

Development of an assessment tool to measure students' perceptions of respiratory care education programs: Item generation, item reduction, and preliminary validation

Ghazi Alotaibi, Adel Youssef¹

Departments of Respiratory Care, ¹Health Information Management and Technology, College of Applied Medical Sciences, University of Dammam, Kingdom of Saudi Arabia

Address for correspondence: Dr. Ghazi Alotaibi, Department of Respiratory Care, College of Applied Medical Sciences, University of Dammam, P.O. Box 40269, Dammam 31952, Kingdom of Saudi Arabia.. E-mail: galotaibi@ud.edu.sa

ABSTRACT

Objectives: Students who perceived their learning environment positively are more likely to develop effective learning strategies, and adopt a deep learning approach. Currently, there is no validated instrument for measuring the educational environment of educational programs on respiratory care (RC). The aim of this study was to develop an instrument to measure students' perception of the RC educational environment. **Materials and Methods:** Based on the literature review and an assessment of content validity by multiple focus groups of RC educationalists, potential items of the instrument relevant to RC educational environment construct were generated by the research group. The initial 71 item questionnaire was then field-tested on all students from the 3 RC programs in Saudi Arabia and was subjected to multi-trait scaling analysis. Cronbach's alpha was used to assess internal consistency reliabilities. **Results:** Two hundred and twelve students (100%) completed the survey. The initial instrument of 71 items was reduced to 65 across 5 scales. Convergent and discriminant validity assessment demonstrated that the majority of items correlated more highly with their intended scale than a competing one. Cronbach's alpha exceeded the standard criterion of >0.70 in all scales except one. There was no floor or ceiling effect for scale or overall score. **Conclusions:** This instrument is the first assessment tool developed to measure the RC educational environment. There was evidence of its good feasibility, validity, and reliability. This first validation of the instrument supports its use by RC students to evaluate educational environment.

Key words: Educational environment, instrument, perception, respiratory care, Saudi Arabia

INTRODUCTION

Health education has undergone fundamental changes in the last two decades. The emphasis is now student-centered rather than teacher oriented and content-centered. This change in focus has necessitated other changes in curriculum design and delivery to meet this shift.

Engagement of students in the learning process is, perhaps, the most important manifestation of the recent educational reform.^[1] In the era of quality assurance and requirements for accreditation, the curriculum is frequently evaluated to ensure optimum student learning experience. One-way of assessing students' learning experience is to collect information about how students perceive their educational environment. The 'educational environment' is everything that is happening in the academic institution.^[2,3] It is the character, spirit, and culture of the educational institution.^[4] The 'educational environment' is also defined as the climate of the institution as experienced and perceived by students. It is believed by many educational authorities that the 'educational environment' is the most significant reflection and the central component of the curriculum.^[1,5]

Access this article online	
Quick Response Code: 	Website: www.jfcmonline.com
	DOI: 10.4103/2230-8229.114770

Research studies have shown that a positive and supportive environment is essential for successful learning. Roff *et al.*^[6] found that the educational environment makes an impact on students' learning experiences and outcomes. It has also been reported that students who perceive their learning environment positively are more likely to develop effective learning strategies and adopt a deep learning approach.^[7] Many educationalists use students' perception of the learning environment as a diagnostic tool to identify weaknesses and strengths of the curriculum.^[8]

Interest in assessing the educational environment is not new. In 1958, Pace and Stern developed an assessment tool to study the educational environment of medical schools, namely medical environment index.^[9] Since then, a plethora of assessment tools have been developed to evaluate students' perceptions of their educational climate. Among the assessment tools that gained popularity is the Dundee ready educational environment measure (DREEM). DREEM was developed to assess the learning environment of medical and health-care related education programs. DREEM has been described as a "culture free" instrument.^[10] Other assessment tools for measuring the educational environment in postgraduate residency, surgery clerkship, and anesthesia have also become available.^[4,11,12]

Respiratory care (RC) is a health-care profession concerned with the assessment and treatment of patients with cardiopulmonary disorders. It usually takes 4-5 years to graduate with a Bachelor's Degree in RC Education is unique in that it comprises lectures and simulated laboratory teaching, interspersed with clinical rotations in hospitals. To assess the learning environment of such a curriculum model, we believe that existing assessment tools may be unsuitable for two reasons. First, the nature of curriculum of RC programs that amalgamate theory, hands-on skills, and clinical practice requires a specially designed instrument to take into account features and specialty-specific components of the curriculum. Secondly, some research studies have reported that the original inventory was modified to suit the specific educational situation or the cultural setting of the institution.^[13,14] Although those studies that used seven inventories had been shown to be valid and reliable, it seems that the nature of the RC profession and cultural settings may require the use of a modified or even a new assessment tool to ensure a reflection of reality.

Currently, to the best of our knowledge, there is no validated inventory dedicated to measuring the educational environment of RC Educational programs. Therefore, we conducted a research project to develop an assessment tool to measure the educational environment of RC programs.

In this publication, we present the process by which the instrument was developed and the preliminary evaluation of the psychometric properties of the new instrument.

MATERIALS AND METHODS

This project used the established principles of instrument design and evaluation proposed by Streiner and Norman:^[15] (1) item generation; (2) item reduction and preliminary assessment of the psychometric properties in the form of validity and reliability; (3) assessment of the factorial structure of the instrument. This article will report on the first two phases of the project. Figure 1 presents an overview of the overall process.

Subjects and setting

The study was conducted in the 3 RC programs in the Kingdom of Saudi Arabia. The first program (program 1) was at the College of Applied Medical Sciences, University of Dammam. This college with both male and female students is the largest in the country. The Baccalaureate program in RC started in 1999 has so far graduated approximately 400 respiratory therapists. The other two programs (programs 2 and 3) are for males only. The structures in all programs are similar consisting of 4 years of didactic, laboratory and clinical teaching followed by 1 year of internship training. All students from the three programs at all levels of the study were invited to participate in the questionnaire validation.

Questionnaire development

Initial item generation

The content of the questionnaire was developed by an interactive process conducted by the 5-lead research team of this project. Literature review was carried out by the research team to retrieve relevant published instruments, identify the common domains and explore issues relevant to the assessment tool under study. After the extensive review, the research team decided to make up the questionnaire to include five domains; perception of clinical rounds, perception of teaching and teaching environment, perception of laboratory teaching, and perception of the RC profession. The item generation stage led to the development of an initial pool of 165 items, which were then reduced to 105 items all in Arabic.

Item review

To assess the instrument's content validity, the item pool was reviewed by a panel of experts. Using three focus groups, 5 RC faculty members provided their comments on the clarity and completeness of the items and relevance to RC education. In a quantitative evaluation, 3 RC faculty members from our institution and six external RC educationalists were asked to rate the relevance of each item on the scale from

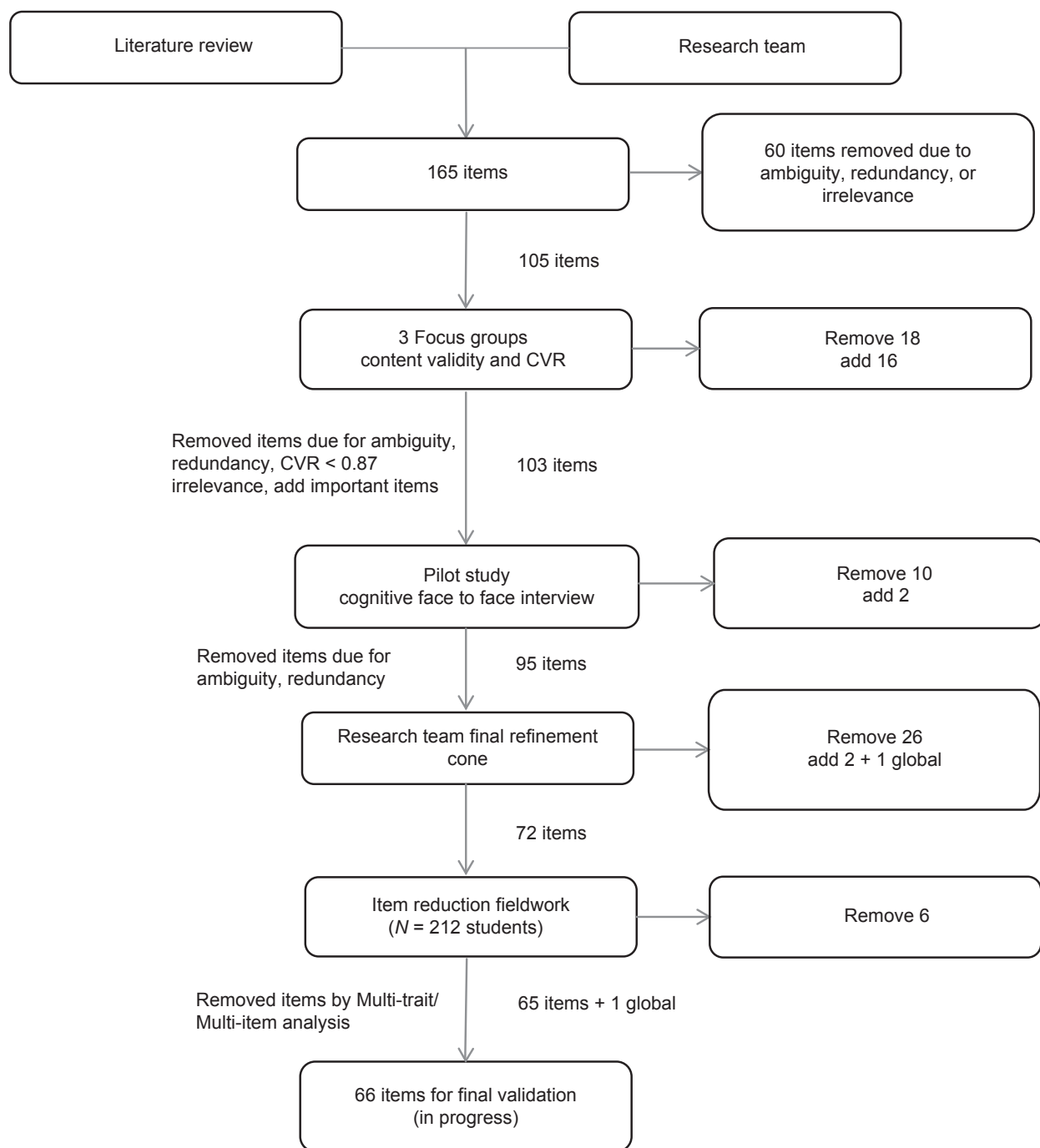


Figure 1: Overview of study main steps

one (not relevant) to four (very relevant). A content validity ratio (CVR) was then derived for each item by calculating the proportion of experts who rated the item “relevant” or “very relevant” and for the whole instrument by calculating the proportion of the total number items that were rated valid.^[16] The RC panel of experts was also asked if they felt there was any content area the questionnaire had failed to deal with. For this number of experts in the panel, values higher than 0.78 for CVR was considered satisfactory as suggested by Lawshe.^[17] The instrument with the content

valid items was then converted to a question format and calibrated on 5-point Likert scale (from strongly disagree to strongly agree) and given to 10 students (five male and five female students from the five levels of the RC program) to complete although face-to-face cognitive interviews were being conducted. The interviews used a think-aloud process to reveal students’ thought processes although completing the questionnaire, and to discover the rationale for the choice of each answer.^[18] Each student also commented on simplicity and clarity of the questions on first reading,

and the relevance of each question to the assigned domain. The research team closely examined the questionnaire for remaining redundant items. Similar items were either combined, vague ones left out or the more specific ones retained.

The item review step resulted in 72-item preliminary questionnaire (71 items plus one global satisfaction item) based on the hypothesized structure of five domains: Perception of clinical rounds (17 questions); perception of teaching and learning (26 questions); perception of program management (14 questions); perception of laboratory teaching (9 questions), and perception of RC profession (5 questions). Item design was based on a 5-point Likert scale in which '0' corresponded to "strongly disagree," 1 – "disagree," 2 – "not sure," 3 – "agree," and 4 – "strongly agree." A global scale was added as a proxy measure to assess students' overall satisfaction with the program.

Field test and item reduction

The preliminary questionnaire (72 items) was distributed by the research team to all students in the 3 RC programs in the Kingdom. None of the students were excluded. Students from each program were asked to gather in a classroom. The survey was distributed to students, who were then given a general overview of the survey. The questionnaire included instructions on the objectives of the questionnaire, guidelines for answering questions, and an assurance of anonymity. It also included questions on demographic characteristics, name of the program and the student's level in the program. The questionnaire was re-administered to a randomly selected 50 students one week after the first round to assess the test-retest reliability of the instrument.

Statistical analysis

Quality and completeness of item responses were assessed for all received questionnaires. The response distribution was examined and items with endorsement rates (percentage of respondents who checked the same response category) of >80% were considered for exclusion.^[19] The overall score of the questionnaire was obtained by adding scores from all items, and the score for each domain was obtained by adding the scores for the items in that scale.

The proportion of items and domains that were missing were calculated with acceptable values below 20%. A high number of missing items or a high percentage of missing data throughout the questionnaire could indicate that the items were either confusing or that the questionnaire layout was problematic.^[20] The rates of floor and ceiling effects were calculated as the proportion of students who obtained the lowest and the highest possible scores respectively for any of the items or domains with expected values below 30%. Pearson's correlation was used in this study.

The construct related validity was assessed by evaluating the item convergent and item discriminant validity. Item convergent validity refers to the extent to which items within a particular domain correlate with each other. Convergence was assessed by evaluating correlations between items within each domain, and also between each item and the overall sum-score for their domain when the item of interest is eliminated from the calculation of the sum-score (item total corrected for overlap). The correlation of each item with its own scale sum-score was considered satisfactory if it was >0.30.^[21] Item discriminant validity on the other hand studies the assumption that in an instrument with more than one domain, the correlation between an item and its own domain is expected to be significantly higher than the correlations of the item with other domains. Scaling success rate was calculated, as suggested by McHorney *et al.*,^[22] as the percentage of items within each domain that met item convergent and item discriminate validity criteria.

Internal consistency reliability was tested by Cronbach's alpha coefficient for each domain and for the entire questionnaire with an acceptable value of ≥ 0.70 ; however, a scale of ≥ 0.60 was regarded as acceptable in a newly developed scales.^[23] Cronbach's alpha was also examined when individual items were deleted. Items that reduced the Cronbach's alpha value of their domain were considered for exclusion.

To examine further whether the five domains measured different aspects of student satisfaction with the program, the Cronbach's alpha of each domain was compared with the domain correlation coefficient of other domains. A Cronbach's alpha of the domain that was higher than the domain's correlation with other domains, would indicate that the domain scores represented different aspects of students' satisfaction.^[21] The inter-domain correlations were also expected to be lower than 0.70.^[24]

All statistical analyses were conducted with Stata 12 (Stata Corp. College Station, TX, USA).

RESULTS

Sample characteristics

A total of 212 students (100% response) completed the 72 item questionnaire. Item response rate ranged from 97.6% to 100%. The majority of the students, 65.1% ($n = 138$) were from program 1. The remaining students came from the two other programs, 21.2% ($n = 45$) and 13.7% ($n = 29$). The students included were equally distributed among the 3 years of the program (35.8%, 30.7%, and 33.4%). Male students made up 68% of the

study population. All students were surveyed at the same time in the academic year 2010-2011 [Table 1].

Item and scale statistical characteristics

The percentage of missing items by dimension was low (range 0.5-2.4%) and scale and total scale scores could be computed for 100% of the sample. Item means were roughly similar and item standard deviations were almost equivalent and around one. Item frequency distribution was variable with an acceptable skew. Ceiling effect ranged from 38.2-76.4% in 21 items. Only one item had a floor effect at 46.2% [Table 2]. In contrast, none of the five domains had a floor or ceiling effect [Table 3].

Multi-trait/Multi-item correlation analysis

The correlation coefficients for item convergent validity ranged from 0.01-0.58. Of the 71 items, 8 had a corrected item-total correlation of <0.30 . Of the remaining 63 items, 60 had a correlation of ≥ 0.4 [Table 4]. Only one domain “perception of laboratory training” was found

to have convergent validity success rate that was $<80\%$. Furthermore, shows in Table 4, a total of 28 items out of 428 item correlations failed to show a successful discriminant validity resulting in an overall success rate of 92%.

Scale internal consistency reliability

As shown in Table 5, reliability of the scale internal consistency was adequate as determined by Cronbach's alpha. Four out of five domains exceeded the desired Cronbach's alpha of 0.70. Only “perception of laboratory teaching” domain was at the margin of the acceptable level of internal consistency reliability (0.58). Table 5 also presents inter-scale correlations (range 0.28-0.64), which was lower than Cronbach's alpha for each scale.

DISCUSSION

To the best of our knowledge, this is the first instrument for assessing educational environment of RC programs. In this paper, we focused on the first two phases of psychometric evaluation and scaling performance of the instrument under development. Further, psychometric evaluations are ongoing. Future reports will focus on more refinement of the questionnaire factor structure and reliability of test-retest.

The results of our preliminary analysis on the psychometric properties of the new instrument showed an overall satisfactory evidence of acceptability, reliability, and validity of the included questions. Good acceptability was indicated by negligible missing data. The high rate of completeness

Table 1: Students characteristics

Characteristic	RC program 1 (n=138)	RC program 2 (n=45)	RC program 3 (n=29)
Sex			
Male	70 (50.7)	45 (100)	29 (100)
Female	68 (49.3)	0 (0)	0 (0)
Year			
2 nd	49 (35.5)	27 (60.0)	0 (0)
3 rd	40 (29.0)	11 (24.4)	14 (48.3)
4 th	49 (35.5)	7 (15.6)	15 (51.7)

RC: Respiratory care

Table 2: Item level statistical characteristics

	Item [†] n	Missing values [‡] %	Ranges of item mean	Ranges of item SD	Ranges of item skewness	Floor effect		Ceiling effect	
						N [§]	%	N [#]	% ^{**}
Perception of clinical training	17	2.4	2.0, 3.3	0.6, 1.2	-1.1, 0.1	0	0	9	58.0
Perception of teaching and learning	26	1.0	2.6, 3.1	0.7, 1.2	-1.1, 0.7	0	0	3	38.2
Perception of program management	14	1.4	1.6, 3.1	0.8, 1.2	-1.3, 0.7	1	46.2	3	42.5
Perception of laboratory teaching	9	0.5	2.6, 3.5	0.5, 1.1	-1.2, 0.9	0	0	3	40.6
Perception of RC profession	5	1.9	2.7, 3.7	0.5, 1.0	-2, -0.4	0	0	3	76.4

Item 72 (global item) is not included in the analysis of this table, [†]Number of items in each domain, [‡]percent of items missing in each domain, [§]Number of items with floor effect in each domain, ^{||}Highest percent of floor effect achieved by items within the domain, [#]Number of items with ceiling effect in each domain, ^{**}Highest percent of ceiling effect achieved by items within the domain, RC: Respiratory care

Table 3: Scale level statistical characteristics

	Item [†] (n)	Mean (SD)	Range [‡]	% at floor [§]	% at ceiling
Perception of clinical training	17	49.6 (7.1)	22-65	0	0
Perception of teaching and learning	26	72.4 (11.8)	41-96	0	0
Perception of program management	14	34.3 (7.2)	13-54	0	0
Perception of laboratory teaching	9	25.9 (4.3)	14-35	0	0
Perception of RC profession	5	16.1 (2.9)	6-20	0	9.9

Item 72 (global item) is not included in the analysis of this table, [†]Number of items in each domain, [‡]Domain score range, [§]Percent of students at the lowest possible score of the domain, ^{||}Percent of students at the highest possible score of the domain, RC: Respiratory care

Table 4: Summary results of multi-trait/multi-item correlation matrix

	Item-scale convergent validity			Item-scale divergent validity		
	Range of item-scale correlations [†]	Number of item-scale correlations [‡]	Success rate (%)	Range of item correlations with other scales [§]	Number of items with higher correlation with own scale	Success rate (%)
Perception of clinical training	0.16-0.53	15/17	88	0.07-0.46	78/85	92.9
Perception of teaching and learning	0.13-0.58	24/26	92	0.02-0.49	128/130	98.5
Perception of program management	0.10-0.52	12/14	86	0.02-0.45	59/70	84.3
Perception of laboratory teaching	0.01-0.42	7/9	78	0.01-0.47	37/45	82.3
Perception of RC profession	0.37-0.57	5/5	100	0.10-0.33	25/25	100

Item 72 (global item) is not included in the analysis of this table, [†]Range of correlations between items and hypothesized domain corrected for overlap, [‡]Number items with correlations exceeding 0.30/number of correlations tested, [§]Range of correlations between items and other domains, ^{||}Number of items with higher correlation between item and hypothesized domain compared to item correlation with other domains/number of correlations tested, RC: Respiratory care

Table 5: Inter-scale correlations and internal consistency reliability

	Inter-scale correlations					Cronbach alpha
	Perception of clinical training	Perception of teaching and learning	Perception of program management	Perception of laboratory teaching	Perception of RC profession	
Perception of clinical training	1	-	-	-	-	0.83
Perception of teaching and learning	0.62	1	-	-	-	0.88
Perception of program management	0.58	0.64	1	-	-	0.88
Perception of laboratory teaching	0.61	0.57	0.58	1	-	0.58
Perception of RC profession	0.36	0.31	0.30	0.28	1	0.75

All correlations are at $P < 0.001$, RC: Respiratory care

of the questionnaires reflects the practicality and feasibility of administering this survey to a group of students. Face validity was confirmed by asking students during the cognitive interviews whether the items looked reasonable at face value. Construct related validity was confirmed by the mostly successful item and scale validity. Of the 71 items, item scale criteria were unsatisfactory in only 12 items. It is possible to explain the failure of the discriminant validity of some of these items; for example the item that asked about “the organization between information provided in the laboratory teaching and class teaching” made it difficult to separate the two domains “perception of laboratory teaching” and “perception of teaching and learning”. Answer to questions in one domain may be a coalescence of the effects of both domains. The same was true for items on “utilization of clinical training in the information received in the laboratory teaching.” It was probably difficult to differentiate between the two domains “perception of lab teaching” and “perception of clinical training”

In the process of item reduction, 6 items out of the 12 with unsatisfactory properties were excluded from the instrument. The other six were left after the research team decided on their high content validity and slight problems with their construct validity. For example, items such as “faculty were partial to some students, which gave me a feeling of unfairness” were considered to be important to keep in the questionnaire (corrected item-total correlation of 0.23 and floor effect of 17%).

Scale reliability that was assessed by internal consistency reliability using Cronbach’s alpha was above the recommended value in four out of the five domains. The slightly lower internal consistency reliability of the “perception of lab training” dimension (Cronbach’s alpha 0.58) may suggest the need to add more items to this scale in future developments of this instrument. The high internal consistency reliability by the Cronbach’s alpha and moderate inter-scale correlations further supported the validity of the internal construct indicating that each of the five domains measured concepts that were related, but distinct.

Some items obtained the highest score (ceiling effect) in a high proportion of the students, which may suggest that it is impossible to test sensitivity or verify improved perception of these items by students, over time or as a result of intervention programs. However, it is important to note that none of the dimension scales showed ceiling or floor effect suggesting that the instrument will continue to be effective in detecting the difference between groups and have good sensitivity to detect changes in student satisfaction using the dimensional and total scores.

This study has some limitations

First, the sample size was relatively small; it was possible that with a larger sample size, there would be stronger evidence of reliability and validity and the performance of the instrument in various subgroups could be investigated. Secondly, the cross-sectional design of the study did

not allow the assessment of the responsiveness of the instrument to change or intervention. Thirdly, without a similar published instrument to assess student perception of RC programs, no comparison of our validation results can be made.

Despite these limitations, this instrument is the first for measuring student satisfaction with RC program educational environment. Item generation and reduction of this study resulted in an instrument with 65 items and five domains.

Overall scaling, validity, and reliability characteristics were very encouraging at this preliminary assessment of the questionnaire. Only a few items had difficulties and these were excluded except when the research team felt it was still desirable to retain them for content validity. The authors of this study are currently in the process of conducting more conclusive psychometric evaluation of this instrument including factor analysis, structural equation modeling, and test-retest reliability analysis.^[25,26]

CONCLUSION

The instrument under study is being rigorously developed as the first validated instrument for measuring students' perception and satisfaction in RC educational programs. This assessment tool is a potentially valid and reliable instrument for use in future studies to assess RC students' perception and satisfaction with the educational environment. Future longitudinal studies are needed to assess the responsiveness and predictive validity of this instrument.

ACKNOWLEDGMENT

The authors are grateful to Ms. Amal Alamer, Ms. Bashair Alfozan, and Ms. Esraa Makhdom for their input and advice on the drafting of the questionnaire. We also would like to thank the expert panel for their time and contribution to the process of developing this questionnaire.

REFERENCES

- Hutchinson L. Educational environment. *BMJ* 2003;326:810-2.
- Genn JM. AMEE medical education guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education-a unifying perspective. *Med Teach* 2001;23:445-54.
- Roff S, McAleer S. What is educational climate? *Med Teach* 2001;23:333-4.
- Holt MC, Roff S. Development and validation of the anaesthetic theatre educational environment measure (ATEEM). *Med Teach* 2004;26:553-8.
- Bassaw B, Roff S, McAleer S, Roopnarinesingh S, De Lisle J, Teelucksingh S, *et al.* Students' perspectives on the educational environment, Faculty of Medical Sciences, Trinidad. *Med Teach* 2003;25:522-6.
- Roff S, McAleer S, Ifere OS, Bhattacharya S. A global diagnostic tool

- for measuring educational environment: Comparing Nigeria and Nepal. *Med Teach* 2001;23:378-82.
- Mayya S, Roff S. Students' perceptions of educational environment: A comparison of academic achievers and under-achievers at kasturba medical college, India. *Educ Health (Abingdon)* 2004;17:280-91.
- Till H. Identifying the perceived weaknesses of a new curriculum by means of the Dundee ready education environment measure (DREEM) inventory. *Med Teach* 2004;26:39-45.
- Pace CR, Stern GG. An approach to the measurement of psychological characteristics of college environments. *J Educ Psychol* 1958;49:269-77.
- Roff S. The Dundee ready educational environment measure (DREEM) – A generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Med Teach* 2005;27:322-5.
- Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees. *Med Teach* 2004;26:260-4.
- Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Med Teach* 2005;27:326-31.
- Gooneratne IK, Munasinghe SR, Siriwardena C, Olupeliyawa AM, Karunathilake I. Assessment of psychometric properties of a modified PHEEM questionnaire. *Ann Acad Med Singapore* 2008;37:993-7.
- Palmgren PJ, Chandratilake M. Perception of educational environment among undergraduate students in a chiropractic training institution. *J Chiropr Educ* 2011;25:151-63.
- Schnittler J, Carledge CM. Item-analysis programs: A comparative investigation of performance. *Educ Psychol Meas* 1976;36:183-7.
- Di Iorio CK. *Measurement in Health Behavior Methods for Research and Education*. San Francisco, CA: Jossey-Bass; 2005.
- Lawshe CH. A quantitative approach to content validity. *Pers Psychol* 1975;28:563-75.
- Willis GB. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Sage Publications: Thousand Oaks; 2005.
- Streiner DL, Norman GR. *Health Measurement Scales: a Practical Guide to Their Development and Use*. 4th ed. Oxford University Press: Oxford; 2008.
- Gandek B, Ware JE Jr, Aaronson NK, Alonso J, Apolone G, Bjorner J, *et al.* Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: Results from the IQOLA project. *International quality of life assessment. J Clin Epidemiol* 1998;51:1149-58.
- Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: The IQOLA project approach. *International quality of life assessment. J Clin Epidemiol* 1998;51:945-52.
- McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;31:247-63.
- Ware JE, Brook RH, Davies AR, Williams KN, Stewart AL, Rogers WH. Conceptualization and Measurement of Health for Adults in the Health Insurance Study. *Model of Health and Methodology*. Vol. I. Doc. no. R-1987/1-HEW. Santa Monica, CA: RAND Corporation; 1980.
- Nunnally JC. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill; 1994.
- Al-Rubaish AM, Rahim SI, Abumadani MS, Wosornu L. Academic job satisfaction questionnaire: Construction and validation in Saudi Arabia. *J Family Community Med* 2011;18:1-7.
- Aspy CB, Hamm RM, Schauf KJ, Mold JW, Flocke S. Interpreting the psychometric properties of the components of primary care instrument in an elderly population. *J Family Community Med* 2012;19:119-24.

How to cite this article: Alotaibi G, Youssef A. Development of an assessment tool to measure students' perceptions of respiratory care education programs: Item generation, item reduction, and preliminary validation. *J Fam Community Med* 2013;20:116-22.
Source of Support: Nil, **Conflict of Interest:** None declared