



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

A Mokken scale analysis of the peer physical examination questionnaire

This is the Published version of the following publication

Vaughan, Brett and Grace, S (2018) A Mokken scale analysis of the peer physical examination questionnaire. *Chiropractic and Manual Therapies*, 26 (6). ISSN 2045-709X

The publisher's official version can be found at
<https://chiromt.biomedcentral.com/articles/10.1186/s12998-018-0176-0>
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/36792/>

RESEARCH

Open Access



A Mokken scale analysis of the peer physical examination questionnaire

Brett Vaughan^{1,2,3*}  and Sandra Grace^{1,3}

Abstract

Background: Peer physical examination (PPE) is a teaching and learning strategy utilised in most health profession education programs. Perceptions of participating in PPE have been described in the literature, focusing on areas of the body students are willing, or unwilling, to examine. A small number of questionnaires exist to evaluate these perceptions, however none have described the measurement properties that may allow them to be used longitudinally. The present study undertook a Mokken scale analysis of the Peer Physical Examination Questionnaire (PPEQ) to evaluate its dimensionality and structure when used with Australian osteopathy students.

Methods: Students enrolled in Year 1 of the osteopathy programs at Victoria University (Melbourne, Australia) and Southern Cross University (Lismore, Australia) were invited to complete the PPEQ prior to their first practical skills examination class. R, an open-source statistics program, was used to generate the descriptive statistics and perform a Mokken scale analysis. Mokken scale analysis is a non-parametric item response theory approach that is used to cluster items measuring a latent construct.

Results: Initial analysis suggested the PPEQ did not form a single scale. Further analysis identified three subscales: 'comfort', 'concern', and 'professionalism and education'. The properties of each subscale suggested they were unidimensional with variable internal structures. The 'comfort' subscale was the strongest of the three identified. All subscales demonstrated acceptable reliability estimation statistics (McDonald's omega > 0.75) supporting the calculation of a sum score for each subscale.

Conclusion: The subscales identified are consistent with the literature. The 'comfort' subscale may be useful to longitudinally evaluate student perceptions of PPE. Further research is required to evaluate changes with PPE and the utility of the questionnaire with other health profession education programs.

Keywords: Evaluation, Item response theory, Osteopathy, Osteopathic medicine, Internal structure

Background

Peer physical examination (PPE) is a key pedagogical strategy in health professional education. Its benefits are well documented and include developing respect and sensitivity to patient's needs [1] and understanding the patient experience 'from the inside' as students undress, disclose personal information and act as models for examination by fellow students [2]. PPE also provides the opportunity to develop the level of clinical competence required for practising on real patients [3] and different body types [4].

A number of tools have been developed to assess attitudes to PPE. The most common approach is by survey of students. The 'Examining Fellow Students' questionnaire or a variation of it has been frequently used [2, 5–10]. The questionnaire was first developed in 1998 by O'Neill et al. [11] and further developed to a whole of body approach by Chang and Power [12] who also evaluated its face validity and pilot tested it with a convenience sample. Using this whole of body approach both sensitive regions (e.g. groin, rectal and external genitals) and non-sensitive regions (e.g. hands, feet, head) could be included. The survey was also adapted by Rees et al. [13] and used to ask medical students to indicate which of eleven body parts

* Correspondence: brett.vaughan@vu.edu.au

¹College of Health & Biomedicine, Victoria University, Melbourne, Australia

²Institute of Sport, Exercise and Active Living, Victoria University, Melbourne, Australia

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



they would not be willing to examine or have examined by a same or opposite gender peer.

Chang and Power [12] developed a 25-item questionnaire to explore students' perceptions of PPE addressing 3 a-priori domains: (1) comfort with PPE; (2) professionalism, appropriateness, and perceived value of PPE; and (3) willingness to perform peer breast, genital, and rectal examinations. Consorti et al. [8] developed the 'Peer Physical Examination Questionnaire' (PPEQ) and used it in conjunction with the 'Examining Fellow Students' questionnaire as a point of reference for evaluating criterion validity. Other surveys used to assess students' attitudes to PPE include one that explored nursing students' confidence in measuring blood pressure on their peers [14] and others that focused on the pedagogy of PPE, including student perceptions of dyad learning [15], partnering [16], and informed consent [17, 18]. Pols et al. [19] surveyed medical students about the incidence and consequences of medical problems discovered during PPE.

Other researchers surveyed clinical educators. In one study clinical educators were sent a list of 14 common invasive and non-invasive clinical procedures and asked if they allowed students to practise them on each other, and if so, what level of consent was used [20]. Another study sought medical clinical educators' responses to two questions: (1) did their students participate in PPE, and (2) did they have a PPE policy. Respondents were invited to send a copy of their policy to the researchers for further analysis.

Despite the number of tools for assessing attitudes to PPE and the frequency of their use reported in the literature, evidence for the validity of the scores derived from these questionnaires is limited. Wismeijer et al. [21] suggest that "given the uncertainty about the true dimensionality of the data, a second opinion provided by a conceptually different method may be very useful" (p. 324). The aim of this study was to investigate the dimensionality and structure of one of these evaluation tools – the Consorti et al. [8] 16-item PPEQ.

Methods

The present study forms part of a larger investigation into PPE perceptions in two Australian osteopathy programs: one at Victoria University (VU) (Melbourne, Australia), and the other at Southern Cross University (SCU) (Lismore, Australia). Human Research Ethics Committees at both institutions approved the study.

Participants

Students enrolled in Year 1 of the osteopathy programs at VU and SCU were invited to participate in the study as part of their first practical skills class in both the 2015 and 2016 academic years. Students were informed by email and notice on the respective university learning

management systems that the study was taking place. The Information to Participants sheets were made available on the learning management systems and in the practical skills classes.

Measure

Participants completed a demographic questionnaire, the 'Examining Fellow Students' questionnaire and the PPEQ. Responses were anonymous, however students self-generated a code in order to match their responses as part of the larger longitudinal study. The PPEQ is the focus of the current evaluation and responses to the 'Examining Fellow Students' will be reported elsewhere. The PPEQ utilises a 5-point Likert-type scale (0 – strongly disagree, 4 – strongly agree) to evaluate students' perceptions of PPE. Principal components analysis of the PPEQ identified three components: 'appropriateness and usefulness,' 'sexual implications' and 'passive role' accounting for 62.8% of the variance and a Cronbach's alpha of 0.86 [8].

Data analysis

Data were entered into SPSS (IBM Corp, USA) with items 3 to 7 and item 12 recoded as per Consorti et al. [8]. After rescoring, data were exported to *R* [22] for analysis using the *psych* [23] and *mokken* [24] packages. Descriptive data analysis was undertaken using the *psych* package.

Mokken scale analysis

The *mokken* package was used to perform a Mokken scale analysis (MSA) following the procedures and steps described by Stochl et al. [25] and Van Der Ark [24]. MSA is a non-parametric item response theory (IRT) approach that is used to evaluate dimensionality (number of concepts measured by the questionnaire) and the internal structure (relationship between the items) of a scale or questionnaire [26, 27]. IRT models attempt to fit the data to 'S' or sigmoid-shaped curves as this shape represents the expected responses to an individual questionnaire item. MSA is not as restrictive with respect to fit of data to sigmoid-shaped item response curves, particularly when compared to more restrictive IRT models such as Rasch analysis [25]. This approach (MSA) may lead to an increase in the number of useful items being retained that would otherwise be removed in a more restrictive IRT model. The creation of a Mokken scale is based on the assumptions of unidimensionality (the data measures a single latent construct), monotonicity (the item characteristic curve demonstrates higher values on the curve for higher values of the latent construct), and local independence (the responses to items measuring the latent construct are independent of the response to every other item on the scale) [25]. For ordinal data, the double monotonicity model (DMM) for Mokken scaling is used and adds the additional assumption of non-

intersecting item characteristic curves [25]. To meet the requirements of a Mokken scale for this model, the basic assumptions described previously must be met. When these requirements are met, the total score can be said to represent the latent construct or the concept theoretically being measured by the items in the questionnaire. Further, the DMM model may also allow for the identification of invariant item ordering, that is, the ordering of the items on the latent construct continuum does not change regardless of the level of the latent construct demonstrated by the respondent.

The following describes how the MSA was performed in the current study in *R*. The first step in MSA is to evaluate the items that may form Mokken scales using the automated item search function (*aisp*) [28]. All 16 PPEQ items were evaluated to identify potential Mokken scales using the initial cut off of 0.3, then increasing incrementally in 0.05 steps until the scales could no longer be logically explained and the analysis resulted in two or more scales [28], as would be consistent with the PPE literature. Once a scale was identified, the scalability coefficients for all items creating a scale (H), the individual items in that scale (H_i), and the item pairs (H_{ij}) were calculated along with the standard error [29]. A scale is said to be ‘weak’ if H is less than 0.3, ‘moderate’ if H is between 0.4–0.5, and ‘strong’ if H is greater than 0.5. Items are thought to be suitable for inclusion in a Mokken scale if they demonstrate a H_i value greater than 0.3, and the H_{ij} value is greater than 0. Next, local dependence was checked using the conditional association procedure [30]. Where an item pair was identified as locally dependent, the item with the lower H_i value was removed and the data set reanalysed. Monotonicity was then checked using both graphical and numerical approaches to ensure that each item demonstrated a monotonically increasing item response function. Invariant item ordering (H^T) was then evaluated with values less than 0.3 suggesting the items could not be meaningfully ordered, with items of between 0.4–0.5 demonstrating moderate ordering, and with items with H^T greater than 0.5 demonstrating strong item ordering. Once a scale had been finalised, Mokken’s *rho* was evaluated as one of the reliability estimations with a value over 0.7 being acceptable [31].

Reliability estimation

In addition to Mokken’s *rho*, McDonald’s omega (ω) [32–34] was the selected reliability estimation method and calculated using the *psych* package. Green and Yang [35] have suggested that the use of Cronbach’s alpha is limited given the propensity for data in the educational and psychological sciences to violate the assumptions underlying its proper use. The *psych* package [23] in *R* [22] presents ω as hierarchical ω_h and total ω_t . ω_t

provides an indication as to the reliability of the general factor, and ω_h is the proportion of the total score variance that can only be attributed to the general factor [36]. Zinbarg et al. [33] suggest that ω_h is the most accurate reliability estimate in most situations. High ω_h values suggest that the general factor accounts for the total score variance, and supports unidimensionality [37]. ω_h values greater than 0.7 support calculation of the total score [36]. Both McDonald’s omega total (ω_t) and omega hierarchical (ω_h) were calculated: ω_t is the reliability of the total score; and, ω_h is the proportion of variance due to the latent construct and provides an indication about the extent to which the total questionnaire score estimates the latent construct [32] – perception of PPE.

Results

Three hundred and fourteen students ($N = 314$) completed the PPEQ at the start of the first teaching period in 2015 ($n = 153$) and 2016 ($n = 161$) representing an overall 90% response rate. Two hundred and thirty responses were from VU ($n = 230$, 73.2%). Descriptive statistics for the PPEQ after rescaling are presented in Table 1.

Mokken scale analysis

Prior to the analysis responses from eight students were removed due to missing data. Scalability of the full 16-item PPEQ is presented in Table 1. Individual item scalability (H_i) for items 5, 6, 7, and 12 were below the accepted cut-off of 0.30 [25], with item 13 falling below this cut-off if the standard error is taken into account. The H coefficient of 0.40 (0.03) suggests a ‘weak’ scale and is likely multidimensional.

The *aisp* function in the *mokken* package [24] was used to identify potential Mokken or unidimensional scales. As per Stochl et al. [25] the initial lower bound was set at 0.30 (*minvi* size) and subsequent analyses undertaken in increasing 0.05 steps (Table 2).

The lower bound values suggested a three-scale structure was most appropriate and that item 12 should be removed. The three subscales were identified as ‘comfort’, ‘concern’, and ‘professionalism and education’. Each scale was evaluated for unidimensionality, monotonicity, and invariant item ordering (IIO) as per Van der Ark [24] and Stochl et al. [25].

Comfort subscale

The ‘comfort’ subscale consisted of PPEQ items 1–4 and 8–11. The H -coefficient was 0.61 (± 0.03) and H_i coefficients were greater than 0.50 and all H_{ij} coefficients were non-negative. One non-significant monotonicity violation was identified for each of items 3

Table 1 Descriptive statistics and initial scalability (H_i and standard error) for the Peer Physical Examination Questionnaire (PPEQ) items

PPEQ item	Mean (SD)	Median	H_i (standard error)
1. In general, I feel comfortable when performing PPE on a colleague of mine	3.24 (0.72)	3	0.47 (0.03)
2. In general, I feel comfortable when a colleague performs PPE on me	3.18 (0.75)	3	0.47 (0.03)
3. I feel embarrassed if I am undressed for PPE in front of my group of colleagues ^a	2.47 (1.11)	2	0.44 (0.03)
4. I feel embarrassed if I am undressed for PPE in front of my teacher or tutor ^a	2.60 (1.10)	3	0.42 (0.03)
5. I am concerned of being a possible object of sexual interest during PPE ^a	3.33 (0.82)	4	0.26 (0.03)
6. I am concerned of experiencing possible sexual interest for my colleagues during PPE ^a	3.39 (0.83)	4	0.27 (0.04)
7. I am concerned of experiencing possible sexual interest for my teacher or tutor during PPE ^a	3.61 (0.72)	4	0.27 (0.05)
8. I feel comfortable when performing PPE on a colleague of my same sex	3.44 (0.67)	4	0.45 (0.04)
9. I feel comfortable when performing PPE on a colleague of the opposite sex than mine	3.21 (0.80)	3	0.48 (0.04)
10. I feel comfortable when PPE is performed on me by a colleague of my same sex	3.44 (0.63)	4	0.47 (0.03)
11. I feel comfortable when PPE is performed on me by a colleague of the opposite sex than mine	3.22 (0.76)	3	0.50 (0.03)
12. It is inappropriate to perform PPE on persons that will be my future colleagues ^a	3.37 (0.93)	4	0.24 (0.05)
13. To perform PPE is an appropriate practice for the education of a medical doctor (osteopath)	3.69 (0.67)	4	0.34 (0.06)
14. To undergo PPE is an appropriate practice for the education of a medical doctor (osteopath)	3.73 (0.56)	4	0.40 (0.04)
15. In performing PPE I get useful feedback from my colleagues about my skill	3.70 (0.50)	4	0.44 (0.04)
16. It is a sign of professionalism as a student to accept to perform and undergo PPE	3.57 (0.66)	4	0.35 (0.04)

^aitems were rescored as per Consorti et al. [8]

and 4. Non-significant IIO was identified for items 1, 2, 9 and 11 however backward selection did not suggest removing any item. H^T was 0.61 indicating high accuracy of the item ordering [38]. These results provide evidence for the unidimensionality of the 'comfort' subscale that meets the requirements of a Mokken scale.

Concern subscale

The 'concern' subscale consisted of PPEQ items 5 to 7. The H -coefficient was 0.71 (± 0.05), all H_i coefficients were greater than 0.60 and all H_{ij} coefficients were non-negative. No violations of monotonicity were identified. Non-significant IIO was identified for items 5 and 6, and backward selection did not suggest removing either

Table 2 Identification of Mokken scales for the Peer Physical Examination Questionnaire (PPEQ) using the *aisp* function

PPEQ item	<i>minvi</i> size				
	0.30	0.35	0.40	0.45	0.5
1. In general, I feel comfortable when performing PPE on a colleague of mine	1	1	1	1	1
2. In general, I feel comfortable when a colleague performs PPE on me	1	1	1	1	1
3. I feel embarrassed if I am undressed for PPE in front of my group of colleagues	1	1	1	1	1
4. I feel embarrassed if I am undressed for PPE in front of my teacher or tutor	1	1	1	1	1
5. I am concerned of being a possible object of sexual interest during PPE	1	2	2	2	2
6. I am concerned of experiencing possible sexual interest for my colleagues during PPE	2	2	2	2	2
7. I am concerned of experiencing possible sexual interest for my teacher or tutor during PPE	2	2	2	2	2
8. I feel comfortable when performing PPE on a colleague of my same sex	1	1	1	1	1
9. I feel comfortable when performing PPE on a colleague of the opposite sex than mine	1	1	1	1	1
10. I feel comfortable when PPE is performed on me by a colleague of my same sex	1	1	1	1	1
11. I feel comfortable when PPE is performed on me by a colleague of the opposite sex than mine	1	1	1	1	1
12. It is inappropriate to perform PPE on persons that will be my future colleagues	x	x	x	x	x
13. To perform PPE is an appropriate practice for the education of a medical doctor (osteopath)	1	1	3	3	3
14. To undergo PPE is an appropriate practice for the education of a medical doctor (osteopath)	1	1	1	3	3
15. In performing PPE I get useful feedback from my colleagues about my skill	1	1	1	3	3
16. It is a sign of professionalism as a student to accept to perform and undergo PPE	1	1	3	3	3

1 = scale one, 2 = scale two, 3 = scale three

item. H^T was 0.35 indicating low accuracy of the item ordering [38]. These results suggest the ‘concern’ subscale is unidimensional and meets the requirements of a Mokken scale, although the ability to discern between levels of student concern with participating in PPE is limited.

Professionalism and education subscale

PPEQ items 13 to 16 comprise the ‘professionalism and education’ subscale. The H -coefficient was 0.71 (0.05), all H_i coefficients were greater than 0.60 and all H_{ij} coefficients were non-negative. No violations of monotonicity or IIO were identified. H^T was 0.17 indicating limited accuracy of the item ordering. These results suggest the ‘professionalism and education’ subscale is unidimensional and meets the requirements of a Mokken scale, although the ability to discern between different levels of perceived ‘professionalism and education’ associated with PPE is negligible.

Reliability estimation

McDonald’s omega was calculated as the reliability estimate and the results are presented in Table 3. All values for both ω_t and ω_h where above an acceptable level. ω_h values obtained in the present study provide support for the calculation of a total score for each subscale [36] given that over three-quarters of the variance in the summed score for each subscale is attributable to the latent constructs of ‘comfort’, ‘concern’, and ‘professionalism and education’ respectively.

Discussion

The present study sought to evaluate the dimensionality and structure of the PPEQ developed by Consorti et al. [8]. Research by the current authors has previously identified the need to explore the properties of the questionnaire to provide evidence for its ongoing use as a PPE evaluation tool [39]. Consorti et al. [8] calculated a total score for the PPEQ and used this score as part of their analyses in Italian medical and osteopathy students. However, Mokken scale analysis of the full 16-item PPEQ suggested that the calculation of a total score for the PPEQ is not valid in an Australian osteopathy student population.

Consorti et al. [8] used a Principal Components Analysis (PCA) to evaluate the internal structure of the PPEQ. Wismeijer et al. [21] suggest the use of a number

of analytic approaches, including Mokken scale analysis, as a complementary data analysis approach to evaluate dimensionality and different levels of an underlying latent construct given this information is not obtainable from a PCA. Consorti et al. [8] identified a 3-component structure for the PPE in their study (‘appropriateness and usefulness’, ‘sexual implications’ and ‘passive role’) however the same structure was not identified in the present study. Through the Mokken scale analysis, three Mokken scales were identified in the present study: ‘comfort’, ‘concern’, and ‘professionalism and education’. The present study provides evidence for an alternative psychometric structure for the PPEQ that comprises three unidimensional subscales where the summed score for each subscale represents the respective latent construct. This alternative structure has similarities with the *a-priori* domains of ‘comfort with PPE’, and ‘professionalism, appropriateness, and perceived value’ of PPE as described by Chang and Power [12]. Given the similarities, it may be that these are two key themes in the evaluation of PPE.

To create the three PPEQ subscales, it is suggested that item 12 be removed as it did not fit into any of the three Mokken scales identified in the analysis. This item, ‘It is inappropriate to perform PPE on persons that will be my future colleagues’, does not appear to measure a construct consistent with the other PPEQ items. The use of the term ‘inappropriate’ may account for this as it is the only item with this term. More frequently used terms in the PPEQ are ‘comfortable’, ‘concerned’ and ‘embarrassed’. Further, students in the present study may ascribe a different meaning to the term ‘inappropriate’. Students in the present study completed the PPEQ on their first day before they had participated in a practical skills class. Their frame of reference for what constitutes ‘inappropriate’ is likely to be different from other students. Another possibility is the translation of the term to ‘inappropriate’ from the initial validation study with Italian osteopathy and medical students [8]. While the item may have had a particular meaning in the initial study, its meaning within the context of the item has been lost.

‘Comfort’ subscale

‘Comfort’ with PPE needs to be evaluated as students are often expected to participate in such activities during their pre-professional program [7, 40]. This subscale consisted of items 1–4 and 8–11. Items 1 to 4 gauge the students’ perception with regard to performance of PPE and exposure of their body. Students appear to be generally comfortable with participating in PPE [12, 39]. Items 8 to 11 specifically address comfort with PPE based on sex. The literature suggests that gender has a significant influence on student perceptions of PPE. Discomfort

Table 3 Reliability estimates for the three Peer Physical Examination Questionnaire (PPEQ) subscales using McDonald’s omega (ω)

Scale	Omega total (ω_t)	Omega hierarchal (ω_h)
Comfort	0.95	0.85
Concern	0.90	0.84
Professionalism & education	0.89	0.75

with examining students of a different gender has been identified by females due to fear of sexual exploitation, but also by males for fear of accusations of harassment [40]. Students' perceptions of these issues are likely to be captured in this PPEQ subscale. Given the increasing awareness of gender diversity it may be that these concerns are not limited to different gender interactions. This has yet to be considered in the literature and provides an avenue for further research. The 'comfort' subscale is the strongest of the three PPEQ subscales from a scalability and item ordering standpoint. This suggests that the *comfort* subscale could potentially measure changes in a students' perception of comfort with PPE over a period of time.

'Concern' subscale

Much of the literature on PPE relates to *concern* about participating in such activities. The focus of the PPEQ items in this subscale relates specifically to 'sexual interest' from not only other students but also the academic and clinical teaching staff. Female students have been reported to be more likely than males to fear critical and teasing comments, and sexual objectification [41]. As highlighted in the discussion of the 'comfort' subscale, it may be that this unease extends beyond the reported female/male interaction, however such an assertion has not been described in the literature. Concern has also been expressed about the "immaturity" of fellow students and about potential sexual harassment [42]. Wearn et al. [2] also suggest that issues may arise where students may be (or have previously been) close friends, housemates, or a sexual partner of their peer examiner/examinee, and therefore have blurred boundaries within the context of PPE. From a psychometric standpoint, the item ordering value (H^T) was low suggesting that the ability to discern between different levels of perceived concern associated with PPE is negligible albeit a total score can be calculated for this subscale.

'Professionalism and education' subscale

'Professionalism and education' are key components in the evaluation of PPE, a position supported by the inclusion of this domain in the study by Chang & Power [12]. Setting professional behaviour standards [12], undertaking a formal PPE participation consent process [18] and creating a positive education environment may contribute to a positive perception of PPE [39]. This subscale also captures the students' perceptions about the need to participate in PPE, a theme consistent with other work [6]. As with the 'concern' subscale, the item ordering value was low suggesting that the ability to discern between different levels of perceived professionalism and education associated with PPE is negligible. This assertion is potentially supported by Vaughan & Grace [39]

who reported no difference in perception of first year osteopathy students in the 4 items of the PPEQ making up this scale over a 12-week period. Evidence is provided for the calculation of a sum score for the subscale.

Reliability estimations

All three PPEQ subscales demonstrated high scalability coefficients suggesting they are unidimensional and the present study provides support for their measurement of the underlying latent constructs, namely, 'comfort', 'concern', and 'professionalism and education' associated with PPE. For all three subscales the ω_t values were approximately 0.90 suggesting that 90% of the variance in the total score for each subscale is accounted for by the respective latent construct. Calculation of a total score for each subscale (after rescaling where required) is supported by the high ω_h values [36]. This subscale score calculation is at odds with Consorti et al. [8] who calculated a total score for the PPEQ and such a position is not supported by the data in the current study.

Study limitations

There are a number of limitations in the present study. First, the study only explored the opinions of students in two Australian osteopathy programs. Therefore the generalisability of these results to other osteopathy programs, and other health professions is limited. Another limitation is that the study did not evaluate whether the structure of the questionnaire was different with different cultural groups. Acceptance and potentially participation in PPE is known to vary with different cultures [40] and this could result in a different questionnaire structure. It is also possible that a degree of bias was introduced with approximately three-quarters of the data in the present study obtained from one institution.

Future research and questionnaire use

The PPEQ as described in the current study has a number of uses both in the classroom and research settings. The questionnaire has the potential to be used as an evaluation of the learning environment to identify systemic concerns with participating in PPE beyond the individual student level. Systemic concerns could be addressed using Grace et al.'s [43] strategies for improving the experience of PPE, including the use of written consent forms and formal feedback. The PPEQ can also be used as part of a formal feedback strategy that could be used to inform educators about changes that can be made to improve or modify the PPE experience. The PPEQ also has the potential to be used in longitudinal evaluations of PPE experiences to evaluate changes in perception over time, particularly where students are participating in PPE on a regular basis. Grace et al. [43]

suggest that "...providing students with written information about what to expect, and about the pedagogical benefits of experiential learning [PPE], and discussing ethical issues that could be associated with experiential learning could be readily implemented in practical classes" (p. 29). The PPEQ could be used to evaluate perceptions about PPE before and after the provision of this information.

The influence of gender diversity and sexuality as factors influencing PPE participation is an avenue for further research given the literature has thus far only considered female/male interactions. Further work could also explore whether the PPEQ measurement properties are retained in different cultural and gender diverse populations.

Conclusion

This study provides evidence for the dimensionality and structure of a 15-item version of the 'Peer Physical Examination Questionnaire'. The current research has identified three subscales within the PPEQ ('comfort', 'concern', and 'professionalism and education') that can be used to explore students' perceptions of PPE. Calculation of a total score for the modified 15-item scale is not supported, however it is possible to calculate a sum score for each of the three subscales. These subscales have the potential to provide an avenue for further research, including longitudinal changes in perception, particularly as the subscale themes are consistent with the literature. The current research has strengthened the psychometric properties of the PPEQ and others are encouraged to explore the use of the modified questionnaire in their student cohorts.

Acknowledgements

Not applicable

Funding

No funding was received to conduct this study.

Availability of data and materials

The datasets generated during and/or analysed during the current study are available in the figshare repository, <https://doi.org/10.6084/m9.figshare.3556413>.

Authors' contributions

BV and SG designed the study. BV undertook the data analysis. BV and SG undertook the literature review and developed the manuscript. BV and SG read and approved the final manuscript.

Authors' information

Brett Vaughan is a lecturer in the College of Health & Biomedicine, Victoria University, Melbourne, Australia and a Professional Fellow in the School of Health & Human Sciences at Southern Cross University, Lismore, New South Wales, Australia. His interests centre on assessment and evaluation in allied health professions education and clinical education. Sandra Grace is Director of Research at the School of Health and Human Sciences, Southern Cross University, Adjunct Associate Professor at the Education for Practice Institute, Charles Sturt University and Visiting Associate Professor at the College of Health & Biomedicine, Victoria University. She has extensive experience in private practice as a chiropractor and osteopath, and as a lecturer and curriculum designer. Her research interests are in health services research and interprofessional practice and education.

Ethics approval and consent to participate

Ethics approval was granted by both the Victoria University and Southern Cross University Human Research Ethics Committee's (ECN16-031). Consent to participate in the study was implied by the return of the questionnaire. No identifying details were collected from the students and students could complete the questionnaire at a time and location of their choosing.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹College of Health & Biomedicine, Victoria University, Melbourne, Australia.

²Institute of Sport, Exercise and Active Living, Victoria University, Melbourne, Australia.

³School of Health & Human Sciences, Southern Cross University, Lismore, Australia.

Received: 11 September 2017 Accepted: 19 January 2018

Published online: 01 March 2018

References

- Braunack-Mayer A. Should medical students act as surrogate patients for each other? *Med Educ*. 2001;35:681–6.
- Wearn AM, Rees CE, Bradley P, Vnuk AK. Understanding student concerns about peer physical examination using an activity theory framework. *Med Educ*. 2008;42:1218–26.
- Bindless L. The use of patients in health care education: the need for ethical justification. *J Med Ethics*. 1998;24:314–9.
- Koehler N, McMenamin C. The need for a peer physical examination policy within Australian medical schools. *Med Teach*. 2014;36:430–3.
- Rees CE, Bradley P, McLachlan JC. Exploring medical students' attitudes towards peer physical examination. *Med Teach*. 2004;26:86–8.
- Chen J, Yip A, Lam C, Patil N. Does medical student willingness to practise peer physical examination translate into action? *Med Teach*. 2011;33:e528–40.
- Rees CE, Wearn A, Vnuk A, Sato T. Medical students' attitudes towards peer physical examination: findings from an international cross-sectional and longitudinal study. *Adv Health Sci Educ*. 2009;4:103–21.
- Consorti F, Mancuso R, Piccolo A, Consorti G, Zurlo J. Evaluation of the acceptability of peer physical examination (PPE) in medical and osteopathic students: a cross sectional survey. *BMC Med Educ*. 2013;13:111.
- Wearn A, Rees CE, Bhoopalkar H, Bradley P, Lam C, McLachlan J, Patil N, Sato T, Vnuk A. 'What not to Touch': medical students from six schools report on peer physical examination in clinical skills and anatomy learning. *Focus Health Prof Educ*. 2008;10:24–5.
- Wearn A, Bhoopalkar H, Mathew T, Stewart L. Exploration of the attitudes of nursing students to peer physical examination and physical examination of patients. *Nurse Educ Today*. 2013;33:884–8.
- O'Neill P, Larcombe C, Duffy K, Dorman T. Medical students' willingness and reactions to learning basic skills through examining fellow students. *Med Teach*. 1998;20:433–437.
- Chang E, Power D. Are medical students comfortable with practicing physical examinations on each other? *Acad Med*. 2000;75:384–9.
- Rees CE, Bradley P, Collett T, McLachlan J. "over my dead body?": the influence of demographics on students' willingness to participate in peer physical examination. *Med Teach*. 2005;27:599–605.
- Baillie L, Curzio J. A survey of first year student nurses' experiences of learning blood pressure measurement. *Nurse Educ Pract*. 2009;9:61–71.
- Tolsgaard MG, Rasmussen MB, Bjorck S, Gustafsson A, Ringsted CV. Medical students' perception of dyad practice. *Perspectives in Medical Education*. 2014;3:500–7.
- Barnette J, Kreitter C, Schuldt S. Student attitudes towards same-gender versus mixed-gender partnering in practicing physical examination skills. *Eval Health Prof*. 2000;23:360–70.
- Redford D, Klein T. Informed consent in the nursing skills laboratory: an exploratory study. *Nursing Education*. 2003;42:131–3.

18. Wearn A, Bhoopatkar H. Evaluation of consent for peer physical examination: students reflect on their clinical skills learning experience. *Med Educ.* 2006;40:957–64.
19. Pols J, Boendermaker P, Muntinghe H. Incidence of and sequels to medical problems discovered in medical students during study-related activities. *Med Educ.* 2003;37:889–94.
20. Hilton P, Barrett D. An investigation into students' performance of invasive and non-invasive procedures on each other in classroom settings. *Nurse Educ Pract* 2009;9:45-52.
21. Wismeijer AA, Sijtsma K, van Assen MA, Vingerhoets AJ. A comparative study of the dimensionality of the self-concealment scale using principal components analysis and Mokken scale analysis. *J Pers Assess.* 2008;90:323–34.
22. R Core Team. R: A language and environment for statistical computing. www.R-project.org/. Accessed June 20 2016.
23. Revelle W. psych; Procedures for Personality and Psychological Research. <http://CRAN.R-project.org/package=psych>. Accessed June 20 2016.
24. New V d ALA. Developments in Mokken scale analysis in R. *J Stat Softw.* 2012;48:1–27.
25. Stochl J, Jones PB, Croudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Med Res Methodol.* 2012;12:1–16.
26. Mokken RJ. A theory and procedure of scale analysis: with applications in political research, vol. 1: Walter de Gruyter; 1971.
27. Sijtsma K, Meijer RR, van der Ark LA. Mokken scale analysis as time goes by: an update for scaling practitioners. *Pers Individ Dif.* 2011;50:31–7.
28. Sijtsma K, Van der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *Br J Math Stat Psychol.* 2017;70:137–58.
29. Kuijpers RE, Van der Ark LA, Croon MA. Standard errors and confidence intervals for scalability coefficients in Mokken scale analysis using marginal models. *Sociol Methodol.* 2013;43:42–69.
30. Straat JH, van der Ark LA, Sijtsma K. Using conditional association to identify locally independent item sets. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences.* 2016;12:117.
31. Sijtsma K, Molenaar IW. Reliability of test scores in nonparametric item response theory. *Psychometrika.* 1987;52:79–97.
32. Revelle W, Zinbarg RE. Coefficients alpha, beta, omega, and the glb: comments on Sijtsma. *Psychometrika.* 2009;74:145–54.
33. Zinbarg RE, Revelle W, Yovel I, Li W. Cronbach's α , Revelle's β , and McDonald's ω H: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika.* 2005;70:123–33.
34. Zinbarg RE, Yovel I, Revelle W, RP MD. Estimating generalizability to a latent variable common to all of a scale's indicators: a comparison of estimators for ω h. *Appl Psychol Meas.* 2006;30:121–44.
35. Green SB, Yang Y. Commentary on coefficient alpha: a cautionary tale. *Psychometrika.* 2009;74:121–35.
36. Hermesen LA, Leone SS, Smalbrugge M, Knol DL, van der Horst HE, Dekker J. Exploring the aggregation of four functional measures in a population of older adults with joint pain and comorbidity. *BMC Geriatr.* 2013;13:119.
37. Reise SP. The rediscovery of bifactor measurement models. *Multivariate Behav Res.* 2012;47:667–96.
38. Ligtoet R, Van der Ark LA, te Marvelde JM, Sijtsma K. Investigating an invariant item ordering for polytomously scored items. *Educ Psychol Meas.* 2010;70:575–98.
39. Vaughan B, Grace S. Perception of peer physical examination in two Australian osteopathy programs. *Chiropr Man Therap.* 2016;24:21.
40. Rees CE, Wearn AM, Vnuk AK, Bradley PA. Don't want to show fellow students my naughty bits: medical students' anxieties about peer examination of intimate body regions at six schools across UK, Australasia and far-East Asia. *Med Teach.* 2009;31:921–7.
41. Rees CE. The influence of gender on student willingness to engage in peer physical examination: the practical implications of feminist theory of body image. *Med Educ.* 2007;41:801–7.
42. Outram S, Nair BR. Peer physical examination: time to revisit. *Med J Aust.* 2008;189:274–6.
43. Grace S, Innes E, Patton N, Stockhausen L. Ethical experiential learning in medical, nursing and allied health education: a narrative review. *Nurse Educ Today.* 2017;51:23–33.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

