



EDUCATION FOR HEALTH

ORIGINAL RESEARCH PAPER

Validation of the Greek Translation of the Dundee Ready Education Environment Measure (DREEM)

IDK Dimoliatis¹, E Vasilaki¹, P Anastassopoulos², JPA Ioannidis^{1,3}, S Roff⁴

¹*Department of Hygiene & Epidemiology, University of Ioannina School of Medicine,
Greece*

²*Fetal Medicine, Harris Birthright Research Centre, King's College Hospital, London,
UK*

³*Department of Medicine, Tufts University School of Medicine, Boston, USA*

⁴*Centre for Medical Education, Dundee University Medical School, Dundee, UK*

Published: April 2010

Dimoliatis IDK, Vasilaki E, Anastassopoulos P, Ioannidis JPA, Roff S

Validation of the Greek Translation of the Dundee Ready Education Environment Measure (DREEM)
Education for Health, Volume 23, issue 1, 2010

Available from: <http://www.educationforhealth.net/>

A B S T R A C T

Context: The educational environment makes an important contribution to student learning. The DREEM questionnaire is a validated tool assessing the environment.

Objectives: To translate and validate the DREEM into Greek.

Methods: Forward translations from English were produced by three independent Greek translators and then back translations by five independent bilingual translators. The Greek DREEM.v0 that was produced was administered to 831 undergraduate students from six Greek medical schools. Cronbach's alpha and test-retest correlation were used to evaluate reliability and factor analysis was used to assess validity. Questions that increased alpha if deleted and/or sorted unexpectedly in factor analysis were further checked through two focus groups.

Findings: Questionnaires were returned by 487 respondents (59%), who were representative of all surveyed students by gender but not by year of study or medical school. The instrument's overall alpha was 0.90, and for the learning, teachers, academic, atmosphere and social subscales the alphas were 0.79 (expected 0.69), 0.78 (0.67), 0.69 (0.60), 0.68 (0.69), 0.48 (0.57),



respectively. In a subset of the whole sample, test and retest alphas were both 0.90, and mean item scores highly correlated ($p < 0.001$). Factor analysis produced meaningful subscales but not always matching the original ones. Focus group evaluation revealed possible misunderstanding for questions 17, 25, 29 and 38, which were revised in the DREEM.Gr.v1. The group mean overall scale score was 107.7 (SD 20.2), with significant differences across medical schools ($p < 0.001$).

Conclusion: Alphas and test-retest correlation suggest the Greek translated and validated DREEM scale is a reliable tool for assessing the medical education environment and for informing policy. Factor analysis and focus group input suggest it is a valid tool. Reasonable school differences suggest the instrument's sensitivity.

Keywords: Educational environment/climate, undergraduate students/education, Greece, DREEM questionnaire/inventory, translation, validation, reliability, validity, sensitivity, responsiveness.

Introduction

Aside from the formal curriculum, students and teachers are well aware of the educational 'environment' or 'climate' of their institution. Is it competitive? Authoritarian? Relaxed? Stressful? Does it vary among courses within the curriculum? Does it motivate? Demotivate? All students respond to these elements in their learning environment¹⁻⁴.

Roff et al.⁵ developed the Dundee Ready Education Environment Measure (DREEM), an international, culturally non-specific, generic instrument that provides global readings and diagnostic analyses of undergraduate educational environments within health professions institutions⁶. It generates a profile of a particular institution's environmental strengths and weaknesses. The measure can be used by faculty and evaluators to compare students' perceptions of the education environments within and between institutions or student cohorts. It can be used to assess relationships between environments and students' academic achievements and can serve as a predictive tool for identifying achievers and underachievers⁷⁻¹¹.

This valuable tool was originally developed in English and has been translated into about twenty languages, but not yet into Greek. The aim of this study was to translate the DREEM into the Greek language and to validate the translated instrument.

Methods

The DREEM Questionnaire

The DREEM inventory (questionnaire, scale, instrument, tool) consists of 41 positively worded statements (items, questions) each scored 0 to 4, and nine negative ones scored in reverse (4 to 0). It generates an overall score and five subscale scores regarding students' perceptions of learning, perceptions of teachers, academic self-perceptions, perceptions of the atmosphere, and social self-perceptions. On all measures (items, subscales, overall score), high scores indicate a good environment - the higher the better^{5,6}. Items and subscales can be seen in Table 1.



Table 1: The original DREEM questionnaire, and the overall, subscale and question mean scores obtained by DREEM.Gr.v0 and their interpretation.

QUESTION / SUBSCALE #	MEAN SCORE (SD) [†]	INTERPRETATION [‡]
OVERALL	107.7 (20.2)	P>N
25. <i>The teaching over-emphasizes factual learning.</i>	1.5 (0.9)	PA-EMC
13. The teaching is student centred.	1.7 (0.9)	PA-EMC
44. The teaching encourages me to be an active learner.	1.8 (1.0)	PA-EMC
22. The teaching is sufficiently concerned to develop my confidence.	1.8 (1.0)	PA-EMC
48. <i>The teaching is to teacher-centred.</i>	1.9 (0.9)	PA-EMC
24. The teaching time is put to good use.	1.9 (1.0)	PA-EMC
47. Long term learning emphasizes over short term.	1.9 (1.0)	PA-EMC
20. The teaching is well focused.	1.9 (0.8)	PA-EMC
1. I am encouraged to participate in class.	2.0 (1.0)	PA-EMC
38. I am clear about the learning objectives of the course.	2.1 (0.9)	CBE
16. The teaching is sufficiently concerned to develop my competence.	2.1 (0.9)	CBE
7. The teaching is often simulating.	2.4 (0.9)	CBE
Subscale I: students' perceptions of learning	22.8 (6.2)	TVN
32. The teachers provide constructive criticism here.	1.7 (0.9)	PA-EMC
29. The teachers are good at providing feedback to students.	1.9 (0.8)	PA-EMC
8. <i>The teachers ridicule the students.</i>	2.3 (0.9)	CBE
37. The teachers give clear examples.	2.4 (0.8)	CBE
6. The teachers are patient with patients.	2.4 (0.8)	CBE
18. The teachers have good communications skills with patients.	2.4 (0.7)	CBE
39. <i>The teachers get angry in class.</i>	2.4 (0.9)	CBE
50. <i>The students irritate the teachers.</i>	2.4 (1.0)	CBE
40. The teachers are well prepared for their classes.	2.5 (0.9)	CBE
9. <i>The teachers are authoritarian.</i>	2.6 (0.9)	CBE
2. The teachers are knowledgeable.	2.8 (0.7)	CBE
Subscale II: Students' perceptions of teachers	25.4 (5.4)	MRD
27. I am able to memorize all I need.	1.7 (1.0)	PA-EMC
41. My problem solving skills are being well developed here.	1.8 (0.9)	PA-EMC
21. I feel I am being well prepared for my profession.	2.0 (0.9)	PA-EMC
5. Learning strategies which worked for me before continue to work for me now.	2.1 (1.1)	CBE
10. I am confident about my passing this year.	2.2 (0.9)	CBE
31. I have learned a lot about empathy in my profession.	2.3 (1.1)	CBE
26. Last year's work has been a good preparation for this year's work.	2.3 (1.0)	CBE
45. Much of what I have to learn seems relevant to a career in healthcare.	2.5 (0.9)	CBE
Subscale III: Students' academic self-perceptions	16.7 (4.5)	FPS
42. The enjoyment outweighs the stress of the course.	1.5 (1.0)	PA-EMC
17. <i>Cheating is a problem in this school</i>	1.8 (1.2)	PA-EMC
12. This school is well time-tabled.	1.8 (1.2)	PA-EMC
43. The atmosphere motivates me as a learner.	1.8 (1.0)	PA-EMC
49. I feel able to ask the questions I want.	2.3 (1.1)	CBE
30. There are opportunities for me to develop interpersonal skills.	2.4 (0.9)	CBE
36. I am able to concentrate well.	2.4 (1.0)	CBE
11. The atmosphere is relaxed during the ward teaching.	2.5 (0.9)	CBE
23. The atmosphere is relaxed during lectures.	2.7 (0.9)	CBE
35. <i>I find the experience disappointing.</i>	2.7 (1.0)	CBE
34. The atmosphere is relaxed during seminars/tutorials.	2.8 (0.8)	CBE
33. I feel comfortable in class socially.	2.9 (0.9)	CBE



Table 1: cont'd

QUESTION / SUBSCALE [#]	MEAN SCORE (SD) [†]	INTERPRETATION [‡]
Subscale IV: Students' perceptions of atmosphere	27.3 (5.8)	MPA
3. There is a good support system for students who get stressed.	1.0 (0.9)	PA-EMC
<i>14. I am rarely bored on this course.</i>	1.1 (0.9)	PA-EMC
4. I am too tired to enjoy the course.	1.6 (1.1)	PA-EMC
28. I seldom feel lonely.	2.4 (1.1)	CBE
19. My social life is good.	3.0 (0.8)	CBE
15. I have good friends in this school.	3.1 (0.9)	?
46. My accommodation is pleasant.	3.4 (0.8)	?
Subscale V: Students' social self-perceptions	15.6 (3.3)	NTB

[#] Items in italics are the negative statements.

[†] WARNING: This is a validation study of the Greek translation of DREEM and not a study of the educational environment in Greek medical schools. The sample does not represent the Greek medical students. The table is given to help readers on which are the items and the subscales. Though given scores can serve as a good starting point (Dimoliatis, 2009), they should be interpreted with caution; especially questions 17, 25, 29 and 38 with great caution (see text).

[‡] According to McAleer & Roff (2002): P>N = more positive than negative (101-150), TVN = teaching is viewed negatively (13-24), MRD = moving in the right direction (23-33), FPS = feeling more on the positive side (17-24), MPA = a more positive atmosphere (25-36), NTB = not too bad (15-21), PA-EMC = problem areas, examine more closely (0-2), CBE = aspects of the climate that could be enhanced (between 2 and 3), ? = not clear, since "a mean score of 3.5 and over are real positive points".

Translation

Three researchers (ID, EV, PA) independently translated the original English version of the DREEM into Greek. Via face-to-face and email interactive sessions, they discussed and resolved the differences in their translations and reached consensus on an initial best wording. This draft was piloted with 50 third-year Ioannina University medical students, and their comments were incorporated into an improved version which was piloted with 14 more volunteers from the same school. Pilot testing showed that students had some difficulties grasping the meaning of two questions (17, 29) and because almost all Greek medical students are competent in English, the English terms "cheating" and "feedback" within these items were kept in the Greek translation, in parentheses, to help in students' understanding. Advice was also sought from eight interested faculty members. The open-ended question on the instrument, "Comments", was replaced by the more specific "If you could change three things about the medical school, what would they be?" after Whittle et al.¹¹.

The product was back-translated by five independent bilingual translators who were unaware of the original English version. All back translators got back to the original meaning and for many statements to the exact original wording. Two of us (ID, EV) refereed wording differences, and the DREEM.Gr.v0 was thus produced.

Validation

We conducted a survey among Greek medical students approved by the Body of Directors of the Ioannina University Medical School. With this data, we tested reliability, validity, sensitivity and responsiveness of the translated instrument.

Participants

The DREEM.Gr.v0 was transformed to an anonymous, scannable form and distributed by interested faculty members to a convenience sample of students within six of the seven Greek medical schools during November and December 2007. The



questionnaire was retested under the same conditions 3.5 weeks later with the same Ioannina students, who were asked to respond without trying to recall their previous responses. The first two years of medical school in Greece are preclinical, the third is transitional, and the last three are clinical.

All completed forms were scanned by a highly reliable optical mark recognition scanner (OpScan iNSIGHT™, Pearson NCS). Using the QuickTesting software, produced by Anova Consulting (www.anova.gr), an electronic data file was obtained, which was checked against the original completed questionnaires. The nine negative questions were reverse-coded prior to any calculations.

Reliability

Cronbach's alpha, alpha if item deleted, and test-retest reliability metrics were calculated. Since alpha depends on both the length of the scale (the number of questions) and the correlation of the items within the scale (actual reliability), the Spearman-Browne formula,

$$\alpha_{\text{subscale}} = k\alpha_{\text{scale}} / (1 + (k-1)\alpha_{\text{scale}})$$

where k is the number of items of the subscale divided by the number of items of the overall scale^{12,13}, was used to estimate expected subscale alphas. For good reliability, the observed scale alpha should be greater than 0.70¹³ if not greater than 0.80¹⁴, and observed subscale alphas should be greater than expected.

Responses from tested and retested Ioannina students (both anonymous) were compared by checking Cronbach's alphas and item mean score correlation.

Validity

Content validity was addressed by the original DREEM. Not willing to limit our ability to compare results internationally, we did not delete or add any items, change items' original (randomly arranged) order, or rearrange subscales. We followed standard instrument translation/validation methodology^{13,14}.

We checked whether the original five subscales fit our data (construct validity) using confirmatory factor analysis (CFA) under the same conditions as used in the instrument's original development, i.e., five factors and requiring loadings of 0.3 or greater⁵. To see whether different factor clustering operated in our data, we also performed exploratory factor analysis (EFA) under the conditions of eigenvalues > 1 and loadings ≥ 0.3.

Questions whose deletion increased the overall alpha, and those that loaded less than 0.3 on the expected factor, loaded on two factors or were grouped within an unexpected factor were thoroughly examined in a preliminary focus group (one psychologist, one PhD psychiatry student, three fifth-year and six fourth-year Ioannina medical students) and in a second focus group of 48 third-year Ioannina medical students.

Sensitivity and Responsiveness



We checked differences among groups (genders, years of study, and schools) to test the instrument's ability to detect differences if they really exist (sensitivity). The instrument's responsiveness, reflecting changes within a group over time, was beyond the scope of this study and was not tested.

Statistics & Software

Non-parametric tests were used to compare means and to correlate tested and retested mean item scores. To be consistent with most published relevant work on the DREEM and related instruments, we also considered their parametric equivalents¹⁴⁻¹⁶. Reported p-values for differences according to gender, year of study and medical school were considered significant at $p < 0.05/3 = 0.017$ adjusting for multiple (3) comparisons. Analyses were performed using Microsoft Excel 2003 and SPSS 11.0.

Results

Participants

Eight hundred and thirty-one questionnaires were distributed and 487 students responded. Respondents were representative of Greek medical students in terms of gender but not according to school or year of study. Overall response was 58.6% (Athens 10/284; Crete 43/75; Ioannina 102/102; Thessaly 112/150; Thessaloniki 127/127; Thrace 93/93). Participants' year of study was preclinical 12%, transitional 62%, clinical 26%.

Data

There were 79 invalid scanner readings (0.3% of all possible 24350 readings) within 61 questionnaires (13%). Checking against the original completed questionnaires, we found that invalid readings were caused when participants had chosen two options and either deleted the one by an X (73) or left them both undeleted (6). We therefore corrected the electronic data by either keeping the undeleted option when that choice was clear or by deleting both responses and treating them as missing values when the choice was not clear.

There were 279 missing values (1%) in 129 questionnaires (27%); thus the overall alpha calculations were based on the remaining 358 questionnaires. All questions had missing values, but questions 29, 18, 6 and 11 had 2.7, 4.3, 4.8 and 5.7 times more than the average of the rest of the 46 questions, reflecting either difficulties in capturing the feedback concept (q29) or in participants not having clinical experience yet (q6, q11, q18). The 24071 non-missing values ranged from 'strongly agree' 2016 (8%), 'agree' 8373 (35%), 'uncertain' 7104 (30%), 'disagree' 5077 (21%), to 'strongly disagree' 1501 (6%).

Reliability

Cronbach's alpha (Table 2) was greater than 0.70 when the entire inventory was considered and greater than expected when subscales were considered, except for the social subscale and marginally for the atmosphere subscale. If items 17 and 25 (and marginally 19 and 50) were deleted, overall alpha was increased.

Alphas among tested and retested students were identical (0.90) and both greater than 0.70 (Table 2). The correlation between tested and retested mean item scores was very high (Kendall's tau-b 0.862, $p < 0.001$; Figure 1).



Table 2: Observed and (in parentheses) minimum expected Cronbach's alpha for different scales and subscales.

Scale	Items	N	Cases	Alpha
OVERALL DREEM	50	487	358	0.90
SUBSCALE				
Learning	12	487	454	0.79 (0.69)
Teachers	11	487	432	0.78 (0.67)
Academic	8	487	456	0.69 (0.60)
Atmosphere	12	487	436	0.68 (0.69)
Social	7	487	463	0.48 (0.57)
GENDER				
Male	50	194	148	0.89
Female	50	275	201	0.91
Not declared	50	18	9	0.92
YEAR OF STUDY				
First and second	50	60	38	0.89
Third	50	302	221	0.90
Fourth and over	50	119	96	0.91
Not declared	50	6	3	0.99
SCHOOL				
Athens	50	10	6	0.94
Crete	50	43	27	0.94
Ioannina	50	102	74	0.90
Ioannina retest	50	105	87	0.90
Thessaly	50	112	92	0.91
Thessaloniki	50	127	85	0.88
Thrace	50	93	74	0.89

Items = the number of questions per scale or subscale;

N = the number of questionnaires (participants).

Cases = the number of questionnaires without *any* missing value; only these are used in alpha calculations.

In bold the less than expected observed alphas.

Subscale expected alphas were estimated using the Spearman-Browne formula (see text).

Observed overall alpha (0.9035) was less than 'alpha if item deleted' for items 17 and 25 and marginally less for items 19 and 50 (0.9069, 0.9082, 0.9038, 0.9037 respectively).

When question 25 coding was reversed (see text), observed overall alpha increased to 0.9079 and was less than 'alpha if item deleted' for item 17 (0.9110) and marginally less for items 19, 25, 28 and 50 (0.9083, 0.9082, 0.9080, 0.9081 respectively)

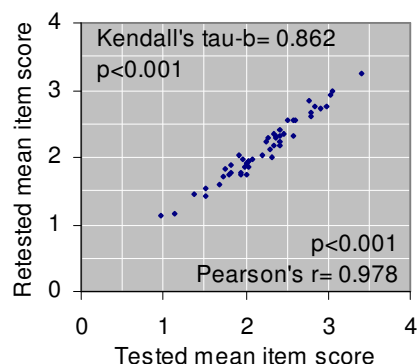


Figure 1: Correlation between tested and retested mean item scores of Ioannina medical students.

Validity

In the confirmatory factor analysis (CFA), eight items did not reach the target factor loading of 0.3 (Tables 3 and 4). The remaining 42 items loaded onto factors *f1* to *f5* somewhat differently than the original DREEM analysis, while six of those items loaded onto two factors (loadings ≥ 0.3). However, the new factors can be interpreted in a meaningful way: *f1* describes the qualities of a good teacher, with the exception of two negative questions (35,48) that might fit better with *f3*; *f2* describes a relaxed atmosphere; *f3* addresses the negative aspects of the climate; *f4* addresses students' learning and coping strategies; and *f5* addresses students' social life.

In the exploratory factor analysis (EFA), six items, mostly the same as in CFA, had loadings less than 0.3 (range 0.19 to 0.29). The remaining 44 spread out in 12 factors. Seven items loaded onto two factors. The new arrangement of items also produced meaningful underlying factors. It seems that *f1* has been divided into F1 (good teachers), F2 (motivating teaching, with the exception of the negative question 35), F8 (developing teaching), and F9 (encouraging and supporting). There are two contradictions with two negative questions: q35 (disappointing experience) fitted into two opposite factors, one about motivating (F2; wrongly) and the other about demotivating (F3; correctly); and q25 (factual learning, a negative item) has been sorted onto the same factor (F12) with q45 (relevant content, a positive item).

In the focus groups, the vast majority of students misunderstood q29 and q38, almost none translated 'cheating' (q17) differently than 'αντιγραφή' (a Greek slang term for copying during the exams), and three in four perceived positively the negative q25. Recoding q25 as positive and rerunning EFA under exactly the same conditions as in Table 4, q25 loaded 0.23 onto F1 'good teachers', suggesting that the whole sample might have understood its concept in reverse.

Sensitivity

The mean overall score for all students was 107.7 (SD 20.2), with no differences according to gender ($p=0.86$) or year of study ($p=0.21$). Mean overall scores did differ across schools ($p<0.001$) even when excluding Athens with so few students: Athens 80.8 (29.8), Crete 105.8 (22.1), Ioannina 109.4 (18.7), Thessaly 108.3 (19.5), Thessaloniki 102.5 (18.8), Thrace 116.1 (17.8). Similar differences among schools were also found in all subscales ($p\leq 0.006$), except for the subscale addressing the social aspects of the environment ($p=0.308$).



Table 3: The original five factors (subscales) and the Confirmatory and Exploratory Factor Analysis of our data.

F	N	ITEMS [†]	INTERPRETATION (students' perceptions of)
Original English DREEM (Roff et al 1997)			
I	12	1, 7, 13, 16, 20, 22, 24, 25, 38, 44, 47, 48	Learning
II	11	2, 6, 8, 9, 18, 29, 32, 37, 39, 40, 50	Teachers
III	8	5, 10, 21, 26, 27, 31, 41, 45	Academic
IV	12	11, 12, 17, 23, 30, 33, 34, 35, 36, 42, 43, 49	Atmosphere
V	7	3, 4, 14, 15, 19, 28, 46	Social
Greek DREEM: Confirmatory Factor Analysis (CFA)[‡]			
f1	27	44, 43, 22, 16, 7, 32, 21, 41 , 29, 38 , 1, 47, 35, 20, 26, 13, 24 , 48, 18, 42, 37, 14, 2, 3, 49, 6, 40	Good teachers ^
f2	5	34, 23, 11, 24, 40	Relaxed atmosphere
f3	5	8, 9, 39, 48 , 35	Negative climate
f4	6	10, 36, 5, 27, 41, 38	Learning / coping strategies
f5	5	19, 28, 33, 15, 30	Social life
f?	(8)	(4), (12), (17), (25), (31), (45), (46), (50)	
Greek DREEM: Exploratory Factor Analysis (EFA)[#]			
F1	8+(1)	2, 37, 40, 29, 49 , 18, 26, 20, (21)	Good teachers
F2	7	44, 43, 7, 22 , 32, 35, 16	Motivating teaching ^^
F3	6+(1)	9, 8, 39, 48, 35 , 49, (50)	Negative (demotivating) climate
F4	7+(1)	10, 27, 36, 14, 42, 41, 5, (38)	Learning / coping strategies
F5	6	19, 28, 33, 15, 30, 46	Social life
F6	2	34, 23	Relaxed teaching
F7	5	12, 5 , 13, 20 , 24	Well organised school ^^
F8	2	16 , 22	Developing teaching
F9	2	3, 1	Encouraging/ supporting
F10	3	11, 18 , 6	Ward (patient) teaching
F11	2	31, 47	Deep learning
F12	1+(1)	45, (25) ^{##}	Relevant teaching / surface learning
F13	(1)	(17)	Cheating
F14	(1)	(4)	Tiredness (fight fatigue)

F = f = factor. N = the number of items per factor. f? = items with loadings less than .3, which were not included in any of the f1 to f5.

[†] Items per factor in descending order by factor loading (except for the original inventory sorted in ascending order by item identification number because loadings are not available), indicating that items towards the beginning of the list are stronger related to corresponding factor than items towards the end; items with loadings <.3 and the weaker of those loaded in two factors tend to the right of the list. *Italics* indicate the negative questions. Items loaded in two factors, i.e. with loadings ≥.3 in both, are presented twice, in bold and plain (higher and lower loadings, respectively). Items in parentheses had loadings less than the criterion of .3 (between .194 and .290).

[‡] Rotated factor matrix, factors 5 (the number of the original subscales), suppress .3 (as the original inventory). Extraction method: Maximum Likelihood. Rotation method: Varimax with Kaiser Normalization. Rotation converged in 29 iterations.

[#] Rotated factor matrix, factors 14 (all those with eigenvalues greater than 1), suppress .3. Extraction method: Maximum Likelihood. Rotation method: Varimax with Kaiser Normalization. Rotation converged in 29 iterations. Items with loadings less than .3 have been put to show which factor they might belong, otherwise F13 and F14 would be empty.

^{##} When question 25 coding was reversed (see text) and EFA was reran, q25 grouped to F1 (with a loading .23<.30), while all others remained unchanged.

^ with only two exceptions, the (negative) questions 48 and 35 which fit better with f3.

^^ with the exception of 35 that fits better with F3 (where, furthermore, it had higher loading)

^^^ with only one exception, the question 5 that fits better with F4.



Table 4: The original five factors and the Confirmatory and Exploratory Factor Matrix of our data set.

ITEM	E	CONFIRMATORY FA ‡					EXPLORATORY FACTOR ANALYSIS (EFA) #													
		f1	f2	f3	f4	f5	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
1	I	.44													.51					
2	II	.36					.60													
3	V	.35													.55					
4	V																			(.28)
5	III				.44				.33			.37								
6	II	.34													.41					
7	I	.60					.41													
8	II			.69				.64												
9	II			.68				.67												
10	III				.45				.53											
11	IV		.36												.63					
12	IV											.46								
13	I	.39										.35								
14	V	.37							.43											
15	V				.45			.51												
16	I	.62					.30						.80							
17	IV																		(.19)	
18	II	.38					.32								.43					
19	V				.65			.63												
20	I	.42					.31					.33								
21	III	.52					.28													
22	I	.64					.39						.35							
23	IV		.72									.73								
24	I	.39	.36									.32								
25	I																	(.25)		
26	III	.40					.32													
27	III				.39				.52											
28	V				.57					.61										
29	II	.47					.45													
30	IV				.37				.42											
31	III															.44				
32	II	.52					.33													
33	IV				.53				.58											
34	IV		.77								.77									
35	IV	.43		.33			.30	.35												
36	IV				.45				.52											
37	II	.37					.56													
38	I	.46			.31				(.29)											
39	II			.49				.61												
40	II	.33	.31				.55													
41	III	.51			.32				.38											
42	IV	.38							.39											
43	IV	.70					.64													
44	I	.71					.81													
45	III																.49			
46	V								.36											
47	I	.44														.34				
48	I	.38		.41				.44												
49	IV	.34					.37	.32												
50	II							(.29)												

E = original English factors (subscales I to V). FA = factor analysis. Items in italics and rose highlighted in the first column are the negative questions. In bold the higher of the two loadings if both were >.3. In parentheses loadings < .3.

‡ Rotated factor matrix, factors 5, suppress .3. Extraction method: Maximum Likelihood. Rotation method: Varimax with Kaiser Normalization. Rotation converged in 29 iterations.



Rotated factor matrix, factors 14 (all those with eigenvalues greater than 1), suppress .3 (items in brackets had loadings less than .3, ranged from .194 to .290). Extraction method: Maximum Likelihood. Rotation method: Varimax with Kaiser Normalization. Rotation converged in 29 iterations.

Discussion

This study's goal was to translate and validate the DREEM in the Greek language. We will discuss four interrelated but also independently important features of the scales from the translation: reliability, validity, sensitivity and responsiveness. We will also discuss the translated instrument's limitations.

Reliability

All overall alphas were much higher than the 0.70¹³ or 0.80¹⁴ thresholds generally considered acceptable for scales and also similar to published studies of the DREEM translation in other languages (Primparyon et al.¹⁷: 0.91; Mayya & Roff¹⁸ 0.92; de Oliveira Filho¹⁹: 0.93; Riquelme et al.²⁰: 0.91). The subscale alphas were higher than expected except for the social (and perhaps the atmosphere).

In our test-retest exercise we found the same alpha and an extremely high correlation between tested and retested mean question scores. We have not seen a previous test-retest assessment of the DREEM in the published literature.

Thus, we can conclude that our translation produced a reliable questionnaire.

Validity

Content validity, a matter of the original DREEM, and face validity are both optimized by involving a wide range of individuals in scale development¹³. The various types and numbers of people involved in the translation, back-translation, consulting, piloting and focus groups indicate that the original content has been successfully transformed into the destination language. That all five alternative answers in all questions were used supports this conclusion²¹. More missing values in 'clinical' questions 6, 11 and 18 were expected (participants were mostly preclinical students), but we did not expect this in q29 about feedback.

Factor analysis (FA) plays a major role in *construct* validation¹³. In general, both confirmatory and exploratory factor analysis produced sensible subscales; however, they did not quite match the original English ones. One reason might be that the originals were largely arrived at by consensus of a qualitative group rather than by statistical methods⁵ as in our case. A second reason might be that participants failed to realize that negative questions should be chosen in reverse because the scanable forms provided empty boxes without a reminder whether each was an 'agree' or a 'disagree' box (a *face* validity issue, corrected in DREEM.Gr.v1). But the most likely reason seems to be the ineffectively translated items spotted by factor analysis and/or 'alpha-if-item-deleted' reliability analysis. Although their translation was appropriate and back-translators had no difficulty, they failed to adjust for the English-Greek cultural differences. Our solution was to give definitions in paraphrases instead of simply translating terms. For example, q38 in DREEM.Gr.v0 read "The learning objectives of the courses are clear to me" and in DREEM.Gr.v1 it became "At the beginning of the course, the teachers clarify what new things I should know or what I should be able to do at the end of the course."²². Finally, though the original 'cheating' might convey a spectrum of cheats, no student in focus groups thought q17 might



be something different than ‘αντιγραφή’ and we adopted it. These solutions made keeping English terms in parentheses unnecessary.

There is no other ‘gold standard’ or well-established instrument in Greek and this is not a prognostic study; thus, *concurrent* and *prognostic validity* remain unchecked.

Sensitivity and Responsiveness

In addition to reliability usually being a prerequisite for sensitivity¹³, the differences found among schools might be a good index of the translated instrument’s ability to detect real differences among groups (sensitivity). It seems reasonable that these differences really do exist. The fact that there were no differences between schools for the social subscale reinforces rather than vitiates this claim; it seems reasonable that the social life is equally good in all cities. The same applies to understanding the lack of differences among students of different genders or year of study: it seems reasonable that all students perceive the educational environment of their schools comparably, whether they are male or female, or in the preclinical, transitional or clinical stage of training. Alternatively, the scale might have missed real differences between gender and year of training: we cannot know that from this study’s data.

Testing the ability of our tool to detect changes over time within a group was beyond the scope of this study. Thus we have no indication of its responsiveness other than its sensitivity: a highly sensitive scale will usually also be highly responsive¹³.

Losing information

Almost one in three answers (30%) were neither in the ‘agree’ nor in the ‘disagree’ side, but in the middle ‘uncertain’, making their interpretation almost impossible. What does ‘uncertain’ really mean? Either, the participant has no experience, is unable to understand what is being asked, and thus prefers a ‘not applicable’ option. These questions include 6, 11, 18 for preclinical students. Or the student does have the experience, does understand what is asked for, but either has not decided yet (the really ‘uncertain’ option) or has decided to stand exactly in the middle (‘not agree nor disagree’ option).

The solution adopted in DREEM.Gr.v1 was to split the ‘uncertain’ option into ‘rather agree’ and ‘rather disagree’ options (slightly agree/disagree), and to prompt students to answer all questions except when they really have no knowledge or experience. They could then leave the question unanswered. While offering more options to participants, the solution is closer to ‘five to nine’ options, an ideal congruent with both respondent preferences and reliability statistics¹⁴. This also prevents the ‘central tendency bias’ that results in loss of reliability and sensitivity¹⁴.

Conclusion

After this validation study successfully debugged the DREEM.Gr.v0, the final DREEM.Gr.v1 has been created (Table 5). A reliable, valid, sensitive and probably responsive Greek version of the DREEM is now available for prospectively evaluating and monitoring the medical educational environment. Although no one can prove scale validity in any *absolute* sense,^{13,16} we expect that its future comprehensive application will help educators and evaluators obtain a good picture of the medical educational environment to inform policies and interventions.



Table 5: The final version of the Greek DREEM after this validation study (DREEM.Gr.v1)

Μέτρηση Εκπαιδευτικού Περιβάλλοντος με το Εργαλείο DREEM
Dundee Ready Education Environment Measure

Φύλο Α Θ Ηλικία [____] Εξάμηνο Σπουδών [____] Σχολή

Παρακαλούμε, σημειώστε αν συμφωνείτε απόλυτα (ΣΑ), συμφωνείτε (Σ), μάλλον συμφωνείτε (ΜΣ), μάλλον διαφωνείτε (ΜΔ), διαφωνείτε (Δ), ή διαφωνείτε απόλυτα (ΔΑ) με καθεμιά από τις παρακάτω προτάσεις. Πρόκειται για το πώς νιώθετε το όλο εκπαιδευτικό περιβάλλον της Σχολής σας ΕΣΕΙΣ προσωπικά.

- Απαντήστε **και** τις 50 ερωτήσεις (γυρίστε και στην **πίσω** σελίδα).
- Κυκλώστε **μία** μόνον επιλογή, αυτήν που τώρα σας εκφράζει περισσότερο.
- **Μόνον** αν σας είναι αδύνατο να τοποθετηθείτε, αφήστε την ερώτηση αναπάντητη.

1	Οι καθηγητές με ενθαρρύνουν να συμμετέχω στο μάθημα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
2	Οι καθηγητές γνωρίζουν καλά το αντικείμενό τους.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
3	Υπάρχει ένα καλό σύστημα υποστήριξης των φοιτητών που αγχώνονται.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
4	Είμαι υπερβολικά κουρασμένος/η για να μπορώ ν' απολαύσω το μάθημα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
5	Οι τεχνικές μάθησης που μου απέδιδαν στο παρελθόν συνεχίζουν να μου αποδίδουν και τώρα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
6	Οι καθηγητές έχουν υπομονή με τους ασθενείς.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
7	Η διδασκαλία συχνά μου εξάπτει το ενδιαφέρον.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
8	Οι καθηγητές προσβάλλουν τους φοιτητές.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
9	Οι καθηγητές είναι αυταρχικοί.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
10	Έχω την πεποίθηση ότι αυτή τη χρονιά θα περάσω τα μαθήματά μου.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
11	Επικρατεί ήρεμη ατμόσφαιρα κατά τη διδασκαλία στους θαλάμους των κλινικών.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
12	Αυτή η σχολή έχει καλό ωρολόγιο πρόγραμμα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
13	Η διδασκαλία είναι επικεντρωμένη στο φοιτητή.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
14	Σπάνια βαριέμαι στο μάθημα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
15	Έχω καλούς φίλους σ' αυτή τη σχολή.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
16	Η όλη εκπαίδευση με βοηθά ν' αναπτύξω τις ικανότητές μου.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
17	Η αντιγραφή είναι πρόβλημα σ' αυτή τη σχολή.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
18	Οι καθηγητές έχουν την ικανότητα να επικοινωνούν καλά με τους ασθενείς.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
19	Η κοινωνική μου ζωή είναι καλή.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
20	Η διδασκαλία είναι καλά εστιασμένη.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
21	Έχω την αίσθηση ότι με ετοιμάζουν καλά για το επάγγελμά μου.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
22	Η εκπαίδευση εδώ με βοηθά ν' αναπτύξω την αυτοπεποίθησή μου.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
23	Επικρατεί ήρεμη ατμόσφαιρα στις παραδόσεις.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
24	Γίνεται καλή χρήση του διδακτικού χρόνου.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
25	Η διδασκαλία δίνει περισσότερη έμφαση στην απομνημόνευση δεδομένων παρά στη μάθηση γενικών αρχών, εννοιών, μηχανισμών και μεθόδων.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
26	Η δουλειά της περασμένης χρονιάς είναι μια καλή προετοιμασία για τη δουλειά αυτής της χρονιάς.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
27	Μπορώ ν' απομνημονεύω όλα όσα χρειάζομαι.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
28	Σπάνια νιώθω μοναξιά.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
29	Οι καθηγητές είναι καλοί στο να μας πληροφορούν συστηματικά για την καθημερινή μας πρόοδο.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
30	Υπάρχουν ευκαιρίες για μένα ν' αναπτύξω δεξιότητες για διαπροσωπικές σχέσεις.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
31	Έχω μάθει καλά πόσο σημαντικό είναι στο επάγγελμά μου το "να μπαίνω στη θέση του άλλου".	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
32	Οι καθηγητές εδώ παρέχουν εποικοδομητική κριτική.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
33	Νιώθω κοινωνικά άνετα στην τάξη.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
34	Επικρατεί ήρεμη ατμόσφαιρα κατά τη διάρκεια των εργαστηρίων/ σεμιναρίων/ φροντιστηρίων/ ασκήσεων.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
35	Βρίσκω την όλη εμπειρία των σπουδών μου απογοητευτική.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
36	Μπορώ να συγκεντρωθώ καλά.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
37	Οι καθηγητές δίνουν σαφή και εύστοχα παραδείγματα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ



Table 5: cont'd

38	Οι καθηγητές ξεκαθαρίζουν από την αρχή του μαθήματος τι καινούργιο θα πρέπει να ξέρω ή να μπορώ να κάνω στο τέλος του.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
39	Οι καθηγητές θυμώνουν στο μάθημα.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
40	Οι καθηγητές είναι καλά προετοιμασμένοι για το μάθημά τους.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
41	Η ικανότητά μου να λύνω προβλήματα αναπτύσσεται καλά εδώ στη σχολή.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
42	Η ευχαρίστησή μου από τα μαθήματα ξεπερνά το στρες από αυτά.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
43	Η ατμόσφαιρα με παρακινεί να μαθαίνω.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
44	Η εκπαιδευτική διαδικασία με ενθαρρύνει να είμαι ένας ενεργός μαθητής.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
45	Πολλά από αυτά που έχω να μάθω μού φαίνονται σχετικά με μια σταδιοδρομία στον τομέα υγείας.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
46	Το σπίτι που μένω κατά τη διάρκεια των σπουδών μου είναι ευχάριστο.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
47	Οι καθηγητές δίνουν περισσότερη έμφαση στη μακροπρόθεσμη παρά στη βραχυπρόθεσμη μάθηση.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
48	Η διδασκαλία παραείναι δασκαλοκεντρική.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
49	Νιώθω ότι μου δίνεται η δυνατότητα να ρωτήσω όποια απορία έχω.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ
50	Οι φοιτητές εκνευρίζουν τους καθηγητές.	ΣΑ	Σ	ΜΣ	ΜΔ	Δ	ΔΑ

Αν μπορούσατε ν' αλλάζετε τρία πράγματα στη σχολή, ποια θα ήταν αυτά; Να είστε όσο πιο συγκεκριμένοι γίνεται:

Και ποια δεν θα αλλάζατε; Να είστε όσο πιο συγκεκριμένοι γίνεται:

ΒΕΒΑΙΩΘΕΙΤΕ ΟΤΙ ΕΧΕΤΕ ΑΠΑΝΤΗΣΕΙ ΟΛΕΣ ΤΙΣ ΕΡΩΤΗΣΕΙΣ ΚΑΙ ΟΤΙ ΕΧΕΤΕ ΔΩΣΕΙ ΜΙΑ ΜΟΝΟ ΑΠΑΝΤΗΣΗ ΣΕ ΚΑΘΕ ΕΡΩΤΗΣΗ

Σας ευχαριστούμε για τη συνεργασία σας. Ημερομηνία/...../200...

This study highlights the importance of assessing and reporting details of scale reliability and factor loadings, done best through the use of a combination of criteria²³, even when assessing a tool already validated in another language. This study also suggests that when translated instruments have items that do not load onto any factor, they may be suffering from poor translation, especially a failure to adequately address cultural differences between the peoples speaking the original and target languages. Providing subjects with definitions (glossary) of confusing terms, instead of attempting to find single-word substitutes in the new language, can sometimes be the solution.

Limitations

Although this is an almost national administration of the DREEM, our sample is not representative of all medical schools and years of study. Various sampling schemes, response rates, and cohort sizes across medical schools and years of study make the data more appropriate for testing the translated instrument than it does for detecting educational environment quality or comparing schools. Although our mean overall score (107.7) was quite similar to those from other countries, ranging from 97 to 143 in a recent review²⁴, reported scores should be interpreted as a starting point and with caution, especially the ill-translated questions discussed above.

Acknowledgements

Virginia Draina, Didi Vergados, Mary O'Connor, Elpiniki Pappu and Lena Cavanagh for back translation; Mary Gouva for her help in reliability and factor analysis; George Tsianos for his comments; colleagues Antony Kafatos, Victoria Kalapothaki, Anastassios Germetis, Alexis Benos, and Theodore Constantinidis for their invaluable help in data collection. We also would like to thank the *Education for Health* editors and the unknown reviewers whose comments greatly improved our work.



References

1. Genn J. AMEE Medical Education Guide No 23 (Part I): Curriculum, environment, climate, quality and change in medical education - a unifying perspective. *Medical Teacher*. 2001; 23(4), 337-344.
2. Genn J, Harden RM. What is medical education here really like? Suggestions for action research studies of climates of medical education environments. *Medical Teacher*. 1986; 8:111-124.
3. Hutchinson L. The ABC of learning and teaching: educational environment. *British Medical Journal*. 2003; 326, 810-812.
4. Roff S. *Educational Environment and Climate*. University of Dundee, Centre for Medical Education, Theme Curriculum Development CD10; 2005.
5. Roff S, McAleer S, Harden RM, Al-Qahtani M, Uddin AA, Deza H, Groenen G, Primparyon P. Development and Validation of the Dundee Ready Education Environment Measure (DREEM). *Medical Teacher*. 1997; 19(4):295-299.
6. McAleer S, Roff S. Part 3; A practical guide to using the Dundee Ready Education Measure (DREEM). In, J. M. Genn (ed), *AMEE Medical Education Guide No.23 Curriculum, environment, climate, quality and change in medical education; a unifying perspective*. Dundee, UK: Association of Medical Education in Europe; 2002.
7. Roff S. The Dundee Ready Educational Measurement (DREEM): a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Medical Teacher*. 2005; 27(4):322-325.
8. Dunne F, McAleer S, Roff S. Assessment of the undergraduate medical education environment in a large UK medical school. *Health Education Journal*. 2006; 65:149-158.
9. Avalos G, Freeman C, Dunne F. Determining the quality of the medical educational environment at an Irish medical school using the DREEM inventory. *Irish Medical Journal*. 2007; 100(7):522-525.
10. Miles S, Leinster SJ. Medical students' perceptions of their educational environment: expected versus actual perceptions. *Medical Education*. 2007; 41(3):265-72.
11. Whittle SR, Whelan B, Murdoch-Eaton DG. DREEM and beyond; studies of the educational environment as a means for its enhancement. *Education for Health*. 2007; 20(1):7.
12. Norusis/SPSS Inc. *SPSS Professional Statistics™ 7.5*. Chapter 13 Measuring scales: reliability analysis examples; 1997. p. 103-111.
13. Fayers PM, Machin D. *Quality of life: assessment, analysis and interpretation*. John Willey & Sons. England; 2000.
14. Streiner DL, Norman GR. *Health Measurement Scales - a practical guide to their development and use*. 4th edition. Oxford University Press. Oxford; 2008.



15. Berk RA. *Thirteen strategies to measure college teaching*. Stylus publishing LLC, Sterling, Virginia, USA; 2006. p. 188.
16. Norman G, Eva KW. *Quantitative research methods in medical education*. Association for the Study of Medical Education booklet series Understanding Medical Education; 2008.
17. Primpayon P, Roff S, McAleer S, Poonchai B, Pemba S. Educational environment, student approaches to learning and academic achievement in a Thai nursing school. *Medical Teacher*. 2000; 22:359-365.
18. Mayya S, Roff S. Students' Perceptions of Educational Environment: A Comparison of Academic Achievers and Under-Achievers at Kasturba Medical College, India. *Education for Health*. 2004; 17(3), 280-291.
19. de Oliveira Filho GR, Vieira JE, Schonhorst L. Psychometric properties of the Dundee Ready Educational Environment Measure (DREEM) applied to medical residents. *Medical Teacher*. 2005; 27(4):343-347.
20. Riquelme A, Oporto M, Oporto J, Mendez J, Viviani P, Salech F, Chianale J, Moreno R, Sanchez I. Measuring Students' Perceptions of the Educational Climate of the New Curriculum at the Pontificia Universidad Catolica de Chile: Performance of the Spanish Translation of the Dundee Ready Education Environment Measure (DREEM). *Education for Health* 9 (online), 2009: 112. Available from: <http://www.educationforhealth.net/articles/subviewnew.asp?ArticleID=112>
21. Pappa E, Kontodimopoulos N, Niakas D. Psychometric evaluation and normative data for the Greek SF-36 health survey using a large urban population sample. *Archives of Hellenic Medicine*. 2006; 23(2):159-166.
22. Vassilopoulos D. *Medical Education*. Self-edition, Athens (in Greek); 2006. p. 27-43.
23. Schönrock-Adema J, Heijne-Penninga M, Van Hell EA, Cohen-Schotanus J. Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Medical Teacher*. 2009; 31(6):e226-232.
24. Dimoliatis I. D. K. (accepted 6 July 2009). The Dundee Ready Education Environment Measure (DREEM) instrument in Greek: how to use and preliminary results for the Greek medical educational environment. *Archives of Hellenic Medicine* (in press). In Greek with an English abstract. Available from www.mednet.gr.
-