Online publication ahead of print, AJMHS Vol 70, 2026

Validation of Subscales from the Peer Physical Examination Questionnaire and the Examining Fellow Students Questionnaire: Evidence of Scalability and Convergent Validity

Piersante Sestini[1*], Paolo Cameli[1], Behar Çekorja[1], Silvana Bala[2], Lutfi Alia[3], Elena Bargagli[1]

[1] Respiratory Diseases Unit, Department of Medicine, Surgery and Neurosciences, University of Siena, Siena, Italy
[2] Department of Pneumology, Faculty of Medicine, University of Medicine, Tirana, Albania
[3] Centre of Oncological Prevention (Centre of excellence), Siena, Italy

Abstract

Background: Peer physical examination (PPE) is a cost-effective preclinical method for learning physical examination skills, but its implementation requires careful evaluation of students' acceptability.

Methods: We analyzed the responses of 2,085 undergraduate students of medicine and of BS programs in different health professions (mostly nursing students) to two questionnaires widely used to assess PPE acceptability: the Peer Physical Examination Questionnaire (PPEQ) and the Examining Fellow Students (EFS) questionnaire. Subscales were derived using Mokken scale analysis and Principal Component Analysis, and concurrent validity was examined.

Results: The PPEQ yielded three subscales - *Emotional* (PPEQ-E, H = 0.50), *Attitudinal* (PPEQ-A, H = 0.68), and *Concerns about Sexual Connotation* (PPEQ-SC, H = 0.52). The EFS could be partitioned into Sensitive Areas (EFS-SA, H = 0.68) and *Less Sensitive Areas* (EFS-LSA, H = 0.59). All subscales showed good scalability and internal consistency, and were strongly intercorrelated. Each PPEQ subscale was independently associated with the EFS score after adjustment for sex and course type ($p < 0.001$). The subscales were able to detect differences in response patterns across sex and course type: female students reported lower emotional and attitudinal scores and higher sexual concerns on the PPEQ, and lower willingness to be examined in sensitive areas on

Address for correspondence: Piersante Sestini*, Department of Medicine, Surgery and Neurosciences, University of Siena, Viale Bracci 16, 53100, Siena, Italy. E-mail: Piersante.sestini@unisi.it

the EFS ($p < 0.001$). Students of BS programs showed lower PPEQ scores than medical students, but also greater reluctance toward sensitive areas, while medical students reported fewer sexual concerns.

Conclusions: The PPEQ and EFS subscales provide complementary perspectives on students' attitudes toward PPE. Their combined use enables a nuanced assessment of acceptability and its determinants across different student populations. These tools may support monitoring of trends over time and contribute to the design of educational activities that maximize learning benefits while minimizing discomfort.

Keywords: Peer Physical Examination; Medical Education, undergraduate; Nursing education, undergraduate; Questionnaires; Acceptability

INTRODUCTION

Peer Physical Examination (PPE) is a simulation-based teaching technique grounded in role play, in that students practice physical examination procedures on one another before engaging with real or standardized patients. PPE was formally introduced in the Skills Lab at Maastricht University in the 1970s (1), and first appeared in the medical literature in 1982 as a method for learning gross anatomy (2). Following a pivotal study by O'Neill in 1998 (3), numerous quantitative and qualitative studies flourished, addressing the many facets of its acceptability in teaching the physical examination (4,5). Its popularity as a tool for learning clinical examination skills grew rapidly, to the point that a 2018 survey of US medical schools reported that about 30% of physical examination teaching time involved PPE (6).

From the outset, however, it became clear that while the practice was generally acceptable to most medical students when more intimate body areas were excluded, a minority of students found it less acceptable, often linked to sexual, cultural or religious factors (3,7,8). Therefore, acceptability should be evaluated in each student population, so that policies and procedures can be adapted accordingly (4,9).

The Examining Fellow Students (EFS (3)) and the Peer Physical Examination Questionnaire (PPEQ (10,11)) are the best-documented instruments to assess PPE acceptability. While the EFS primarily focuses on behavioral intention, the PPEQ provides detailed information on positive and negative feelings, beliefs, and students' attitudes toward the technique, that may be useful in organizing activities to maximize benefits and minimize discomfort. In this context, "acceptability" can be conceptualized as a multidimensional construct, potentially encompassing emotional comfort, moral or professional appropriateness, and behavioral intention. A study in chiropractic students suggested that the questionnaire could be segmented into three subscales addressing comfort, concerns, and "professional and educational" aspects (11).

Focusing on chest examination, our aim was to reproduce these findings in a broader, more heterogeneous population of medical and other health professional students, and to validate them by examining convergent validity with the EFS questionnaire.

METHODS

Study design

The Peer-to-Peer Physical Examination of the Chest module was regularly performed at the University of Siena from 2015 to 2022, when it was integrated into the core curriculum. The module was designed as a multiprofessional Observable Practice Activity (12) with the specific objective of enabling students to perform chest physical examination procedures in a technically accurate and relationally sensitive manner. It was therefore a pre-clinical practical activity in preparation for clinical clerkships. The module was open to all students of the Medicine

graduating course and of any bachelor's degree program in the health professions.

Participation in the session required completion of both the EFS and PPEQ questionnaires, administered online through the Moodle platform of the course.

The EFS questionnaire (3) consists of yes/no questions on the willingness to examine, or to be examined on, nine different body regions (head and neck, arms, chest, breasts, abdomen, dorsolumbar region, inguinal region, genital region, legs) by a male or female peer student, for a total of 36 items. The EFS score was calculated as the sum of positive answers, and linearly rescaled to a 0–100 range.

The PPEQ (10,11) includes 18 items rated on a 5-point Likert scale (0 = nothing, 1 = little, 2 = fair, 3 = much, 4 = very much) assessing agreement with statements on various aspects, ranging from feelings about visiting or being visited by peers, participation with students and teachers of different sex, to attitudes and beliefs about PPE. The PPEQ score was computed as the sum of all items (after reversing negatively worded ones), with a maximum of 72, and linearly rescaled to 0–100 for comparability.

All questionnaires stored on the university server were retrieved.

Statistical analysis

Neither EFS nor PPEQ scores were normally distributed and were analyzed using non-parametric methods. Correlation was assessed using Spearman's rank test. Differences by sex and course were analyzed using rank-based ANOVA on percentile ranks (100–0). For the analysis of changes in scores between preliminary and post-activity follow-up questionnaires, score differences were ranked according to their absolute value, and the original sign was then restored to preserve direction. Estimates were obtained using ordinary least squares with robust standard errors. Scores of the retained principal components were analyzed using generalized linear models assuming a normal distribution. When included in rank-based models, they were added as continuous covariates (rank-based ANCOVA (13)).

Mokken scale analysis (MSA) was performed separately on all 18 PPEQ items (reversing their value when needed) and on all 36 items of the EFS questionnaire, with an initial cutoff value for the scalability coefficient of 0.3 (14). Principal component analysis (PCA) was performed on the same items, using the tetrachoric correlation matrix for the dichotomous EFS items. Only components with an eigenvalue >1, identified on the scree plot, were retained.

Unsupervised cluster analysis was conducted using the k-means partition method on all the retained components from PCA. The optimal number of clusters (k) was determined by comparing solutions with different k values, utilizing scree plots of within-cluster sum-of-squares (WSS), its logarithm (log WSS), the η^2 coefficient, and the proportional reduction of error (PRE) (15).

Data are presented as percentages, means, or regression coefficients with 95% confidence

intervals (CIs). Two-tailed p values <0.05 were considered statistically significant; p values ≥0.05 are reported as not significant (NS). All analyses were performed using Stata/SE version 19 (StataCorp, College Station, TX, USA).

RESULTS

A total of 2085 baseline questionnaires were available, 80% from medical students and 67% from females: 52% from female medical students, 28% from male medical students, 16% from female health-professions students, and 4% from male health-professions students. Among health-professions students, 83% were nursing students, 11% perfusionists, 4% dieticians, and 2% speech therapists. The mean number of questionnaires per year was 261 (range, 95–570).

EFS and PPEQ scores were strongly correlated (Spearman ρ=0.52, p<0.0001). Both scores by undergraduate course and sex are shown in Figure 1. EFS scores were lower in health-professions students (BS) than in medical students (MD) and in females than in males in both groups. Overall, scores were good, with at least 75% of students of each group in the upper half of the scale.

The percentage of students willing to practice chest examination and to accept examination by a peer of the same or opposite sex is shown in Figure 2. Percentages were higher in MD students and in males. Only 4% of students reported unwillingness to visit a peer.

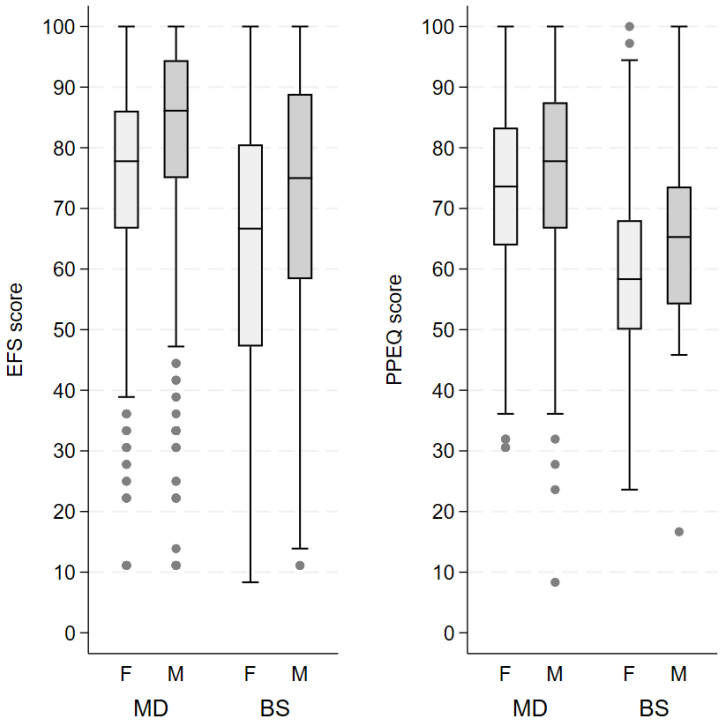

Figure 1. Box plots of EFS and PPEQ scores according to sex (F : M) and course type (MD or BS)

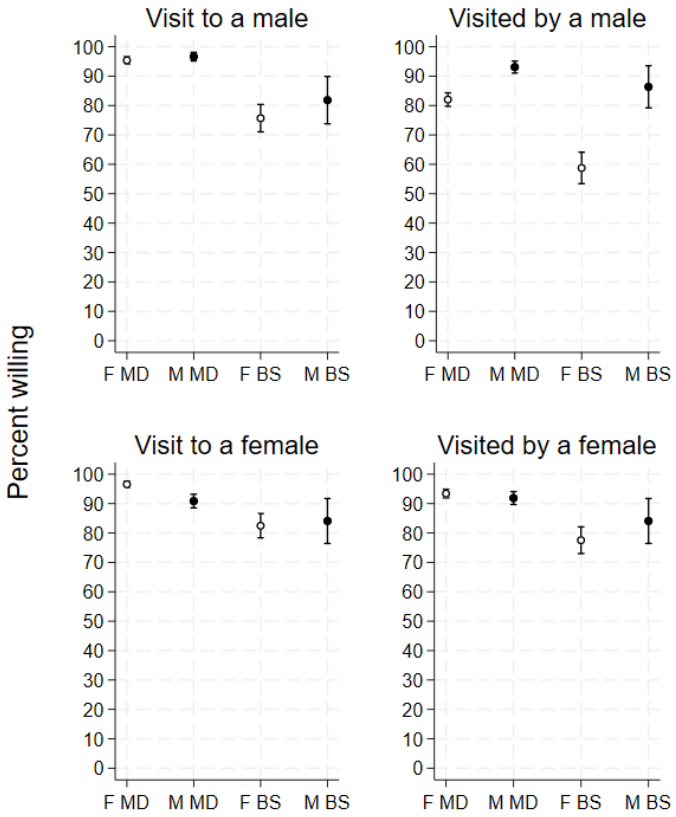

Figure 2. Percentage of students willing to perform or be subject of chest examination with a male or female peer, according to sex and course type

PPEQ questionnaire Mokken scale analysis (MSA) of PPEQ items at a threshold of 0.3 yielded a single scale with weak scalability (Loevinger's H = 0.375 ± 0.009) and good internal consistency (Cronbach's α = 0.896). Monotonicity checks revealed 4 violations out of 4,622 comparisons (0.09%), none statistically significant, confirming that monotonicity was met for all items.

Raising the threshold split items into two scales, both with acceptable H values. The split, however, reflected item polarity rather than construct: positively worded items (confidence, professionalism, learning) loaded on one scale, while negatively worded items (worries, inappropriateness) loaded on the other. For instance, the item on the inappropriateness of practicing PPE clustered separately from those on appropriateness of performing or undergoing PPE. These subscales, therefore, were discarded. After MSA confirmed scalability and consistency of the whole PPEQ questionnaire, PCA was performed to identify specific underlying dimensions of PPE acceptability. PCA identified 22 components, of which four with eigenvalues > 1 were retained. Table 1 reports loadings of each item on the first four components.

	PPEQ Question	PC1	PC 2	PC 3	PC 4
1	Comfortable visiting a peer	**0.265**	0.182	**-0.090**	0.140
2	Comfortable being visited by a peer	**0.267**	0.168	**-0.205**	0.062
3	Embarrassed by peers	**-0.216**	0.147	**0.440**	0.211
4	Embarrassed by teacher	**-0.223**	0.174	**0.406**	0.183
5	Worried to be object of sexual interest	-0.195	**0.338**	0.018	-0.127
6	Worried of sexual interest for peers	-0.155	**0.369**	-0.203	**-0.384**
7	Worried of sexual interest for teacher	-0.142	**0.359**	-0.229	**-0.377**
8	Comfortable visiting same-sex peer	**0.249**	0.222	0.101	**0.342**
9	Comfortable visiting diff-sex peer	**0.271**	0.188	-0.156	**0.212**
10	Comfortable being visited by same-sex peer	**0.262**	0.233	-0.024	**0.290**
11	Comfortable being visited by diff-sex peer	**0.272**	0.173	-0.288	0.109
12	PPE on future collegues inappropriate	**-0.201**	**0.282**	-0.093	0.043
13	Appropriate to perform PPE	**0.250**	0.120	**0.408**	-0.183
14	Appropriate to be subject of PPE	**0.276**	0.134	**0.268**	**-0.268**
15	With PPE I obtain useful feedback	**0.254**	0.131	**0.327**	**-0.275**
16	PPE sign of student professionalism	**0.253**	0.116	0.143	**-0.255**
17	Diff-sex peers inappropriate	**-0.211**	**0.322**	-0.033	0.193
18	Diff-sex teacher inappropriate	**-0.222**	**0.303**	0.070	0.225

	EFS Principal Component 1				**EFS Principal Component 2**			
EFS item	To male	by male	to female	by female	to male	by male	to female	by female
Head & neck	0.17	0.16	0.17	0.16	-0.19	-0.21	-0.21	-0.22
Arms	0.17	0.17	0.17	0.16	-0.17	-0.17	-0.20	-0.20
Chest	0.18	0.17	0.18	0.18	-0.02	0.03	-0.02	-0.02
Breast	0.17	0.15	0.15	0.17	**0.07**	**0.14**	**0.19**	**0.15**
Abdomen	0.18	0.18	0.19	0.18	-0.06	0.00	-0.03	-0.03
Dorsolumbar	0.19	0.19	0.19	0.19	-0.05	-0.02	-0.07	-0.06
Groin	0.17	0.15	0.16	0.16	**0.17**	**0.21**	**0.19**	**0.20**
Genitals	**0.11**	**0.10**	**0.11**	**0.11**	**0.34**	**0.35**	**0.31**	**0.32**
Legs	0.18	0.18	0.18	0.18	-0.03	-0.03	-0.04	-0.03

Table 1. Loadings of PPEQ items (top) on the first four principal components and of EFS items (bottom) on the first two principal components. Higher absolute loadings indicate stronger contribution of the item to the component. Loadings that most characterize each component are shown in bold. For EFS items, bold values highlight body regions contributing most to the differentiation between components, particularly in PC2, which separates more sensitive (genital, groin, breast) from less sensitive areas.

PPEQ-PC1 reflected professional confidence and acceptance: higher scores indicated greater confidence, appropriateness, and professionalism, and lower embarrassment. Males scored higher than females, and BS students lower than MD students ($p<0.001$ for both).

PPEQ-PC2 reflected concerns about sexual connotation and cross-sex interactions, with higher scores indicating stronger perceptions of inappropriateness and greater worries about sexual interest from peers or teachers. BS students scored higher than MD students ($p=0.006$), with no sex difference.

PPEQ-PC3 appeared to reflect a combination of feelings of personal inadequacy and embarrassment in peer/teacher interactions (not sexually related), together with endorsement of PPE appropriateness and usefulness. As this component accounted for a smaller proportion of total variance, its interpretation should remain cautious. Females scored higher than males ($p<0.001$) and MD students higher than BS students ($p=0.001$).

PPEQ-PC4 was characterized by loadings suggestive of more general negative convictions toward PPE (weaker endorsement of appropriateness, usefulness, and professionalism, without links to sexual concerns or major discomfort). As this component explained a limited proportion of variance, it may represent a secondary dimension emerging from the variance structure of the questionnaire rather than a fully distinct psychological construct. Females scored higher than males ($p<0.001$), with no difference between BS and MD students.

Exploratory clustering based on the four PCs revealed no clear structure: indices such as the elbow method, η^2, and proportional reduction of error showed only gradual changes, suggesting that students are better represented along a continuum of confidence and acceptability rather than discrete clusters. When participants were grouped into four clusters, profiles followed the same gradient already captured by EFS, PPEQ, and PC1. Differences in PC2–PC4 across clusters were inconsistent, indicating no additional typologies beyond those identified by component scores.

EFS questionnaire MSA of the EFS questionnaire yielded two scales: one containing the four genital-region questions (visiting or visited, same or different sex; $H = 0.79$) and one containing all others ($H = 0.64$).

PCA of EFS responses yielded two components (Table 1, bottom). The first presented positive loadings for all items, slightly smaller for genital-region questions. The second showed negative loadings for all items except for intimate regions (breast, groin, genitals), which showed positive loadings.

Scores of both components were higher in MD than BS students ($p<0.0001$ in both) and, among BS students only, component 2 was higher in males than in females ($p=0.049$).

Based on homogeneous loading profiles, PPEQ loadings suggested partitioning the total score into three subscales: PPEQ-E (Emotions),

reflecting comfort/discomfort (items 1–4, 8–11, with 3–4 reversed); PPEQ-A (Attitudes), reflecting beliefs on professional appropriateness and usefulness (items 13–16); and PPEQ-SC (Sexual Concerns), reflecting worries and beliefs related to sex issues (items 5–7, 12, 17–18). These subscales showed good scalability, with H = 0.50, 0.68, and 0.52, respectively.

Similarly, EFS loadings suggested two subscales by body-area sensitivity, the same distinction identified by MCA: EFS-SA (Sensitive Areas), including 12 items on breast, groin, and genitals, and EFS-LSA (Less Sensitive Areas), including the remaining 24. Scalability was good (H = 0.68 and 0.59). EEFS-SA scores were widely distributed (median 42%, IQR 8–67), whereas EFS-LSA scores clustered near the maximum (median 100%, IQR 92–100), with only eight students (0.4%) scoring below 17%.

Convergent validity

Convergent validity analyses tested associations between principal components and subscales of each questionnaire and the total score of the other. Table 2 shows associations between the percentile rank of EFS score and PPEQ components and subscales (top) and between percentile rank of PPEQ score and EFS components and subscales (bottom), analyzed with rank-based ANCOVA.

Predictor	Univariate	R^2	Multivariable
PPEQ components			
PPEQ-PC1	5.8 (5.5–6.2) *<0.0001*	0.27	5.8 (5.5–6.2) *<0.0001*
PPEQ-PC2	-0.2 (-1.1–0.6) *NS*	0.0002	-0.3 (-1.0–0.5) *NS*
PPEQ-PC3	-4.0 (-5.0–-3.0) *<0.0001*	0.03	-4.1 (-4.9–-3.2) *<0.0001*
PPEQ-PC4	-1.8 (-2.9–-0.6) *0.003*	0.01	-1.7 (-2.7–-0.8) *<0.0001*
PPEQ subscales			
PPEQ-E	19.8 (18.5–21.1) *<0.0001*	0.27	16.5 (14.7–18.2) *<0.0001*
PPEQ-A	13 (11.6–14.5) *<0.0001*	0.13	2.6 (0.9–4.3) *0.003*
PPEQ-SC	-2.4 (-2.7–-2.1) *<0.0001*	0.12	-5.4 (-7.3–-3.5) *<0.0001*
EFS components			
EFS-PC1	16.4 (14.2–18.5) *<0.0001*	0.12	7.2 (4.7–9.6) *<0.0001*
EFS-PC2	17.1 (15.5–18.6) *<0.0001*	0.17	13.1 (11.2–15) *<0.0001*
EFS subscales			
EFS-SA	0.8 (0.8–0.9) *<0.0001*	0.90	0.7 (0.7–0.7) *<0.0001*
EFS-LSA	1.1 (1.0–1.2) *<0.0001*	0.44	0.5 (0.4–0.5) *<0.0001*

Table 2. Associations between PPEQ and EFS components and subscales in univariate and multivariable models. Coefficients (95% CI) represent changes in percentile rank of the outcome score per one-unit increase in predictor score. Positive coefficients indicate higher acceptability; negative coefficients indicate lower acceptability. R^2 indicates the proportion of variance explained.

Of PPEQ parameters, PC1 was positively associated with EFS, PC3 strongly and PC4 weakly negatively associated, while PC2 showed no association. As expected for PCA-derived features, regression coefficients were similar in uni- and multivariable models. PPEQ components explained 31% of EFS variability. PC1 alone accounted for most of the explained variance, as its scores varied over a wider range. As shown in Figure 3, the slope of EFS vs. PC1 was similar across the four sex-by-course groups despite different mean levels.

Sex and course together explained 9.7% of EFS variability. When both were entered with the four PPEQ PCs, their effects were attenuated but remained significant, indicating that some acceptability aspects related to sex and course were not captured by PPEQ dimensions. Nonetheless, the explained variability rose only from 31% to 32%, showing that PPEQ PC1 closely mirrors overall acceptability as measured by EFS, with other components reflecting secondary aspects such as sexual concerns, embarrassment, and same- vs. opposite-sex comfort.

Significant associations were also observed between PPEQ subscores and ranked EFS score, with R^2 ranging 0.12–0.27 in univariate and 0.29 in multivariable models. Both EFS components (Table 2, bottom) were associated with PPEQ score in both analyses. EFS-PC1 explained 12% of PPEQ variation, EFS-PC2 16.5%, and together 18%.

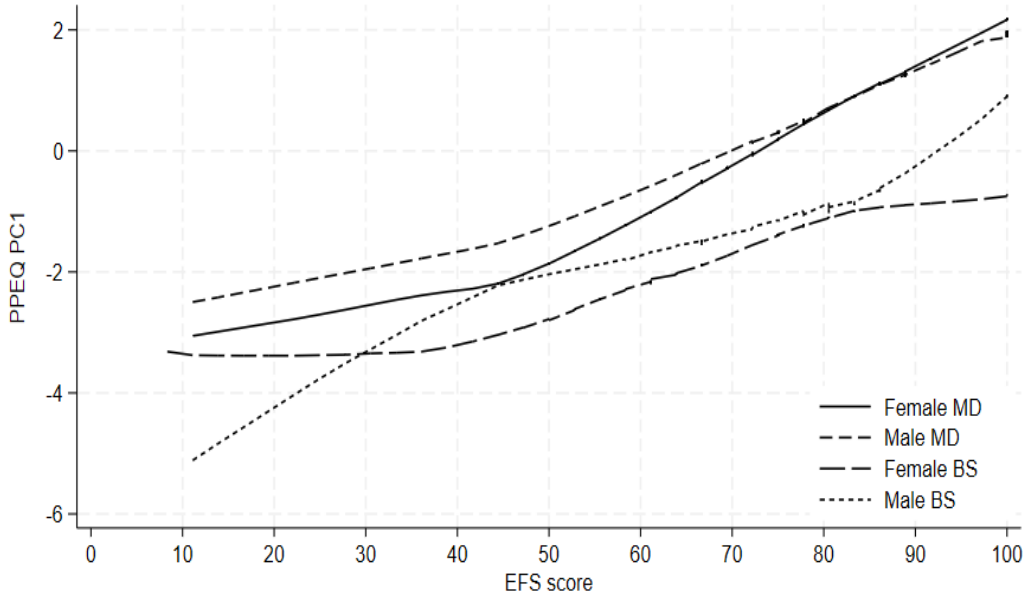

Figure 3. LOWESS regression of the first principal component (PC1) on the EFS score, stratified by sex and type of course. Despite different mean levels across groups, the positive trend with EFS was consistent in all subgroups

The two EFS subscales were also associated with ranked PPEQ scores, with R^2 values up to 0.90 (Table 2, bottom). All associations remained significant ($p < 0.0001$) after adjustment for sex and course type.

DISCUSSION

The acceptability of using PPE to practice physical examination procedures before entering clinical training, and the favorable attitude of medical and other health profession students towards this approach, along with its limitations, has been documented in several studies worldwide (2,3,7,8,10,16–28) and recently reviewed elsewhere (5). Our findings are consistent with these observations, including the well-established lower favor among females compared to males, and extend them to a large multiprofessional student population. We also found that in our institution BS students, mostly nursing students, reported lower acceptability scores than MD students, both on the EFS and PPEQ questionnaires, irrespective of sex. To our knowledge, no previous studies have directly compared MD and nursing students. A small study from New Zealand reported good EFS scores among nursing students, similar to those observed in MD students in the same and other institutions, but without a direct comparison. Exploring the reasons for this possible discrepancy is beyond the scope of this article; nevertheless, several factors could differently influence the perspective of medical and nursing students towards PPE, including the different timing and opportunity of patient contact (with nursing students routinely involved in intimate care from the beginning of training, and perceiving greater opportunity to practice directly on patients), different gender roles within the professional community (5), different socioeconomic backgrounds (29), and the comparatively limited role of physical examination in nursing practice (30). For nursing students in particular, the combination of early and extensive exposure to patient care and the lesser centrality of physical examination in their profession, at least in our institution, may make PPE appear less important or less necessary. Whatever the underlying reason, in our population, acceptability remained high in both groups, and the heterogeneity of the student population further strengthens our results.

Using MSA, we were unable to replicate the results obtained in a more homogeneous group of Australian chiropractic students (11), where three subscales (‘comfort’, ‘concern’, and ‘professionalism and education’) had been identified. No clear separation emerged with this method. Nonetheless, PCA allowed us to identify three subscales with striking similarities to theirs. Specifically, our PPEQ-E (Emotions) and PPEQ-A (Attitudes) subscales were identical in content to their “comfort” and “professionalism and education” subscales. Our PPEQ-SC (Sexual Concerns, concerns about sexual connotations) subscale largely overlapped with their “concern” subscale. Notably, it also included item 12 (“inappropriateness of performing PPE on

persons that will be my future colleagues"), that had not been retained in their analysis, and two additional questions about practicing PPE in the presence of teachers or students of the opposite sex, that were not part of their instrument but are, especially the latter, relevant in shaping perceptions and organizational aspects of the activity.

The PCA results were also broadly consistent with those obtained with factor analysis by Consorti et al. in Italian medical and chiropractic students (28). In that study, a factor labeled "appropriateness and usefulness" (like our PPEQ-PC1) loaded both on comfort and on positive attitudes and beliefs, while a factor "sexual implications" (like our PPEQ-PC2) was driven by concerns about sexuality and again by item 12. Also in that population, item 12 clustered with sexual concerns rather than with attitudes and beliefs.

Our results further refined this picture, detecting additional components: one reflecting uneasiness not specifically related to interactions with the opposite sex (potentially including low self-confidence, body image, or feelings of unpreparedness), and another reflecting conceptual or ideological objections, not directly linked to uneasiness or sexual worries.

The correlation between EFS and PPEQ scores was moderately strong (Spearman's ρ=0.52), supporting our strategy of cross-validating the subscales of each questionnaire through concurrent validity. The four PPEQ components accounted for 31% of the variability of the EFS score, with the first component alone explaining 27%.

The three PPEQ subscales, which closely overlapped with those proposed by Vaughan and Grace (11), were all independently associated with the EFS score, accounting together for 29% of its variance. PPEQ-E alone, however, contributed 27%, emerging as the strongest predictor of the EFS score. The PPEQ-SC subscore was also associated with the EFS, whereas the score of the PCA-derived component (PPEQ-PC2) from which it was extracted was not. In PCA, the variance linked to discomfort and negative attitudes was probably absorbed into PPEQ-PC1, where feelings, beliefs and attitudes were intermingled. In contrast, the PPEQ-SC subscale, computed as the sum of the relevant items of that component, preserved their conceptual meaning and therefore showed the expected association with the EFS.

We are not aware of any previous MSA or PCA of the EFS questionnaire. MSA and PCA both distinguished items relating to more sensitive from those relating to less sensitive body areas, but differed in the way they defined this separation. MSA isolated only the items concerning the genital region, whereas PCA grouped together all the more sensitive areas, including breast and groin. We considered the latter solution more conceptually appropriate. Items about examining the male breast, not usually regarded as embarrassing, clustered with the other sensitive areas, possibly reflecting a sense of reciprocity.

The two main instruments used to assess the acceptability of PPE are based on different principles. The EFS is pragmatic, as it simply investigates willingness to examine or to be examined in different body areas, a logical prerequisite for any PPE activity. The PPEQ, instead, explores the underlying feelings, attitudes and beliefs. Their results are therefore complementary. The practical value of the EFS decreases, however, when moving from the total score or the EFS-SA subscale to the EFS-LSA subscale. The percentage of PPEQ variance explained is higher for the EFS-SA or total score, that include sensitive areas usually excluded from PPE (except in motivated groups of students (31)), than for the EFS-LSA, that only covers areas directly involved in the activity. In this study, which focused on chest examination—a body area of intermediate sensitivity between the intimate regions of the EFS-SA and other less sensitive areas (32)—availability to examine a fellow student's chest was nearly 100%, despite the generally low EFS-SA. That 4% of students enrolling in a chest PPE module reported unwillingness to examine the chest of either sex should not be too surprising: these questionnaires are better seen as a poll than as a formal commitment, and students are not always fully consistent with their responses (33). Thus, while the EFS or EFS-SA scores are better indicators of general disposition towards PPE, the EFS-LSA or single-item responses on the body areas actually involved may be more useful to identify the few outliers who could have difficulties. These cases can usually be anticipated and successfully managed within a healthy learning community, supported by a written policy that ensures sensitive, respectful, inclusive and pragmatic choices. It is plausible that the relative contribution of the identified subscales would differ in modules involving body regions of higher or lower sensitivity. For example, dimensions related to sexual concerns or emotional discomfort might play a more prominent role in activities involving intimate examinations, whereas attitudinal components could predominate in less sensitive contexts. Further studies applying the same instruments to different anatomical modules would help clarify the stability and generalizability of these dimensions.

Although we confirmed internal consistency, scalability, and convergent validity, overall construct validity — including external or criterion-related aspects and a better definition of the concept of "acceptability" — remains to be demonstrated. In addition, the study was conducted within a single institution. Cultural norms, institutional policies, and the organization of PPE within a given curriculum may substantially influence students' perceptions and reported acceptability. Therefore, caution is warranted when generalizing these findings to different sociocultural or educational contexts.

Nevertheless, these subscales could be useful not only to assess the general acceptability of the practice in different contexts, but also to explore the factors associated with it and to monitor

progress and evaluate programs. Since the early diffusion of PPE in the teaching of physical examination, concerns have been periodically raised about important ethical issues that must be carefully addressed but are sometimes overlooked, resulting in policies—or even the absence of policies—and practices insufficiently respectful of the basic principles of beneficence, non-maleficence, autonomy and justice of all the students (4,9,34–40). It has recently been recommended that a framework for ethical experiential learning should, among other aspects, contribute to curriculum quality assurance and review cycles, incorporating feedback from both students and staff (4). Longitudinal analysis of PPEQ subscales could help to ensure the quality of the learning experience and to adapt policies to evolving needs.

Acknowledgement: None declared.

Conflict of Interest Statement: The authors declare that they have no conflict of interest.

REFERENCES

1. Servant VFC. Revolutions and Re-Iterations: An Intellectual History of Problem-Based Learning. PhD thesis. Erasmus University Rotterdam; 2016. https:/repub.eur.nl/pub94113

2. Metcalf NF, Prentice ED, Metcalf WK, Stinson WW. Peer group models in examination instruction as an integral part of medical gross anatomy. J Med Educ 1982;57(8):641-644. doi:10.1097/00001888-198208000-00013

3. 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(5):433-437.

4. Grace S, Stockhausen L, Patton N, Innes E. Experiential learning in nursing and allied health education: Do we need a national framework to guide ethical practice? Nurse Educ Pract 2019;34:56-62. doi:10.1016/j.nepr.2018.11.003

5. Al-Chalabi N. Practicing Peer Physical Examination by Medical and Allied Health Professions Students: Why the Hesitation and Where are We Now? A Narrative Literature Review. Adv Med Educ Pract 2024;15:513-525. doi:10.2147/AMEP.S462588

6. Uchida T, Achike FI, Blood AD, et al. Resources Used to Teach the Physical Exam to Preclerkship Medical Students: Results of a National Survey. Acad Med 2018;93(5):736-741. doi:10.1097/ACM.0000000000002051

7. Hamo I, DAS N, Lanphear J. The Skill Laboratory as Viewed by Medical Students. Annals of Community-Oriented Education 1994;7:209-315.

8. Chang EH, Power DV. Are medical students comfortable with practicing physical examinations on each other? Acad Med 2000;75(4):384-389.

9. Suzuki E, Nawa N, Okada E, et al. An action research study on the needs and perspectives of medical students and junior residents on peer

physical examination guidelines to respect medical students' autonomy and privacy in Japan. BMC Med Educ 2025;25(1):174. doi:10.1186/s12909-025-06649-2

10. Consorti F, Mancuso R, Milazzo F, Notarangelo M, Piccolo A. Acceptability among Italian medical students of Peer Physical Examination (PPE): validation of tools of measure and first results of a survey. Peer Physical Examination for Italian students. Tutor 2012;12(1):16-25.

11. Vaughan B, Grace S. A Mokken scale analysis of the peer physical examination questionnaire. Chiropr Man Therap 2018;26:6. doi:10.1186/s12998-018-0176-0

12. Warm EJ, Mathis BR, Held JD, et al. Entrustment and mapping of observable practice activities for resident assessment. J Gen Intern Med 2014;29(8):1177-1182. doi:10.1007/s11606-014-2801-5

13. Conover WJ, Iman RL. Analysis of covariance using the rank transformation. Biometrics 1982;38(3):715-724.

14. Hardouin JB, Bonnaud-Antignac A, Sebille V. Nonparametric item response theory using Stata. Stata J 2011;11(1):30-51.

15. Markles A. Stata tip 110: How to get the optimal k-means cluster solution. The Stata Journal 2012;12(2):247-351.

16. Rees CE, Bradley P, McLachlan JC. Exploring medical students' attitudes towards peer physical examination. Med Teach 2004;26(1):86-88. doi:10.1080/0142159031000164-2984

17. Rees CE, Wearn AM, Vnuk AK, Sato TJ. Medical students' attitudes towards peer physical examination: findings from an international cross-sectional and longitudinal study. Adv Health Sci Educ Theory Pract 2009;14(1):103-121. doi:10.1007/s10459-007-9094-y

18. Chinnah TI, de Bere SR, Collett T. Students' views on the impact of peer physical examination and palpation as a pedagogic tool for teaching and learning living human anatomy. Med Teach 2011;33(1):e27-36. doi:10.3109/0142159X.2011.530313

19. Reid KJ, Kgakololo M, Sutherland RM, Elliott SL, Dodds AE. First-year medical students' willingness to participate in peer physical examination. Teach Learn Med 2012;24(1):55-62. doi:10.1080/10401334.2012.641489

20. Wearn AM, Bhoopatkar H, Mathew TK, Stewart L. Exploration of the attitudes of nursing students to peer physical examination and physical examination of patients. Nurse Educ Today 2013;33(8):884-888. doi:10.1016/j.nedt.2012.08.012

21. Manjunath H, Venkatesh D, Taklikar R, Vijayanath V. Medical Student's Perception And Preferences About Peer Physical Examination (PPE). Asian Journal of Health Sciences 2014;2(1):42-44. doi:10.15419/ajhs.v2i1.423

22. Vaughan B, Grace S. Perception of peer physical examination in two Australian osteopathy programs. Chiropr Man Therap 2016;24:21. doi:10.1186/s12998-016-0102-2

23. Burggraf M, Kristin J, Wegner A, et al. Willingness of medical students to be examined in a physical examination course. BMC Med Educ 2018;18(1):246. doi:10.1186/s12909-018-1353-5

24. Soqia J, Shamaa MA, Alhomsi D, et al. Syrian medical students' acceptance of peer physical examination and its associating factors: a cross-sectional study. BMC Med Educ 2022;22(1):898. doi:10.1186/s12909-022-03985-5

25. Gupta V, Begum Y, Singh A, Agrawal D. Perception of medical students towards teaching basic clinical skills in otorhinolaryngology through peer physical examination (PPE). J Educ Health Promot 2023;12:172. doi:10.4103/jehp.jehp_1165_21

26. Khadour YA, Khadour AM, Ebrahem BM, Khadour FA. Acceptance of peer physical examination in medical students in Wuhan, China: A cross-sectional study. Health Sci Rep 2024;7(3):e1959. doi:10.1002/hsr2.1959

27. Suzuki E, Nawa N, Okada E, et al. Mixed-methods study of medical students' attitudes toward peer physical examinations in Japan. BMC Med Educ 2024;24(1):681. doi:10.1186/s12909-024-05635-4

28. 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. doi:10.1186/1472-6920-13-111

29. Suzuki Y, Tsunekawa K, Takeda Y, Cleland J, Saiki T. Impact of medical students' socioeconomic backgrounds on medical school application, admission and migration in Japan: a web-based survey. BMJ Open 2023;13(9):e073559. doi:10.1136/bmjopen-2023-073559

30. Douglas C, Windsor C, Lewis P. Too much knowledge for a nurse? Use of physical assessment by final-semester nursing students. Nurs Health Sci 2015;17(4):492-499. doi:10.1111/nhs.12223

31. Politis M, El Brown M, Huser CA, Crawford L, Pope L. "I wouldn't know what to do with the breasts": the impact of patient gender on medical student confidence and comfort in clinical skills. Educ Prim Care 2022;33(6):316-326. doi:10.1080/14739879.2022.2129460

32. Reid K, Sutherland R, Dodds A, McNair R, Pierce D. Abdomen and chest examinations in peer physical examination: Variation in participation by gender. Focus on Health Professional Education: A Multi-Professional Journal 2015;16(2):13-26. doi:10.11157/fohpe.v16i2.66

33. Chen JY, Yip ALM, Lam CLK, Patil NG. Does medical student willingness to practise peer physical examination translate into action? Med Teach 2011;33(10):e528-540. doi:10.3109/0142159X.2011.599893

34. Braunack-Mayer AJ. Should medical students act as surrogate patients for each other? Med Educ 2001;35(7):681-686. doi:10.1046/j.1365-2923.2001.00970.x

35. Outram S, Nair BR. Peer physical examination: time to revisit? Med J Aust 2008;189(5):274-276.

36. Delany C, Frawley H. We need a new model for obtaining students' consent to conduct peer physical examinations. Acad Med 2011;86(5):539; author reply 539. doi:10.1097/ACM.0b013e318212eb2c

37. Koehler N, McMenamin C. The need for a peer physical examination policy within Australian medical schools. Med Teach 2014;36(5):430-433. doi:10.3109/0142159X.2013.874551

38. Wearn A, Bhoopatkar H. Experience of a peer physical examination policy within a New Zealand medical programme. Med Teach 2014;36(9):826-827. doi:10.3109/0142159X.2014.940878

39. Koehler N, Currey J, McMenamin C. What Should Be Included in a Peer Physical Examination Policy and Procedure? Med Sci Educ 2014;24(4):379-385. doi:10.1007/s40670-014-0068-4

40. Abdelrahman A, Whitney T, Salas NM, Barrett E, Okanlami FO. Changing Policy for Inclusion: Peer-to-Peer Physical Exam Practice in Medical School. Teaching and learning in medicine 2025;37(2). doi:10.1080/10401334.2023.2298865