LUND UNIVERSITY
Faculty of Medicine

LUP

Lund University Publications

Institutional Repository of Lund University

Psychometric evaluation of the Dundee Ready Educational Environment Measure (DREEM): Swedish version

Running head: Psychometric evaluation of the DREEM instrument

Ulf Jakobsson [1], Nils Danielsen [2], Gudrun Edgren [3]

[1]Department of Clinical Sciences-Malmo, Lund University, Sweden, [2]Department of Experimental Medical Science, Lund University, Sweden and [3]Centre for Teaching and Learning, Faculty of Medicine, Lund University, Sweden

Address for correspondence;
Ulf Jakobsson,

Centre for Primary Health Care Research, Faculty of Medicine, Lund University.

CRC, entrance 72, building 28, floor 11

Malmö University Hospital

SE-205 02, Malmö, Sweden

E-mail: Ulf.Jakobsson@med.lu.se

Telephone: +46 40 391317, Fax: +46 46 2221934

ABSTRACT

Background: The Dundee Ready Education Environment Measure (DREEM) has been used in various studies to evaluate the educational environment. However, psychometric evaluations of the instrument seem sparse, for all known versions of the instrument.

Aim: The aim was to psychometrically evaluate the Swedish version of the DREEM instrument.

Method: A total of 503 students (undergraduate medicine), aged 19-46 years, in semesters 2, 6 and 10 were included in the study. Validity was evaluated through analysis of construct validity and reliability.

Results: The instrument had in general both acceptable validity and reliability. Due to a rather poor model fit in the confirmatory factor analysis an explorative factor analysis was also employed which suggested a new five factor solution for the instrument.

Conclusions: The Swedish version of the DREEM instrument is shown to be valid and reliable, except for the factor structure. The new five factor solution found in this study is not proven to be a superior measurement model compared with the original, but could be seen as an alternative model to the original, where the strong and weak areas are somewhat more easily identified.

INTRODUCTION

The Dundee Ready Edcucational Environment Measure (DREEM) was developed and validated for international use within health professions' curricula about 15 years ago (Roff et al 1997). It has since been extensively used to measure students' perceptions of educational environment with different purposes. Schools and sites within schools (Roff et al 2001, Al-Hazimi et al 2004, Varma et al 2005, McKendree 2009) and academic achievers and under-achievers (Mayya & Roff 2004) have been compared. Information about students' perceptions of ideal environment (Till 2005) and expectations of the climate (Miles & Leinster 2007) has been collected. DREEM has also been used as an instrument for improvement (Whittle et al 2007) and to follow what happens during curriculum reform (Jiffry et al 2005, Bouhaimed et al 2009).

The instrument has been translated into several languages (Roff 2005). Despite this extensive research very few psychometric studies have been published since the original development of the instrument (Roff et al 1997). Among those are a Portuguese version for use in training of residents (de Oliveira Filho et al 2005), a Chinese version for nursing (Wang et al 2009), and a greek translation (Dimoliatis et al 2010) The DREEM instrument contains 50 items and a Likert-scale is used giving a global maximum score of 200. Factor analysis has resulted in five sub-scales (Roff et al 1997). The original sub-scale structure was not maintained in the three translated versions (de Oliveira Filho et al 2005, Wang et al 2009, Dimoliatis et al 2010). We have translated the instrument into Swedish and used it to measure the educational environment in an undergraduate medical programme (Edgren et al 2010). A professional translation of the instrument does not guarantee that the original reliability and validity of the instrument remain when the instrument is use in a novel environment. Therefore a psychometric test is warranted.

The aim of this study was thus to psychometrically evaluate the Swedish version of the Dundee Ready Education Environment Measure (DREEM).

METHODS

Context

The medical program at Lund University is five and a half years with undergraduate entry. At the time of the study 96 students were enrolled twice a year (two semesters) and 50-60 % of the students were female.

Questionnaire

The questionnaire (DREEM) contains 50 items with a five point Likert response scale ranging from 0=strongly disagree to 4=strongly agree (Roff et al., 1997). Factor analysis has revealed five subscales: 1) perceptions of learning (max score 48), 2) perceptions of teachers (max score 44), 3) academic self-perceptions (max score 32), 4) perceptions of atmosphere (max score 48), and 5) social self-perceptions (max score 28). The total maximum score is 200. Some items are negative and the scores are reversed before calculations.

The DREEM questionnaire was translated into Swedish by a teacher proficient in English and then translated back into English by a professional translator. Differences were discussed between the translator and two of the authors. The instrument was piloted with a group of 20 students in the 4th year of study.

Sample

Data was collected through the DREEM questionnaire which was distributed to all students (at lectures) in semesters 2, 6 and 10, respectively, in 2003 and 2005. Students who were present at the lecture got the form immediately and those who did not attend the lecture received it from course administrative staff. The cohort of semester 2 in 2003 and semester 6 in 2005 was thus the same, as were the cohorts of semester 6 in 2003 and semester 10 in 2005. The cohorts of semester 2 in 2005 differed slightly from the 2003 cohort, since they started their medical programme with an introductory course (8-weeks, focussing on an overview of the programme, PBL training and study techniques) which the cohort of semester 2 in 2003 did not. Else there were no changes in the curriculum during the study time.

Statistical analyses / Data analysis

Data were analysed and compared across cohorts as well as for the total sample. The analyses were performed using the Student´s t-test and the Chi-square test. The Chi-square test was used for comparisons of nominal data, while the Student´s t-test was employed for interval/ratio data.

The instrument was tested for construct validity and reliability (Nunnally & Bernstein 1994). Validity was evaluated through analysis of item response rate, floor- and ceiling effect, corrected item-total correlations, and factor structure. Item response rate was calculated as the number of (percent of) respondents that had completed all items the total instrument as well as in the subscales. A figure of 90 % and above is considered as satisfactory. The proportion of floor and ceiling effects (i.e. people obtaining minimum and maximum scores respectively) was studied among the subscales. The corrected item-total correlations (i.e. the correlation between each item and the total score of the remaining items in the hypothesized scale) should

be 0.30 or greater to be considered as acceptable and interpreted as evidence supporting construct validity.

To evaluate the DREEM instrument's factor structure a confirmatory factor analyses (LISREL) were performed, based on Spearman's rank order correlation matrix. The factors were defined in concordance with the sub-scales in the instrument, while correlation between the latent variables, factor loadings on the latent variables, and the residuals were set as free parameters. The chi-square goodness-of-fit test, the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), the normed fit index (NFI), the non-normed fit index (NNFI) and the root mean squared error of approximation (RMSEA) were used to assess goodness-of-fit. If the factor model fits, the chi-square test should be non-significant (i.e. $p > 0.05$). The GFI, AGFI and NFI are scaled to range between 0 and 1, with values above 0.90 suggesting a good model fit while values ≥ 0.85 can be considered to indicate acceptable model fit (Schermelleh-Engel & Moosbrugger, 2003). An exception to this is the NNFI, which is not scaled to range between 0 and 1, although a higher value still indicates better fit (Schermelleh-Engel & Moosbrugger, 2003). Whereas RMSEA values ≤ 0.08 (preferably ≤ 0.05) are considered to indicate acceptable fit. The reason why the goodness-of-fit of the factor model was assessed by various descriptive fit indices was because this statistic is sensitive to sample size (e.g. the Chi-square test). Furthermore, when the number of free parameters increases in the model, the better it will fit e.g. in GFI (Nunnally & Bernstein, 1994). However, the AGFI and the NNFI adjust (downwards) the GFI and NFI respectively, based on the number of parameters/degrees of freedom.

In order to further investigate the factor structure of the DREEM instrument an explorative factor analysis (i.e. principal component analysis) with varimax rotation was also performed.

6

In the explorative factor analysis both the Kaiser´s eigenvalue criterion (Eigenvalue ≥1), and a scree plot was used to determine the number of factors to extract (Ferguson & Cox, 1993). Missing values were excluded listwise. The quality of the factor analysis models was assessed using Bartlett´s test for sphericity and the Kaiser-Meyer-Olkin (KMO) test, together with evaluation of communalities, factor loadings and eigenvalues. Bartlett´s test is a measure of the probability that the initial correlation matrix is an identity matrix and should be <0.05. The KMO test measures the degree of multicollinearity (based on partial correlations) and varies between 0 and 1 (should be greater than 0.50-0.60).

The reliability of the instrument was analysed by means of tests for internal consistency, employing Cronbach's alpha (Cronbach, 1951). The alpha values preferably range between 0.7 and 0.9 even if figures as low as 0.6 may be acceptable (Nunnally & Bernstein, 1994). Alpha values lower than 0.7–0.6 indicate too high heterogeneity and values higher than 0.9 indicate that the items may be too similar.

The items (no. 4, 8, 9, 17, 25, 35, 39, 48, 50) were revised so that all items ranged from positive to negative replies before the factor analyses, correlations and Cronbach's alpha were performed. Data were analyzed using SPSS 15.0 for Windows and LISREL 8.51.

RESULTS

The study comprised 395 individuals with a mean age of 25.5 (95% CI 25.0-26.0), range 19-46 years (Table 1). The overall response rate was 82% (201/246) in 2003 and 75% (194/257) in 2005. Of the students that responded 58.9% were females and 41.1% males. Mean results and SD are presented in Table I.

The rate of incomplete questionnaires seemed at first rather high. Of the 103 persons (26 %) who failed to complete all 50 items, the number of missing responses in each item ranged between 1 (0.3 %) and 73 (18.5 %). However, it was actually item number six ("The teacher are patient with the patients") that had the highest internal missing (n=73, 18.5%), and the missing responses were all from the students in semester 2 (i.e. students in the beginning of the program containing core pre-clinical studies and have only had a few encounters with patients). When not considering this item the missing responses ranged between 1 (0.3 %) and 6 (1.5 %). Missing responses in the subscales (when including all items) were subscale I: 3.8%, subscale II: 20.5%, subscale III: 4.6%, subscale IV: 17.0% and subscale V: 2.3%. Analyses of floor- and ceiling effects showed no major problems (Figure 1). No floor effects were identified and only minor ceiling effects were detected; subscale II had two respondents (0.5%) that reached the maximum value, subscale IV had four respondents (1.0%) reaching the maximum, and in subscale V three respondents (0.8%) did so. The corrected item-total correlations indicated that the instrument had acceptable construct validity (Table 2). All correlations were r≥0.3, except item number 48 ("The teaching is too teacher-centred") in subscale I, item number 17 ("Cheating is a problem in this school") in subscale IV, and item number 46 ("My accommodation is pleasant") in subscale V.

Factor loadings from the confirmatory factor analysis ranged from 0.028 to 0.762 (Table 3) and goodness-of-fit indices did not fully support the hypothesized items-to-summary measures structure (Table 3). The chi-square test (p<0.001), GFI (=0.75), AGFI (=0.73), produced by the LISREL program all indicated a rather low goodness-of-fit, while NFI (=0.89), NNFI (=0.93) and the RMSEA (=0.068) indicated an acceptable fit. All items had factor loadings of ≥0.3 except item number 17 ("Cheating is a problem in this school"), 23

("The atmosphere is relaxed during lectures"), 39 ("The teachers get angry in class"), 46 ("My accomodation is pleasant"). Item 39 had an especially low factor loading.

The instrument was also found to have acceptable Cronbach's alpha for all scales even if the alpha value for subscale V was a little bit low (but was approx. α=0.7). Cronbach's alpha for the five subscales (I-V) were I: 0.807; II: 0.785; III: 0.720; IV: 0.786 and V: 0.689. The alpha value for the whole DREEM instrument (all items) was 0.930.

Due to the rather poor model fit in the confirmatory factor analysis an exploratory factor analysis was also performed. Various factor solutions (ranging between a five to a nine factor solution) were suggested. However, only the five factor structure could give a logical classification of the items (Table 4). This new five factor solution explained 42% of the variance, and the communalities varied between 0.119 and 0.627. Two items (no. 26 & 46) had communality values <0.2 (Table 4). Each new factor was reviewed and given new headings based on its contents (Appendix). In the review two items were found, based on its content, to be more suitable in another factor, and these two changed place. Item 26 was moved from F2 to F4 and item 11 from F4 to F2. The decision to make these changes was due to on the one hand that the item better matched the content in the new factor, and on the other hand the factor loadings were approximately the same for these items in both factors. A description of the items for each factor is presented in the appendix, and descriptive statistics for the new factors is presented in table 1. Cronbach´s alpha was calculated for the new five factors and ranged between 0.705-0.892 (Table 4). Missing responses in the subscales (for the new five factor solution) were subscale I: 6.6%, subscale II: 22.3%, subscale III: 1.8%, subscale IV: 2.8% and subscale V: 2.0%. No floor or ceiling effects were identified.

Students could write comments on other things that they thought were important for the environment, and a few of them commented that the buildings (lecture halls, group rooms, self study rooms) are important.

DISCUSSION

The Swedish version of the DREEM instrument was found to have both acceptable validity and reliability. The validity was tested in terms of construct validity and reliability in terms of internal consistency. Also the feasibility was assessed through analysis of item response rate.

The overall response rate of 82% and 75% respectively can be considered acceptable. The proportion of female and male respondents fitted well with the figures in the whole group. Item 17 ("Cheating is a problem in this school") and item 46 ("My accommodation is pleasant") appear to have both low correlation and low factor loading. Do the students misinterpret the item about cheating? Does it mean only at exams or otherwise too? The item about accommodation is probably not relevant in a Swedish context. It has the highest score in the instrument (mean 3.46 SD 0.79). Swedish standard of living is generally high, and most importantly it is not the responsibility of the universities to provide accommodation for students.

The other items that had low factor loadings were number 23 ("The atmosphere is relaxed during lectures") and 39 ("The teachers get angry in class") which had an especially low factor loading. Maybe these items do not appear relevant to the Swedish students where the relation between students and teachers are very egalitarian, and they therefore make different interpretations than students in other contexts.

Due to a rather poor model fit in the confirmatory factor analysis an explorative factor analysis was employed which suggested a new five factor solution for the instrument. The new five factor solution had slightly better Cronbach´s alpha values than the original factor model. Neither the confirmatory factor analysis nor the exploratory analysis provided clear support for the internal construct validity (factor structure) of the proposed measurement model Roff et al 1997). This has also been seen in the other psychometric analyses of the instrument that have been published (de Oliveira Filho et al., 2005¸ Wang et al 2009¸ Dimoliatis et al 2010). Dimoliatis et al (2010) presented both confirmative and exploratory factor analyses, and neither completely confirmed the original model. According to these authors (Dimoliatis et al 2010) the original model was developed by a qualitative method, and this could explain the differences. Wang et al (2009) proposed yet another five factor solution in their study from an undergraduate nursing school, and cultural differences are suggested as an explanation. The Portuguese version was used for residents, which could explain some differences, since the instrument was developed for undergraduate students (de Oliveira Filho et al., 2005). A reason for the poor model fit in some of the tests, for the confirmatory factor analysis in this study, might be due to the fact that there were few observations in relation to the number of parameters estimated. The recommended ration is 20 observations for each estimated parameter (i.e. $\approx 1/20$) (Hair et al, 2008). In this study this ratio was 50/395 ($\approx 1/4$), while the number of respondent should be around 1000 according to this rule of thumb. Another reason to the poor model fit may be the influence of some items (i.e. item number 17, 23, 39, 46, 48) that had low correlations with the summary score of the subscale to which they were intended to belong. However, despite these limitations, half of the goodness-of-fit tests indicated an acceptable model fit and the majority (92%) of factor loadings were ≥ 0.3 which could be seen as an indication for acceptable construct validity of the DREEM instrument

anyway. Furthermore, the analysis of item-total correlation (Table 2) supported this assumption.

Both the confirmatory factor analysis and the item-total correlation identified some items to have low correlation with the factor. These items were item number 17, 39, 46 (Table 2, 3). One explanation mentioned above might be that these questions are not as relevant for the Swedish context. However, this problem was not found in the factor model derived from the explorative factor analysis (except for item number 46 which had a communality of 0.156). Thus, the low factor loadings may still be a result of a poor model fit and if the items are to be included in the Swedish version of the DREEM instrument then the factor structure suggested by the explorative factor analysis is probably a better option.

The suggested five factor solutions was not in accordance with the hypothesized (original) measurement model of the DREEM instrument and is not proven to be a superior measurement model compared with the original, but could be seen as an alternative model which can be further explored/tested. The new solution however added some value to the interpretation of the results. There was a somewhat bigger difference in the average scores on these subscales (0,68-0,82% of the maximum) than on the original ones (0,70-0,77% of the maximum) (Table 1). The strong and weak areas, respectively, are somewhat more easily identified, and can be subjected to development efforts. In our case it became obvious that areas where students have a stronger influence (e g "psychosocial situation" 82% of max) were more highly rated than areas where the teachers have a more obvious responsibility ("communication" and "teaching organisation and progression" both 68% of max) (cf Table 1).

It was interesting that some students commented that buildings and rooms are also a part of the environment and contribute to the perceptions of the students. It seems reasonable to believe that rooms for teaching and learning also influence the climate, and an item about those could be added to the instrument.

The results of this study and others (De Oliveira Filho et al 2005; Wang et al, 2009, Dimoliatis et al 2010) indicate that the DREEM may not be as culturally independent as was originally stated (Roff et al 1997). For future use in Sweden some items should probably be removed, at least item 46. Item 17 and 39 should perhaps also be taken out or subjected to an analysis of how the students interpret these items. When we used the instrument for measurement of the climate in Lund university medical programme (Edgren et al 2010) we made used of the total score, scores on individual items, and comparisons between years and semesters. Our conclusions in that study should not be affected by the potential weaknesses discovered in this study.

Study limitations

The relatively small number of students is a limitation of this study and also that it has been undertaken in a single context.

ACKNOWLEDGEMENTS

13

PRACTICE POINTS

- The Swedish version of the DREEM instrument is in general a psychometrically valid and reliable instrument which can be used to identify problem areas in medical education programmes.

- A new five factor solution found in this study that could be seen as an alternative model to the original measurement model where the strong and weak areas are somewhat more easily identified.

NOTES ON CONTRIBUTORS

Gudrun Edgren, PhD, MMedEd, is senior lecturer and head of the Centre for Teaching and Learning at the Faculty of Medicine, Lund University. At the centre she teaches courses and run workshops in medical education and takes part in evaluations and research projects in education.

Nils Danielsen, MD, PhD, is senior lecturer and is the chairman for the Committee for Biomedical, Medical and Public Health Education, Faculty of Medicine, Lund University. This committee makes all strategic decisions related to teaching including financing. ND teaches anatomy and is enrolled as a tutor in PBL.

Ulf Jakobsson, PhD, is senior lecturer at the Faculty of Medicine, Lund University. He is a teacher in research methodology and statistics at undergraduate and postgraduate levels at the Faculty of Medicine, Lund University.

REFERENCES

AL-HAZIMI, A., ZAINI, R., AL-HYIANI, A., HASSAN, N., GUNAID, A., PONNAMPERUMA, G., KARUNATHILAKE, I., ROFF, S, MCALEER, S. & DAVIS, M. (2004) Educational environment in traditional and innovative medical schools: a study in four undergraduate medical schools. *Education for Health*, 17(2), pp. 192-203

BOUHAIMED, M., THALIB, L. & DOI, S.A.R. (2009) Perception of the educational environment by medical students undergoing a curriculum transition in Kuwait. *Medical Principles and Practice*, 18(3), pp. 204-208

CRONBACH, L.J. (1951). Coefficient alpha and the internal structures of tests. *Psychometrika*, 3, pp. 297-334.

DE OLIVEIRA FILHO, G.R., EDSON VIEIRA, J. & SCHONHORST, L. (2005). Psychometric properties of the Dundee Ready Educational Environment Measure (DREEM) applied to medical residents. *Medical Teacher*, 27 (4), pp. 343-347.

DIMOLIATIS, I.D.K., VASILAKI, E., ANASTASSOPOULOS, J.P.A.& ROFF, S. (2010). Validation of a greek translation of the Dundee Ready Education Environment measure (DREEM). *Education for Health*, 23 (1), pp. ??. [Epub ahead of print]

EDGREN, G., HAFFLING, A.C., JAKOBSSON, U., MCALEER, S. & DANIELSEN, N. (2010). Comparing the educational environment (as measured by DREEM) at two different stages of curriculum reform. *Medical Teacher*, 32 (6), pp. e233-238.

FERGUSON, E. & COX, T. (1993). Exploratory factor analysis: a users´ guide. *International Journal of Selection Assessment*, 1, pp. 84-94.

HAIR, J.F., BLACK, B., BABIN, B.J., ANDERSON, R.E. & TATHAM, R.L. (2008). *Multivariate data analysis* (6th ed.). New Jersey: Prentice-Hall Inc.

JIFFRY, M.T.M., McALEER, S., FERNANDO, S. & MARASINGHE, R.B. (2005). Using the DREEM questionnaire to gather baseline information on an evolving medical school in Sri Lanka. *Medical Teacher*, 27(4), pp. 348-352.

MAYYA, S.S. & ROFF, S. (2004). Students' perception of educational environment: a comparison of academic achievers and under-achievers at Kasturba Medical College, India. *Education for Health*, 17(3), pp. 280-291.

MCKENDREE, J. (2009). Can we create an equivalent educational experience on a two campus medical school? *Medical Teacher*, 31, e202-e205.

MILES, S. & LEINSTER, S.J. (2007). Medical students' perceptions of their educational environment: expected versus actual perceptions. *Medical Education*, 41(3), pp. 265-272.

NUNNALLY, J.C. & BERNSTEIN, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill Inc.

ROFF, S., MCALEER, S., HARDEN, R.M., AL-QAHTANI, M., AHMED, A.U., DEZA, H., GROENEN, G. & PRIMPARYON, P. (1997). Development and validation of the Dundee Ready Education Environment Measure (DREEM), *Medical Teacher,* 19(4), pp. 295-299.

ROFF, S., McALEER, S., IFERE, O.S. & BHATTACHARYA, S. (2001) A global diagnostic tool for measuring educational environment: comparing Nigeria and Nepal. *Medical Teacher,* 23(4), pp. 378-382.

ROFF, S. (2005). The Dundee Ready Educational Environment Measure (DREEM) – a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Medical Teacher*, 27(4), pp. 322-325.

SCHERMELLEH-ENGEL, K. & MOOSBRUGGER, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, pp. 23-74.

TILL, H. (2005). Climate studies: can students´ perceptions of the ideal educational environment be of use for institutional planning and resource utilization? *Medical Teacher*, 27(4), pp. 332-337.

VARMA, R., TIYAGI, E. & GUPTA, J.K. (2005). Determining the quality of educational climate across multiple undergraduate teaching sites using the DREEM inventory. *BMC Medical Education*, 5(1), p. 8.

WANG, J., ZANG, S. & SHAN, T. (2009). Dundee Ready Educational Environment Measure psychometric testing with Chinese nursing students. *Journal of Advanced Nursing*, 65(12), pp. 2701-2709.

WHITTLE, S., WHELAN, B. & MURDOCH-EATON, D.G. (2007). DREEM and beyond: studies of the educational environment as a means for its enhancement. *Education for Health*, 20(1), p. 7.
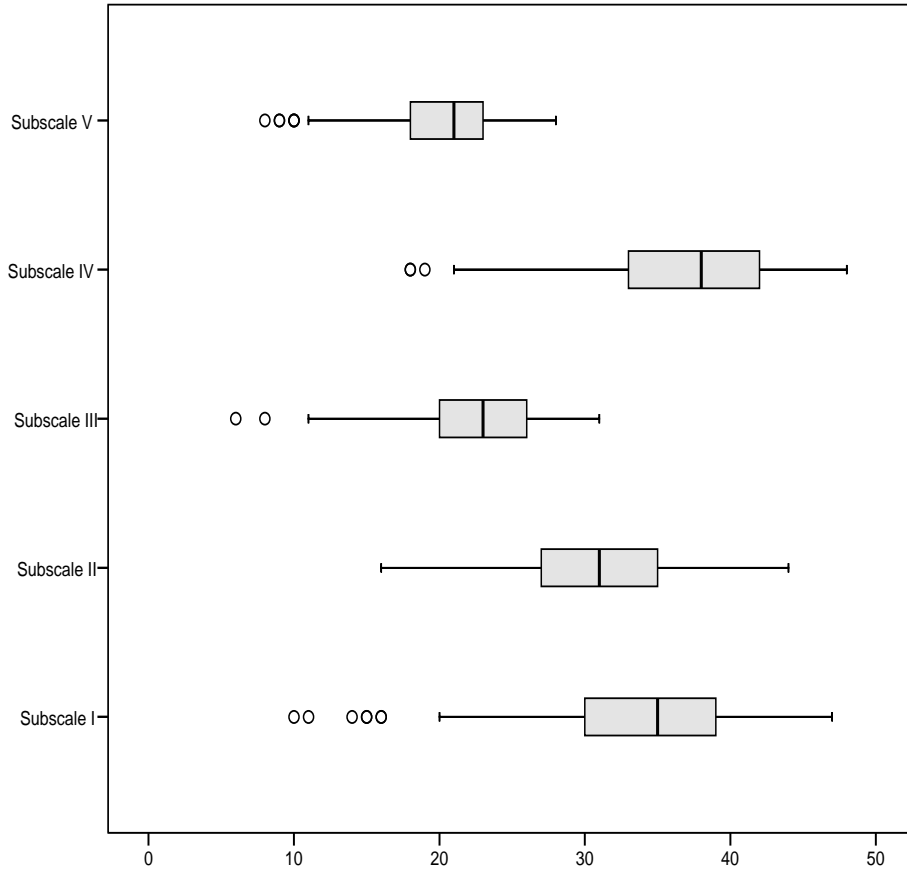
Figure 1. The distribution of the responses for each factor

Table 1. Description of the sample and comparison between the two cohorts

	Total sample n=395	Cohort 2003 n=201	Cohort 2005 n=194	P-value (comparisons across cohorts)
Age, mean (SD) [a]	25.5 (4.8)	25.1 (4.6)	26.0 (5.0)	0.063
Women % [b]	58.9	58.7	59.1	0.942
DREEM total score, mean (SD) [a]	145.2 (22.2)	144.1 (21.8)	146.2 (22.7)	0.417
DREEM subscale scores, mean (SD) [a]				
1) Perceptions of learning	33.6 (6.7)	33.6 (6.5)	34.2 (6.7)	0.393
2) Perception of teachers	31.0 (5.5)	30.5 (5.3)	31.4 (5.6)	0.160
3) Academic self-perceptions	23.0 (4.3)	22.5 (4.2)	22.4 (4.7)	0.699
4) Perceptions of atmosphere	37.1 (6.1)	37.0 (6.0)	37.6 (6.0)	0.356
5) Social self-perceptions	20.5 (4.0)	20.1 (3.6)	20.8 (4.2)	0.111
Mean score divided with the total score, for each sub scale in DREEM [b]				
1) Perceptions of learning	0.70	0.70	0.71	
2) Perception of teachers	0.70	0.69	0.71	
3) Academic self-perceptions	0.72	0.70	0.70	
4) Perceptions of atmosphere	0.77	0.77	0.78	
5) Social self-perceptions	0.73	0.72	0.74	
DREEM subscale scores from the **new five factor solution**, mean (SD) [a]				
1) Learning and motivation	48.1 (9.6)	47.4 (9.4)	48.9 (9.7)	0.139
2) Communication	21.6 (4.6)	21.3 (4.3)	21.9 (4.9)	0.298
3) Psychosocial situation	32.8 (5.1)	32.4 (5.0)	33.3 (5.2)	0.103
4) Teaching org. and progression	18.9 (4.1)	19.0 (4.0)	18.8 (4.2)	0.645
5) Bad teaching	23.8 (4.2)	23.7 (4.1)	24.0 (4.4)	0.506
Mean score divided with the total score, for each sub scale in the **new five factor solution version** [b]				
1) Learning and motivation	0.71	0.70	0.72	
2) Communication	0.68	0.67	0.68	
3) Psychosocial situation	0.82	0.81	0.83	
4) Teaching org. and progression	0.68	0.68	0.67	
5) Bad teaching	0.74	0.74	0.75	

[a] Student´s t-test, [b] Chi-square test

Table 2. Corrected item-total correlations for each item (divided in the predefined subscales)

Subscale I	Item-total correlations	Subscale II	Item-total correlations	Subscale III	Item-total correlations	Subscale IV	Item-total correlations	Subscale V	Item-total correlations
- Item 1	0.523	- Item 2	0.330	- Item 5	0.407	- Item 11	0.420	- Item 3	0.332
- Item 7	0.348	- Item 6	0.438	- Item 10	0.405	- Item 12	0.448	- Item 4	0.375
- Item 13	0.539	- Item 8	0.414	- Item 21	0.527	- Item 17	0.155	- Item 14	0.393
- Item 16	0.594	- Item 9	0.364	- Item 26	0.319	- Item 23	0.428	- Item 15	0.451
- Item 20	0.606	- Item 18	0.460	- Item 27	0.385	- Item 30	0.369	- Item 19	0.532
- Item 22	0.545	- Item 29	0.504	- Item 31	0.388	- Item 33	0.483	- Item 28	0.541
- Item 24	0.532	- Item 32	0.535	- Item 41	0.512	- Item 34	0.478	- Item 46	0.204
- Item 25	0.214	- Item 37	0.545	- Item 45	0.430	- Item 35	0.476		
- Item 38	0.451	- Item 39	0.419			- Item 36	0.406		
- Item 44	0.599	- Item 40	0.468			- Item 42	0.552		
- Item 47	0.364	- Item 50	0.415			- Item 43	0.585		
- Item 48	0.270					- Item 49	0.344		

Table 3. Result from the confirmatory factor analysis (the items are divided in the predefined subscales)

Subscale I	Factor loading	Subscale II	Factor loading	Subscale III	Factor loading	Subscale IV	Factor loading	Subscale V	Factor loading
- Item 1	0.486	- Item 2	0.288	- Item 5	0.459	- Item 11	0.407	- Item 3	0.380
- Item 7	0.762	- Item 6	0.438	- Item 10	0.443	- Item 12	0.580	- Item 4	0.584
- Item 13	0.569	- Item 8	0.373	- Item 21	0.548	- Item 17	0.156	- Item 14	0.649
- Item 16	0.480	- Item 9	0.370	- Item 26	0.364	- Item 23	0.381	- Item 15	0.428
- Item 20	0.545	- Item 18	0.381	- Item 27	0.463	- Item 30	0.387	- Item 19	0.477
- Item 22	0.631	- Item 29	0.544	- Item 31	0.463	- Item 33	0.372	- Item 28	0.620
- Item 24	0.569	- Item 32	0.562	- Item 41	0.531	- Item 34	0.407	- Item 46	0.263
- Item 25	0.244	- Item 37	0.492	- Item 45	0.389	- Item 35	0.551		
- Item 38	0.469	- Item 39	0.028			- Item 36	0.408		
- Item 44	0.617	- Item 40	0.418			- Item 42	0.557		
- Item 47	0.424	- Item 50	0.485			- Item 43	0.637		
- Item 48	0.302					- Item 49	0.453		

Chi-square: $p<0.001$; GFI=0.750; AGFI=0.726; NFI=0.888; NNFI=0.926; RMSEA=0.068 $p<0.001$

Table 4. Principal component analysis with varimax rotation for the DREEM instrument

Variable	F 1	F 2	F 3	F 4	F5	Communalities
			Factor loadings			
Item no. 1	**0.587**	0.195	0.183	0.047	0.107	0.430
Item no. 2	**0.582**	0.068	-0.022	0.017	0.215	0.390
Item no. 3	**0.321**	0.201	0.144	0.151	0.165	0.214
Item no. 4	0.078	-0.035	**0.425**	0.361	0.207	0.361
Item no. 5	0.222	0.091	-0.006	**0.588**	0.214	0.449
Item no. 6	0.158	**0.587**	-0.023	0.120	0.207	0.427
Item no. 7	**0.531**	0.051	0.063	-0.053	-0.097	0.301
Item no. 8	0.152	0.093	0.097	0.004	**0.622**	0.428
Item no. 9	-0.160	0.140	-0.023	0.098	**0.673**	0.508
Item no. 10	0.001	0.155	0.202	**0.615**	0.013	0.443
Item no. 11	0.178	0.301	0.072	**0.340**	0.160	0.268
Item no. 12	**0.521**	0.142	0.116	0.285	0.068	0.391
Item no. 13	**0.531**	0.283	0.092	0.164	0.228	0.449
Item no. 14	**0.639**	0.130	0.289	0.017	-0.077	0.515
Item no. 15	0.269	-0.088	**0.653**	-0.043	0.050	0.510
Item no. 16	**0.610**	0.150	0.198	0.314	0.034	0.534
Item no. 17	0.109	-0.216	-0.101	0.273	**0.461**	0.355
Item no. 18	0.147	**0.636**	-0.011	0.112	0.181	0.472
Item no. 19	0.110	0.104	**0.631**	0.298	0.005	0.510
Item no. 20	**0.636**	0.146	0.064	0.348	0.142	0.572
Item no. 21	**0.526**	0.244	0.179	0.382	0.063	0.518
Item no. 22	**0.492**	0.348	0.218	0.243	0.098	0.479
Item no. 23	0.255	0.200	**0.290**	0.086	0.226	0.248
Item no. 24	**0.690**	0.130	0.048	0.067	0.069	0.505
Item no. 25	0.006	0.070	0.009	0.227	**0.407**	0.222
Item no. 26	0.021	**0.213**	0.177	0.201	-0.039	0.119
Item no. 27	0.179	0.124	0.183	**0.469**	0.001	0.301
Item no. 28	0.184	0.136	**0.636**	0.143	0.085	0.484
Item no. 29	0.284	**0.593**	0.086	0.084	0.151	0.470
Item no. 30	0.225	**0.616**	0.213	0.067	0.002	0.479
Item no. 31	0.188	**0.548**	0.102	0.184	-0.059	0.383
Item no. 32	0.147	**0.721**	0.147	0.041	0.149	0.587
Item no. 33	0.180	0.050	**0.717**	0.051	0.165	0.579
Item no. 34	0.182	0.094	**0.485**	0.057	0.322	0.384
Item no. 35	**0.477**	0.079	0.196	0.265	0.214	0.389
Item no. 36	0.074	0.127	0.269	**0.617**	0.062	0.479
Item no. 37	0.335	**0.383**	-0.108	0.325	0.348	0.497
Item no. 38	0.382	0.007	0.042	**0.526**	0.107	0.436
Item no. 39	0.073	0.075	0.181	-0.020	**0.673**	0.497
Item no. 40	**0.461**	0.206	-0.002	0.003	0.377	0.397
Item no. 41	**0.434**	0.382	0.175	0.317	-0.072	0.470
Item no. 42	0.274	0.160	**0.453**	0.392	0.023	0.460
Item no. 43	**0.646**	0.140	0.334	0.279	0.005	0.627
Item no. 44	**0.609**	0.101	0.224	0.369	0.067	0.572
Item no. 45	0.264	0.280	**0.344**	0.238	0.153	0.347
Item no. 46	-0.010	0.167	**0.343**	0.101	0.001	0.156
Item no. 47	0.220	0.143	0.037	**0.392**	0.128	0.240

Item no. 48	0.145	0.067	0.244	0.086	**0.451**	0.295
Item no. 49	-0.003	0.352	0.272	0.167	**0.365**	0.359
Item no. 50	0.198	0.203	0.177	-0.074	**0.462**	0.331
Eigenvalues after rotation	6.405	3.813	3.706	3.692	3.221	
% explained variance	12.809	7.626	7.411	7.384	6.442	
Cumulative %	12.809	20.436	27.847	35.231	41.673	
Cronbach's alpha	0.892	0.766	0.787	0.709	0.704	

Kaiser-Meyer-Olkin´s test: 0.894
Bartlett´s test of sphericity: p<0.001

Appendix

Roman numerals indicate the original subscale of the items.

F1, Learning and motivation
1   I am encouraged to participate in class I
2   The teachers are knowledgeable II
3   There is a good support system for students who get stressed V
7   The teaching is often stimulating I
12   This school is well time-tabled IV
13   The teaching is student centred I
14   I am rarely bored on this course V
16   The teaching is sufficiently concerned to develop my competence I
20   The teaching is well focused I
21   I feel I am being well prepared for my profession III
22   The teaching is sufficiently concerned to develop my confidence I
24   The teaching time is put to good use I
35   I find the experience disappointing IV
40   The teachers are well prepared for their classes II
41   My problem solving skills are being well developed here III
43   The atmosphere motivates me as a learner IV
44   The teaching encourages me to be an active learner I

F2, Communication
6   The teachers are patient with patients II
11   The atmosphere is relaxed during the ward teaching IV
18   The teachers have good communication skills with patients II
29   The teachers are good at providing feedback to students II
30   There are opportunities for me to develop interpersonal skills IV
31   I have learned a lot about empathy in my profession III
32   The teachers provide constructive criticism here II
37   The teachers give clear examples II

F3, Psychosocial situation
4   I am too tired to enjoy this course V
15   I have good friends in this school V
19   My social life is good V
23   The atmosphere is relaxed during lectures IV
28   I seldom feel lonely V
33   I feel comfortable in class socially IV
34   The atmosphere is relaxed during seminars/tutorials IV
42   The enjoyment outweighs the stress of studying medicine IV
45   Much of what I have to learn seems relevant to a career in medicine III

46   My accommodation is pleasant V

F4, Teaching organisation and progression
5   Learning strategies which worked for me before continue to work for me now III
10   I am confident about my passing this year III
26   Last year's work has been a good preparation for this year's work III
27   I am able to memorise all I need III
36   I am able to concentrate well IV
38   I am clear about the learning objectives of the course I
47   Long term learning is emphasised over short term I

F5, Bad teaching
8   The teachers ridicule the students II
9   The teachers are authoritarian II
17   Cheating is a problem in this school IV
25   The teaching over-emphasises factual learning I
39   The teachers get angry in class II
48   The teaching is too teacher-centred I
49   I feel able to ask the questions I want IV
50   The students irritate the teachers II