



Measurement of perceptions of educational environment in evidence-based medicine

Anne-Marie Bergh,¹ Jackie Grimbeek,¹ Win May,²
A Metin Gülmezoglu,³ Khalid S Khan,⁴ Regina Kulier,³
Robert C Pattinson¹

10.1136/eb-2014-101726

¹Medical Research Council Unit for Maternal and Infant Health Care Strategies, University of Pretoria, Pretoria, South Africa

²Department of Medical Education, Keck School of Medicine, University of Southern California, Los Angeles, California, USA

³World Health Organization, Geneva, Switzerland

⁴Women's Health Research Unit, The Blizard Institute, Barts and The London School of Medicine, Queen Mary, University of London, London, UK

Correspondence to:

Dr Anne-Marie Bergh

Medical Research Council Unit for Maternal and Infant Health Care Strategies, University of Pretoria, Private Bag X323, Arcadia, Pretoria 0007, South Africa;
anne-marie.bergh@up.ac.za



Editor's choice
Scan to access more
free content

Abstract

In recent years, there has been a renewed interest in measuring perceptions regarding different aspects of the medical educational environment. A reliable tool was developed for measuring perceptions of the educational environment as it relates to evidence-based medicine as part of a multicountry randomised controlled trial to evaluate the effectiveness of a clinically integrated evidence-based medicine course. Participants from 10 specialties completed the questionnaire. A working dataset of 518 observations was available. Two independent subsets of data were created for conducting an exploratory factor analysis ($n=244$) and a confirmatory factor analysis ($n=274$), respectively. The exploratory factor analysis yielded five 67-item definitive instruments, with five to nine dimensions; all resulted in acceptable explanations of the total variance (range 56.6–65.9%). In the confirmatory factor analysis phase, all goodness-of-fit measures were acceptable for all models (root mean square error of approximation ≤ 0.047 ; comparative fit index ≥ 0.980 ; normed $\chi^2 \leq 1.647$; Bentler-Bonett normed fit index ≥ 0.951). The authors selected the factorisation with seven dimensions (factor-7 instrument) as the most useful on pragmatic grounds and named it Evidence-Based Medicine Educational Environment Measure 67 (EBMEEM-67). Cronbach's α for subscales ranged between 0.81 and 0.93. The subscales are: 'Knowledge and learning materials'; 'Learner support'; 'General relationships and support'; 'Institutional focus on EBM'; 'Education, training and supervision'; 'EBM application opportunities'; and 'Affirmation of EBM environment'. The EBMEEM-67 can be a useful diagnostic and benchmarking tool for evaluating the perceptions of residents of the environment in which evidence-based medicine education takes place.

Introduction

Educational environment in medical education can be described as the context in which clinical staff and students teach and learn.¹ This environment has also been associated with educational or learning climate^{2–5} and educational culture.⁵ According to Genn,⁴ climate is a manifestation of the concept of environment. In the past 20 years, several publications have appeared on the measurement of students' perceptions of various types of medical educational environments. Besides more general descriptions of the medical education environment,^{4–7} environments have also been delimited to specific situations such as the operating theatre,^{8–10} general practice training¹¹ and undergraduate¹² and postgraduate³ training. An implicit aim of all education

in healthcare settings is to produce an environment conducive to advanced and in-service learning.

Measuring students' or doctors' perceptions of the medical educational environment has a long history. Some of the older, widely used instruments are the Learning Environment Questionnaire (LEQ) and the Medical School Learning Environment Survey (MSLES).⁷ According to Schönrock-Adema *et al*,¹⁴ the lack of consensus about which concepts to measure may be explained by the absence of a common theoretical framework. They propose a framework with three broad domains for developing medical educational environment measures: goal orientation (content and aims of education in relation to personal development); relationships (open and friendly atmosphere and affiliation); and organisation/regulation (system maintenance and change). An instrument that has in the past 15 years been applied in a variety of settings across the world and that also caught our attention is the so-called Dundee Ready Education Environment Measure (DREEM).¹² The DREEM sparked the development and validation of other tools measuring more specific postgraduate educational environments.⁸

The measurement of the educational environment was a secondary outcome of a randomised controlled trial conducted between March 2009 and November 2011 to evaluate the effectiveness of a clinically integrated EBM course for obstetrics and gynaecology residents.²³ Ethical approval for the validation of the tool was received as part of the trial protocol. One of the points of departure in the trial was that the application of EBM was also influenced by the workplace climate or environment. Any teaching in EBM should therefore also be measured in terms of its ability to facilitate evidence-based practice in the broader clinical environment, going beyond imparting knowledge and skills. This paper reports on the development and validation of such a tool, which was administered before and after the intervention.

Method

As the measurement of educational environment in the EBM education trial would be based on the perceptions of participants in postgraduate education, a survey design was considered appropriate for measuring attitudes and opinions.²⁴ The study design comprised two phases. The first was to develop a draft instrument for measuring residents' perceptions of their educational environment as it related to EBM, whereas the second focused on the validation of this tool as a secondary outcome measure in the EBM education trial. The development and validation process is depicted in figure 1.

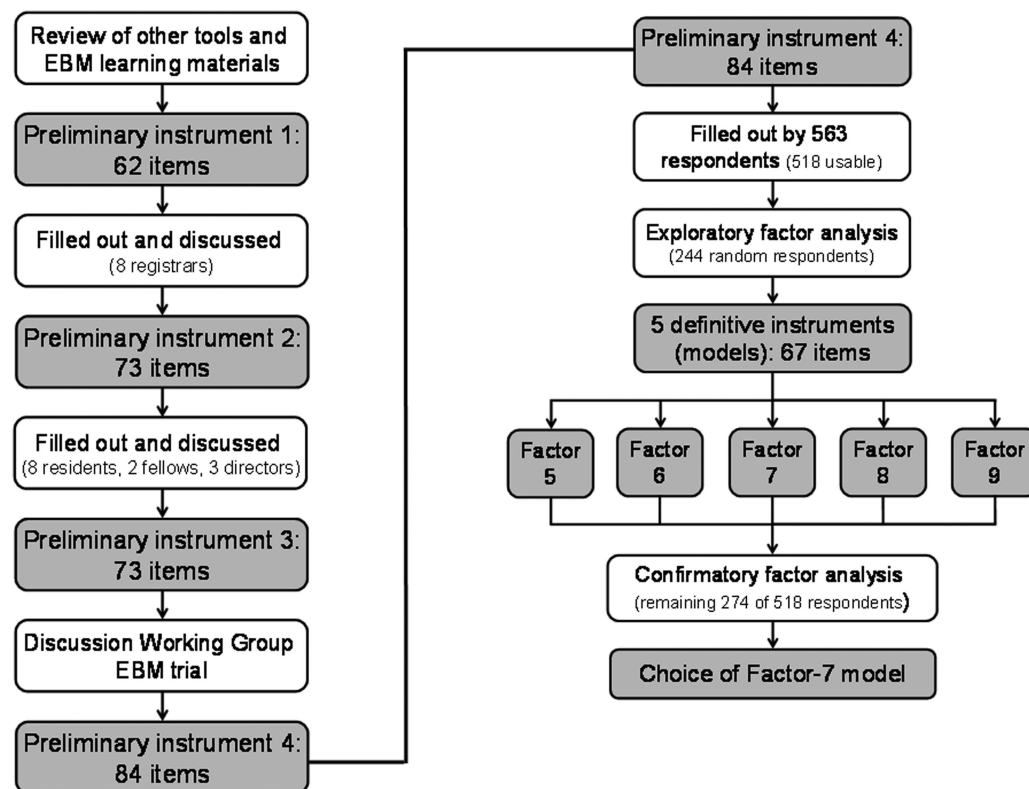


Figure 1 Development and testing of the instrument.

Development of scales and items for a draft questionnaire

Two investigators (WM and A-MB) reviewed the literature on the development and validation of educational environment tools to identify and adapt potentially useful scales and items. They formulated provisional themes for taking the process forward, namely perceptions of: (1) learning opportunities, (2) self-learning ('EBM competence'), (3) availability of learning resources, (4) teachers and teaching, supervision and support ('EBM specific'), (5) *EBM* practice ('EBM atmosphere'), (6) general atmosphere. These themes correspond to some extent with the three domains proposed by Schönrock-Adema *et al*¹⁴ to measure the medical educational environment. Themes 2 and 5 relate to goal direction; theme 6 to relationships and themes 1, 3 and 4 to organisation/regulation.

The manuals of the e-course of the randomised education trial were then studied and more items that would give feedback on the actual *EBM* practice of an institution were generated. Investigators in the trial also interacted with further inputs and comments. Two senior consultants involved in the teaching of *EBM* in obstetrics and gynaecology and in internal medicine at the University of Pretoria, South Africa were then requested to comment on the scope of the scales and items. The result was a preliminary instrument with 62 items. These were presented randomly in a questionnaire given for completion to five registrars (residents) in obstetrics and gynaecology and three in internal medicine at the University of Pretoria. An item-by-item discussion checked for any ambiguities or unclear statements and

ranked the items according to importance. In this process, some items were dropped and others were split into more than one item to enhance clarity, yielding the second preliminary instrument with 73 items. These items were submitted to a group of five residents from obstetrics and gynaecology, three chief residents from internal medicine, two paediatric pulmonary fellows, two programme directors and one assistant programme director at the Keck School of Medicine of the University of Southern California. The 73 items of the previous version remained, with some further refinement in the wording. The English version of the third preliminary instrument was then presented at a working group of the *EBM* education trial²³ investigators for further discussion. Eleven more items were added, of which 8 pertained to access, use and usefulness of the WHO Reproductive Health Library (RHL)¹—amounting to 84 items. It was agreed that the RHL-related items would be used for trial purposes only and not during the data analysis exercise for developing an instrument.

An electronic version and a paper-based version of the questionnaire were generated to encourage greater participation in the validation process. Respondents could participate anonymously and participants with difficulty in accessing the Internet could complete the paper-based version. For the purposes of the trial, experts familiar with the *EBM* environment also translated the questionnaire into French, Spanish and Portuguese. One of the challenges in the development of the instrument was the different terminologies used in different countries. Eventually, we decided on the following as alternatives in the instrument: 'registrars'/

'residents' and 'faculty'/'consultants'. For the purposes of this paper, registrars and residents will be called 'residents' and consultants will be called 'faculty'. The final instrument contains both sets of terminologies for readers wishing to adapt the tool according to their context (see online Supplement 1).

Study participants in the administration of the preliminary instrument

Using the factor analysis rule of thumb of 10 participants per item,²⁵ 760 participants were needed for this study and more participants had to be recruited beyond the trial candidates. To be able to generate a tool for use wider than obstetrics and gynaecology (the field of specialisation targeted in the randomised education trial²³), we recruited 'non-trial' participants from other specialties also, namely anaesthesiology, otolaryngology, family medicine, general surgery, internal medicine, neurology, paediatrics, psychiatry and radiology. Trial participants came from Argentina, Brazil, the Democratic Republic of the Congo, India, Philippines, South Africa and Thailand; only their preintervention data were used. Non-trial data came from India, the Netherlands, Philippines, South Africa and the UK. There were participants from all specialty years, between years 1 and 6. Table 1 provides a summary of respondents according to country, specialty and year of study specialisation.

Non-trial participants were recruited through three initiatives—a web-based questionnaire administered at the WHO (n=13), a paper-based initiative in the UK (n=34) and a paper-based initiative in South Africa (n=106) (December 2008–May 2009). The rest of the responses came from the trial dataset (n=410; March 2009–November 2010). The final set of raw data had 563 observations.

Preparation of data

After administering the preliminary instrument to trial and non-trial participants, the data were consolidated and scores were allocated as follows:

Strongly disagree=0; disagree=1; neutral=2; agree=3 and strongly agree=4.

Scores for items formulated in the negative were reversed and those pertaining to the use of the RHL excluded. The remaining 76 items were kept for analysis and the development of the instrument.

Participants responding to 46 or more of the 76 items were retained after performing a binomial analysis,²⁶ which was applied to calculate the threshold for the minimum number of items to have significantly more than 50% of item responses per questionnaire. This resulted in the removal of 45 cases, leaving 518 observations in the final dataset. Data imputation was conducted in three consecutive phases to create a complete dataset:

- On a detailed level ((*country*) by (*year of specialisation*) by (*specialty*)), missing data were imputed by using the *mode* of the available data.
- On a second level, the imputation used the *mode* of the available data for the combinations ((*country*) by (*year of specialisation*)).
- On a more general level, all data still missing after step (B) were imputed by using the *mode* of the data available at the (*country*) level as substitution.

Outline of the data analysis process

Two independent datasets were created from the raw dataset of 518 observations through stratification according to (*country*) by (*year of specialisation*) by (*specialty*). Random systematic sampling resulted in 244 and 274 observations for an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA),²⁷ respectively. As the data were measured on a Likert scale, polychoric correlations were calculated separately for the EFA and CFA datasets.²⁸

The EFA dataset served as a basis to identify possible factor-constructs for which Cronbach's α was calculated to determine internal consistency and to drop items not correlating highly enough with the dimension under consideration. The factor models identified during the EFA and item analysis phases were subjected to a structural equation modelling (SEM) analysis during the CFA for possibly removing item(s) from dimension(s) and/or reallocating items to other dimension(s).

Table 1 Summary of respondents included in the analysis

Country	Sample size			Specialty	Sample size			Year of study	Sample size		
	Sum	[t]	(nt)		Sum	[t]	(nt)		Sum	[t]	(nt)
Argentina	29	[29]		Anaesthesiology	15		(15)	Year 1	113	[79]	(34)
Brazil	151	[151]		Otolaryngology	4		(4)	Year 2	236	[213]	(23)
DRC	7	[7]		Family Medicine	2		(2)	Year 3	78	[45]	(33)
India	43	[37]	(6)	General Surgery	7		(7)	Year 4	29	[8]	(21)
The Netherlands	14		(14)	Internal Medicine	22		(22)	Year 5	11	[1]	(10)
Philippines	41	[41]		Neurology	3		(3)	Year 6 +	27		(27)
South Africa	131	[33]	(98)	Obgyn	436	[366]	(70)	Unknown	24	[20]	(4)
Thailand	60	[60]		Paediatrics	15		(15)				
United Kingdom	34		(34)	Psychiatry	11		(11)				
Unknown	8	[8]		Radiology	3		(3)				
Total	518	[366]	(152)	Total	518	[366]	(152)	Total	518	[366]	(152)

(nt), non-trial participants; [t], trial participants; Obgyn, obstetrics and gynaecology.

Descriptive statistics, polychoric correlations, EFA, Cronbach's α and generalisability coefficients were calculated using SAS V.9.3 software.²⁹ EQS V.6.1³⁰ was mainly used for the CFA. Robust statistics (Satorra and Bentler³¹ from EQS) were used in place of the statistics based on normal theory, because of the large multivariate kurtosis as measured by the normalised estimate of Mardia's coefficient of kurtosis.³²

Results

Exploratory factor analysis

We applied oblique varimax rotation²⁷ to the EFA dataset. Factor-constructs with 5, 6, 7, 8 and 9 dimensions were identified as models with practical application value (henceforth called 'factor-5 model', 'factor-6 model', etc), all of which presented with an acceptable explanation of the total variance (range 56.6–65.9%). The preliminary subscales (division of items) only had a few factor loadings below 0.50; the vast majority was above 0.50 and some even above 0.90. In a further investigation of all the factor models, items with a too small loading (<0.45) or not fitting logically under any dimension were removed.

Cronbach's α was calculated per dimension on the remaining items for each factor, which led to the further removal of items. At this stage, all factor models had 67 items, with Cronbach's α ranging between 0.76 and 0.96, which is well above the suggested rule of thumb value of 0.70.^{27–25} Details of Cronbach's α are provided in online Supplement 2. There was a generally decreasing trend in mean Cronbach's α from the factor-5 to factor-9 models (0.882, 0.868, 0.870, 0.855 and 0.841).

Confirmatory factor analysis

All five factor models (5–9) identified during the EFA phase were investigated by application of SEM during the CFA phase. Tests for goodness-of-fit^{27–25–33} and other aspects of the model resulted in very acceptable values (see table 2). The normed χ^2 was calculated at about 1.6 for the factor-5 to factor-8 models and at 1.3 for the factor-9 model, which can be regarded as a satisfactory fit. The Bentler-Bonett normed fit index (BBNFI) and the comparative fit index (CFI) were both relatively high. The estimated value and 90% CI of the root mean square error of approximation were also within the accepted limits (<0.05). Therefore, the goodness-of-fit measures were all acceptable. The maximum absolute

standardised residuals observed were 0.43, 0.50, 0.35, 0.37 and 0.28 for the factor-5 to factor-9 models, respectively, with the residuals in general following a bell-shaped distribution. Using robust statistics, all items of all models loaded significantly on the dimension to which an item was allocated ($p \leq 0.05$).

The Lagrange multiplier test for reallocating items from one dimension to another was applied simultaneously with the Wald test for dropping parameters. For all dimensions of all models, some items were reallocated, but none could be removed. Cronbach's α at this stage varied between 0.75 and 0.96.

The models fitted explained the data to a very good degree, and it can be assumed that the EFA procedure followed by the CFA modelling was very successful in developing a measurement model based on the available data.

Labelling of dimensions

Dimensions were named differently across the different models. The researchers formulated the label names to reflect the content of the items included under each label. A summary of dimension labels is given in table 3. Three dimension labels (A–C in table 3) feature in all models: 'General relationships and support', 'EBM application opportunities' and 'Affirmation of EBM environment'. There was some difficulty in finding a construct for appropriately naming the last dimension above, as all (except one) of the negatively phrased items clustered together here. This dimension was therefore interpreted as an affirmation of a respondent's perception of the EBM environment.

One dimension (D) appears in four of the five models ('Education, training and supervision') and three (E, I, M) in three models ('Institutional focus on EBM', 'Knowledge and learning materials' and 'Resources and communication'). Two dimensions, 'Teachers' (F) (2 models) and 'Learner support' (G) (1 model), are related to the dimension 'Education, training and supervision'. The dimensions that were more fragmented or where items were grouped in different ways in different models relate to EBM knowledge, learning materials and resources, and communication (dimension labels H–M).

Discussion

To the best of our knowledge, this is the first study to report on the validation of an educational environment or learning climate questionnaire that yielded more than

Table 2 Goodness-of-fit summary for SEM measurement models

Measure	Number of dimensions					Norm
	5	6	7	8	9	
Satorra-Bentler scaled χ^2	3435.9	3503.8	3421.5	3400.0	2761.5	
df	2132	2128	2123	2116	2099	
Normed $\chi^2 = (\chi^2)/df$	1.612	1.647	1.612	1.607	1.316	<2.0 (or <5.0)
BBNFI	0.951	0.951	0.951	0.952	0.959	>0.90 (or >0.95)
CFI	0.981	0.980	0.981	0.981	0.990	>0.90 (or >0.95)
RMSEA	0.047	0.047	0.047	0.047	0.034	<0.05
90% Lower	0.044	0.044	0.044	0.044	0.030	
90% Upper	0.050	0.050	0.050	0.050	0.037	(<0.05, <0.10)

BBNFI, Bentler-Bonett normed fit index; CFI, comparative fit index; df, degrees of freedom; Norm, norms/rules suggested by various authors; RMSEA, root mean square error of approximation; SEM, structural equation modelling.

one acceptable instrument for use and where the choice of model to propose for general use was dictated by pragmatic considerations. The goodness-of-fit results for all five SEMs were very satisfactory and no single model fitted the data markedly 'better' than the others (table 2). The factor-9 model may be considered the best statistically, but dimensions labelled as 'Communication and resources' and 'Access to learning resources' in this model have a semantic overlap and the division appears artificial (table 3). The same goes for the dimensions 'Teachers' and 'Education, training and supervision'.

The factor-7 and factor-8 models appear to be the favoured models from a practical point of view, because of a more balanced distribution of items across the different dimensions. If the dimension labels of the factor-7 and factor-8 models are compared, the labels of the factor-7 model appear to be more 'neatly' divided, whereas there is some semantic overlap between three dimensions in the factor-8 model, namely 'Knowledge and learning materials', 'Access to learning materials and teachers' and 'Resources and communication'. The factor-5 and factor-6 models have too many items in some of the dimensions.

We therefore propose a tool to measure the education environment based on the factor-7 model, which also had a slightly improved Cronbach's α mean value (0.870) compared to the factor-6 model (0.868). Following the naming of other instruments, the 67-item tool is called the Evidence-Based Medicine Educational Environment Measure 67 (EBMEEM-67) and has the following subscales:

- 1 Knowledge and learning materials (8 items)
- 2 Learner support (10 items)
- 3 General relationships and support (8 items)
- 4 Institutional focus on *EBM* (14 items)
- 5 Education, training and supervision (9 items)
- 6 *EBM* application opportunities (12 items)
- 7 Affirmation of *EBM* environment (6 items)

Table 4 contains a summary of the subscales with their items. A user-friendly format of the complete tool

and the instructions for use are attached as an online supplementary file (Supplement 1).

Closer investigation of the subscales and individual items revealed a large degree of correspondence with the three-domain framework of Schönrock-Adema *et al.*¹⁴ Table 5 gives an overview of the similarities.

Compared to other instruments measuring some aspect of the medical educational environment, the EBMEEM-67 has a higher internal consistency for its subscales, with Cronbach's α ranging from 0.81 to 0.93. The psychometric properties of the instruments reported in the literature are summarised in an online supplementary file. The internal consistency of the instrument supports the rigorous process followed in the development and validation of the tool. Items were developed and refined by means of a review of the existing tools and reviews by residents, field experts and the trial study group. The study used the EFA and CFA procedures, with all results for both phases being statistically acceptable and at least significant at a 5% level, where applicable. The EFA was accomplished by factor analysis followed by a varimax oblique rotation to enhance the interpretation of the different dimensions. Cronbach's α demonstrated internal consistency. SEM was successfully applied to the CFA data. Polychoric correlations formed the basis of the correlation matrices for the EFA and CFA analyses. The results of a generalisability study^{34 35} for the EBMEEM-67 showed that mean absolute and relative coefficients of above 0.80 may be expected, which confirmed its effectiveness in measuring the study populations.

Potential strengths and limitations of the study

Respondents were recruited from nine different countries spanning 10 different specialties and 6 years of study, which strengthens the generalisation of the instrument. The wider application of the instrument was confirmed by the generalisability study. Our study fell somewhat short of the general rule requiring 5–10 respondents per item to successfully apply an EFA or CFA procedure.²⁵ Calculating polychoric correlations and using robust statistics during the CFA phase somewhat rectified the

Table 3 Dimension labels and number of items for the different factor models

Dimension label	Factor-5	Factor-6	Factor-7	Factor-8	Factor-9
A General relationships and support	28	10	8	7	5
B <i>EBM</i> application opportunities	10	25	10	11	17
C Affirmation of <i>EBM</i> environment	13	12	8	15	7
D Education, training and supervision		8	14	6	7
E Institutional focus on <i>EBM</i>	10		9		5
F Teachers				5	11
G Learner support			12		
H Knowledge of calculations					5
I Knowledge and learning materials	6	6	6		
J Knowledge of calculations and appraisal				6	
K Access to learning materials and teachers				10	
L Access to learning resources					6
M Resources and communication		6		7	4
Total	67	67	67	67	67

Table 4 Subscales and items per subscale for the EBMEEM-67 tool (CFA data, n=274)

Item	Mean	SD
<i>Subscale 1: Knowledge and learning materials</i>		
14 I know how to calculate Number Needed to Treat when reviewing an article	1.58	1.10
23 Sufficient learning material is available on how to review the literature	2.19	1.11
26 I know how to calculate Relative Risk Reduction when reviewing an article	1.72	1.13
31 Sufficient learning material is available on how to critique articles	2.06	1.04
38 I know how to calculate Absolute Risk Reduction when reviewing an article	1.60	1.08
40 Sufficient learning material is available on how to ask a clinical question	2.22	1.03
53 I know how to calculate Relative Risk when reviewing an article	1.91	1.12
65 I know how to calculate Odds Ratio when reviewing an article	1.75	1.22
<i>Subscale 2: Learner support</i>		
10 Faculty/the consultants set clear expectations about how my review of evidence should support my clinical decisions	2.24	1.06
15 I am given time to participate in educational activities related to <i>EBM</i>	2.05	1.15
28 Residents/registrar are adequately supervised for their application of <i>EBM</i> in their clinical care	1.99	1.10
30 I have the opportunity to apply evidence-based principles in my clinical care	3.01	0.86
33 Faculty/the consultants are enthusiastic about teaching me <i>EBM</i>	2.57	1.10
42 Faculty/my mentor(s) provided me with internet <i>EBM</i> resources	2.38	1.18
48 Faculty/the consultants serve as role models for using <i>EBM</i> in clinical practice	2.27	1.04
50 There are good channels of communication in the hospital(s) and clinic(s) in which I work	2.36	1.05
55 I have online access to the most important <i>EBM</i> sources	2.46	1.15
58 Faculty/the consultants help me to increase my competence in <i>EBM</i>	2.34	0.97
<i>Subscale 3: General relationships and support</i>		
2 Consultants are easily available for assistance in any of the steps required for critical appraisal	2.57	0.92
9 Faculty/the consultants are approachable and friendly	3.12	0.85
17 There are good channels of communication in my department	2.50	1.01
21 I am able to ask 'why?' with regard to clinical practice issues to faculty/consultants	2.75	0.95
43 I have good rapport with faculty/the consultants	3.02	0.82
45 I feel comfortable in my programme	2.64	0.85
54 Working here, I feel part of a team	2.90	0.89
60 My views are respected by faculty/consultants	2.54	0.83
<i>Subscale 4: Institutional focus on EBM</i>		
1 Feedback from review of practice or from morbidity and mortality conferences is used to change practice in my department	2.59	0.92
18 <i>EBM</i> gets sufficient attention in my department	2.39	1.06
20 Faculty/the head of department/head of the firm conducts 'evidence-based' ward rounds with residents/registrar	2.21	1.12
24 Faculty/The consultants are willing to challenge, question and explore	2.69	0.90
27 Faculty/the consultants address clinical questions through a formal review of the evidence	2.31	0.92
29 Case discussions in my hospital/unit/department emphasise critical appraisal of articles	2.66	0.98
34 Using <i>EBM</i> is a routine practice in my institution	2.19	1.10
36 I have access to evidence-based clinical advice from the consultants	2.57	0.91
41 Best practices that are communicated in my institution are normally implemented	2.70	0.83
47 Faculty/the consultants routinely review evidence in morbidity and mortality conferences and in case discussions	2.57	1.08
49 Senior faculty/consultants in my department listen when someone brings evidence that should lead to change of practice	2.75	0.94
51 Reviews of practice are carried out routinely in my department to improve the quality of care	2.45	1.03
63 The evidence from Cochrane Systematic Reviews is used in my department to develop new clinical practice guidelines	2.41	1.07
67 Faculty/the consultants usually use <i>EBM</i> principles when considering treatment for patients	2.57	0.99
<i>Subscale 5: Education, training and supervision</i>		
3 My workload does not allow me sufficient time to review evidence for my case presentations*	1.48	1.13
5 There is systematic, structured training in <i>EBM</i> at my institution	1.91	1.17
7 I receive regular feedback from faculty/seniors on my application of <i>EBM</i>	1.82	1.03

Continued

Table 4 Continued

Item	Mean	SD
16 Residents/registrar address clinical questions through a formal review of the evidence	2.23	1.06
35 Faculty/the consultants taught me how to read and critique the scientific literature	1.92	1.08
39 The <i>EBM</i> materials on the computer at work are up-to-date	2.14	1.03
62 Residents/registrar get a good foundation in <i>EBM</i>	1.85	1.05
64 I have attended seminars or discussions on searching the literature and the critical appraisal of articles during my residency/since starting to work as a resident/registrar	2.32	1.15
66 Faculty/consultants give me regular feedback on my <i>EBM</i> knowledge and skills	2.05	0.99
<i>Subscale 6: EBM application opportunities</i>	2.46	0.60
4 For the patients that I am taking care of, I am able to assess the validity of the evidence that I have found	2.36	0.92
11 I have the opportunity to identify my knowledge gaps in <i>EBM</i>	2.39	1.01
13 For the patients that I am taking care of, I am able to figure out what I need to know	2.86	0.73
19 For the patients that I am taking care of, I am able to apply the valid evidence that I have found to the patient	2.78	0.69
25 I am allocated sufficient time for applying <i>EBM</i> principles in my clinical cases	1.89	1.07
32 I was taught how to interpret a meta-analysis of studies	1.84	1.11
37 When I take care of patients, I have the opportunity to review relevant literature	2.31	1.00
44 I feel comfortable in generating clinical questions	2.36	0.96
46 There are ample opportunities for me to apply my knowledge of <i>EBM</i> in patient care	2.46	0.95
56 For the patients that I am taking care of, I am able to find the evidence needed to apply to the patient	2.62	0.83
59 When I take care of patients, I have the opportunity to ask a focused clinical question	2.84	0.79
61 My seniors expect me to review the evidence when taking care of a patient	2.60	0.93
<i>Subscale 7: Affirmation of EBM environment</i>	2.44	0.75
6 It does not really help to review evidence when taking care of a patient*	2.95	0.95
8 Faculty/consultants do not promote <i>EBM</i> among residents/registrar*	2.30	1.13
12 The environment in which I work discourages initiative*	2.50	1.08
22 <i>EBM</i> in clinical practice is not valued in my department*	2.61	1.09
52 I seldom have the opportunity to use <i>EBM</i> when I am presenting cases*	1.99	1.10
57 My fellow residents/registrar are not keen on using <i>EBM</i> principles*	2.30	0.95

*Reversed scores for items formulated in the negative.

CFA, confirmatory factor analysis; EBMEEM-67, Evidence-Based Medicine Educational Environment Measure 67.

lack of a larger sample. No statistical explanation could be found for virtually all the negatively formulated items loading together on the seventh subscale. King³⁶ cites examples of studies with a similar tendency. It also appears as if subscale 7 may contain a few items

(eg, 6 and 57) that do not fit well into the theoretical framework proposed by Schönrock-Adema *et al.*¹⁴ Lastly, the sample was biased towards a very large representation of participants with obstetrics and gynaecology.

Table 5 Comparison of an existing theoretical framework with the subscales and items of the EBMEEM-67 instrument

Schönrock <i>et al.</i> ¹⁴	EBMEEM-67	
Domain	Subscale	Items*
1. Goal orientation (content and aims of education in relation to personal development)	6. <i>EBM</i> application opportunities	All items
	1. Knowledge and learning materials	More than half of the items (14, 26, 38, 53, 65)
2. Relationships (open and friendly atmosphere and affiliation)	3. General relationships and support	All items
3. Organisation/regulation (system maintenance and change)	2. Learner support	All itemst
	4. Institutional focus on <i>EBM</i>	All itemst
	5. Education, training and supervision	All itemst
	7. Affirmation of <i>EBM</i> environment	Most items (except 6 and 57)
	1. Knowledge and learning materials	Three items (23, 31, 40)

*Items are only listed in terms of the most prominent domain they relate to.

tSome items have an overlap with goal orientation.

EBMEEM-67, Evidence-Based Medicine Educational Environment Measure 67.

Conclusion

A tool, the EBMEEM-67, was successfully developed for soliciting the perceptions of residents on their *EBM* educational environment. This tool can be recommended for use, especially with residents in obstetrics and gynaecology. The EBMEEM-67 can be applied within one institution or department or across sites for benchmarking purposes, cross-sectionally or longitudinally. Cross-sectional investigations can be undertaken within one department across all years of study or by comparing results from different departments or from different training sites at a particular point in time. Longitudinally, the EBMEEM-67 can be used for before-and-after comparison in *EBM* education intervention studies, or for following the same cohort of residents over the different years of study. Administering the EBMEEM-67 in combination with other tools measuring the educational or learning environments, such as the PHEEM¹⁵ or D-RECT,³ could serve a useful diagnostic purpose.

The number of items remaining in the EBMEEM-67 is quite high for use in settings where time and motivation to complete an instrument of this nature are of the essence. A statistical process of reducing the number of items is currently underway. It is proposed that a thorough generalisability and decision (G&D) study^{34–35} be conducted during this follow-up. Further analyses should also be carried out to interpret the results of our study in relation to the theory behind the measurement of the education environment.¹⁴

Acknowledgements The authors wish to thank all the trial and non-trial participants and the investigators of the trial for their continued support. The editorial support by Barbara English from the office of the Deputy Dean: Research of the University of Pretoria's Faculty of Health Sciences is acknowledged with appreciation.

Funding The randomised controlled trial to evaluate the effectiveness of a clinically integrated evidence-based medicine course was funded by the UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction, Department of Reproductive Health and Research, WHO.

Competing interests A technical report on the validation process of the five definitive instruments was submitted to JAMA as part of their requirements for publishing the results of the e-learning trial. Some of the data for this study were derived from that report. JG was remunerated by the WHO and the MRC Unit for Maternal and Infant Health Care Strategies for statistical services.

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/eb-2014-101726>).

References

1. Kulier R, Khan KS, Gülmezoglu AM, *et al.* A cluster randomized controlled trial to evaluate the effectiveness of the clinically

integrated RHL evidence-based medicine course. *Reprod Health* 2010;7:8.

2. Roff S, McAleer S. What is education climate? *Med Teach* 2001;23:333–4.
3. Boor K, Van der Vleuten C, Teunissen P, *et al.* Development and analysis of D-RECT, an instrument measuring residents' learning climate. *Med Teach* 2011;33:820–7.
4. Genn JM. AMEE Medical Education Guide No. 23 (Part 1): curriculum, environment, climate, quality and change in medical education—a unifying perspective. *Med Teach* 2001;23:337–44.
5. Genn JM. AMEE Medical Education Guide No. 23 (Part 2): curriculum, environment, climate, quality and change in medical education—a unifying perspective. *Med Teach* 2001;23:345–54.
6. Soemantri D, Herrera C, Riquelme A. Measuring the educational environment in health professions studies: a systematic review. *Med Teach* 2010;32:947–52.
7. Stewart TJ. Learning environments in medical education. *Med Teach* 2006;28:387–9.
8. Holt MC, Roff S. Development and validation of the Anaesthetic Theatre Educational Environment Measure (ATEEM). *Med Teach* 2004;26:553–8.
9. Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees. *Med Teach* 2004;26:260–4.
10. Kanashiro J, McAleer S, Roff S. Assessing the educational environment in the operating room—a measure of resident perception at one Canadian institution. *Surgery* 2006;139:150–8.
11. Mulrooney A. Development of an instrument to measure the practice vocational training environment in Ireland. *Med Teach* 2005;27:338–42.
12. Roff S, McAleer S, Harden R, *et al.* Development and validation of the Dundee Ready Education Environment measure (DREEM). *Med Teach* 1997;19:295–9.
13. Roff S. The Dundee Ready Educational Environment Measure (DREEM)—a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Med Teach* 2005;27:322–5.
14. Schönrock-Adema J, Bouwkamp-Timmer T, Van Hell EA, *et al.* Key elements in assessing the educational environment: where is the theory? *Adv in Health Sci Educ* 2012;17:727–42.
15. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Med Teach* 2005;27:326–31.
16. Clapham M, Wall D, Batchelor A. Educational environment in intensive care medicine—use of Postgraduate Hospital Educational Environment Measure (PHEEM). *Med Teach* 2007;29:e184–91.
17. Aspegren K, Bastholt L, Basted KM, *et al.* Validation of the PHEEM instrument in a Danish hospital. *Med Teach* 2007;29:504–6.
18. Mallin B, Mortensen LS, Scherpbier AJJ, *et al.* Educational climate seems unrelated to leadership skills of clinical consultants responsible of postgraduate medical education in clinical departments. *BMC Med Educ* 2012;10:62.
19. Schönrock-Adema J, Heijne-Penninga M, Van Hell EA, *et al.* Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Med Teach* 2008;31:e226–32.
20. Riquelme A, Herrera C, Aranis C, *et al.* Psychometric analyses and internal consistency of the PHEEM questionnaire to measure the clinical learning environment in the clerkship of a Medical School in Chile. *Med Teach* 2009;31:e221–5.

21. Wall D, Clapham M, Riquelme A, *et al.* Is PHEEM a multi-dimensional instrument? An international perspective. *Med Teach* 2009;31:e521–7.
22. Nagraj S, Wall D, Jones E. Can STEEM be used to measure the educational environment within the operating theatre for undergraduate medical students? *Med Teach* 2006;28: 642–7.
23. Kulier R, Gülmezoglu AM, Zamora J, *et al.* Effectiveness of a clinically integrated e-learning course in evidence-based medicine for reproductive health training: a randomized controlled trial. *JAMA* 2012;308:2218–25.
24. Black TR. *Doing quantitative research in the social sciences. An integrated approach to research design, measurement and statistics.* London: Sage, 1999.
25. Hair JF, Anderson RE, Tatham RL, *et al.* *Multivariate data analysis.* 5th edn. Upper Saddle River, NJ: Prentice-Hall, Inc., 1998.
26. Steyn AGW, Smit CF, Du Toit SHC, *et al.* *Modern statistics in practice.* 5th edn. Pretoria: JL van Schaik, 1994.
27. Hatcher L. *A step-by-step approach to using SAS for factor analysis and structural equation modeling.* Cary, NC: SAS Institute, Inc, 1994.
28. Drasgow F. Polychoric and polyserial correlations. In: Kotz S, Johnson NL. eds *Encyclopedia of statistical sciences.* vol 7. New York: John Wiley & Sons, 1986:68–74.
29. SAS Institute, Inc. *SAS/STAT® 9.3 User's Guide.* Cary, NC: SAS Institute, Inc., 2011.
30. EQS 6.1 for Windows, Multivariate Software, Inc.
31. Satorra A, Bentler PM. Scaling corrections for chi-square statistics in covariance structure analysis. In: *American Statistical Association 1988 Proceedings of the Business and Economics Section.* Alexandria, VA: American Statistical Association, 1988:308–13.
32. Mardia KV. Measures of multivariate skewness and kurtosis with applications. *Biometrika* 1970;57:519–30.
33. Kline RB. *Principles and practice of structural equation modeling.* 2nd edn. New York: The Guilford Press, 2005:143–4, 321–2.
34. Shavelson RJ, Webb NM. *Generalizability theory: a primer.* Thousand Oaks, CA: Sage, 1991.
35. Bloch R, Norman G. Generalizability theory for the perplexed: a practical introduction and guide: AMEE Guide No. 68. *Med Teach* 2012;34:960–92.
36. King CV. Factor analysis and negatively worded items. *Populus.* <http://csi.ufs.ac.za/resres/files/King.pdf> (accessed 12 Feb 2014).