Australian Journal of Educational & Developmental Psychology. Vol 4, 2004, pp 110-125

Use of Multitrait-Multimethod Modelling to Validate Actual and Preferred Forms of theTechnology-Rich Outcomes-Focused Learning Environment Inventory (Troflei)

Jill M. Aldridge
Curtin University of Technology
Perth
Australia

Jeffrey P. Dorman[1]
Australian Catholic University
Brisbane
Australia

&

Barry J. Fraser
Curtin University of Technology
Perth
Australia

ABSTRACT

This paper describes the validation of actual and preferred forms of a new classroom environment instrument – the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI) – using multitrait-multimethod modelling. The 80-item TROFLEI assesses 10 classroom environment dimensions: Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, Equity, Differentiation, Computer Usage and Young Adult Ethos. A sample of 1,249 high school students from Western Australia and Tasmania responded to actual and preferred forms of the TROFLEI. Separate exploratory factor analyses for the actual and preferred forms supported the 10 scale *a priori* structure of the instrument. The use of multitrait-multimethod modelling with the 10 scales as traits and the two forms of the instrument as methods supported the TROFLEI's construct validity. The results of this research provide strong evidence of the sound psychometric properties of this new learning environment instrument.

Over the past 35 years, the study of classroom environments has received increased attention by researchers, teachers, school administrators and administrators of school systems. The concept of environment, as applied to educational settings, refers to the atmosphere, ambience, tone, or climate that pervades the particular setting. Because classrooms are essentially about people, research on classroom environments has focussed historically on its

[1] Contact details
Associate Professor Jeffrey P. Dorman
School of Education
Australian Catholic University Limited
P.O. Box 456
VIRGINIA 4014
Australia

ISSN 1446-5442 Web site: http://www.newcastle.edu.au/journal/ajedp/

psychosocial dimensions – those aspects of the environment that focus on human behaviour in origin or outcome (Boy & Pine, 1988). Reviews of classroom environment research by Fraser (1998b), Dorman (2002) and Goh and Khine (2002) have delineated at least 10 domains (e.g., linking the quality of a classroom environment to student cognitive and attitudinal outcomes, using classroom environment assessments to assist teachers improve their classrooms). A review of these domains is outside the scope of this article.

One of the strongest traditions of classroom environment research has been the development of valid instruments for use in classrooms throughout the world (Fraser 1998a). Overwhelmingly, these high-inference measures have assessed students' perceptions of the classroom environment by seeking summary judgments after a period of involvement in classroom events. This approach contrasts with the use of low-inference measures, which focus on specific classroom phenomena (see Dunkin & Biddle, 1974). To a large extent, the truly international status of the learning environment research field is due to the development and reporting of such instruments, including: Classroom Environment Scale (CES; Moos & Trickett, 1987), Individualised Classroom Environment Questionnaire (ICEQ; Fraser, 1990), College and University Classroom Environment Inventory (CUCEI; Fraser & Treagust, 1986), Science Laboratory Environment Inventory (SLEI; Fraser, Giddings, & McRobbie, 1995), Constructivist Learning Environment Survey (CLES; Taylor, Fraser, & Fisher, 1997), Questionnaire on Teacher Interaction (QTI; Wubbels & Levy, 1993), and the What Is Happening In this Class? questionnaire. (WIHIC; Aldridge & Fraser, 2000).

The purpose of the present article is to report the development and validation of a new classroom environment instrument, the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI). The TROFLEI builds upon and extends existing learning environment instrumentation through the use of the WIHIC as a basis for the development of a comprehensive instrument that includes a focus on technology and outcomes in secondary school classrooms. Thus, the unique nature of the TROFLEI is that it has specific dimensions to assess technology and outcomes dimensions of the learning environment. Another distinctive characteristic of the present study is that it employed multitrait-multimethod (MTMM) modelling within a confirmatory factor analysis (CFA) framework. Few validation studies of classroom environment instruments have employed CFA to validate scale structure and there is no history of using MTMM modelling. Before providing details of the present study, background information of the development of classroom environment instruments and the use of multitrait-multimethod modelling are provided.

DEVELOPMENT OF CLASSROOM ENVIRONMENT INSTRUMENTS

Four approaches to the development of classroom environment instruments can be identified: intuitive-rational, intuitive-theoretical, factor analytic and empirical group discriminative (see Fraser, 1986; Hase & Goldberg, 1967). While intuitive-rational and intuitive-theoretical scales rely on the nomination of items to tentative scales prior to questionnaire administration, factor analytic scales employ factor analysis to group items solely on the responses of a sample of the target population being investigated. Empirical group discriminative scales also require test administration prior to scale formation but they are aligned with an external criterion by selecting items that maximise discrimination between groups of respondents (Fraser, 1986). Within the above delineation of approaches, the instrument development reported in this article utilised an initial intuitive-rational approach complemented by the use of exploratory and confirmatory factor analyses. Such analyses can be performed with desktop computers using computer packages that are now widely available (e.g. LISREL: Jöreskog & Sörbom, 1993).

The validity of intuitive-rational scales rests partly on the subjective opinions of the investigators and other experts and usually involves three steps: identification of salient dimensions, writing test items, each of which is linked conceptually with one salient dimension, and field testing the questionnaire. Identifying salient dimensions usually involves a review of literature and the subjective opinions of expert researchers and practitioners. For classroom environment instruments, it is important that these dimensions provide coverage of Moos's (1979) three general categories of human environments (viz. Relationship – the nature and intensity of personal relationships; Personal Growth – opportunities for personal growth and self-enhancement; and System Maintenance and System Change – the extent to which the environment is orderly, clear in expectations, maintains control and is responsive to change). For each dimension, a set of tentative items is written and reviewed by experts in item analysis and classroom environment research. Field testing involves administration of the questionnaire to a sample of the target population and studying the internal consistency (usually employing the Cronbach α coefficient as an index) and discriminant validity (sometimes using the mean correlation of a scale with the remaining scales as a convenient index) of each scale.

MULTITRAIT-MULTIMETHOD MODELLING

In conducting research employing instruments that have multiple scales (or traits) and multiple methods of data collection, there are three central assumptions concerning construct validity: different assessment methods concur in their assessment of the same trait (the issue of convergent validity), different methods diverge in their assessment of the same traits (the issue of discriminant validity), and trait measurements are free of method bias (see Byrne, 1998; Byrne & Bazana, 1996). Byrne and Goffin (1993) reported four approaches to analysing data for such multitrait-multimethod modelling: the Campbell-Fiske model (Campbell & Fiske, 1959), the General CFA model (Widaman, 1985), the Correlated Uniqueness model (Marsh, 1988), and the Composite Direct Model (Browne, 1984). All of these approaches have advantages and limitations and a detailed discussion of these issues is beyond the scope of the present article.

The approach employed in the present paper was the General CFA model which has received considerable attention since Widaman's (1985) article. In this article, Widaman proposed a series of nested CFA analyses to address the construct validity issues described above. Previous successful attempts to use CFA in MTMM modelling have focussed mainly on psychological constructs. Reports on the construct validity of scales to assess self-efficacy (Bong & Hocevar, 2002), depression and anxiety (Cole, Martin, Peeke, Henderson, & Harwell, 1998), and self-concept (Marsh & Byrne, 1993) are examples of this analytic approach. The following section describes the present research.

THE PRESENT STUDY

Sample and Procedures

The sample of students was drawn from high schools in two Australian states (Western Australia and Tasmania). As shown in Table 1, the sample from Western Australia consisted of 772 students from 79 classes in one large metropolitan school. In Tasmania, 477 students from 48 classes in eight schools responded to the questionnaire. Because each student was asked to respond to actual and preferred forms of the TROFLEI, the following section describes these forms in greater detail.

Development of the TROFLEI

As indicated above, the TROFLEI was developed using an intuitive-rational approach complemented by exploratory and confirmatory factor analyses. The first stage of the development of the TROFLEI was made much simpler by using an existing classroom environment instrument, the What Is Happening In this Class? (WIHIC) questionnaire as a starting point. The WIHIC was originally developed by Fraser, McRobbie and Fisher (1996) and attempted to incorporate those scales that previous studies had shown to be predictors of student outcomes. Both *personal forms* and *class forms* of the WIHIC have been developed. The personal form uses the same scales and comparable items as the class form, but is worded

Table 1: Description of Sample

Gender	Sample Size				
	Western Australia		Tasmania		Total
	Year 11	Year 12	Year 11	Year 12	
Male	310	96	157	78	641
Female	275	91	154	88	608
Total	585	187	311	166	1,249

to elicit the student's perceptions of his/her individual role within the classroom, as opposed to the student's perceptions of the class as a whole (Fraser, 1994, 1998a, 1998b; Fraser, Giddings & McRobbie, 1995; Fraser & McRobbie, 1995; Fraser, McRobbie & Fisher, 1996). The personal form is concordant with the constructivist theory of learning (Bruner, 1986; Tobin, 1993; von Glasersfeld, 1989). Based on the assumption that individuals construct their own meaning and knowledge of the world, rather than attaining it from external sources, the personal form enables students to provide individual interpretations of their environment. The personal form was used in the present study.

The robust nature of the What Is Happening In this Class? (WIHIC) questionnaire, in terms of reliability and validity, has been widely reported in studies that have used the instrument in different subject areas, at different age levels and in nine different countries. Since the initial development of the WIHIC, the questionnaire has been used successfully in studies to assess the learning environment in Singapore (Fraser & Chionh, 2000), Australia and Taiwan (Aldridge & Fraser, 2000), Brunei (Khine & Fisher, 2001), Canada (Zandvliet & Fraser, in press), Australia (Dorman, 2001), Indonesia (Adolphe, Fraser & Aldridge, 2003), Korea (Kim, Fisher & Fraser, 2000), the United States (Allen & Fraser, 2002) and Canada, Britain and the US (Dorman, 2003). Within these countries, the WIHIC has been used to assess a range of subjects including high school science (Aldridge & Fraser, 2000), mathematics (Margianti, Fraser & Aldridge, 2001), mathematics and science (Raaflaub & Fraser, 2002) and mathematics and geography (Fraser & Chionh, 2000).

Using a sample of 3980 high school students from Australia, Britain and Canada, confirmatory factor analysis was used to support the seven-scale *a priori* structure of the WIHIC (Dorman, 2003). In this study, Dorman found that all items loaded strongly on their *a priori* scale, although model fit indices revealed a degree of scale overlap. Overall, the study strongly supported the international applicability of the WIHIC as a valid measure of the classroom psychosocial environment.

The robust nature of the WIHIC made it a sensible choice as a starting point for the present study. All seven of the original WIHIC scales were included in the new instrument, namely, Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation and Equity. Three new scales of educational importance were developed for the purpose of this study. To capture the individualised nature of an outcomes-based program, a Differentiation scale was adapted from the Individualised Classroom Environment Questionnaire (ICEQ; Fraser, 1990). This scale assesses the extent to which the teacher provides opportunities for students to choose the topics on which they would like to work and to work at their own pace. Because technology-rich learning environments require students to use computers in a range of ways, the Computer Usage scale was developed to provide information about the extent to which students used a computers in various ways (e.g. email, accessing the internet, discussion forums). Finally, a Young Adult Ethos scale was developed to assess the extent to which teachers give their students responsibility for their own learning.

Historically, negatively-worded items have been used to guard against passive responses. However, Barnette (2000) questions the utility of such items, as they cannot be considered direct opposites of their positively-worded counterparts. In addition, studies reveal that positively-worded items improve response accuracy and internal consistency (Chamberlain & Cummings, 1984; Schreisheim, Eisenbach & Hill, 1991; Schriesheim & Hill, 1981). It was considered appropriate, therefore, to use only items with a positive scoring direction in our study.

Description of Actual and Preferred Forms of the TROFLEI

The TROFLEI consists of 80 items assigned to 10 underlying scales (8 items per scale). Table 2 shows scale names and descriptions. Students respond to items using a five-point frequency response format (viz. Almost Never, Seldom, Sometimes, Often, Almost Always). To provide contextual cues and to minimise confusion to students, it was considered appropriate to group together in blocks items that belong to the same scale instead of arranging them randomly or cyclically (Aldridge, Fraser, Taylor, & Chen, 2000). Scale scores for each respondent are obtained by aggregating scores for the eight items for that scale. Of particular relevance to the present study is the distinction between *actual* and *preferred* forms of the TROFLEI. While the actual form elicits information on what students perceive to be the current classroom environment, the preferred form assesses students' perceptions of what environment they would like in the classroom. For example, the actual form TROFLEI item *I explain my ideas to other students* has a corresponding preferred form item *I like explaining my ideas to other students*. For each item of the TROFLEI, students record their perceptions of actual and preferred environments on adjacent response scales.

Data Analysis

There were three distinct components to the analyses conducted in the present study. First, the mean, standard deviation, Cronbach alpha coefficient (as an index of internal consistency reliability) and mean correlation of a scale with the remaining scales (as an index of discriminant validity) were computed for each of the 10 *a priori* scales for both actual and preferred forms of the TROFLEI. The second component of data analysis consisted of separate exploratory factor analyses on the actual and preferred forms of the TROFLEI.

The third component of the data analysis was a multitrait-multimethod (MTMM) analysis within a CFA framework using LISREL 8.3 (Jöreskog & Sörbom, 1993). In the present study, the 10 TROFLEI scales were traits and the two forms of the instrument were taken to be methods, an arrangement that falls within Marsh and Grayson's (1997) general MTMM definitions. Widaman's (1985) theory on MTMM compares a hypothesised baseline MTMM model with a series of nested, more restrictive models. To perform these analyses, some preliminary data manipulation was required. Separate exploratory factor analyses were performed for each set of *a priori* scale items. These 20 factor analyses computed factor score coefficients and subsequent factor scores which were taken as scale scores for each of the 20 observed variables. According to Holmes-Smith and Rowe (1994), scale reliability is maximised if the weight on each item (i.e. coefficient) is the corresponding factor score coefficient. In the baseline model, each TROFLEI scale (or trait) is assessed using actual and preferred forms (method). Figure 1 clarifies this arrangement with the 20 observed variables shown in rectangles and the latent variables which are not measured directly shown in ellipses.

Table 2: Descriptive Information and Scale Statistics for 10 TROFLEI Scales

Scale Name	Scale Description	Cronbach α		Mean Correlation		Scale Mean		Scale Standard Deviation	
		Act.	Pref.	Act.	Pref.	Act.	Pref.	Act.	Pref.
Student Cohesiveness	The extent to which students know, help and are supportive of one another.	.88	.90	.34	.42	31.34	33.79	5.41	5.52
Teacher Support	The extent to which the teacher helps, befriends, trusts and is interested in students.	.92	.92	.35	.43	28.70	31.31	6.74	6.38
Involvement	The extent to which students have attentive interest, participate in discussions, do additional work and enjoy the class.	.89	.92	.38	.45	25.69	28.26	6.20	6.67
Investigation	The extent to which skills and processes of inquiry and their use in problem solving and investigation are emphasised.	.88	.94	.34	.43	31.34	35.52	5.60	5.48
Task Orientation	The extent to which it is important to complete activities planned and to stay on the subject matter.	.93	.95	.32	.42	23.47	27.85	6.92	7.86
Cooperation	The extent to which students cooperate rather than compete with one another on learning tasks.	.91	.94	.38	.47	30.12	32.61	6.42	6.63
Equity	The extent to which students are treated equally by the teacher.	.94	.95	.34	.41	32.63	34.80	6.80	6.02
Differentiation	The extent to which teachers cater for students differently on the basis of ability, rates of learning and interests.	.77	.84	.17	.23	23.80	26.48	6.33	7.03
Computer Usage	The extent to which students use their computers as a tool to communicate with others and to access information.	.88	.90	.16	.25	23.79	27.37	7.93	8.04
Young Adult Ethos	The extent to which teachers give students responsibility and treat them as young adults.	.94	.94	.33	.38	33.13	35.54	6.24	5.41

Note. Act. = Actual form of TROFLEI, Pref. = Preferred form of TROFLEI

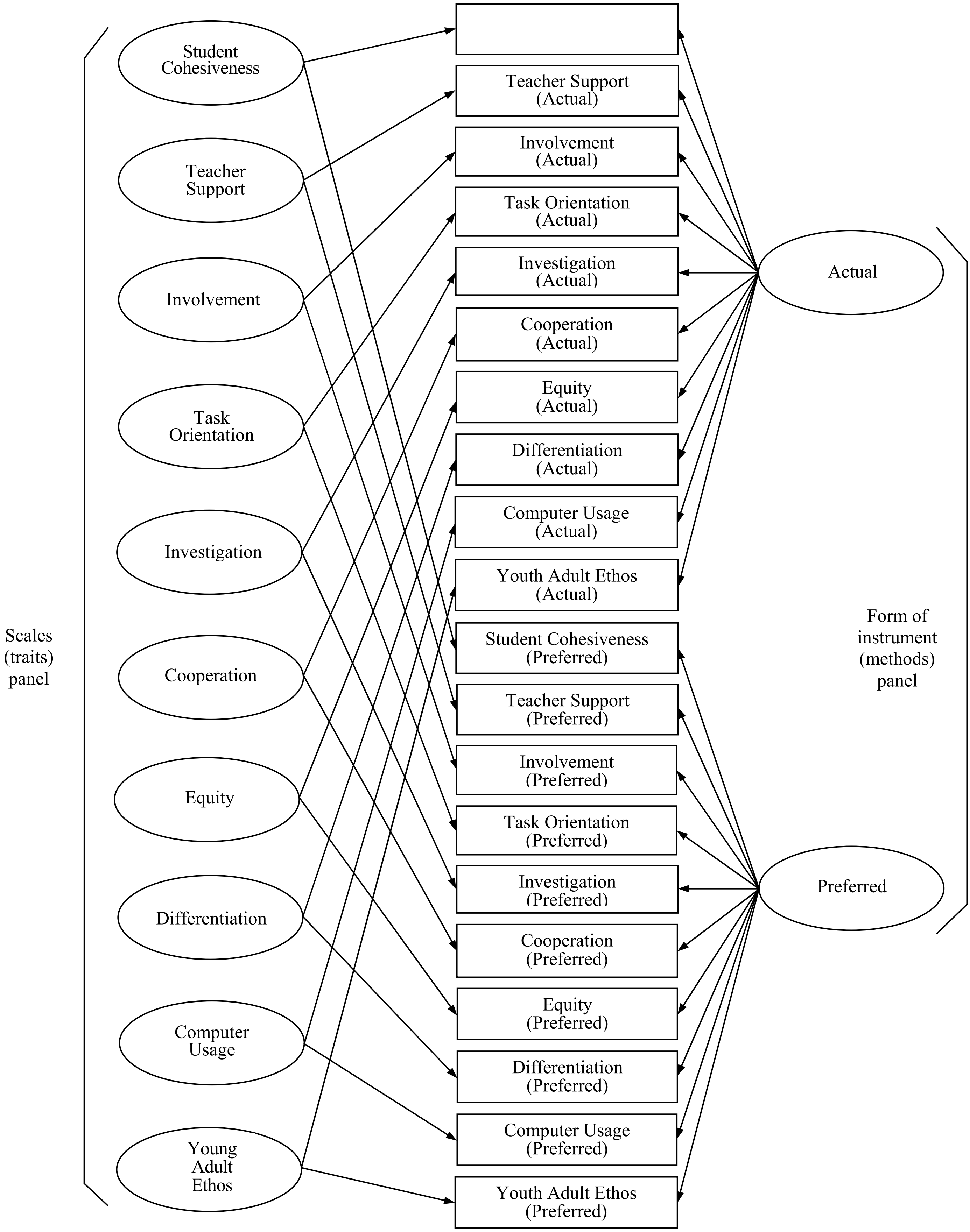

Figure 1. Hypothesised Baseline MTMM Model
Note. Paths showing intercorrelations among latent variables have been omitted.

Of the many indices available to report model fit, model comparison and model parsimony, seven indices are reported in this component of the data analysis. For model fit, the Root Mean Square Error of Approximation (RMSEA), Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI) are provided. Model comparison is represented by the Tucker-Lewis Index (TLI) and the Comparative Fit Index (CFI). The Parsimony Goodness of Fit Index (PGFI) and the Parsimony Normed Fit Index (PNFI) are indices of model parsimony. To interpret these indices, the following rules which are generally accepted in CFA literature were accepted. RSMEA should be below .05 with perfect fit indicated by an index of zero. GFI, AGFI, TLI, CFI should be above .90 with perfect fit indicated by an index of 1.00. PGFI and PNFI should be above .50. Acceptable values of these indices have been discussed by Schumacker and Lomax (1996), Byrne (1998) and Kelloway (1998). Finally, the coefficient of determination, which provides an indication of the quality of the overall model, was computed for each form of the TROFLEI.

Apart from this first model to be tested (Model 1), three more restrictive models were tested. In Model 2, there are no scales (traits) and the forms of the TROFLEI (methods) are freely correlated. Comparing Models 1 and 2 allows an assessment of convergent validity. Model 3 has perfectly correlated scales (traits) and freely correlated forms of the TROFLEI (methods). Comparing Models 1 and 3 allows an assessment of discriminant validity across the scales. Model 4 has freely correlated scales (traits) and uncorrelated forms of the TROFLEI (methods). Comparing Models 1 and 4 provides an indication of discriminant validity across the forms of the TROFLEI (methods). In addition to these comparisons at the model level, individual estimates of parameters were calculated. These analyses allowed the amount of variance explained by scales of the TROFLEI (traits) and forms (methods). The square of the factor loadings represents the percentage of variance explained by trait and method. As noted by Byrne and Bazana (1996), the sum of these two variance components equals the squared multiple correlation for a scale of the TROFLEI.

RESULTS

Descriptive Statistics

The results presented in Table 2 indicate that all scales of the actual and preferred forms of the TROFLEI had good internal consistency reliability. Cronbach coefficient alpha ranged from .77 for the actual form of the Differentiation scale to .95 for the preferred form of the Task Orientation and Equity scales. Discriminant validity (using the mean correlation of a scale with the remaining nine scales as an index) ranges from .16 for the actual form of the Computer Usage scale to .47 for the preferred form of the Cooperation scale. These values suggest that the TROFLEI's scales are distinct but tend to overlap. Given that the TROFLEI has 10 scales, this result is not surprising. For all scales, the minimum recorded score was 8 (the lowest possible score) and the maximum recorded score was 40 (the highest possible score). A ceiling effect was evident for most scales. Apart from actual Involvement and actual Investigation, all scales had significant negative skewness ($p<.05$). This was especially so for the preferred Task Orientation and preferred Young Adult Ethos scales, both of which had medians of 38. Additionally, kurtosis for 15 of the 20 scales differed significantly from zero ($p<.05$).

Table 3 shows the results of exploratory factor analyses (EFA) on the Actual and Preferred Forms of the TROFLEI. A principal components analysis with varimax rotation yielded 10 factors for both the actual and preferred forms of the TROFLEI. These factors accounted for 62.8% and 69.3% of variance in scores on the actual and preferred forms of the TROFLEI, respectively. All items had loadings of at least .41 with the factor corresponding to their *a priori* scale and below .35 with other factors.

Table 3: Results of Exploratory Factor Analyses for Actual and Preferred Forms of the TROFLEI

	Loading of apriori Scale Item on Factor							
Factor	1	2	3	4	5	6	7	8
Student Cohesiveness	.73 (.75)	.67 (.73)	.62 (.73)	.78 (.78)	.61 (.69)	.50 (.59)	.74 (.74)	.50 (.47)
Teacher Support	.75 (.73)	.78 (.77)	.73 (.73)	.60 (.62)	.71 (.73)	.74 (.74)	.72 (.71)	.67 (.62)
Involvement	.78 (.73)	.79 (.72)	.58 (.71)	.66 (.71)	.54 (.59)	.66 (.69)	.52 (.62)	.58 (.68)
Task Orientation	.73 (.73)	.55 (.72)	.77 (.81)	.64 (.78)	.84 (.85)	.81 (.84)	.82 (.82)	.78 (.80)
Investigation	.65 (.63)	.64 (.70)	.66 (.67)	.63 (.68)	.74 (.76)	.63 (.72)	.60 (.70)	.64 (.73)
Cooperation	.67 (.63)	.67 (.69)	.67 (.59)	.77 (.74)	.69 (.69)	.75 (.75)	.71 (.70)	.68 (.68)
Equity	.70 (.66)	.71 (.71)	.73 (.74)	.78 (.75)	.79 (.77)	.76 (.75)	.71 (.74)	.77 (.74)
Differentiation	.41 (.42)	.43 (.56)	.63 (.62)	.77 (.78)	.44 (.51)	.81 (.83)	.78 (.79)	.76 (.78)
Computer Usage	.56 (.61)	.79 (.79)	.81 (.81)	.78 (.80)	.80 (.84)	.81 (.81)	.60 (.66)	.54 (.53)
Young Adult Ethos	.77 (.77)	.79 (.75)	.76 (.72)	.79 (.79)	.68 (.76)	.71 (.73)	.77 (.76)	.74 (.74)

Note. Loadings for the Preferred form of the TROFLEI are shown in parentheses.

Multitrait-Multimethod Analysis

As indicated earlier in this paper, TROFLEI analyses were conducted at two levels. At the model or global matrix level, four models were fitted to the data and differences between these models examined. In the second set of analyses, individual parameter estimates were reviewed so as to provide a more exact assessment of variance related to scale (trait) and form of the TROFLEI (method).

Comparisons at the Model Level. To investigate the construct validity of the TROFLEI, LISREL was employed to compute goodness-of-fit indices for the four MTMM models described above (see Table 4). Model 1, which has freely-correlated scales and freely-correlated forms of the instrument, was the baseline model against which the other three models were compared. It had the best overall model fit to the data. Models 2 and 3 had very poor fit with TLI indices of .42 and .50, respectively. Results in Table 4 show that Model 4 (freely-correlated scales and uncorrelated forms of instrument) had good fit and compared favourably with Model 1.

To summarise differences among the models, Table 5 has been assembled. Model 2 has no scales and freely-correlated forms of the TROFLEI. Comparing Models 1 and 2 assesses convergent validity. A significant $\Delta\chi^2$ for Δdf and sizeable ΔCFI and ΔGFI indicate strong convergent validity. As $\Delta\chi^2$ (71, N = 1,249) = 6,012.88 (p<.001) and ΔCFI and ΔGFI were very large (.49 and .40, respectively), it can be concluded that strong convergent validity exists.

Table 4: Summary of Goodness-of fit-Indices for MTMM Models

	Model	χ^2	df	RMSEA	CFI	GFI	TLI	AGFI
1	Freely correlated scales (traits); Freely correlated forms of instrument (methods)	341.84	98	.05	.98	.96	.98	.91
2	No scales (traits); Freely correlated forms of instrument (methods)	6354.72	169	.21	.49	.56	.42	.45
3	Perfectly correlated scales (traits); Freely correlated forms of instrument (methods)	4853.15	149	.19	.61	.65	.50	.51
4	Freely correlated scales (traits); Uncorrelated forms of instrument (methods)	356.13	105	.05	.98	.96	.95	.92

Table 5: Differential Goodness-of-Fit Indices for MTMM Nested Model Comparisons

Model Comparisons	$\Delta\chi^2$	Δdf	ΔCFI	ΔGFI
Test of Convergent Validity				
Model 1 versus Model 2 (scales)	6012.88	71	.49	.40
Tests of Discriminant Validity				
Model 1 versus Model 3 (scales)	4511.31	51	.37	.31
Model 1 versus Model 4 (forms)	14.29	7	.00	.00

Model 3 has perfectly-correlated scales (traits) and freely-correlated forms of the instrument (methods). Comparing Models 1 and 3 allows an assessment of discriminant validity of the scales. A significant $\Delta\chi^2$ for Δdf and sizeable ΔCFI and ΔGFI indicate strong discriminant validity. As $\Delta\chi^2$ (51, N = 1,249) = 4,511.31 (p<.001) and ΔCFI and ΔGFI were large (.37 and .31, respectively), it can be concluded that strong discriminant validity of the scales exists. Model 4 has freely-correlated scales (traits) and uncorrelated forms of the TROFLEI (methods). Comparing Models 1 and 4 provides an indication of discriminant validity of the forms of the TROFLEI (methods) (i.e. method effects). In this comparison, small differences in χ^2, CFI and GFI indicate discrimination across the forms of the TROFLEI. It can be observed from Table 5 that, while $\Delta\chi^2$ was significant [$\Delta\chi^2$ (7, N = 1,249) = 14.29 (p<.001)], the differences in the CFI and GFI between Models 1 and 4 were very small. Overall, there is evidence of some method discrimination.

Examination of Individual Parameter Estimates. As noted earlier in this article, convergent validity at the parameter level can be assessed by considering variance proportions computed from factor loadings for scale and form of the TROFLEI. Table 6 shows these results. Clearly, variance explained by scales exceeded that explained by form of the TROFLEI for all 20 scales. The largest difference was for the actual form of the Cooperation scale for which 88% of variance was explained by the scale but only 1% was explained by form of the TROFLEI. An inspection of Table 6 data suggests that there was stronger convergent validity for the actual form compared to the preferred form of the TROFLEI. Overall, these data provide strong support for the convergent validity findings that were identified in the matrix-level modelling.

Table 6: Variance explained by Scales (Traits), Forms (Methods) and Error for Model 1

	Scale (Trait)										Form (Method)		Error
	SC	TS	IN	TO	IV	CO	E	DI	CU	YAE	Act.	Pref.	
Actual Form													
SC	.81										.01		.18
TS		.55									.31		.14
IN			.72								.02		.26
TO				.64							.04		.32
IV					.85						.01		.14
CO						.88					.01		.11
E							.55				.29		.16
DI								.71			.06		.23
CU									.72		.02		.26
YAE										.58	.18		.24
Pref. Form													
SC	.71											.15	.14
TS		.77										.13	.10
IN			.67									.19	.14
TO				.58								.25	.17
IV					.62							.29	.09
CO						.69						.23	.08
E							.76					.18	.06
DI								.74				.14	.12
CU									.72			.16	.12
YAE										.58		.20	.22

Note. SC: Student Cohesiveness, TS: Teacher Support, IN: Involvement, TO: Task Orientation, IV: Investigation, CO: Cooperation, E: Equity, DI: Differentiation, CU: Computer Usage, YAE: Young Adult Ethos.

Discriminant validity at the parameter level was investigated with factor correlation matrices. The results shown in Table 7 show moderate to strong positive correlations among the scales with correlations ranging from -.11 for Computer Usage with Equity to .72 for Cooperation with Student Cohesiveness. Clearly there is scale overlap. These results confirm the widely-held view that classroom environment instruments tend to have conceptually-distinct but empirically-overlapping scales (Fraser, 1998a). The correlation of .24 between the actual and preferred forms of the TROFLEI indicates a sound level of method discrimination.

Table 7: Scale (Trait) and Form (Method) Correlations for MTMM Model 1

	Scale (Trait)										Forms (Methods)	
Measures	SC	TS	IN	TO	IV	CO	E	DI	CU	YAE	Act.	Pref.
Scale (Trait)												
SC	1.00											
TS	.50	1.00										
IN	.59	.64	1.00									
TO	.54	.56	.48	1.00								
IV	.37	.37	.60	.45	1.00							
CO	.72	.47	.58	.57	.47	1.00						
E	.44	.61	.47	.68	.31	.51	1.00					
DI	.08	.10	.23	-.02	.33	.15	-.08	1.00				
CU	.16	.05	.19	-.06	.22	.21	-.11	.42	1.00			
YAE	.39	.52	.39	.70	.32	.51	.71	-.03	.01	1.00		
Form (Method)												
Actual											1.00	
Preferred											.24	1.00

Note. SC: Student Cohesiveness, TS: Teacher Support, IN: Involvement, TO: Task Orientation, IV: Investigation, CO: Cooperation, E: Equity, DI: Differentiation, CU: Computer Usage, YAE: Young Adult Ethos.

DISCUSSION

The research reported in this article is important to the assessment of learning environment for three reasons. First, the specific findings of these analyses have shown the TROFLEI to have very sound structural characteristics. A large amount of true-score variance was explained by the scales (traits) rather than forms of instrument (methods). That is, there was little attenuation of scales by form of instrument. This is an important characteristic and supports the view that the TROFLEI is an important addition to the suite of high-inference classroom environment instruments developed over the past 35 years.

Second, this research is one of the few reported attempts to employ CFA in validating the structure of learning environment instruments. Various reviews of learning environment research and instruments (e.g. Fraser, 1994, 1998b) and validation studies of specific instruments (e.g. Fisher & Waldrip, 2002; Thomas, 2003) have typically used exploratory factor analysis to establish factor structure. Few of these studies have reported convergent validity data at both the matrix and parameter levels.

Third, no learning environment studies to date have employed multitrait-multimethod modelling within a CFA framework to study convergent and discriminant validity when more than one form of an instrument or more than one method of data collection have been employed. That is, the present study breaks new ground in the learning environment field. Of particular note is the fact that classroom environment data are often collected from more than one source with different forms of an instrument. For example, researching the differences between the perceptions of students and teachers has been one of the enduring lines of classroom environment research. MTMM could be used to study the construct validity of student and teacher forms of the one classroom environment instrument. Another possible use of MTMM modelling concerns the use of low-inference and high-inference measures of the same classroom environment dimensions. Whereas low-inference measures focus on discrete observable classroom phenomena, high-inference measures require students and teachers to make summary judgments based on long-term immersion in the environment. Of particular interest here is the issue of method bias and whether scales are free of systematic measurement error, halo effects and other sources of rater bias (see Pedhazur & Schmelkin, 1991).

CONCLUSION

This article has reported the validation of a new classroom environment instrument, the Technology-Rich Outcomes-Focused Learning Environment Inventory (TROFLEI), using a multitrait-multimethod approach within a confirmatory factor analysis framework. Using a sample of male and female high school students in Western Australia and Tasmania, this study had shown the TROFLEI to be a valid measure of classroom environment. As teachers, administrators, and learning environment researchers fulfill important professional roles, they need valid instruments to assess contemporary classroom environments. This research has provided substantive validation of the actual and preferred forms of the TROFLEI, although further validation work with the TROFLEI should be conducted in other countries.

REFERENCES

Adolphe, F. S. G., Fraser, B. J., & Aldridge, J. M. (2003, January). A cross-national study of classroom environment and attitudes among junior secondary science students in Australia and Indonesia. In D. Fisher & T. Marsh (Eds.), *Science, mathematics and technology education for all: Proceedings of the Third International Conference on Science, Mathematics and Technology Education* (pp. 435-446). Perth, Australia: Curtin University of Technology.

Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research, 3*, 101-134.

Aldridge, J. M., Fraser, B. J., Taylor, P. C., & Chen, C.-C. (2000). Constructivist learning environments in a cross-national study in Taiwan and Australia. *International Journal of Science Education, 22,* 37-55.

Allen, D., & Fraser, B. J. (2002). *Parent and student perceptions of the classroom learning environment and its influence on student outcomes.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*, 361-70.

Bong, M., & Hocevar, D. (2002). Measuring self-efficacy: A multitrait-multimethod comparison of scaling procedures. *Applied Measurement in Education, 15,* 143-171.

Boy, A. V., & Pine, G. J. (1988). *Fostering psychosocial development in the classroom.* Springfield, IL: Charles C. Thomas.

Browne, M. W. (1984). The decomposition of multitrait-multimethod matrices. *British Journal of Mathematical and Statistical Psychology, 37*, 1-21.

Bruner, J. S. (1986). *Actual minds, possible worlds.* Boston, MA: Harvard University Press.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications and programming.* Mahwah, NJ: Erlbaum.

Byrne, B. M., & Bazana, P. G. (1996). Investigating the measurement of social and academic competencies for early/late preadolescents and adolescents: A multitrait-multimethod analysis. *Applied Measurement in Education, 9*, 113-132.

Byrne, B. M., & Goffin, R. D. (1993). Modeling MTMM data from additive and multiplicative covariance structures: An audit of construct validity concordance. *Multivariate Behavioral Research, 28*, 67-96.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Chamberlain, V. M., & Cummings, M. N. (1984). Development of an instructor/course evaluation instrument. *College Student Journal, 18,* 246-250.

Cole, D. A., Martin, J. M., Peeke, L., Henderson, A., & Harwell, J. (1998). Validation of depression and anxiety measures in white and black youths: Multitrait-multimethod analyses. *Psychological Assessment, 10,* 261-276.

Dorman, J. P. (2001). Associations between classroom environment and academic efficacy. *Learning Environments Research, 4*, 243-257.

Dorman, J. P. (2002). Classroom environment research: Progress and possibilities. *Queensland Journal of Educational Research, 18,* 112-140.

Dorman, J. P. (2003). Cross national validation of the What Is Happening In This Class questionnaire using confirmatory factor analysis. *Learning Environments Research, 6,* 231-245.

Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching.* New York: Holt, Rinehart & Winston.

Fisher, D. L., & Waldrip, B. G. (2002). Measuring culturally sensitive factors of classroom learning environments with the CLEQ. In S. C. Goh & M. S. Khine (Eds.), *Studies in educational learning environments: An international perspective* (pp. 27-48). Singapore: World Scientific.

Fraser, B. J. (1986). *Classroom environment*. London: Croom Helm.

Fraser, B. J. (1990). *Individualised Classroom Environment Questionnaire*. Melbourne: Australian Council for Educational Research.

Fraser, B. J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 493-541). New York: Macmillan.

Fraser, B. J. (1998a). Classroom environment instruments: Development, Validity, and applications. *Learning Environments Research, 1*, 7-33.

Fraser, B. J. (1998b). Science learning environments: Assessments, 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., & Chionh, Y. H. (2000, April). *Classroom environment, self-esteem, achievement, and attitudes in geography and mathematics in Singapore.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. *Journal of Research in Science Teaching, 32*, 399-422.

Fraser, B. J., & McRobbie, C. J. (1995). Science laboratory classrooms at schools and universities: A cross-national study. *Educational Research and Evaluation: An International Journal on Theory and Practice, 1*(4), 1-25.

Fraser, B. J., McRobbie, C. J., & Fisher, D. L. (1996, April). *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.
Fraser, B. J., & Treagust, D. F. (1986). Validity and use of an instrument for assessing classroom psychological environment in higher education. *Higher Education, 15*, 37-57.
Goh, S. C., & Khine, M. S. (Eds.). (2002). *Studies in educational learning environments: An international perspective*. Singapore: World Scientific.
Hase, H. D., & Goldberg, L. G. (1967). Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin, 67*, 231-248.
Holmes-Smith, P., & Rowe, K. J. (1994, January). *The development and use of congeneric measurement models in school effectiveness research: Improving the reliability and validity of composite and latent variables for fitting multilevel and structural equation models.* Paper presented at the International Congress for School Effectiveness and Improvement, Melbourne, Australia.
Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.
Kelloway, E. K. (1998). *Using LISREL for structural equation modeling: A researcher's guide*. Thousand Oaks, CA: Sage.
Khine, M. S., & Fisher, D. L. (Eds.). (2003). *Technology-rich learning environments: A future perspective*. Singapore: World Scientific.
Kim, H., Fisher, D., & Fraser, B. (2000). Classroom environment and teacher interpersonal behaviour in secondary science classes in Korea. Evaluation and Research in Education, 14, 3-22.
Margianti, E. S., Fraser, B. J., & Aldridge, J. M. (2001, April). *Classroom environment and students' outcomes among university computing students in Indonesia.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
Marsh, H. W. (1988). Multitrait-multimethod analysis. In J. P. Keeves (Ed.), *Educational research methodology, Measurement, and evaluation: An international handbook* (pp. 570-578). Oxford, England: Pergamon.
Marsh, H. W., & Byrne, B. M. (1993). Confirmatory factor analysis of multitrait-multimethod self-concept data: Between-group and within-group invariance constraints. *Multivariate Behavioral Research, 28,* 313-349.
Marsh, H. W., & Grayson, D. (1997). Multitrait-multimethod analysis. In J. P. Keeves (Ed.), *Educational research, Methodology, and measurement: An international handbook* (2nd ed., pp. 600-607). Oxford, England: Pergamon.
Moos, R. H. (1979). *Evaluating educational environments: Procedures, measures, findings and policy implications*. San Francisco: Jossey Bass.
Moos, R. H., & Trickett, E. J. (1987). *Classroom environment scale manual* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.
Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design and analysis: An integrated approach.* Hillsdale, NJ: Erlbaum.
Raaflaub, C. A., & Fraser, B. J. (2002, April). *Investigating the learning environment in Canadian mathematics and science classes in which computers are used.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Erlbaum.
Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement, 51,* 67-78.
Schriesheim, C. A., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. *Educational and Psychological Measurement,* 41, 1101-1114.

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

Thomas, G. P. (2003). Conceptualisation, Development and validation of an instrument for investigating the metacognitive orientation of science classroom learning environments: The Metacognitive Orientation Learning Environment Scale - Science (MOLES-S). *Learning Environments Research, 6*, 175-197.

Tobin, K. (Ed.). (1993). *The practice of constructivism in science education.* Washington: AAAS.

von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Syntheses, 80,* 121-140.

Widaman, K. F. (1985). Hierarchically tested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9*, 1-26.

Wubbels, T., & Levy, J. (Eds.). (1993). *Do you know what you look like? Interpersonal relationships in education*. London: Falmer.

Zandvliet, D. B., & Fraser, B. J. (in press). Learning environments in IT classrooms. *Technology, Pedagogy and Education.*