ORIGINAL PAPER

Investigating the use of student perception data for teacher reflection and classroom improvement

Lisa M. Bell · Jill M. Aldridge

Abstract   This article reports a study involving: (1) developing and validating a learning environment instrument, (2) investigating whether reflection on feedback generated using the instrument led to improvements in the learning environment and (3) examining how teachers used the feedback. For the new learning environment instrument, validity was supported for a sample of 10,345 secondary students over 3 years (2,042 students in 2008, 4,467 in 2009 and 3,836 in 2010). To investigate whether reflection on feedback from this instrument led to improvements in the learning environment, multivariate analysis of variance and effect sizes were used to examine pre–post differences. Statistically significant improvements in students' views of the learning environment were found for 8 of the 11 scales. Finally, we examined how teachers used the student feedback. Of the 459 teachers, 45 focus teachers used the feedback as part of a formal action research approach involving entries in reflective journals, written reports, discussions and participation at a forum. These data were analysed to provide an overview of the ways in which the information was used, and to provide a more detailed account of the journey of one of the teachers.

L. M. Bell
University of Western Australia, Crawley, WA 6009, Australia
e-mail: lisa.bell@uwa.edu.au

J. M. Aldridge (✉)
Curtin University, Perth, WA 6845, Australia
e-mail: j.aldridge@curtin.edu.au

Introduction

The results of studies conducted over the past 40 years have shown that the quality of the learning environment has a significant influence on student learning (Fraser 2007, 2012). That is, students learn better when they perceive the classroom environment more positively. By evaluating students' perceptions of their actual and preferred classroom learning environments, teachers can utilise an action research process in which they reflect on student feedback, plan for improvement, implement strategies and re-assess the learning environment to gauge the impact of their efforts and plan for improvement. Given that action research begins with teachers deliberately and consciously reflecting on their teaching practices, it was hypothesised that feedback, based on students' perceptions of the learning environment, would provide a useful means of doing just that. The purpose of this article is to report the development, validation and use of a new classroom learning environment instrument. The following sections outline the specific objectives of the study, provide a background to the study, describe the design and procedures used for the study and report the results of the study. Finally, the article provides a discussion of the results and a conclusion.

Objectives

The aims of the research reported in this article were to:

1. Develop and validate an instrument to assess students' perceptions of the classroom learning environment in terms of relationships, assessment and instructional delivery.
2. Examine the extent to which teachers' reflection on student perception data helped teachers to improve their classroom environments.
3. Investigate whether improvements in the learning environment differed for teachers who used the student feedback for reflection and those who used it as part of a more formalised action research process.
4. Investigate the processes used by teachers during the action research process.

Background

The field of learning environments: a means of providing teachers with student perceptions

The study reported in this article drew on and extended research carried out in the field of learning environments (Aldridge and Fraser 2008; Fraser 1998a, 2007; Goh and Khine 2002; Fisher and Khine 2006). Contemporary research within the field of learning environments, in part, was inspired by Lewin's (1936) ground-breaking work in non-educational settings, which recognised that both the environment and its interaction with characteristics of the individual are potent determinants of human behaviour. Since then, the notion of person-environment fit has been elucidated in education by Stern (1970), and Walberg (1981) has proposed a model of educational productivity in which the educational environment is one of nine determinants of student outcomes. Research specifically on classroom learning environments took off over 40 years ago with the work of Anderson and Walberg (1968) and Moos (1974) which spawned many, diverse research programs

around the world (Fisher and Khine 2006; Fraser 1998a) and the creation of *Learning Environments Research: An International Journal*. Past research on learning environments provides numerous research traditions, conceptual models and research methods that are relevant to our study.

This study draws on the rich resource of diverse, valid, economical and widely-applicable assessment instruments available in the field of learning environments (Fraser 1998b) as a starting point for developing the questionnaire described in this article. Typically, studies involving the use of feedback from a learning environment survey to guide improvements have used a five-step procedure (Fraser 2007):

1. assessing the actual and preferred learning environment
2. providing feedback to teachers based on students' responses
3. reflection and discussion based on feedback from the survey in order to identify which aspects of the learning environment might be chosen for change and for considering teaching strategies that might be used
4. implementing an intervention over a period of time in an attempt to change the learning environment
5. re-administration of the survey to students at the end of the intervention period to determine whether students perceive their learning environment differently from before.

As this five-step procedure has been used successfully at a range of educational levels, the teachers involved in this study made use of the same procedure.

Reflection and teacher action research

Reflection has been described as a way of thinking about a problematic situation that needs to be resolved. According to Fullan (1999), it is only through reflection that teachers begin to question and think differently about their teaching practices. In this respect, reflection provides the opportunity for teachers to be aware that a problem exists. Teacher reflection is considered "essential to identifying, analysing, and solving the complex problems that characterise classroom thinking" (Spalding and Wilson 2002, p. 1394) which results in teacher actions that are better considered and which bring benefits to the teacher and his or her students (Spalding and Wilson 2002).

Because Schön (1987) claims that an understanding of alternative perspectives about one's teaching lies at the heart of professional development, it is possible that the perspectives of students can provide a teacher with a valuable source of data for personal reflection. Seeking students' perspectives can help a teacher to question assumptions and to view his or her own practice through the eyes of others, both of which are recognised as important to successful teacher change (Aldridge and Fraser 2008; Bustingorry 2008; Hoban and Hastings 2006; Rhine 1998). The present study examined the use of students' perceptions, using a learning environment survey as a tool for reflection, in teacher action research.

Design and procedures

Three main assumptions underpinned our research: firstly, teachers generally are interested in improving their practice; secondly, conscious and purposeful reflection on one's teaching can provide a catalyst for change and guide improvements in the way in which

Table 1 Sample of the number of teachers, classes, student responses and schools over a 3-year period

Year	Teachers		Classes			Student responses		Schools
	Pre	Post	Pre	Post	Focus classes	Pre	Post	
2008	107	91	147	122	15	2,042	1,182	9
2009	247	210	298	248	15	4,467	2,749	25
2010	194	158	239	190	22	3,836	2,176	21
Total	548	459	684	560	52	10,345	6,107	29

one teaches; and, thirdly, because students are major stakeholders in the education process, they are worth listening to.

Our study utilised different 'grain sizes' (Fraser 1999) for collecting data over a 3-year period. At the largest grain size, a sample of 10,345 student responses (2,042 student responses in 2008, 4,467 in 2009 and 3,836 in 2010) in 684 classes (147 classes in 2008, 298 in 2009 and 239 in 2010) in 29 coeducational high schools in Western Australia was used. A breakdown of the sample, in terms of the number of teachers, classes and students involved is provided in Table 1. The sample included mostly upper-secondary school students from four schools located in regional areas and 25 schools in metropolitan areas. The teachers involved in the study ranged in both age and teaching experience and the sample included classes from all learning areas (English, science, mathematics, society and environment, technology and enterprise, the arts and LOTE). Teacher and student participation in the study was voluntary; all students who responded to the survey instrument were provided with information about the study and signed a consent form. This large sample was used to investigate the validity and reliability of the new instrument (Research Objective 1).

The second research objective was to examine the extent to which teachers' reflection on student feedback data helped them to improve their classroom environment. To address this objective, the classes of 459 teachers (some of whom selected more than one class) were used to examine pre–post changes. Only data for those students who were present for both the pretest and posttest were involved and this sample comprised 6,107 student responses (1,182 student responses in 2008, 2,749 in 2009 and 2,176 in 2010) in 560 classes (122 classes in 2008, 248 in 2009 and 190 classes in 2010). This subsample was part of the larger sample described previously and, as with the large sample, represented the full range of learning areas.

The third research objective involved whether improvement in students' perceptions of the classroom environment differed for teachers who used the student feedback for reflection and those who used it as part of a more formalised action research approach. To address the objective, 45 of the teachers volunteered to be focus teachers. Unlike the other teachers (who were provided with data for reflection), these focus teachers (12 teachers in 2008, 13 in 2009 and 20 in 2010) used the student feedback data as part of a more formal action research process during which they recorded their intended strategies and kept a reflective journal (described below). These teachers were monitored more closely than other teachers in the study.

The fourth research objective involved how the focus teachers used the feedback data as they engaged in action research. To address this objective, the 45 teachers described above documented their action research activities using a written plan, entries in a reflective journal and a summary or report that outlined the processes undertaken. Information

included the types of strategies that they implemented and their views of the effectiveness of these practices and the effects that these strategies had on their classroom environment and their teaching practices. In addition to this, focus teachers were given opportunities to discuss their activities with the researchers or their colleagues at different stages of their action research and semi-structured interviews were held with at least five focus teachers in each year of the study. At the end of the year, focus teachers were invited to present their findings at a forum during which further information was gathered. This group of teachers was considered to be representative of the larger population in that they ranged in age, level of experience and included teachers from all learning areas.

Analyses and results

Development of the Constructivist-Oriented Learning Environment Survey (COLES)

The first objective of the study focused on the development of an instrument to assess students' perceptions of the classroom learning environment. During this process, we endeavoured to include scales with a constructivist orientation. With this in mind, we identified principles relevant to student-centred classrooms and which are consistent with a more constructivist approach. These principles were used to identify dimensions that could be used as a basis for identifying and developing specific scales that would give an indication of the extent to which practices consistent with these principles were present. As a result of these efforts, a widely-applicable instrument for assessing students' perceptions of their learning environment has been developed.

Six of the eleven scales are based on the widely-used What Is Happening In this Class? (WIHIC, Fraser et al. 1996) questionnaire that has been cross-validated across many subject areas and in numerous countries (Aldridge et al. 1999; Afari et al. 2013; Khoo and Fraser 2008). In addition to these six scales, the new instrument incorporated the Young Adult Ethos scale (adapted from the Technology-Rich Outcomes-Focused Learning Environment Instrument, Aldridge and Fraser 2008), the Differentiation scale (modified from the Individualised Classroom Environment Questionnaire, Rentoul and Fraser 1979) and the Personal Relevance scale (modified from the widely-used Constructivist Learning Environment Survey, Taylor et al. 1997). A description of each of the scales and the reasons for their inclusion can be found in Aldridge et al. (2012).

As previous learning environment instruments had not included aspects that gauged students' perceptions of the assessment practices used by the teacher, two additional scales were developed, namely, Formative Assessment and Clarity of Assessment Criteria. Assessment, particularly formative assessment, is critical for knowing how a student is performing academically and how best to support that student in the classroom (Allen et al. 2009). Making clear the students' goals and providing useful feedback on their progress provide a powerful means for improving student learning. Such goals encourage students to reflect on their current understandings and to plan for success in their future learning (O'Donovan et al. 2004).

The Formative Assessment scale assesses the extent to which students feel that assessment tasks make a positive contribution to their learning. Formative assessment relates to assessment practices that are part of the learning process rather than a separate (summative) process conducted at the end of the learning journey (Bell and Cowie 2001). Formative assessment practices can provide opportunities for students to discuss their progress with their teachers, as well as to engage in peer and self assessment as ways of

monitoring and reflecting on their learning. More importantly, if students are able to recognise the link between their learning and assessment, then assessment becomes educative, improves student learning, enables students to plan for future learning and can foster self-directed learning practices (Bell and Cowie 2001; Black and Wiliam 1998). This new scale was developed to tap into these important aspects.

The Clarity of Assessment Criteria scale assesses the extent to which the assessment criteria are explicit so that the basis for judgements is clear and public. While clear and public criteria can support and enhance a teachers' ability to make professional judgements about student learning, as a matter of fairness, students should also know the criteria by which they are being assessed. Therefore, this scale assesses the students' views on this important aspect.

The initial version of the COLES was comprised of 11 scales, each with eight items (a total of 88 items). In 2008, after consideration of a number of issues (related to survey fatigue and low reading levels), the instrument was refined in consultation with classroom practitioners who had used the survey. As a result, individual items were refined and the number of items in each scale was reduced. The final version of the COLES, reported in this article, has 11 scales with six items in each (and, in one case, seven) providing a total of 67 items. Although it is acknowledged that an instrument comprising 11 scales cannot assess every aspect of the learning environment, the selected scales were all considered to be relevant to teachers aiming to create more constructivist-oriented learning environments in a range of contexts. Importantly, many of these scales have also been shown to be good predictors of student outcomes.

The 11 dimensions of the COLES can be grouped into three broad categories: Relationships (Student Cohesiveness, Teacher Support, Equity and Young Adult Ethos), Assessment (Clarity of Assessment Criteria and Formative Assessment) and Delivery (Task Orientation, Differentiation, Personal Relevance, Involvement, and Cooperation). A description of each scale and a sample item is included in Table 2.

Students respond to the items using a five-point frequency scale consisting of Almost Always, Often, Sometimes, Seldom and Almost Never. Importantly, this instrument enables students to provide information about the learning environment that is currently present in the classroom (the *actual* environment) as well as information about the learning environment that they would like (their *preferred* environment). To achieve this, the instrument makes use of a side-by-side response format in which students are able to respond to each item with respect to both their actual and preferred responses (see Fig. 1).

Validity and reliability of the COLES

Data collected from 10,345 students in 684 secondary high school classes, over 3 years (described earlier), were analysed in various ways to support the validity and reliability of the actual and preferred versions of the new instrument at the secondary or high-school level (Pallant 2011; Trochim and Donnelly 2008). When the researcher's goal is to construct a multiscale questionnaire, factor analysis provides a means of determining whether items within the same scale are tapping into the same construct and whether each scale is assessing a distinct construct. Principal axis factor analysis with oblique rotation (used because it can be assumed that dimensions of a learning environment overlap) was used to examine the factor structure for actual and preferred data for each of the 3 years. The two criteria for retaining any item were that it must have a factor loading of at least 0.40 on its own scale and less than 0.40 on any of the other scales (Field 2005; Stevens 1992; Thompson 2004).

Table 2 Description, example item and origin for each COLES scale

Scale	Scale description	Example of item	Origin of scale
	The extent to which…		
Relationships			
Student Cohesiveness	…students know, help and are supportive of one another	Members of this class are my friends	What is happening in this class? (WIHIC)
Teacher Support	…the teacher helps, befriends, trusts and is interested in students	The teacher moves around the class to talk with me	What is happening in this class? (WIHIC)
Equity	…students are treated equally by the teacher	I get the same amount of help from the teacher as do other students	What is happening in this class? (WIHIC)
Young Adult Ethos	…teachers give students responsibility and treat them as young adults	I am given the opportunity to be independent	Technology-rich outcomes-focused learning environment instrument (TROFLEI)
Assessment			
Formative Assessment	…students feel that the assessment tasks given to them make a positive contribution to their learning	Assessment tasks help me to monitor my learning	Developed for the COLES
Assessment Criteria	…the assessment criteria are explicit so that the basis for judgments is clear and public	I understand how the teacher judges my work	
Delivery			
Involvement	…students have attentive interest, participate in discussions, ask questions and share ideas	I explain my ideas to other students	What is happening in this class? (WIHIC)
Task Orientation	…it is important to complete activities planned and to stay on the subject matter	I pay attention during this class	What is happening in this class? (WIHIC)
Personal Relevance	…subject is relevant to students' everyday out-of-school experiences	I relate what I learn in this class to my life outside of school	Constructivist Learning Environment Survey (CLES)
Cooperation	…students cooperate with one another on learning tasks	When I work in groups in this class, there is teamwork	What is happening in this class? (WIHIC)
Differentiation	…teachers cater for students differently on the basis of ability, rates of learning and interests	I am able to work at the speed which suits my ability	Individualised Classroom Environment Questionnaire (ICEQ)

		ACTUAL					PREFERRED			
Formative Assessment	Almost Never	**Seldom**	Some times	Often	Almost Always	Almost Never	**Seldom**	Some times	Often	Almost Always
14. Assessment tasks help me to monitor my learning.	1	2	3	4	5	1	2	3	4	5

Fig. 1 Illustration of side-by-side response format for actual and preferred responses

For all 3 years, all of the items for the actual version, with the exceptions of three, had a loading of at least 0.40 on their a priori scale and no other scale for all 3 years. The three exceptions (Item 32 from the Clarity of Assessment scale in 2010 only, Item 49 from the Task Orientation scale in 2008 only and Item 62 from the Differentiation scale for all 3 years) all had a loading of less than 0.40 on their own scale and all other scales. For all 3 years, all items for the preferred version, with the exception two, had a loading of at least 0.40 on their a priori scales and no other scale. These two exceptions, Item 13 from the Equity scale (2010 only) and Item 62 from the Differentiation scale (all 3 years), both had a loading of less than 0.40 on their own scale and all other scales.

For the actual version, the percentage of variance ranged from 1.17 to 26.02 % for different scales, with the total variance accounted for being 70.19 % in 2008, 68.63 % in 2009 and 64.32 % in 2010. For the preferred version, the percentage of variance ranged from 1.19 to 22.64 % for different scales, with the total variance accounted for being 66.80 % in 2008, 65.20 % in 2009 and 62.87 % in 2010.

Two further indices of scale reliability and validity were examined separately for each year, namely, the internal consistency reliability and the ability of each COLES scale to differentiate between the perceptions of students in different classrooms. Table 3 reports the internal consistency of each COLES scale, using Cronbach's alpha coefficient for two units of analysis (the individual student and the class mean), separately for each year, and separately for the actual and preferred versions. Using the individual as the unit of analysis, scale reliability estimates for the actual form ranged from 0.74 to 0.92 in 2008, 0.70 to 0.93 in 2009 and 0.80 to 0.92 in 2010. For the preferred form, scale reliabilities ranged from 0.75 to 0.92 in 2008, 0.73 to 0.92 in 2009 and 0.80 to 0.92 in 2010. Generally reliability figures were higher with the class mean as the unit of analysis. For the actual form, scale reliabilities ranged from 0.83 to 0.97 in 2008, 0.79 to 0.97 in 2009 and 0.84 to 0.97 in 2010. For the preferred form, scale reliabilities ranged from 0.76 to 0.96 in 2008, 0.75 to 0.96 in 2009 and 0.74 to 0.96 in 2010. These alpha reliability estimates support the internal consistency of all scales of the COLES for each year of administration.

A one-way analysis of variance (ANOVA), with class membership as the independent variable, was computed to determine the degree to which each COLES scale was able to differentiate between the perceptions of students in different classes. The proportion of variance accounted for by class membership was calculated using the eta^2 statistic (the ratio of 'between' to 'total' sums of squares). The results, reported in Table 3, show that all 11 COLES scales differentiated significantly between classes ($p < 0.01$), indicating that students within the same class perceived the environment in a relatively similar manner, while the within-class mean perceptions of the students varied between classes. The eta^2 statistic (an estimate of the strength of association between class membership and the dependent variable) for different COLES scales ranged from 0.14 to 0.31 in 2008, 0.13 to 0.30 in 2009 and 0.13 to 0.30 in 2010. These results indicate that all scales in the COLES scales are sufficiently sensitive to distinguish between the learning environments of different classrooms.

Table 3 Internal consistency reliability (Cronbach alpha coefficient) for actual and preferred responses and ability to differentiate between classrooms (ANOVA) for actual version of the COLES

Scale	Unit of analysis	Alpha reliability Actual			Alpha reliability Preferred			ANOVA (Eta^2) Actual only		
		2008	2009	2010	2008	2009	2010	2008	2009	2010
Student Cohesiveness	Individual	0.85	0.85	0.82	0.88	0.88	0.85	0.14**	0.17**	0.19**
	Class mean	0.89	0.90	0.88	0.92	0.92	0.90			
Teacher Support	Individual	0.91	0.90	0.90	0.88	0.88	0.87	0.31**	0.30**	0.27**
	Class mean	0.96	0.96	0.96	0.93	0.92	0.91			
Equity	Individual	0.92	0.93	0.92	0.92	0.91	0.90	0.23**	0.24**	0.28**
	Class mean	0.97	0.97	0.97	0.95	0.96	0.94			
Young Adult Ethos	Individual	0.92	0.90	0.89	0.92	0.90	0.89	0.20**	0.19**	0.30**
	Class mean	0.96	0.95	0.94	0.96	0.94	0.91			
Formative Assessment	Individual	0.90	0.88	0.86	0.90	0.89	0.88	0.17**	0.17**	0.18**
	Class mean	0.94	0.94	0.92	0.95	0.94	0.92			
Clarity of Assessment	Individual	0.90	0.90	0.87	0.92	0.92	0.89	0.18**	0.19**	0.23**
	Class mean	0.95	0.95	0.92	0.96	0.97	0.95			
Involvement	Individual	0.87	0.86	0.90	0.90	0.90	0.92	0.17**	0.15**	0.20**
	Class mean	0.91	0.90	0.94	0.93	0.93	0.96			
Task Orientation	Individual	0.84	0.82	0.83	0.88	0.88	0.87	0.15**	0.13**	0.13**
	Class mean	0.88	0.88	0.84	0.93	0.94	0.91			
Personal	Individual	0.91	0.90	0.90	0.92	0.92	0.92	0.22**	0.20**	0.23**
Relevance	Class mean	0.95	0.95	0.93	0.96	0.95	0.94			
Cooperation	Individual	0.88	0.87	0.83	0.89	0.89	0.87	0.18**	0.16**	0.19**
	Class mean	0.94	0.93	0.86	0.94	0.93	0.92			
Differentiation	Individual	0.74	0.70	0.80	0.75	0.73	0.80	0.20**	0.17**	0.19**
	Class mean	0.83	0.79	0.84	0.76	0.75	0.74			

The sample consisted of 2,042 students in 147 classes in 2008, 4,467 students in 298 classes in 2009 and 3,836 students in 239 classes in 2010

The eta^2 statistic (which is the ratio of 'between' to 'total' sums of squares) represents the proportion of variance explained by class membership

** $p < 0.01$

In summary, for our sample of 10 345 students ($N = 147$ classes for 2008, 298 classes in 2009 and 239 classes in 2010), we found satisfactory factorial validity, internal consistency reliability (Cronbach alpha coefficient) and ability of each scale to differentiate between the perceptions of the students in different classrooms. The results reported in Table 3, in conjunction with the factor analysis results, support the COLES as a valid and reliable instrument for assessing students' perceptions of their classroom psychosocial environments at the high-school level.

Pre–post changes in students' perceptions of the learning environment

The second research objective involved the extent to which the teachers' reflections on feedback data from students guided improvements in the classroom learning environment.

Over a three-year period, a total of 459 teachers (working with 560 classes) were involved in a pre-post design. All of these teachers selected at least one class with which they would like to work. The learning environment survey, described above, was administered to students in the selected classes and teachers were provided with feedback based on students' responses to the COLES (see below for examples of profiles). After 6–8 weeks, the COLES was readministered to the same classes. Only students present for both the pretest and posttest were included in the analyses.

Multivariate analysis of variance (MANOVA) with repeated measures (using the class mean as the unit of analysis) was used to investigate whether differences between pretest and posttest scores were statistically significant. When the multivariate test (Wilks' lambda) revealed significant pre–post differences overall, the ANOVA with repeated measures was interpreted for each COLES scale (see Table 4). Pre-post differences were statistically significant ($p < 0.01$) for eight of the 11 COLES scales, namely, Student Cohesiveness, Teacher Support, Young Adult Ethos, Clarity of Assessment, Involvement, Personal Relevance, Cooperation and Differentiation. For those scales with a statistically significant difference, students perceived the learning environment more favourably over time.

To examine the magnitudes of these pre–post differences, as recommended by Thompson (2001), effect sizes were calculated in terms of the differences in means divided by the pooled standard deviation. The effect sizes for those scales with statistically significant differences, reported in Table 4, ranged between 0.06 and 0.24 of a standard deviation. These results suggest moderately important educationally differences between learners' perceptions of the classroom learning environment before and after the teachers' reflection on the feedback data.

Posttest changes for reflection teachers and action-research teachers

Of the 459 teachers, 45 teachers volunteered to be 'focus teachers' and to be involved in the action research component on a more formal basis. These teachers all agreed to develop a plan, keep a reflective journal (in which they recorded the strategies that they intended to use and the outcomes of using these strategies as they implemented them in the classroom) and provide a summary of the research. Because there were no statistically significant differences for any of the COLES scales between pretest scores of teachers who used the data for reflection and for those teachers who used the data as part of the action research process, it was considered acceptable to examine the differences between posttest scores only for each of the groups.

As for Research Objective 2, MANOVA with repeated measures (using the class mean as the unit of analysis) was used to investigate whether differences between the posttest scores of the two groups of teachers were statistically significant. When the multivariate test (Wilks' lambda) revealed a statistically significant posttest difference overall, the ANOVA with repeated measures was interpreted for each COLES scale (see Table 5). The posttest differences for these two groups of teachers, reported in Table 5, were statistically significant ($p < 0.05$) for six of the eleven COLES scales, namely, Teacher Support, Equity, Young Adult Ethos, Formative Assessment, Task Orientation, and Differentiation. In all cases for which there was a statistically significant difference, the results indicated that students perceived the learning environment more favourably in the classrooms of the focus teachers.

The effect sizes for those scales with statistically significant differences, calculated to examine the magnitudes of the posttest differences (as recommended by Thompson 2001), ranged between 0.15 and 0.20 standard deviations. These data suggest moderately important educationally differences (Cohen 1992) between learners' perceptions of the classroom learning environment.

Table 4 Average item mean, average item standard deviation, effect size and MANOVA results for differences between pretest and posttest scores using the class mean as the unit of analysis

Scale	Average item mean[a]		Average item standard deviation		Difference	
	Pretest	Posttest	Pretest	Posttest	Effect size	F
Student Cohesiveness	4.18	4.25	0.29	0.30	0.12	2.70**
Teacher Support	3.88	3.95	0.49	0.50	0.07	2.33**
Equity	4.10	4.09	0.46	0.48	−0.01	0.89
Young Adult Ethos	4.16	4.21	0.38	0.38	0.07	2.03**
Formative Assessment	3.96	3.97	0.36	0.41	0.01	0.94
Clarity of assessment	3.86	3.96	0.40	0.38	0.13	2.74**
Involvement	3.29	3.49	0.40	0.42	0.24	3.65**
Task Orientation	4.03	4.02	0.29	0.35	−0.02	0.68
Personal Relevance	3.22	3.41	0.52	0.49	0.18	3.28**
Cooperation	3.78	3.83	0.38	0.41	0.06	1.99**
Differentiation	3.38	3.48	0.39	0.44	0.12	2.74**

$N = 560$ classes

Effect size was calculated using formula of $d = M_1 - M_2/\sqrt{[(\sigma_1^2 + \sigma_2^2)/2]}$

** $p < 0.01$

[a] Average item mean = Scale mean divided by the number of items in that scale

Using student feedback data as the basis for teacher action research

The fourth research objective involved how the focus teachers used the feedback generated from the student perceptual data and the types of activities that were undertaken as part of the action research process. This section provides an outline of how the teachers, in general, used the data and how one of the teachers, Michael, used the data to reflect, plan and implement improvement strategies.

In all cases, the COLES was administered to classes by an external party. These data were used to generate a teacher feedback package that included two profiles and complementary data for each scale. The first profile involved a simple comparison of the average item means for actual and preferred responses for each scale (see Fig. 2). The second profile was a box plot, including the mean, median and range of responses for each scale. Finally, the package provided teachers with a bar graph which depicted the mean for actual and preferred responses for each item.

Teachers were shown how to interpret the results for their class, either through a workshop, a small-group meeting or a one-on-one session. Teachers were then asked to reflect on their results and use actual-preferred discrepancies to identify an area or areas on which they felt that they would like to focus. In some cases, teachers discussed the results with their students to help to explain anomalies. One teacher explained:

> After some reflection and discussion with students in the class, I decided that I needed to structure my questions better in order to check for understanding, address issues with equity, increase student involvement and focus on clarifying with students my expectations of them both as individuals and as part of the whole class. (Teacher, Mathematics)

Table 5 Average item mean, average item standard deviation, effect size and MANOVA results (using the class mean as the unit of analysis) for differences in posttest scores for teachers using the results for reflection only and those using the results as part of an action research process

Scale	Average item mean[a]		Average item standard deviation		Difference	
	Reflection only	Action research	Reflection only	Action research	Effect size	F
Student Cohesiveness	4.25	4.27	0.30	0.26	0.04	0.80
Teacher Support	3.94	4.09	0.50	0.48	0.15	1.45*
Equity	4.07	4.25	0.48	0.39	0.20	1.77**
Young Adult Ethos	4.20	4.35	0.38	0.35	0.20	1.73**
Formative Assessment	3.96	4.07	0.42	0.32	0.15	1.49*
Clarity of Assessment	3.96	4.02	0.40	0.37	0.08	1.03
Involvement	3.48	3.53	0.43	0.32	0.07	1.01
Task Orientation	4.02	4.12	0.35	0.26	0.16	1.56*
Personal Relevance	3.41	3.43	0.53	0.50	0.02	0.59
Cooperation	3.83	3.87	0.41	0.35	0.05	0.89
Differentiation	3.48	3.59	0.44	0.48	0.12	1.36*

$N = 560$ classes. Teachers used the data for reflection only in 508 classes and teachers used the data as part of an action research process in 52 classes

Effect size was calculated using formula of $d = M_1 - M_2/\sqrt{[(\sigma_1^2 + \sigma_2^2)/2]}$

** $p < 0.01$

[a] Average item mean = Scale mean divided by the number of items in that scale

In other cases, the teachers discussed the results with colleagues. To this end, one teacher commented:

> I worked with a colleague in Society and Environment and two colleagues in Science; we discussed our current teaching strategies that were focused on fostering cooperative learning. We decided to video-record each other's lessons and to meet later for further discussion about what parts of our selected strategies made the most impact on increasing student-centred learning. (Teacher, Physics)

Some of the teachers reflected on their data individually and planned strategies with which they were already aware. A human biology teacher stated:

> The data indicated to me that I had become too comfortable with using only a few teaching strategies, and that there were numerous effective strategies that I could be implementing. It was just a matter of remembering them and finding ways… to implement them. I decided to modify a couple of assessments so that they became more student directed. This was easy and just a matter of letting the students pick a research topic that they personally found interesting or perhaps was relevant in their lives at that time. (Teacher, Human Biology)

Once teachers had decided upon the scale or scales on which they would like to focus, they worked either individually or collaboratively to devise strategies that they felt would help to reduce the actual-preferred gap. For example, one teacher identified two aspects of her teaching that she wanted to do differently. To this end, she commented:
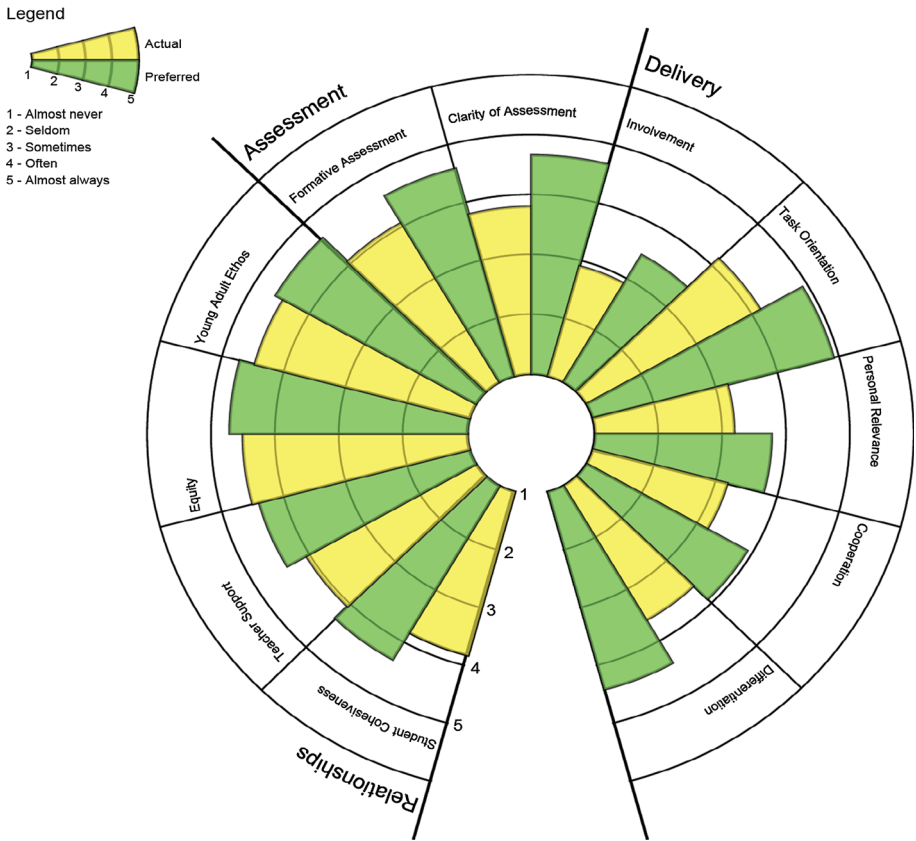
Fig. 2 Pretest: actual and preferred learning environment for the pretest for Michael's class

I identified two aspects of my teaching that I wanted to do differently. Firstly, I needed to make my instructions much clearer for assessment tasks. It was apparent that students were not sure what their assessment tasks should include or what the final product should look like. Secondly, I needed to make my marking keys more transparent because the students were not sure how to achieve the highest possible marks. (Teacher, Career and Enterprise)

After the teachers had implemented their selected strategies over a six-week to eight-week period (comprised of approximately 20–30 lessons ranging from 45 to 60 min in duration), the COLES was then re-administered to allow teachers to examine whether the strategies had led to changes in students' perceptions of the learning environment. All of the focus teachers reported that, while the pretest data helped them to reflect and identify areas to target for improvement, the posttest results provided meaningful information about the success of the strategies that they had implemented. Analysis of data collected using reflective journals and reports (written by focus teachers) indicated that the teachers used a range of strategies and that, in most cases, these strategies were specific to the context of the class from which the student responses were drawn. To illustrate how the student

feedback data were used, a description of how one teacher, Michael, utilised the student feedback data is provided below.

Michael, a teacher with over 22 years of classroom experience, was one of the 45 focus teachers who volunteered to record their action research activities. Michael selected a grade 11 Photography class with 21 students. Although it was a fairly cooperative class with what he considered to be 'nice' kids, he felt that his rapport with these students differed from other classes that he taught and that the students were not generally achieving the results that he expected.

After examining his feedback data (see Fig. 3), Michael decided to focus on the assessment category because of the large actual-preferred difference. His key strategy involved changing the way in which he delivered his feedback to students. To do this, he supplemented his usual written feedback with audio comments. The comments were saved as audio files on the school network and, when he handed assignments back (with his written comments), students were able to spend the first part of the lesson listening to his audio feedback. Michael implemented his strategy over a six-week intervention period during which he provided audio feedback for one major assessment submitted by his students.

When the class was re-surveyed at the end of the intervention period, Michael noted that there were positive shifts in student scores on both of the scales that he had targeted: Formative Assessment and Clarity of Assessment. He was interested to note that scores on the Teacher Support scale also had improved. Michael felt that this improvement in the Teacher Support scale could be attributed to improved rapport with his students that came about as a result of the strategy that he had implemented. In addition to an improved relationship with his students, he also noticed an increase in the number of students who asked questions about their work and who asked for clarification of aspects of the next assessment task. Importantly, it became evident in subsequent assessments that students made a concerted effort to improve areas of their work that he had highlighted in the audio feedback. Michael reported that students had responded positively to his new method of feedback and made note of the fact that three of the students had commented that they felt the feedback was personalised and that they believed that the teacher was really trying to help them to improve.

The 45 focus teachers involved in this study provided valuable feedback about the COLES and the activities that they undertook as part of their action research. Through their written reports and their participation at an end-of-year forum, the teachers provided valuable criticism that was used to fine tune the instrument. In the main, these teachers felt that the data gathered using the COLES provided them with valuable insights into their students' views of the classroom learning environment that provided a good vehicle for reflection. According to these teachers, they were able to effectively use the data to devise strategies to target areas or issues particular to a specific class. Finally, the teachers reported that the posttest data allowed them not to only gauge the success of the strategies implemented, but also to provide further opportunities for reflection on their practice within their own classrooms.

Discussion

Much literature suggests that action research has the capacity to bring about meaningful changes or improvements at the classroom level (Kemmis and McTaggart 1998). As part of the action research cycle, reflection is perceived to be an integral component that alerts
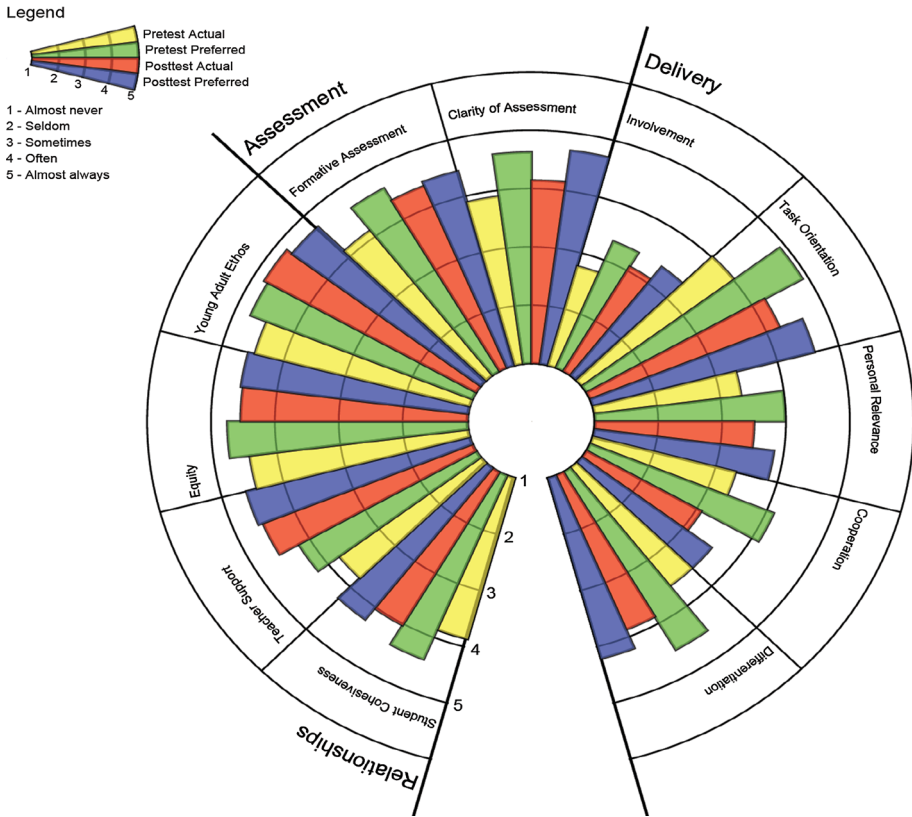
Fig. 3 Posttest: actual and preferred learning environment for the pretest and posttest for Michael's class

teachers that a problem or an issue exists (Bustingorry 2008). The present study considered the viability of using student feedback as a basis for such reflection.

An important contribution of the present study was the development and validation (over 3 years) of the COLES to tap into students' perceptions of important aspects of the learning environment. Feedback packages, generated using students responses to the COLES, were used by teachers to help them to reflect on what was happening in their classroom through the eyes of their students. The refined version of the COLES has 67 items that assess 11 dimensions of the classroom environment, namely, Student Cohesiveness, Teacher Support, Involvement, Personal Relevance, Task Orientation, Cooperation, Equity, Differentiation, Young Adult Ethos, Formative Assessment and Clarity of Assessment Criteria. Incorporated into the questionnaire is a side-by-side response format which enables students to record their views of their actual and preferred learning environment.

The sample of 10 345 student responses (2,042 student responses in 2008, 4,467 in 2009 and 3,836 in 2010) in 684 classes (147 classes in 2008, 298 in 2009 and 239 in 2010) in 29 schools were used to examine the validity and reliability of the instrument over a three-year period. Principal axis factor analysis confirmed the a priori factor structure for the COLES with all items loading on their a priori scale and no other scales, with the exception of three items for the actual version and two items for the preferred version. For these five

exceptions, the items had a loading of less than 0.40 on their own scale and all other scales. For both the actual and preferred versions of the COLES, the internal consistency reliability was found to be satisfactory both at the class mean and individual levels of analysis. Further analyses supported the ability of the actual form of each of the 11 COLES scales to differentiate between classrooms. Given the strong reliability of the COLES, teachers in this and future studies can be confident about the feedback provided to them. Overall, this new survey offers an expedient tool with which feedback information, relevant to constructivist-oriented learning environments, can be obtained. Further, the validity of the COLES across a varied sample of learning areas provides support for the wide applicability of the new survey.

To investigate pre–post changes in students' perceptions of the classroom learning environment for students in 560 classes, a one-way MANOVA was used. There was a statistically significant pre–post difference for eight of the 11 COLES scales. For all of these statistically significant differences, there was a positive change in the classroom environment. These statistically significant changes suggest that providing teachers with student perception data upon which to reflect regarding their teaching could be a powerful means of effecting change. The reflective nature of this aspect of the study, which involved encouraging teachers to examine their teaching practices through the eyes of their students, offers much promise. When teachers reflect on student feedback, they are likely to change the classroom environment in ways that are favourable for students. This component of our study replicates past research which has been successful in stimulating improvements in classroom environments (Aldridge and Fraser 2008; Aldridge et al. 2009; Aldridge et al. 2004; Fraser and Fisher 1986; Sinclair and Fraser 2002; Thorp et al. 1994; Yarrow et al. 1997).

The study went further also examined whether differences existed between the changes made by those teachers who used the student feedback as a basis for reflection (414 teachers and 508 classes) and the 45 focus teachers (52 classes) who used the feedback as part of a formal action research process. Posttest differences for the two groups of students were statistically significantly larger for those teachers who used a more formal approach than for their counterparts who used student feedback as a basis for reflection. These results suggest that, for teachers wishing to improve the classroom environment, it is worthwhile to not only reflect upon the data, but to also engage in a more formal action research approach that involves writing an action plan and reflecting upon each stage of the action research process.

Qualitative data, gathered from the 45 focus teachers, were used to examine how the focus teachers used the data. The results indicated that teachers interpreted the data in ways that were meaningful to the context of their classrooms. In many cases, particularly when teachers were unable to explain actual-preferred differences, they discussed the results with their students. Teachers generally reported that the feedback, based on students' responses, was useful and helped them to decide upon appropriate strategies, with a view to closing the actual-preferred gap. Although our qualitative findings indicated that teachers used a range of strategies, it is recommended that further research involves examining the types of strategies that were most likely to successfully improve the learning environment and whether these differed for different learning environments.

Concluding remarks

The results suggest that the COLES provides a reliable tool that can be used to generate student feedback data. The teachers were able to use the feedback generated using the

COLES for the purposes of reflecting on their teaching, providing them with an opportunity to step back from their teaching and reflect on what they do in a more objective and analytical manner. An initial examination of the reports, summaries and reflective journals collected from the 45 focus teachers who agreed to record their activities during the formal action research process indicated that the COLES provides valuable data that can be used to prompt reflections to guide teachers to implement classroom changes to improve the classroom learning environment. Further analysis of the qualitative data is likely to provide a greater depth of understanding of the ways in which student feedback data can be used by teachers to improve classroom learning environments and its potential for encouraging teacher-driven professional development.

References

Afari, E., Aldridge, J. M., Fraser, B. J., & Myint, S. W. (2013). Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research, 16*, 131–150.

Aldridge, J. M., & Fraser, B. J. (2008). *Outcomes-focused learning environments: Determinants and effects (Advances in Learning Environments Research series)*. Rotterdam, The Netherlands: Sense Publishers.

Aldridge, J. M., Fraser, B. J., Bell, L. M., & Dorman, J. (2012). Using a new learning environment questionnaire for reflection in teacher action research. *Journal of Science Teacher Education, 23*, 259–290.

Aldridge, J. M., Fraser, B. J., & Huang, I. T. C. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *Journal of Educational Research, 93*, 48–62.

Aldridge, J. M., Fraser, B. J., & Ntuli, S. (2009). Utilising learning environment assessments to improve teaching practices among in-service teachers undertaking a distance-education programme. *South African Journal of Education, 29*, 147–170.

Aldridge, J. M., Fraser, B. J., & Sebela, M. P. (2004). Using teacher action research to promote constructivist learning environments in South Africa. *South African Journal of Education, 24*, 245–253.

Allen, D., Ort, S. W., & Schmidt, J. (2009). Supporting classroom assessment practice: Lessons from a small high school. *Theory into Practice, 48*, 72–80.

Anderson, G. L., & Walberg, H. J. (1968). Classroom climate group learning. *International Journal of Educational Sciences, 2*, 175–180.

Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. Dordrecht: Kluwer Academic Publishers.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*, 7–11.

Bustingorry, S. O. (2008). Towards teachers' professional autonomy through action research. *Educational Action Research, 16*, 407–420.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*, 155–159.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications.

Fisher, D. L., & Khine, M. S. (Eds.). (2006). *Contemporary approaches to research on learning environments: Worldviews*. Singapore: World Scientific.

Fraser, B. J. (1998a). Science learning environments: Assessment, effects and determinants. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 527–564). Dordrecht, The Netherlands: Kluwer.

Fraser, B. J. (1998b). Classroom environment instruments: Development, validity and applications. *Learning Environments Research, 1*, 7–33.

Fraser, B. J. (1999). "Grain sizes" in learning environment research: Combining qualitative and quantitative methods. In H. C. Waxman & H. J. Walberg (Eds.), *New directions for teaching practice and research* (pp. 285–296). Berkeley, CA: McCutchan.

Fraser, B. J. (2007). Classroom learning environments. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 103–124). Mahwah, NJ: Lawrence Erlbaum.

Fraser, B. J. (2012). Classroom learning environments: Retrospect, context and prospect. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1191–1239). Dordrecht, The Netherlands: Springer.

Fraser, B.J., & Fisher, D.L. (1986). Using short forms of classroom climate instruments to assess and improve classroom psychosocial environment. Journal of Research in Science Teaching, 5, 387–413.

Fraser, B. J., McRobbie, C. J., & Fisher, D. L. (April 1996). *Development, validation and use of personal and class forms of a new classroom environment instrument*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Fullan, M. (1999). *Change forces: The sequel*. London: Falmer Press.

Goh, S. C., & Khine, M. S. (Eds.). (2002). *Studies in educational learning environments: An international perspective*. Singapore: World Scientific.

Hoban, G., & Hastings, G. (2006). Developing different forms of student feedback to promote teacher reflection: A 10-year collaboration. *Teaching and Teacher Education, 22*, 1006–1019.

Kemmis, S., & McTaggart, R. (Eds.). (1998). *The action research planner*. Geelong, Australia: Deakin University.

Khoo, H. S., & Fraser, B. J. (2008). Using classroom psychosocial environment in the evaluation of adult computer application courses in Singapore. *Technology, Pedagogy and Education, 17*, 53–67.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw.

Moos, R. H. (1974). *The Social Climate Scales: An overview*. Palo Alto, CA: Consulting Psychologists Press.

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education, 9*, 325–335.

Pallant, J. (2011). *SPSS: Survival manual* (4th ed.). Sydney, NSW: Allen and Unwin.

Rentoul, A., & Fraser, B. J. (1979). Conceptualisation of enquiry-based or open classroom learning environments. *Journal of Curriculum Studies, 11*, 233–245.

Rhine, S. (1998). The role of research and teachers' knowledge base in professional development. *Educational Researcher, 27*(5), 27–31.

Schön, D. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass.

Sinclair, B. B., & Fraser, B. J. (2002). Changing classroom environments in urban middle schools. *Learning Environments Research, 5*, 301–328.

Spalding, E., & Wilson, A. (2002). Demystifying reflection: A study of pedagogical strategies that encourage reflective journal writing. *Teachers College Record, 104*, 1393–1421.

Stern, G. G. (1970). *People in context: Measuring person-environment congruence in education and industry*. New York: Wiley.

Stevens, J. P. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research, 27*, 293–302.

Thompson, B. (2001). Significance, effect sizes, stepwise methods and other issues: Strong arguments move the field. *Journal of Experimental Education, 7*, 80–93.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Thorp, H., Burden, R. L., & Fraser, B. J. (1994). Assessing and improving classroom environment. *School Science Review, 75*, 107–113.

Trochim, W., & Donnelly, J. P. (2008). *The research methods knowledge base*. Mason, OH: Cengage Learning.

Walberg, H. J. (1981). A psychological theory of educational productivity. In F. Farley & N. J. Gordon (Eds.), *Psychology and education: The state of the union* (pp. 81–108). Berkeley, CA: McCutchan.

Yarrow, A., Millwater, J., & Fraser, B. J. (1997). Improving university and primary school classroom environments through pre-service teachers' action research. *International Journal of Practical Experiences in Professional Education, 1*(1), 68–93.