PHuSeG in a further investigation.

Regarding the intervention, using various films did not sufficiently affect the targeted constructs. Movies reflecting a specific construct may be more preferable. In addition to watching movies, other interventions might make the change more evident, e.g., a training to strengthen a specific characteristic might be more suitable.

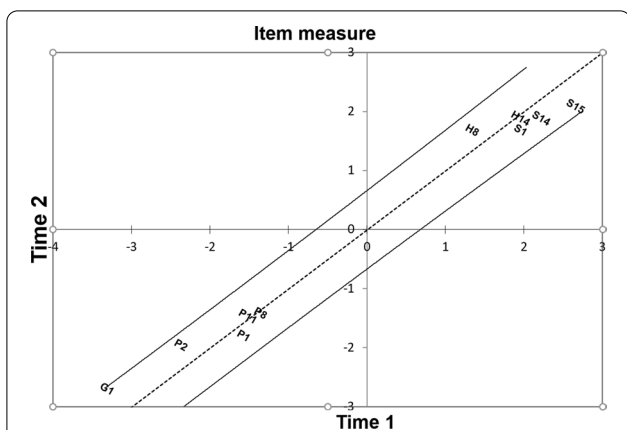


Figure 2 Scatterplot of item difficulty estimates from raked analysis. The line ($y = x$) indicating no change in item difficulty estimates has been drawn to depict change in item difficulty estimates over time. Items above the line represent items estimated as more difficult by the postintervention assessment while items below the line represent items estimated as less difficult by the postintervention assessment. The centered line is the identity line added for reference. The parallel-identity lines are approximate 95% two-sided confidence bands. Item G1 was higher at Time 2 than at Time 1 ($t(78) = -1.77, p = 0.080$), whereas Item S15 was lower ($t(78) = 1.27, p = 0.207$).

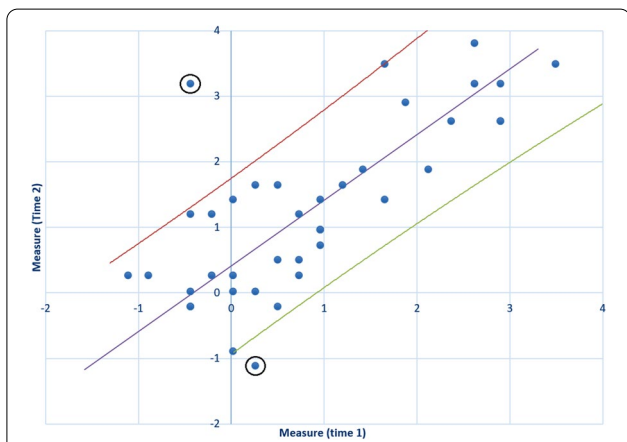


Figure 3 Rasch analysis run on stacked data from the two time points. The person (student) measures at Time 2 are plotted against those at Time 1. Greater measures indicate more positive strengths students. Each dot represents one person. Circled dots indicate statistically significant change at time 2 from time 1. The centered line is the identity line added for reference. The parallel-identity lines are approximate 95% two-sided confidence bands. One person (above the lines) scored significantly higher at Time 2 than at Time 1 ($t(78) = -4.45, p < 0.0001$), whereas the other person (below the lines) significantly scored lower ($t(78) = 2.76, p = 0.007$).

Conclusion

Rasch analyses provide more useful information than other measures. Rasch models allow us to examine the invariance of the instrument across time points, and also provides

some insight regarding individual data. In addition, it helps validate the psychometric property of the instrument used as well.

Limitations

This study was conducted using a small sample size and may not be generalizable to all medical students as only first year medical students were selected.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13104-020-05424-4>.

Additional file 1: Figure S1. Flowchart of the study.

Additional file 2: Table S1. PHuSeG items calibrated at pretest and post-test with mean-square variance–ratio fit statistics.

Abbreviations

INFIT: Information-weighted fit statistics mean square; MNSQ: Mean square; OUTFIT: Outlier-sensitive fit statistics mean square; PCA: Principal Component Analysis; PHuSeG: Prudence, Humility, Self-regulation, and Gratitude.

Acknowledgements

The authors wish to thank all the participants who participated in the study.

Authors’ contributions

SS, SC, CU, SM, TW, NW and DW participated in the concept and design of the study. SS, SC, CU, SM and PK collected and cleaned the data. SS, NW and TW performed the statistical analyses. All authors drafted and edited the manuscript. All authors made substantial contributions to interpret data and revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Funding

This study was self-funded with no support from external agencies.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

This study was approved by the research ethics committee of the Faculty of Medicine, Chiang Mai University. All patients provided written informed consent to participate in the study.

Consent for publication

Consent for publication is not applicable.

Competing interests

All the authors declare that they have no competing interests.

Author details

¹ Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand. ² Psychotherapy unit, Department of Psychiatry, Faculty of Medicine, Chiang Mai University, 110 Intawaroros Rd., T. Sriphum, A. Muang, Chiang Mai, Thailand. ³ Saybrook University, Oakland, California, USA.

Received: 28 June 2020 Accepted: 16 December 2020

Published online: 06 January 2021

References

- Boonyanaruthee V, Chan-Ob T. Good doctor: what the good and bad attributes are. *J Med Assoc Thai.* 2002;85:361–8.

2. Tsai T-C, Lin C-H, Harasym PH, Violato C. Students' perception on medical professionalism: the psychometric perspective. *Med Teacher*. 2007;29:128–34.
3. Baker EH, Dowden JE, Cochran AR, Iannitti DA, Kimchi ET, Staveley-O'Carroll KF, Jeyarajah DR. Qualities and characteristics of successfully matched North American HPB surgery fellowship candidates. *HPB (Oxford)*. 2016;18:479–84.
4. LaRochelle J, Durning SJ, Gilliland W, Henry J, Ottolini M, Reamy B, Ritter J, Dorrance KA. Developing the next generation of physicians. *Mil Med*. 2018;183:225–32.
5. Seligman ME, Csikszentmihalyi M. Positive psychology. an introduction. *Am Psychol*. 2000;55:5–14.
6. Peterson C, Seligman MEP. Introduction to a "Manual of the Sanities" Character strengths and virtues a handbook and classification. New York: Oxford University Press; 2004. p. 3–32.
7. Bryan CS, Babelay AM. Building character: a model for reflective practice. *Acad Med*. 2009;84:1283–8.
8. Kotzee B, Ignatowicz A, Thomas H. Virtue in medical practice: an exploratory study. *HEC Forum*. 2017;29:1–19.
9. Cruess RL, Cruess SR. Professionalism and professional identity formation: the cognitive base. In: Cruess RL, Cruess SR, Steinert Y, editors. *Teaching medical professionalism: supporting the development of a professional identity*. 2nd ed. Cambridge: Cambridge University Press; 2016. p. 5–25.
10. Peterson C, Seligman ME. Character strengths and virtues: a handbook and classification. New York: Oxford University Press; 2004.
11. Hofmeyer A, Taylor R. Strategies and resources for nurse leaders to use to lead with empathy and prudence so they understand and address sources of anxiety among nurses practising in the era of COVID-19. *J Clin Nurs*. 2020. <https://doi.org/10.1111/jocn.15520>.
12. Oc B, Bashshur MR, Daniels MA, Greguras GJ, Diefendorff JM. Leader humility in Singapore. *Leadership Quar*. 2015;26:68–80.
13. Rego A, Owens B, Leal S, Melo AIMPE, Cunha, Gonçalves L, Ribeiro P. How leader humility helps teams to be humbler, psychologically stronger, and more effective: a moderated mediation model. *Leadership Quart*. 2017;28:639–58.
14. Sansone RA, Sansone LA. Gratitude and well being: the benefits of appreciation. *Psychiatry (Edgmont)*. 2010;7:18–22.
15. Pang D, Ruch W. The mutual support model of mindfulness and character strengths. *Mindfulness (N Y)*. 2019;10:1545–59.
16. Li J, Pearce PL, Oktadiana H. Can digital-free tourism build character strengths? *Ann Tour Res*. 2020;85:103037.
17. Wedding D, Niemiec RM. The clinical use of films in psychotherapy. *J Clin Psychol*. 2003;59:207–15.
18. Rufer MD, Jones L. Magic at the movies: positive psychology for children, adolescents and families. 2014.
19. Smithikrai C. Effectiveness of teaching with movies to promote positive characteristics and behaviors. *Procedia Soc Behav Sci*. 2016;217:522–30.
20. Niemiec RM, Wedding D. Positive psychology at the movies: Using films to build virtues and character strengths. Hogrefe Publishing; 2013.
21. Cavanagh RF, Kent DB, Romanoski JT. An illustrative example of the benefits of using a Rasch analysis in an experimental design investigation. Sydney: In The Assessment and Measurement Special Interest Group; 2005.
22. PHuSeG scale. A composite scale of prudence, humility, self-regulation and gratitude. <http://www.pakaranhome.com/index.php?lay=show&ac=article&id=2147599706>.
23. Linacre JM. Winsteps® Rasch measurement computer program User's Guide. Beaver-ton, Oregon: Winsteps.com; 2017.
24. Niemiec RM, Wedding D. Positive psychology at the movies: Using films to build virtues and character strengths. Gottingen: Hogrefe & Huber; 2008.
25. Wright BD. Rack and stack: time 1 vs. time 2 or pre-test vs. post-test. *Rasch Measurement Trans*. 2003;17:905–6.
26. Bond TG. Applying the Rasch model : fundamental measurement in the human sciences/authored by Trevor G. Bond and Christine M. Fox. New York: Routledge, Taylor & Francis Group; 2015.
27. Wright BD, Stone MH. Making measures. Chicago: Phaneron Press; 2004.
28. Linacre JM. Winsteps® Rasch measurement computer program user's guide. 2017.
29. Josephine MN, Fitzpatrick R, Jill D, Jenkinson C. Comparing alternative rasch-based methods vs raw scores in measuring change in health. *Med Care*. 2004;42:125–36.
30. Anselmi P, Vidotto G, Bettinardi O, Bertolotti G. Measurement of change in health status with Rasch models. *Health Quality Life Outcomes*. 2015;13:16.
31. Kahler E, Rogausch A, Brunner E, Himmel W. A parametric analysis of ordinal quality-of-life data can lead to erroneous results. *J Clin Epidemiol*. 2008;61:475–80.
32. Abbakumov D, Desmet P, Van den Noortgate W. Measuring student's proficiency in MOOCs: multiple attempts extensions for the Rasch model. *Heliyon*. 2018;4:e01003.
33. Axon DR, Augustine JM, Warholak T, Lee JK. Improving rating scales: applying Rasch analysis to student pharmacists' attitudes towards herbal medications. *Curr Pharm Teach Learn*. 2019;11:658–63.
34. Elder C, McNamara T, Congdon P. Rasch techniques for detecting bias in performance assessments: an example comparing the performance of native and non-native speakers on a test of academic English. *J Appl Meas*. 2003;4:181–97.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

