The Dundee Ready Educational Environment Measure: A Confirmatory Factor Analysis in a Sample of Malaysian Medical Students

Muhamad Saiful Bahri Yusoff
Medical Education Department
School of Medical Sciences
Universiti Sains Malaysia
16150 Kubang Kerian, Kota Bharu, Kelantan
Malaysia

Abstract

Introduction*: The validity and reliability of the DREEM inventory was established across educational settings. It has been translated into various languages and claimed as a 'cultural-free tool' to measure the educational climate at educational institutions. To the author's knowledge, none of the articles reported its validity and reliability among Malaysian medical students.*

Objective*: This study aims to evaluate the construct validity of DREEM using confirmatory factor analysis, as well as its internal consistency in a sample of Malaysian medical students.*

Method*: A cross-sectional study was carried out on 656 medical students from first, third and fifth year. Confirmatory factor analysis was done using AMOS 19 to assess the construct validity. Item reduction was carried out, based on the modification indices, standardized residual covariance, and standardized factor loadings to select which items are fit to remain in the best model fit. Reliability analysis was performed using SPSS 18 to assess internal consistency of DREEM.*

Result*: A total of 511 (77.9%) completely responded to the DREEM inventory. The proposed five-factor structure of DREEM failed to demonstrate model fit (X^2 = 4650.79, RMSEA = 0.076, RMR = 0.057, GFI = 0.693, AGFI = 0.667, CFI = 0.710, NFI = 0.648, RFI = 0.633, IFI = 0.711, TLI = 0.698). The five-factor structure of the shortened DREEM demonstrated model fit (X^2 = 297.90, RMSEA = 0.058, RMR = 0.027, GFI = 0.935, AGFI = 0.909, CFI = 0.953, NFI = 0.928, RFI = 0.910, IFI = 0.953, TLI = 0.941). The overall Cronbach's values for the original and shortened DREEM were 0.936 and 0.921 respectively. The Cronbach's alpha values for subscales of the original DREEM ranged between 0.58 and 0.82 while the shortened DREEM ranged between 0.53 and 0.82.*

Conclusion*: This study did not support the proposed five-factor structure of the DREEM. The shortened version demonstrated good degree of goodness of fit with the proposed structure and was found as reliable as the original DREEM. Continued research is required to verify and maximize the psychometric credentials of the DREEM across institutions and nationalities.*

Keywords: Validity, Reliability, Educational Environment, DREEM, Medical Students

Introduction

It is widely agreed among medical educators that an optimal educational climate is an important factor for effective learning to occur (Dent & Harden, 2009; Newble, Cannon, & Kapelis, 2001). Indeed, evaluation of the educational climate has been highlighted as key to the delivery of high quality medical education (Dent & Harden, 2009; Newble, et al., 2001). Therefore, to conduct such evaluation, a valid and reliable tool is vital.

Over the past 15 years, medical and allied health educators across places and educational settings have widely used the Dundee Ready Educational Environment Measure (DREEM) to appraise their institutions' educational climate (Al-hazimi, Al-hyiani, & Roff, 2004; Al-Hazimi, Zaini, et al., 2004; Arzuman, Yusoff, & Chit, 2010; Roff, et al., 1997; Said, Rogayah, & Arzuman, 2009; Thomas, Abraham, Alexander, & Ramnarayan, 2009; Varma, Tiyagi, & Gupta, 2005) Please place references in date order throughout the paper.

This valuable tool was originally designed in English (Roff, et al., 1997) and has been translated into various languages such as Swedish, Greek and Spanish (Dimoliatis, Vasilaki, Anastassopoulos, Ioannidis, & Roff, 2010; Jakobsson, Danielsen, & Edgren, 2011; Riquelme, et al., 2009). These papers have shown that DREEM is internationally accepted as a useful tool to provide feedback on strengths and weaknesses of the educational climate at particular educational institutions. One of important implications of DREEM is that it provides a standardized way for international comparisons between medical schools as well as allowing them to benchmark their educational climate (Hammond, O'Rourke, Kelly, Bennett, & O'Flynn, 2012). In addition, it may locate areas of concern shared by the majority of students that might be unintentionally neglected by educators.

Validity is broadly described as the ability of a measurement to measure attributes that it intended to measure (Streiner & Norman, 2008). The initial psychometric evaluation that was carried out by its developer showed that DREEM is a valid tool to measure educational environments. However, three previous studies (Dimoliatis, et al., 2010; Hammond, et al., 2012; Jakobsson, et al., 2011) reported that confirmatory factor analysis did not support the five-factor structure claimed by the DREEM developer. These studies concluded that the construct validity of DREEM was not well supported. These psychometric shortcomings do invite further inspection on the validity aspects of DREEM (Hammond, et al., 2012). To the author's knowledge, no article has reported on the validity of DREEM among Malaysian medical students.

Reliability is broadly described as the consistency or reproducibility of a measurement over time and occasions and it can be gauged in the form of internal consistency and stability (Streiner & Norman, 2008). The internal consistency of a tool is commonly measured and based on a single administration while the stability of a tool is measured based on multiple administrations on different occasions or time (Streiner & Norman, 2008). The DREEM has been reported to have a high level of internal consistency with the overall Cronbach's alpha coefficient being more than 0.7 (Dimoliatis, et al., 2010; Hammond, et al., 2012; Jakobsson, et al., 2011; Khan, Tabasum, Yousafzai, & Fatima, 2011; Riquelme, et al., 2009; Roff, et al., 1997). It was also found to have a high level of stability with a test-retest correlation coefficient of more than 0.8 (Dimoliatis, et al., 2010). Apart from this instance, none of the articles have so far reported its reliability among Malaysian medical students.

This study aimed to evaluate the psychometric properties of DREEM in a sample of Malaysian medical students. This study was designed to answer four questions:

1) Is the original version of DREEM a valid tool to measure the educational climate in a sample of Malaysia medical students?
2) Is the original version of DREEM a reliable tool to measure the educational climate in a sample of Malaysia medical students?
3) What is the best fit model of DREEM to measure the educational climate in the studied population?
4) Does its internal consistency vary across years of study?

The author hypothesized that DREEM demonstrated a high level of overall internal consistency across years of study; however, its construct validity will differ from the original construct proposed by the DREEM developers.

Method

A cross sectional study was carried and purposive sampling method was applied. Based on the best practice of sample size calculation for a validation study, 10 samples per item were considered adequate to obtain a significant result (Costello & Osborne, 2005). After considering the 20% dropout rate, the required sample size for this study was 625. Researchers selected first, third and fifth year medical students (i.e. a total number of 656) in a public medical school as study subjects.

Data was collected by a guided self-administered questionnaire during a face-to-face session. Informed consent was obtained from the respondents prior to the questionnaire administration. Completion of the DREEM inventory was voluntary and students were informed that not returning the inventory would not affect their progress in the medical course. The inventory was immediately returned after completion . Data was analysed by Statistical Package for Social Sciences (SPSS) version 18 and Analysis of Moment Structure software version 19 (AMOS 19). The DREEM inventory was developed as a tool to measure educational climate at educational institutions (McAleer & Roff, 2001; Roff, et al., 1997) and was claimed as a 'cultural-free' instrument (Roff & McAleer, 2001).

There are 50 items measuring five aspects of the educational environment based on students' perception, which include students' perception of learning (SPoL), students' perception of teaching (SPoT), students' academic self-perception (SASP), students' perception of atmosphere (SPoA) and students' social self-perception (SSSP). Each item is rated based on five Likert-scales range between 0 and 4 (0 = strongly disagree, 1 = disagree, 2 = unsure, 3 = agree and 4 = strongly agree). There are 9 negative items that must be scored in a reverse manner prior to analysis and interpretation; item 4, 8, 9, 17, 25, 35, 39, 48 and 50 (Roff, et al., 1997). It has been translated in various languages and the reported overall Cronbach's alpha coefficient ranges between 0.89 and 0.93 (Dimoliatis, et al., 2010; Hammond, et al., 2012; Jakobsson, et al., 2011; Khan, et al., 2011; Riquelme, et al., 2009). The original version of DREEM was used in this study.

The confirmatory factor analysis was performed using AMOS 19. The measurement of model fit with the data was checked with model chi-square goodness-of-fit, and approximate fit indices (Piaw, 2009). Insignificant model chi-square goodness-of-fit (set at 0.05) signifies model fit. For approximate fit indexes, Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Normed fit index (NFI), relative fit index (RFI), incremental fit index (IFI), Tucker-Lewis fit index (TFI) and comparative fit index (CFI) of above 0.9 would indicate model fit (Arbuckle, 1995; Brown, 2006; Kline, 2010; Piaw, 2009). For another approximate fit index, root mean square error of approximation (RMSEA) value less than 0.08 and root mean squared residual (RMR) value less than 0.05 would signify reasonable model fit (Brown, 2006; Piaw, 2009; Stevens, 2009). Significance of standardized regression weighted (standardized loading factor) estimates signify that the indicator variables are significant and representative of their latent variable (Brown, 2006). Significant of estimates of correlations indicates significant two-way correlation between specified variables. Modification indices (MI) suggested correlations between variables and the respective reductions in chi-square values should these correlations added to the model (Brown, 2006). Standardized residual covariances (SRC) is used to estimate a standard normal distribution if the model is correct, so, if the model is correct, most of items should have an SRC value of less than two in absolute value (Arbuckle, 1995; Brown, 2006). So, MI, SRC and standardized regression weighted were used an indicators to select which items fit to be remained in the model (Brown, 2006). Though reduction in chi-square values would improve model fit, following the suggestions in MI, SRC and standardized regression weighted should be based on literature review or theoretical basis (Kline, 2010; Piaw, 2009).

Reliability analysis was applied to determine the internal consistency of the DREEM inventory. Internal consistency of items was evaluated by the Cronbach's alpha, corrected item-total correlation (CITC) and Cronbach's alpha if item deleted (CAID) values. They were analysed by Statistical Package for Social Sciences (SPSS) version 18. The items were considered to represent an acceptable level of internal consistency if the Cronbach's alpha value within 0.5 to 0.7 and a good level if the Cronbach's alpha value more than 0.7 (Nunally, 1978; Streiner & Norman, 2008). An item is considered to highly contributed to the measured construct if CITC value more than 0.3 and CAID value decreased (Yusoff, Rahim, & Yaacob, 2010).

Result

A total of 511 (77.9%) completely responded to the 50 statements of DREEM. Most of them were female (61.1%), third year (38.4%) and Malay (52.1%) medical students (Table 1).

Table 1: Demographic profile of participants

Variables		Frequency (%) (N=511)
Year of study	First year	156 (30.5)
	Third year	196 (38.4)
	Fifth year	159 (31.1)
Sex	Male	175 (34.2)
	Female	312 (61.1)
	Missing data	24 (4.7)
Race	Malay	266 (52.1)
	Chinese	170 (33.3)
	Indian	43 (8.4)
	Other	7 (1.4)
	Missing data	25 (4.9)

Confirmatory factor analysis (CFA) showed that the one factor model of DREEM, consisting 50 items (i.e. the original DREEM) was not fit, indicating it was a multidimensional instrument (Table 2). On further CFA, it appeared that the proposed five factor structure of the original DREEM (Roff, et al., 1997) was not fit, since all the goodness of fit indices did not signify a model fit. Item reduction was done based on MI, SRC and standardized regression weighted values to select which DREEM items should remain in the model (Brown, 2006). As shown in the Table 2, the five factor model of DREEM consisting of 17 items (i.e. the shortened DREEM) was found to be model fit, since all the goodness of fit indices signify a model fit.

Standardized regression weighted values (i.e. standardized factor loading) for the proposed five-factor structure of the original DREEM ranged between 0.03 and 0.79, suggesting that certain items did not represent the construct being measured. Whereas, for the best fit model (i.e. model G), the standardized factor loadings ranged between 0.46 and 0.81 (figure 1), indicating that all items contributed highly to the constructs being measured. The majority of standardized correlation coefficients (r) between the five domains were more than 0.9, except the correlation between SPoA-SPoL (r=0.88), SSSP-SPoT (r=0.84), SPoL-SASP (r=0.86) and SSSP-SASP (r=0.88) (Figure 1), suggesting that they might be assessing similar constructs (Brown, 2006). The inter-item correlation coefficients ranged between 0.24 and 0.66, indicating moderate correlation between the 17 items; it reflects an acceptable degree of overlapping between the items.

Table 2: The results of confirmatory factor analysis

USMaP-i model	X^2 – statistic (df)	*p*-value	Goodness of fit indices								
			RMSEA	RMR	GFI	AGFI	CFI	NFI	RFI	IFI	TLI
Model A: One factor model (50 items)	4650.79 (1175)	< 0.001	0.076	0.057	0.693	0.667	0.710	0.648	0.633	0.711	0.698
Model B: Original five factors model (50 items)	4475.96 (1165)	< 0.001	0.075	0.057	0.709	0.681	0.724	0.661	0.644	0.725	0.710
Model C: Five factors model (42 items)	3266.32 (809)	< 0.001	0.077	0.042	0.745	0.716	0.779	0.727	0.710	0.780	0.765
Model D: Five factors model (38 items)	2947.52 (655)	< 0.001	0.083	0.042	0.745	0.712	0.785	0.741	0.722	0.786	0.769
Model E: Five factors model (34 items)	2476.55 (517)	< 0.001	0.086	0.043	0.756	0.720	0.798	0.759	0.739	0.799	0.781
Model F: Five factors model (28 items)	1705.02 (340)	< 0.001	0.089	0.041	0.792	0.752	0.825	0.792	0.769	0.826	0.806
Model G: Five factors model (17 items)	**297.90 (109)**	**< 0.001**	**0.058**	**0.027**	**0.935**	**0.909**	**0.953**	**0.928**	**0.910**	**0.953**	**0.941**

Best fitting model in bold. Model C (item 8, 9, 14, 17, 35, 39, 48, and 50 were removed from Model B); Model D (item 4, 10, 25, and 28 were removed from Model C); Model E (item 1, 2, 5, and 34 were removed from Model D); Model F (item 7, 11, 12, 13, 23, and 47 were removed from Model E); Model G (item 15, 16, 18, 21, 27, 29, 31, 32, 36, 38, and 44 were removed from Model F).

Reliability analysis shows that the overall Cronbach's alpha values for the original and shortened DREEM were 0.936 and 0.921 respectively (Table 3). Both versions showed a high level of internal consistency in measuring students' perception of educational climate. The Cronbach's alpha values for the five subscales of the original DREEM ranged between 0.58 and 0.82 while for the shortened DREEM ranged between 0.53 and 0.82 (table 3). The subscales for both versions showed acceptable to high level of internal consistency (Nunally, 1978; Streiner & Norman, 2008) in measuring the five aspects of students' perception of educational climate.

Reliability analysis showed that the internal consistency of DREEM for both versions varied across years of study.

The Cronbach's alpha values across years of study for subscales of the original DREEM ranged between 0.53 and 0.87 while for the shortened DREEM ranged between 0.43 and 0.92 (table 3). The SSSP subscale demonstrated the lowest level of internal consistency compared to other subscales.

Reliability analysis shows that the original 50-items DREEM had CITC values ranged between 0.003 and 0.727. This result reflected that certain items p contribute poorly to the constructs being measured as the CITC values less than 0.30 (Yusoff, et al., 2010); item 8, 9,14, 17, 35, 39, 48, 50. In contrast, the CITC values for the 17-item DREEM ranged between 0.458 and 0.751. It indicates that all of the items in the shortened DREEM highly contributed to the constructs being measured as the CITC values more than 0.3 (Yusoff, et al., 2010).

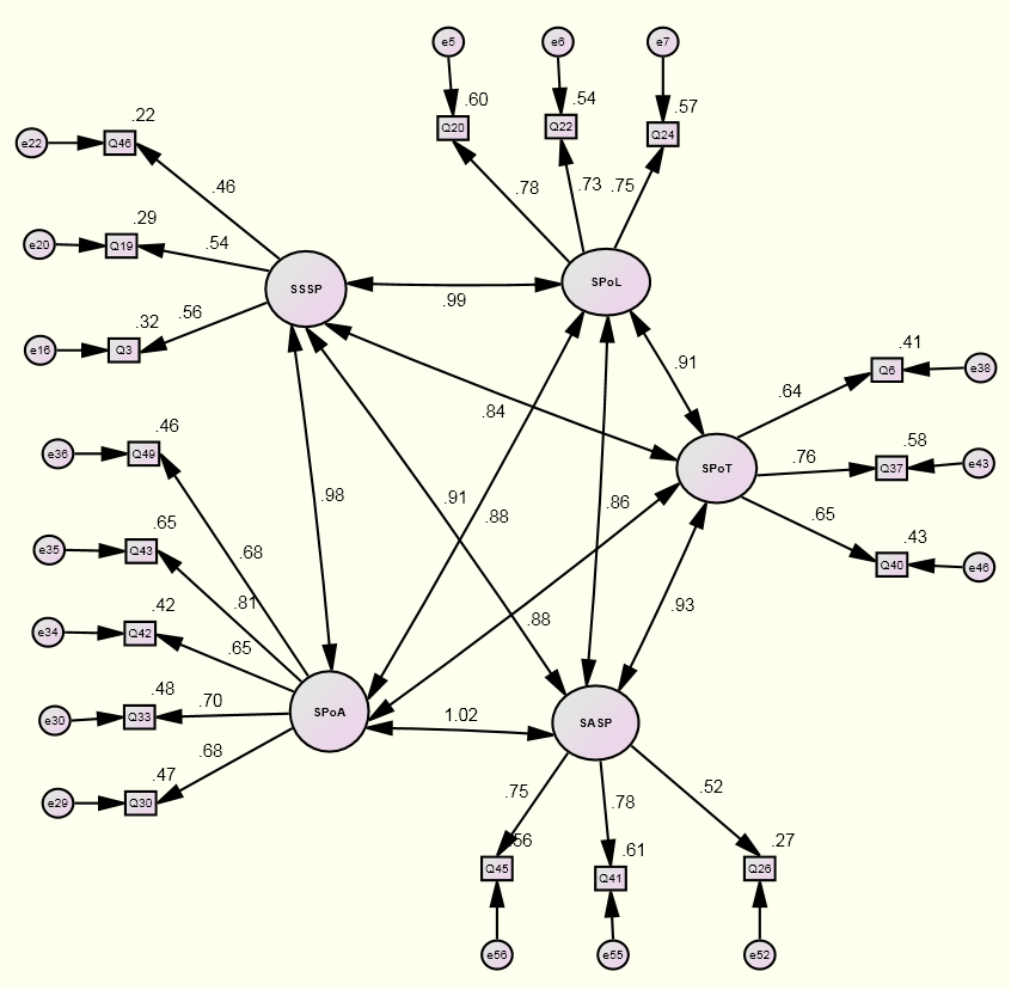

Figure 1: Standardized factor loading for the best fit model of DREEM.

Table 3: Cronbach's alpha values of the original and best fit model of DREEM.

Domain	Cronbach's alpha							
	The original model of DREEM				The best fit model of DREEM			
	Year 1	Year 3	Year 5	Overall year	Year 1	Year 3	Year 5	Overall year
Overall	0.946	0.921	0.917	0.936	0.953	0.890	0.890	0.921
SPoL	0.832	0.745	0.723	0.785	0.904	0.747	0.692	0.798
SPoT	0.719	0.737	0.661	0.732	0.850	0.593	0.644	0.720
SASP	0.867	0.749	0.785	0.814	0.779	0.582	0.685	0.690
SPoA	0.860	0.762	0.796	0.820	0.919	0.728	0.772	0.821
SSSP	0.621	0.529	0.530	0.577	0.610	0.520	0.433	0.533

Table 4: Reliability analysis and mean score of the 50 items of DREEM according to the five domains.

No. & statement	CITC	CAID
Students' Perception of Learning (SPoL)		
Q1. I am encouraged to participate during teaching sessions	0.466	0.935
Q7. The teaching is often stimulating	0.606	0.934
Q13. The teaching is student-centred	0.577	0.934
Q16. The teaching helps to develop my competence	0.613	0.934
Q20. The teaching is well-focused	0.713	0.933
Q22. The teaching helps to develop my confidence	0.667	0.933
Q24. The teaching time is put to good use	0.700	0.933
Q25. The teaching over-emphasizes factual learning*	-0.348	0.940
Q38. I'm clear about the learning objectives of the course	0.613	0.934
Q44. The teaching encourages me to be an active learner	0.692	0.933
Q47. Long-term learning is emphasized over short-term learning	0.532	0.934
Q48. The teaching is too teacher-centred*	0.003	0.938
Students' Perception of Teachers (SPoT)		
Q2. The teachers are knowledgeable	0.476	0.935
Q6. The teachers adopt a patient-centred approach to consulting	0.592	0.934
Q8. The teachers ridicule the students*	0.182	0.937
Q9. The teachers are authoritarian*	0.028	0.938
Q18. The teachers have good communication skills with patients	0.621	0.934
Q29. The teachers are good at providing feedback to students	0.685	0.933
Q32. The teachers provide constructive criticism here	0.492	0.934
Q37. The teachers give clear examples	0.670	0.933
Q39. The teachers get angry in teaching*	0.149	0.937
Q40. The teachers are well-prepared for their teaching sessions	0.568	0.934
Q50. The students irritate the teachers*	0.215	0.937
Students' Academic Self-Perception (SASP)		
Q5. Learning strategies that worked for me before continue to work for me now	0.464	0.935
Q10. I am confident about my passing this year	0.399	0.935
Q21. I fell I am being well prepared for my profession	0.571	0.934
Q26. Last year's work has been a good preparation for this year's work	0.521	0.934
Q27. I am able to memorize all I need	0.541	0.934
Q31. I have learnt a lot about empathy in my profession	0.585	0.934
Q41. My problem-solving skills are being well developed here	0.690	0.933
Q45. Much of what I have to learn seems relevant to a career in healthcare	0.670	0.934
Students' Perception of Atmosphere (SPoA)		
Q11. The atmosphere is relaxed during ward teaching	0.585	0.934
Q12. This school is well time-tabled	0.615	0.933
Q17. Cheating is a problem in this school*	0.129	0.938
Q23. The atmosphere is relaxed during lectures	0.606	0.934
Q30. There are opportunities for me to develop my interpersonal skills	0.677	0.933
Q33. I feel comfortable in class socially	0.670	0.933
Q34. The atmosphere is relaxed during class/seminars/tutorials	0.555	0.934
Q35. I find the experience disappointing*	0.089	0.938
Q36. I am able to concentrate well	0.603	0.934
Q42. The enjoyment outweighs the stress of the course	0.599	0.934
Q43. The atmosphere motivates me as a learner	0.727	0.933
Q49. I feel able to ask the questions I want	0.627	0.933
Students' Social Self-Perception (SSSP)		
Q3. There is a good support system for students who get stressed	0.560	0.934
Q4. I am too tired to enjoy the course*	0.320	0.936
Q14. I am rarely bored in this course	0.125	0.938
Q15. I have good friends in this course	0.442	0.935
Q19. My social life is good	0.499	0.934
Q28. I seldom feel lonely	0.355	0.936
Q46. My accommodation is pleasant	0.408	0.935

* Negative item; CITC = Corrected Item-Total Correlation; CAID = Cronbach's Alpha if Item Deleted; SD = Standard deviation

Table 5: Reliability analysis on individual item of the best fit DREEM model.

Domain	No.	Statement	CITC	CAID
SPoL	Q20	The teaching is well-focused	0.700	0.914
	Q22	The teaching helps to develop my confidence	0.676	0.914
	Q24	The teaching time is put to good use	0.689	0.914
SPoT	Q6	The teachers adopt a patient-centred approach to consulting	0.582	0.917
	Q37	The teachers give clear examples	0.673	0.915
	Q40	The teachers are well-prepared for their teaching sessions	0.559	0.918
SASP	Q26	Last year's work has been a good preparation for this year's work	0.513	0.919
	Q41	My problem-solving skills are being well developed here	0.733	0.914
	Q45	Much of what I have to learn seems relevant to a career in healthcare	0.677	0.915
SPoA	Q30	There are opportunities for me to develop my interpersonal skills	0.649	0.915
	Q33	I feel comfortable in class socially	0.751	0.913
	Q42	The enjoyment outweighs the stress of the course	0.611	0.916
	Q43	The atmosphere motivates me as a learner	0.652	0.915
	Q49	I feel able to ask the questions I want	0.643	0.916
SSSP	Q3	There is a good support system for students who get stressed	0.554	0.919
	Q19	My social life is good	0.486	0.920
	Q46	My accommodation is pleasant	0.458	0.922

CITC = Corrected Item-Total Correlation CAID = Cronbach's Alpha if Item Deleted

Discussion

The author supports the view that the DREEM inventory is a very useful tool for recognizing the educational climate at educational institutions and its widespread use verifies the need for such an instrument (Hammond, et al., 2012). However, our data did not support the five-factor structure of DREEM consisting of 50 items measuring the educational climate. Our finding seems to be consistent with previous studies reporting that its construct validity was not well supported by the empirical data (Dimoliatis, et al., 2010; Hammond, et al., 2012; Jakobsson, et al., 2011). As has been suggested by the previous studies, removal of certain items might improve the goodness of fit of the five-factor structure (Dimoliatis, et al., 2010; Hammond, et al., 2012; Jakobsson, et al., 2011). It is worth mentioning that our findings are based on Malaysian medical students; even so it is unlikely these weaknesses could be attributed to language factors since items of the DREEM were constructed using simple and comprehensible English sentences. It appears obvious that the hypothetical five-factor structure proposed by the DREEM developers is not well supported and perhaps continued efforts focusing on revising and establishing its psychometric properties are required (Dimoliatis, et al., 2010; Hammond, et al., 2012; Jakobsson, et al., 2011).

The author conducted further analysis as an attempt to propose a shortened version of DREEM that met the requirements for a model fit, even though this is not recommended by a previous study (Hammond, et al., 2012). We found that the one factor structure of the 50-item DREEM (i.e. the original DREEM) failed to demonstrate model fit; this suggests that DREEM measures multiple constructs. Our data found that the five-factor structure of the DREEM that consists of 17 items (i.e. the shortened DREEM) demonstrated a model fit, since all of the goodness of fit indices were significant for model fit, except the chi-square value. Based on these finding, it seems that after removal of certain items (perhaps 'poorly represent' the constructs being measured), the five-factor structure proposed by the DREEM developers was supported. However, the 33 items that were removed during CFA perhaps need to be revisited and revised because they might represent important and meaningful constructs of educational climates as mentioned by previous studies (Dimoliatis, et al., 2010; Jakobsson, et al., 2011). Even so, most of the standardized correlations values between the five constructs were more than 0.9, suggesting there were significant overlapping and lack of discrimination between them (Brown, 2006). These findings support the view that either the items need to be restructured to fit the proposed model or the model itself needs to be revised and reconsidered (Hammond, et al., 2012). On the other hand, these findings also suggest there are repetitions of similar items that assessing similar constructs that compromise the construct validity of the DREEM.

In general our data supports a high level of reliability for the DREEM inventory as the overall Cronbach's alpha values were more than 0.7 (Nunally, 1978; Streiner & Norman, 2008), which is in line with previous studies (Hammond, et al., 2012; Jakobsson, et al., 2011; Khan, et al., 2011; Roff, et al., 1997). Our data demonstrated that the Cronbach's alpha values for the DREEM subscales varied across years of study, indicating their internal consistency was compromised once years of study were taken into consideration. As a result, this may compromise the ability of DREEM to give similar results if a similar population is being studied at different times and occasions. Similar concerns were echoed by a previous study (Hammond, et al., 2012). In addition, our data showed that the shortened DREEM was found as reliable as the original DREEM. Apart from that, CITC values of items in the shortened DREEM were more consistently showing high degree of contributions to the internal consistency than the items in the original DREEM (Table 4 and Table 5). Perhaps one of possible reasons for this could be due to certain items measuring different constructs. These findings suggest that a substantial number of items in the original DREEM should be revised and restructured to improve the internal consistency of the DREEM subscales. Perhaps there are more than five constructs being measured by the DREEM as was noted by a previous study (Dimoliatis, et al., 2010)

In conclusion, our findings did support the reliability, but not the construct validity, of the DREEM inventory. Nevertheless, our study has also several limitations that need to be considered for interpretation.

Firstly; our sample was confined to a medical school in Malaysia that might not represent the Malaysian medical student distribution across medical schools. Secondly; the sampling method applied was non-probability that may lead to sampling bias, that may compromise the results obtained.

Our study however has several strengths that could be used to verify the authenticity of our data. Firstly; samples were selected across years of study that may be considered as representing students from different stages of medical education. Secondly; the sample size was calculated based on the recommended ratio of subjects per item. Thirdly; authentic and rigorous analyses were applied in this study to evaluate the psychometric properties of the instrument. Based on these limitations and strengths, data reported in this study should be interpreted with caution and any attempt to generalize the findings should be performed in context.

Conclusion

Our study did not support the proposed five-factor structure of DREEM. The shortened version has demonstrated a better fit with the proposed model and was found to be as reliable as the original version. Continued research is required to verify and maximize the psychometric credentials of the DREEM across institutions and nationalities.

References

Al-hazimi, A., Al-hyiani, A., & Roff, S. (2004). Perceptions of the educational environment of the medical school in King Abdul Aziz University, Saudi Arabia. *Medical Teacher, 26*(6), 570-573.

Al-Hazimi, A., Zaini, R., Al-Hyiani, A., Hassan, N., Gunaid, A., Ponnamperuma, G., et al. (2004). Educational environment in traditional and innovative medical schools: A study in four undergraduate medical schools. *EDUCATION FOR HEALTH-ABINGDON-CARFAX PUBLISHING LIMITED-. 17*(2), 192-203.

Arbuckle, J. L. (1995). IBM SPSS Amos 19 user's guide Retrieved from http://www.amosdevelopment.com/download/amos.pdf

Arzuman, H., Yusoff, M. S. B., & Chit, S. P. (2010). Big Sib Students' Perceptions of the Educational Environment at the School of Medical Sciences, Universiti Sains Malaysia, using Dundee Ready Educational Environment Measure (DREEM) Inventory. *Malaysian Journal of Medical Sciences, 17*(3), 40-47.

Brown, T. A. (2006). *Confirmatory factor analysis for applied sciences*. New York: The Guilford Press.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.

Dent, J. A., & Harden, R. M. (2009). *A practical guide for medical teachers* (3rd ed.): Elsevier Churchill Livingstone.

Dimoliatis, I., Vasilaki, E., Anastassopoulos, P., Ioannidis, J., & Roff, S. (2010). Validation of the Greek translation of the Dundee ready education environment measure (DREEM). *Education for Health, 23*(1), 348.

Hammond, S. M., O'Rourke, M., Kelly, M., Bennett, D., & O'Flynn, S. (2012). A psychometric appraisal of the DREEM. *BMC Medical Education, 12*(1), 2.

Jakobsson, U., Danielsen, N., & Edgren, G. (2011). Psychometric evaluation of the Dundee Ready Educational Environment Measure: Swedish version. *Medical Teacher, 33*(5), 267-274.

Khan, J. S., Tabasum, S., Yousafzai, U. K., & Fatima, M. (2011). DREEM ON: Validation of the Dundee Ready Education Environment Measure in Pakistan. *J Pak Med Assoc, 61*(9), 885-888.

Kline, R. B. (2010). *Principles and Practice of Structural Equation Modeling* (3rd ed.). New York: Guilford Publications.

McAleer, S., & Roff, S. (2001). A practical guide to using the Dundee Ready Education Environment Measure (DREEM). *AMEE Medical Education Guide*(23), 29-33.

Newble, D., Cannon, R. A., & Kapelis, Z. A. (2001). *A handbook for medical teachers*: Kluwer Academic Publishers.

Nunally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw Hill.

Piaw, C. Y. (2009). *Statistik Penyelidikan Lanjutan* (Vol. 5). Malaysia: McGraw Hill.

Riquelme, A., Oporto, M., Oporto, J., Mendez, J. I., Viviani, P., Salech, F., et al. (2009). Measuring students' perceptions of the educational climate of the new curriculum at the Pontificia Universidad Catalica de Chile: performance of the Spanish translation of the Dundee Ready Education Environment Measure (DREEM). *Education for health (Abingdon, England), 22*(1), 112.

Roff, S., & McAleer, S. (2001). What is educational climate? *Medical Teacher, 23*(4), 333.

Roff, S., McAleer, S., Harden, R. M., Al-Qahtani, M., Ahmed, A. U., Deza, H., et al. (1997). Development and validation of the Dundee ready education environment measure (DREEM). *Medical Teacher, 19*(4), 295-299.

Said, N. M., Rogayah, J., & Arzuman, H. (2009). A study of learning environments in the Kulliyyah (Faculty) of Nursing, International Islamic University Malaysia. *Malaysian Journal of Medical Sciences, 16*(4), 15-24.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York: Taylor & Francis Group.

Streiner, L. D., & Norman, G. R. (2008). *Health Measurement Scales: A Practical Guide to Their Development and Use* (4th ed.). New York: Oxford University Press.

Thomas, B. S., Abraham, R. R., Alexander, M., & Ramnarayan, K. (2009). Students' perceptions regarding educational environment in an Indian dental school. *Medical Teacher, 31*(5), 185-188.

Varma, R., Tiyagi, E., & Gupta, J. (2005). Determining the quality of educational climate across multiple undergraduate teaching sites using the DREEM inventory. *BMC Medical Education, 5*(1), 8.

Yusoff, M. S. B., Rahim, A. F. A., & Yaacob, M. (2010). The Development and Validity of the Medical Student Stressor Questionnaire (MSSQ). *ASEAN Journal of Psychiatry, 11*(1).