ANALISIS 7(1&2), 155-174 (2000)

TESTING FACTORIAL VALIDITY OF MOTIVATIONAL CONSTRUCTS USING CONFIRMATORY FACTOR ANALYSIS

ROSNA AWANG HASHIM
School of Cognitive Sciences and Education
Universiti Utara Malaysia

ABSTRACT

The present study examines the validity of the newly translated Malay version of the "State Self-assessment" and "Trait Self-assessment" instruments. These instruments are each composed of three constructs that are considered integral to the social cognitive theories of motivation: self-efficacy, effort, and worry. Data were collected from 360 college undergraduates at Universiti Utara Malaysia. A confirmatory factor analysis was used to test the measurement model. Multiple group invariance analyses were also conducted to examine whether components of the measurement model were invariant across gender and ethnic groups (Malay and Chinese from Malaysia). Findings provide support for the reliability and factorial validity of the constructs. Results from the multi-group analyses suggest that all constructs are structured similarly across gender and ethnic groups.

ABSTRAK

Kajian ini meneliti kesahihan instrumen 'State Self-assessment' dan 'Trait Self-assessment' versi Bahasa Melayu. Setiap instrumen di atas mengandungi tiga konstruk yang dianggap penting dalam teori motivasi kognitif sosial: keberkesanan kendiri, usaha, dan kerisauan. Sampel kajian terdiri daripada 360 pelajar Universiti Utara Malaysia. Analisis 'confirmatory factor analysis' telah digunakan untuk menguji model pengukuran. Analisis invarian pelbagai kumpulan telah juga dijalankan untuk meneliti sama ada komponen dalam model pengukuran adalah invarian berdasarkan jantina dan kumpulan etnik (Melayu dan Cina dari Malaysia). Dapatan kajian menyokong kesahihan faktorial setiap konstruk. Dapatan dari analisis invarian berkumpulan menunjukkan bahawa tiada perbezaan antara jantina dan kumpulan etnik dalam model pengukuran, sekali gus menyokong ketidakbiasan model ini terhadap responden mengikut jantina atau kumpulan etnik.

OBJECTIVE

This paper reports a study on reliability, validity, and equivalence of the newly translated version of self-report measures of trait and state motivational constructs. Specifically, the study aimed to explore cross-cultural validity of the State and Trait Self-Assessment measures for Malaysian college undergraduates. The instruments, adapted from O'Neil and his associates (Huang, 1996; Malpass, 1994; O'Neil & Abedi, 1992, 1996; O'Neil & Herl, 1998), were designed to examine the following motivation constructs: trait effort, trait ability, trait self-efficacy, state self-efficacy, state effort, and state worry. Thus, the purpose of this study is threefold: (1) to obtain evidence of internal consistency reliability of the newly translated instrument; (2) to test factorial validity of the translated instrument; and 3) to examine constructs equivalence across gender and ethnic groups (Malay and Chinese from Malaysia).

THEORETICAL FRAMEWORK

The theoretical framework for the study was based on social cognitive theories of motivation (Bandura, 1997, 1998; Zimmerman, 1996, 1998, 2000) and state-trait anxiety theory (Spielberger, 1975, 1980, Spielberger & Vagg, 1995). Following Spielberger's (1975) state-trait anxiety theory, O'Neil and his colleagues (O'Neil & Abedi, 1992, 1996; O'Neil & Herl, 1998) have further applied these trait-state conceptions to both cognitive and affective measurement. The trait-state distinction is made along stable and unstable dimensions where traits are perceived as relatively enduring predispositions or characteristics of individuals, and states are perceived as temporal and situational specific. The instruments examined in the present study are each composed of three constructs which are considered integral to the social cognitive theories of motivation (Bandura, 1997,1998; Pintrich & Schunk, 1996): self-efficacy, effort, and worry.

Social cognitive approach of motivation seeks to examine the interconnectedness of individual' perception, inferences and interpretations of social experiences to achievement behaviors. The role of personal agency constitutes the key factor in social cognitive models of learning (Bandura, 1986, 1997; Zimmerman, 1989, 1996, 2000). Self-beliefs, such as self-efficacy, have been shown to be significant predictors and/or mediators of motivation-related variables such as effort, persistence, anxiety, and task choice (Bandura, 1997; Pintrich & Schunk, 1996; Rosna, 2000; Schunk & Zimmerman, 1997). For instance, the levels of effort expended may be influenced by perception of efficacy; highly self-efficacious individuals are more likely to invest greater effort than low efficacious individuals (Rosna, 2000; Schunk & Zimmerman, 1997). Worry, on the other hand, was found to be negatively correlated with achievement because its manifestation, such as negative expectation, may demand or take up some or all of the individual cognitive processes in the working memory (Hembree, 1988, 1992; O'Neil & Fukumura, 1992; Rosna, 2000; Zeidner, 1998).

METHOD

Instruments

The study employed Trait and State self-assessment questionnaires which were adapted from various studies (see O'Neil & Herl, 1998) with some modified wording to make them appropriate to the subject area and to the Malaysian context. Traits are relatively stable across situations, but states vary in intensity across situations, depending on specific environmental circumstances. Traits constructs are then seen as stable individual difference variables to respond to intellectual situations with varying degrees of state constructs (O'Neil & Abedi, 1996).

Trait Instruments: The 24-item Self-Assessment Trait Questionnaire (see Appendix A) consists of three subscales – students' perceived self-efficacy, effort, and worry–each of which comprises 8 items. Since traits are considered stable characteristics of individuals' predispositions, they are rated on a frequency dimension (O'Neil & Abedi, 1992). Students respond to each of the trait items along a four-point response scale: 1=Almost Never; 2=Sometimes; 3=Often; 4=Almost Always. Students were asked to describe how they generally think or feel. A brief description of the trait scales follows:

1. Trait Self-efficacy is defined as students' judgment about their ability to accomplish a specific task (8 items). Originally developed by Pintrich and De Groot, the scale was later adapted by O'Neil and his colleagues (see O'Neil & Herl, 1998). An example of the item is: "I'm confident I can learn the most complex material presented by the instructor in this course."
2. Trait Effort is defined as the extent to which one generally works hard on a task (8 items). This scale was adapted from Huang's (1996) study. A sample item of this scale is: " I put forth my best effort on tasks."
3. Trait worry refers to the cognitive concerns about the consequences of failure in a test situation. It is a situation-specific personality trait, particularly in evaluative situations (Spielberger, 1975, 1980). In this study, trait worry (8 items) was defined as a stable personality trait to respond to evaluative situations (O'Neil & Fukumura, 1992). A sample item of this scale is: "During examinations I get so nervous that I forget facts I really know."

State instrument: The 22-item Self-Assessment State Questionnaire (see Appendix B) also consists of three scales–self-efficacy, effort, and worry. Originally designed by O'Neil et al. (1992) as the State Self-regulation Questionnaire, the instruments were then adapted by Malpass (1994) and Huang (1996). In this study, the state effort and state worry scales were adapted from Huang (1996). The state self-efficacy scale adapted five items from Huang (1996), one item from Malpass (1994), and two items from Pintrich et al. (1990). The state questionnaire follows a different format from the trait questionnaire. Since state constructs vary in intensity and change over time, students' responses were measured using an intensity scale: 1= Not At All; 2= Somewhat; 3=

Moderately So, and 4= Very Much So. Students were asked to respond according to their thoughts or feelings they experienced while taking the Applied Statistics exam. A brief description of the three state scales follows:

1. State Self-efficacy is defined as students' temporal judgment about their ability to accomplish a specific task. In this study, state self-efficacy (8 items) refers to students' temporal self-appraisal of their ability to perform in the statistics exam. A sample item is: " I'm sure I did an excellent job on the questions on this test."
2. State Effort is defined as students' temporal mental exertion expended for a specific task. Specifically, in this study effort refers to students' willingness to keep trying on their statistics exam. A sample item is: "I tried to do my best on the exam."
3. State Worry is defined as the level of worry experienced by students while taking a test. Worry is the cognitive component of the anxiety experience. It is also situation specific as it varies in intensity and fluctuates across situations (O'Neil & Fukumura, 1992). A sample item is: " I'm worried that I would get a bad grade."

Criterion Measure

The criterion measures were two Applied Statistics mid-term tests. The tests were problem-solving in nature and lasted for about an hour and a half. As participants came from four different intact classes under two different lecturers, the two mid-term scores were initially standardized within classes and later converted to t-scores ($\underline{M}$= 50; $\underline{SD}$ =10). The mid-term tests, which accounted for 20% of the total grade, were very important to the final grades that students were getting for the Statistics course.

Translation Process

To ensure cross-cultural equivalence of the instruments, both the state and trait questionnaires were translated into the Malay Language, (the national language) and also the primary language of instruction in Malaysia, using the back-translation suggested by Brislin (1970, 1980).

Participants

Participants consisted of 360 undergraduates who were enrolled in undergraduate level Statistics at the Northern University of Malaysia. Using a convenience sampling method four different classes were selected to participate in this study. The participants were 121 (33.6%) males and 239 (66.4%) females of whom 241 (66.9%) were Malays and 119 (33.1%) were Chinese. The mean age was 21.6 years old ($\underline{SD}$ =1.87) with the median age at 21.

Procedure

Trait Self-assessment Questionnaires were group administered to 4 classes a month before the mid-term exam. State Self-assessment Questionnaires were group administered immediately after the mid-term exam. Both questionnaires took about 15 minutes each to complete.

RESULTS

Descriptive Statistics

Descriptive statistics results were computed for the six scales and the statistics mid-term exam using SPSS for Windows v8.0 (1998). Table 1 summarizes means (M) and standard deviations (SD) for each scale and the mid-term exam scores. The item means and standard deviations were obtained from each scale mean and standard deviation by dividing its mean with the number of its item scale. Results indicate that the item means of the state effort scale (M =3.80, SD =.63) was higher than the means of the other scales. This finding suggests that students were more likely to invest higher effort during the statistics mid-term exam, which accounted for 20% of the total course grade. It is, however, not unexpected that students would be likely to put more effort in a high-stakes exam such as the one in this study.

Trait worry, on the other hand, shows the lowest item mean scores (M =2.01, SD =.52) when compared to the rest of the scales, suggesting that students in general did not have high levels of worry. In general, however, the Malaysian undergraduates' trait worry (M =16.15; SD =4.13) is higher than their China (M =13.9; SD =3.9), Italy (M =13.8; SD =4.0), Turkey (M =13.1; SD =2.8), Netherlands (M =11.9; SD =3.9), and their US counterparts (M =14.3; SD =5.3) as summarized in O'Neil et al. (1992). A closer look at the composite means within the Malaysian samples indicated that the Chinese had a lower mean (M =14.32; SD =3.80) than the Malays (M =17.05; SD =4.00).

The statistics mid-term exam scores were converted to T-scores, thus indicating a mean of 50 and a standard deviation of 10. The exam scores were initially standardized (via Z scoring) within classes and then converted to T-scores to control for potential between-class achievement differences.

Table 1
Summary Statistics for Each Scale and Statistics Exam for Total Group

Scale	# of items	Scale		Item	
		M	SD	M	SD
Trait					
Self-efficacy	8	22.58	4.10	2.82	.51
Effort	8	23.97	3.59	3.00	.45
Worry	8	16.15	4.13	2.01	.52
State					
Self-efficacy	5	10.87	3.54	2.17	.71
Effort	5	19.01	1.65	3.80	.33
Worry	7	18.86	4.39	2.69	.63
Statistics Exam	6	50	10		

Note: N (Total)=360

Descriptive statistics were also computed for both ethnic groups (Malay vs. Chinese) and gender (male vs. female) separately. These results are shown in Tables 2 and 3. In general, data indicate that Chinese Malaysian participants had higher trait self-efficacy (M=2.86) compared to other Chinese ethnic students from other nations. For instance, Chinese from Taiwan in Huang's (1996) and Wang's (1997) studies reported a mean of 2.52 and 2.41 for trait self-efficacy respectively. Moreover, Chinese undergraduates in this study reported low state self-efficacy (M=2.13). Similar findings were also reported in Huang's (1996) (M =2.33) and Malpass' (1994) (M =2.30) studies. Malpass' (1994) sample consisted of 60% Asian Americans. Malaysian Chinese also reported dramatically higher state effort (M =3.82) than other Chinese nationals. Huang's (1996), Wang's (1997), Malpass' (1994), and Li's (1994) studies reported a mean of 3.29, 2.59, 3.49, and 3.25 respectively.

Table 2
Summary Statistics and Reliability for Each Scale for Each Gender and Ethnicity

Scale	# of items	Gender				Ethnicity			
		Male (n=121)		Female (n=239)		Malay (n=241)		Chinese (n=119)	
		Mean α	SD	Mean α	SD	Mean α	SD	Mean α	SD
Trait									
Self-efficacy	8	22.56	4.26	22.56	4.04	22.40	3.56	22.87	5.02
		(2.81)	(.53)	(2.82)	(.51)	(2.80)	(.45)	(2.86)	(.63)
		α=.86		α=.85		α=.84		α=.88	
Effort	8	23.31	3.87	24.25	3.40	23.83	3.43	24.18	3.89
		(2.91)	(.48)	(3.03)	(.43)	(2.98)	(.42)	(3.02)	(.48)
		α=.79		α=.76		α=.77		α=.76	

Scale	# of items	Gender				Ethnicity			
		Male (n=121)		Female (n=239)		Malay (n=241)		Chinese (n=119)	
		Mean α	SD	Mean α	SD	Mean α	SD	Mean α	SD
Worry	8	15.38	3.68	16.50	4.30	17.05	4.00	14.32	3.80
		(1.92)	(.46)	(2.06)	(.54)	(2.13)	(.50)	(1.79)	(.47)
		α=.74		α=.79		α=.75		α=.76	
State									
Self-efficacy	5	11.18	3.66	10.65	3.49	10.91	3.37	10.65	3.9
		(2.24)	(.73)	(2.13)	(.69)	(2.18)	(.67)	(2.13)	(.78)
		α=.85		α=.84		α=.82		α=.86	
Effort	5	18.76	1.76	19.14	1.58	18.97	1.70	19.11	1.53
		(3.75)	(.35)	(3.83)	(.32)	(3.79)	(.34)	(3.82)	(.31)
		α=.71		α=.78		α=.77		α=.71	
Worry	7	18.94	4.54	18.87	4.29	19.33	4.28	17.99	4.43
		(2.70)	(.65)	(2.69)	(.61)	(2.76)	(.61)	(2.57)	(.63)
		α=.80		α=.80		α=.80		α=.79	

Note: Parentheses indicate item means and item standard deviations

In terms of worry, the Chinese students in this study reported lower state worry ($\underline{M}$ =2.57) than Huang's (1996) ($\underline{M}$ =2.88), but similar to Malpass' (1994) and Li's (1994) with a mean of 2.44 and 2.60 respectively.

Table 3 shows means for each ethnicity by gender. Similar trends were observed for each gender from both ethnic groups in terms of scales' means. To determine the relationship between gender and ethnicity and the dependent variables, a two-way multivariate analysis of variance (MANOVA) was performed on each set of Trait and State dependent variables separately: self-efficacy, effort, and worry. With the use of the Wilks criterion the combined Trait dependent variables were significantly related to gender (Wilks' λ= .978, $\underline{F}$ [3,350]= 3.27, $\underline{p}$<.05) and ethnicity (Wilks' λ= .972, $\underline{F}$ [3, 350] = 9.94, $\underline{p}$<.05). Their interaction was not significant ($\underline{F}$ [3, 350] = .284, $\underline{p}$> .05). Further univariate analyses indicated significant mean differences between male and female students on Trait Effort ($\underline{F}$ [1, 356] = 5.803, $\underline{p}$<.05), and also between Malay and Chinese students on Trait Worry , ($\underline{F}$ [1, 356] = 29.39, $\underline{p}$<.05). No significant difference was observed for trait self-efficacy. As indicated in Table 2, females generally had higher trait effort than males and Malay students were generally more anxious (trait worry) than their Chinese counterparts.

Table 3
Summary Statistics for Each Scale for Malay and Chinese by Gender

Scale	# of items	Gender				Ethnicity			
		Male (n=74)		Female (n=164)		Malay (n=46)		Chinese (n=73)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Trait									
Self-efficacy	8	22.42	3.4	22.45	3.64	22.96	5.33	22.82	4.86
		(2.80)	(.43)	(2.81)	(.45)	(2.87)	(.67)	(2.85)	(.61)
Effort	8	22.35	3.62	24.09	3.30	22.48	4.28	24.60	3.57
		(2.92)	(.45)	(3.01)	(.42)	(2.93)	(.53)	(3.08)	(.45)
Worry	8	16.36	3.87	17.37	4.02	13.98	2.88	14.53	4.28
		(2.04)	(.48)	(2.17)	(.50)	(1.74)	(.36)	(1.82)	(.54)
State									
Self-efficacy	5	11.30	3.57	10.73	2.15	10.96	3.75	10.70	4.00
		(2.26)	(.71)	(2.15)	(.65)	(2.19)	(.75)	(2.14)	(.80)
Effort	5	18.66	1.86	19.10	1.63	18.89	1.64	19.28	1.43
		(3.73)	(.37)	(3.82)	(.33)	(3.78)	(.33)	(3.86)	(.29)
Worry	7	19.11	4.72	19.33	4.17	18.57	4.20	17.67	4.53
		(2.73)	(.67)	(2.76)	(.60)	(2.65)	(.60)	(2.52)	(.65)

Note: Parentheses indicate item means and item standard deviations

For the State dependent variables (self-efficacy, effort, and worry) significant relationships were also observed between gender (Wilks' λ= .977, $\underline{F}$ [3,341]= 2.73, $\underline{p}$<.05) and ethnicity (Wilks' λ= .980, $\underline{F}$ [3,341] = 2.89, $\underline{p}$<.05). Again, no significant interaction was observed. The univariate analyses indicated significant mean differences between male and female students on State Effort ($\underline{F}$ [1,347] = 4.82, $\underline{p}$<.05), and also between Malay and Chinese students on State Worry ($\underline{F}$ [1,347] = 4.44, $\underline{p}$<.05), but not state self-efficacy. An examination of the means (see Table 2) indicated that female students put more effort into the exam than their male counterparts, and Malay students had higher worry level than Chinese students during the exam. These findings suggested that females in general were more hardworking (trait effort) than males and invested more effort (state effort) during the statistics exam. Females from both ethnic groups also reported higher trait worry, which suggested that females in general were more anxious (trait worry) than males.

RELIABILITY AND ITEM ANALYSIS

A classical measure of reliability, Cronbach's Alpha, was used to examine internal consistency reliability for the items within each scale. A corrected item-total correlation was examined to see how well an item fit within a particular scale. A principal component factor analysis with varimax rotation was

performed on all items on each measure. Based on descriptive, internal consistency measures, and the results of exploratory factor analysis, poor items were identified and removed, thus resulting in a reduced number of items. The items that were deleted had the following characteristics: 1) had the lowest item-total correlations, i.e., less than .30, 2) had the lowest factor loadings, i.e., less than .40, and 3) failed to load on only one factor, e.g., items which did not maintain the distinct nature of the separate factors. However, items were carefully eliminated (one item at each iteration) so that there was no significant reduction in the reliability or validity indices of each scale. The purpose of the above analyses was to ensure the homogeneity and unidimensionality of each scale.

The internal consistency estimates (coefficient alpha) for the six scales are shown in Table 4. The internal consistency estimates were based on the combined sample (n= 360). This analysis yielded acceptable results ranging from Cronbach's a =.75 to a = .85. Very similar reliability estimates were obtained separately from men and women, and from Malay and Chinese samples (see Table 2). The following coefficient alphas were obtained from the Trait Self-Assessment Questionnaire for the total sample: trait self-efficacy (8 items), α = .85; trait effort (8 items), α = .78 ; trait worry (8 items), a = .78. All items in the Trait measures were retained for the study. However, five items from the State Self-Assessment Questionnaire were dropped from the study based on the heuristics established earlier (see Table 5). Three items from the state self-efficacy scale (SSE3: "I expected to do well on this test."; SSE4: "I had doubts about my ability to do well on this test [REVERSE]."; and SSE7: "I'm sure I did an excellent job on the questions on this test".) were deleted resulting in five items with α = .84. Item 6 from the state effort scale ("I did not give up, even though the exam was hard.") was deleted resulting in five items with a = .75. Item 6 from the state worry scale ("I felt the test contained things I didn't study.") was deleted resulting in seven items with α = .80.

A final exploratory factor analysis was run separately for the revised Trait and State Self-Assessment Questionnaires separately. The results of the principal components analysis with varimax rotation indicated three factors for each Trait and State measures with item loadings all over .40. All items loaded on its expected factors and significant at $p<.05$ level. Tables 6 and 7 show final factor loadings for Trait and State measures respectively.

Table 4
Reliability and Item-Total Correlation Indices of the Questionnaires

Scale	Item	Corrected Item-Total Correlation		α if item deleted
Trait Self-efficacy	8	.63		.83
		.49		.84
		...		.84
		.54	.85	.84
		.63		.83
		.62		.83
		.66		.83
		.63		.83
Trait Effort	8	.37		.77
		.55		.74
		.62		.73
		.42	.78	.76
		.51		.74
		.52		.74
		.46		.75
		.37		.77
Trait Worry	8	.46		.75
		.45		.76
		.45		.76
		.42		.76
		.52	.78	.75
		.55		.74
		.48		.75
		.53		.74
State Self-efficacy	5	.67		.80
		.71		.79
		.52	.84	.84
		.66		.81
		.67		.80
State Effort	5	.58		.68
		.56		.70
		.31	.75	.78
		.57		.69
		.62		.67
State Worry	7	.46		.78
		.51		.77
		.59		.76
		.54	.80	.77
		.44		.79
		.56		.77
		.60		.76

Table 5
Initial Factor Loadings of the State Scales

Expected Factor/ Observed Factor	Item	SSE	SE	SWOR
State Self-efficacy (SSE)	SSE2	.83		
	SSE6	.81		
	SSE8	.77		
	SSE1	.76		
	SSE5	.60		
	SSE3	.53	**.42**	
	SSE7	.46		**-.40**
	SSE4			**-.69**
State Effort (SE)	SE2	.74		
	SE5	.73		
	SE1	.68		
	SE4	.67		
	SE3	.52		
	SE6	**.32**		
State Worry (SWOR)	SWOR8			.73
	SWOR5			.71
	SWOR7			.68
	SWOR4			.66
	SWOR3			.66
	SWOR2			.57
	SWOR1			.56
	SWOR6			**.29**

Notes: Deleted items are boldfaced

Table 6
Final Factor Loadings of the Trait Scales

Expected Factor/ Observed Factor	Item	TSE	TE	TWOR
Trait Self-efficacy (TSE)	TSE8	.77		
	TSE1	.76		
	TSE7	.70		
	TSE6	.66		
	TSE5	.66		
	TSE4	.61		
	TSE2	.56		
	TSE3	.56		

Expected Factor/ Observed Factor	Item	TSE	TE	TWOR
Trait Effort (TE)	TE3		.80	
	TE2		.71	
	TE6		.65	
	TE5		.62	
	TE7		.55	
	TE1		.49	
	TE4		.48	
	TE8		.47	
Trait Worry (TWOR)	TWOR8			.67
	TWOR6			.66
	TWOR5			.64
	TWOR7			.61
	TWOR1			.61
	TWOR3			.60
	TWOR2			.60
	TWOR4			.57

Table 7
Final Factor Loadings of the State Scales

Expected Factor/ Observed Factor	Item	SSE	SE	SWOR
State Self-efficacy (SSE)	SSE2	.85		
	SSE6	.83		
	SSE1	.79		
	SSE8	.77		
	SSE5	.55		
State Effort (SE)	SE5		.77	
	SE2		.76	
	SE1		.71	
	SE4		.70	
	SE3		.53	
State Worry (SWOR)	SWOR8			.75
	SWOR5			.74
	SWOR7			.68
	SWOR3			.65
	SWOR4			.65
	SWOR2			.59
	SWOR1			.56

TEST OF THE MEASUREMENT MODEL

A confirmatory factor analysis (CFA) was run to examine the adequacy of the hypothesized factor loading (i.e., the measurement model), the degree of fit, and latent factor intercorrelations. Confirmatory factor analysis is viewed as a subset of the more general structural equation modeling approach (Pedhazur & Schmelkin, 1991), which enables a researcher to test the hypotheses that a particular linkage between the observed variables and their underlying latent factors actually exists in a statistically reliable manner (Bryne, 1994). If the measurement models are adequate, the structural models are tested based on the measurement model.

The measurement model in this study comprised six latent factors with eighteen indicators and eighteen residuals. Following Anderson and Gerbing's (1988) suggestion each latent factor was defined by three indicators. The three indicators were formed via split-thirds, i.e., arbitrarily splitting the items of each factor to the three indicators of its latent factor.

Even though a nonsignificant chi-square is a desired result for not rejecting the model, it is suggested within the structural equation modeling community that chi-square fit statistics is overly strict and sensitive for models with numerous variables and subjects (Bentler, 1980; Bentler & Bonnet, 1980; Newcomb, 1990). The following criteria were established to examine model fit: (1) the three fit indices from the EQS application, i.e., the Normed Fit Index (NFI), Non-normed Fit Index (NNFI), and the Comparative Fit Index (CFI), were greater than .90; (2) the chi-square/degrees-of-freedom ratio was less than 3.0 (Carmines & McIver, 1981); and (3) the Root Mean Square Error of Approximation (RMSEA) was less than .05 (Browne & Cudeck, 1993; Rigdon, 1996).

Based on the criteria delineated above, the measurement model fits quite well for the entire sample ($\chi^2 = 186.69$, $p<.001$, NFI =.93, NNFI = .96, CFI = .97 and $\chi^2/df = 1.56$). The Root Mean Square Error of Approximation (RMSEA [γ]) = .040, with a 90% confidence interval of $.028 \leq \gamma \leq .050$. In addition, all hypothesized factor loadings were substantial (ranging from .607 to .862) and significant in the expected directions (see Table 8).

Table 8
Subscales to Target Factor Loadings[1]

Trait Effort: (TE)	
TEA (items: 2+11)	.616
TEB (items: 5+14+20)	.794
TEC (items: 8+17+23)	.771
Trait Worry: (TWOR)	
TWORA (items: 3+12)	.607
TWORB (items: 6+15+21)	.782
TWORC (items: 9+18+24)	.783

Trait self-efficacy: (TSE)	
TSEA (items: 1+10)	.785
TSEB (items: 4+13+19)	.862
TSEC (items: 7+16+22)	.774
State self-efficacy: (SSE)	
SSEA (item: 6)	.859
SSEB (items: 15+18)	.763
SSEC (items: 3+22)	.733
State Worry: (SWOR)	
SWORA (items: 2+19)	.742
SWORB (items: 5+8+14)	.614
SWORC (items: 11+21)	.775
State effort: (EFFORT)	
SEA (item: 1)	.663
SEB (items: 4+10)	.744
SEC (items: 7+13)	.713

[1]Items can be found in Appendix A (Trait) and Appendix B (State)

The disattenuated interfactor correlations (see table 9), i.e., correlations that had been corrected for measurement error, were smaller than .70, indicating that the six scales measured factorially distinct but related components. The highest interfactor correlation observed was between Trait Self-efficacy and Trait Effort (r=.606, p<.05). Findings from the confirmatory factor analysis indicate that the measured variables were reliable and valid indicators of the latent factors and thus the model was adequate to serve as a measurement model for future structural analyses.

Table 9
Disattenuated Factor Correlations and Statistics Achievement

Factors	1	2	3	4	5	6	7
(1) Trait Effort	1.000						
(2) Trait Self-efficacy	.606	1.000					
(3) Trait Worry	-.053[a]	-.277	1.000				
(4) State Effort	.406	.215	-.047[a]	1.000			
(5) State Self-efficacy	.346	.519	-.163	.420	1.000		
(6) State Worry	-.253	-.388	.29	-.384	-.487	1.000	
(7) Statistics Ach.	.222	.219	-.193	.345	.239	-.409	1.000

Note: [a] not statistically significant

In order to ensure each gender and ethnic group scales were structured similarly, multi-group invariance analyses were computed and discussed next.

TEST OF THE INVARIANT FACTORIAL STRUCTURE

To examine whether components of the measurement model were invariant (i.e., equivalent) across gender and ethnic groups, multiple group invariance analyses were conducted. Separate invariance analyses were conducted for gender and ethnicity. These multiple-group analyses tested hypotheses related to the equivalencies of the Trait and State measurement across groups. Specifically, the researcher examined whether the six latent constructs and the eighteen measured variables used to measure them were equivalent across male and female, and Malay and Chinese college undergraduates in Malaysia.

In testing group invariance several hypotheses were tested in specific order. As discussed in Byrne (1994), Ullman (1996), and Newcomb (1994; Personal Communication, October 1st 1998), the following heuristics were used for this study. First, the researcher established a baseline model for each group separately. In this study the baseline model was identical for all groups (male vs. female, and Malay vs. Chinese). These baseline models were first analyzed simultaneously with no between-group constraints, i.e., all parameter estimates – loadings, factor variances / covariances, and item uniqueness- were free to be estimated.

Second, following the baseline model estimation, factor loadings were constrained to be equal across groups. This step tested the hypothesis that the factor structures were the same across different groups. The researcher examined the chi-squared difference between the restricted (constrained) model and the baseline model. A nonsignificant chi-squared difference is desirable, indicating similar measurement model for both groups.

Third, the researcher tested if the factor covariances were equal across groups. Again, the chi-squared difference was examined between the less restricted model and the more restricted model. If all the constraints (factor loadings and covariances) were equal across groups (due to nonsignificant chi-squared difference), it can be concluded that the two groups (male vs. female, and Malay vs. Chinese) represent samples from the same population.

Results of gender invariance were discussed first, followed by ethnicity. Tables 10 and 11 summarize results from the multiple-group invariance analyses for gender and ethnicity respectively.

Table 10
Summary Statistics of Models Tested for Multi-group Invariance Analyses for Gender

Model	χ^2	df	χ^2/df	NNFI	CFI	RMSEA
1: Estimates Free	335.40	240	1.40	.951	.961	.034
2: Only Loadings Fixed	354.55	258	1.37	.954	.961	.033
3: Factor Loadings and	364.91	272	1.34	.958	.962	.031
Covariances Fixed						
Model Comparisons:	$\Delta\chi^2$	Δdf				
Model 2 vs. Model 1	19.15	18[a]				
Model 3 vs. Model 2	10.36	14[a]				

NOTE. [a] not statistically significant
NNFI=Bentler-Bonnett Nonnormed Fit Index; CFI= Comparative Fit Index; RMSEA= Root Mean Square Error of Approximation

A comparison of the baseline (Model 1) and the restricted model (Model 2) for gender revealed an ever slight difference in fit (NNFI=.951 vs. NNFI=.954; CFI=.961 vs. CFI=.961). The difference in chi-square ($\Delta\chi^2$= 19.15, Ddf =18, p>.05) was not statistically significant. These findings suggested that the specifications of cross-group equivalencies of factor loadings were tenable. A comparison between the less restricted (Model 2) and more restricted model (Model 3) revealed a slight increment in fit (NNFI=.954 vs. NNFI=.958; CFI=.960 vs. NFI=.962). However, the difference in chi-square ($\Delta\chi^2$= 10.36, Δdf=14, p>.05) was again not statistically significant. Thus, it was concluded that the across gender relationships were invariant.

Table 11
Summary Statistics of Models Tested for Multi-Group Invariance Analyses for Ethnicity

Model	χ^2	df	χ^2/df	NNFI	CFI	RMSEA
1: Estimates Free	348.58	240	1.45	.944	.956	.035
2: Loadings Fixed	375.81	258	1.39	.942	.951	.034
3: Factor Loadings and	381.71	272	1.40	.948	.954	.034
Covariances Fixed						
Model Comparisons:	$\Delta\chi^2$	Δdf				
Model 2 vs. Model 1	27.23	18[a]				
Model 3 vs. Model 2	5.90	14[a]				

Note: [a] not statistically significant
NNFI=Bentler-Bonnett Nonnormed Fit Index; CFI= Comparative Fit Index; RMSEA= Root Mean Square Error of Approximation

Similar results were observed in the multi-group analysis for ethnicity. A comparison of the baseline (Model 1) and the restricted model (Model 2) for ethnicity revealed a slight decrement in model fit (NNFI=.944 vs. NNFI=.942; CFI=.956 vs. CFI=.951). However, the difference in chi-square ($\Delta\chi^2$=27.23, Δdf = 18; p>.05) was not statistically significant. These findings suggested that the specification of cross-group equivalencies of factor loadings were tenable.

Next, the researcher fixed both factor loadings and covariances to be equal between groups. Although a comparison between this new model (Model 3) and the less restricted one (Model 2) revealed a slight increment in fit (NNFI=.942 vs. NNFI=.948; CFI=.951 vs. CFI=.954), the difference in chi-square ($\Delta\chi^2$=5.9; Δdf=14, p>.05) was not statistically significant. These findings suggested that the specification of cross-group equivalencies of factor loadings and covariances were tenable.

CONCLUSION

Findings substantiate the utility of the Malay State and Trait Self-Assessment Questionnaires for Malaysian undergraduates. The scales' internal consistency reliability coefficients are comparable to the English and Chinese versions in previous studies, and thus are considered adequate. This is further supported by the confirmatory factor analysis analysis which demonstrated a satisfactory model fit. Low inter-factor correlations indicated adequate discriminant validity, thus support the theoretical orthogonality of the constructs. Findings from the multi-group analyses suggested that all scales were structured similarly for both gender and ethnic groups. Thus, we can conclude that the two groups (Male vs. Female, and Malay vs. Chinese) represent samples from the same population. These multi-group invariance analyses confirmed that the measured variables were reliable indicators of the latent factors and were equivalent across groups, and thus the model was adequate to serve as a measurement model for future structural analyses.

SCIENTIFIC IMPORTANCE OF THE STUDY

Transportability of psychological and educational measuring instruments is vital in cross-cultural research since not all cross-cultural research can be conducted in English. Valid and reliable measures can help extend cross-cultural examination of the theories that have been tested in the west. It is imperative then to investigate whether a certain theory is applicable across cultures (the etic truth) or it is only applicable in a specific culture (the emic truth), because failing to recognize such emics and etics can make us guilty of ethnocentrism. This study provides evidence of reliability and validity of several motivational constructs that have been tested in the West. This examination is necessary to validate western theories in a different context, and thus help enrich our understanding about the world.

REFERENCES

Anderson, J.C., & Gerbing, D.W. (1998). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1993). Perceived Self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.

Bandura, A. (1993). *Self-Efficacy in Changing Society.* Cambridge, NY: Cambridge University Press.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control.* New York: Freeman.

Bandura, A. (1998). Personal and collective efficacy in human adaptation and change. In D. Belanger & J. Adair (Eds.). *Advances in Psychological Science, 1: Social, Personal, and Cultural Aspects,* (57-71). England, UK: Psychology Press / Erlbaum.

Bentler, P.M. (1995). *EQS: Structural Equations Program Manual.* Encino, CA: Multivariate Software, Inc.

Bentler, P.M. (1998). *EQS: Structural Equations Program (Version 5.7b) [Computer Software].* Encino, CA: Multivariate Software, Inc.

Bentler, P.M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31, 419-456.

Bentler, P.M., & Bonnett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

Brislin, R.W. (1970). Back-translation for cross-cultural research. *Journal of Cross-ultural Psychology*, 1 (3), 185-216.

Brislin, R.W. (1980). Translation and content analysis of oral and written materials. In H.C. Triandis & J.W. Berry (Eds.). *Handbook of Cross-cultural Psychology, 2: Methodology.* Boston: Allyn and Bacon.

Brown, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing Structural Equation Models,* (111-135). Baverly Hills, CA: Sage.

Bryne, B. M. (1994). *Structural Equation Modeling with EQS and EQS/Windows.* Thousand Oaks, CA: Sage.

Carmines, E., & McIver, J. (1981). Analysis models with unobserved variables: Analysis of covariance structures. In G. Bohrnstedt & E.F. Borgatta (Eds.), *Social Measurement: Current Issues,* (65-115). Baverly Hills, CA: Sage.

Hembree, R. (1988). Correlates, causes, and treatment of testing anxiety. *Review of Educational Research*, 58, 47-77.

Huang, D. (1996). The role of perceived parental attitude, effort and self-efficacy in the math achievement of high and low track high school students in Taiwan. Unpublished dissertation. Los Angeles: University of Southern California.

Malpass, J.R. (1994). A Structural model of self-efficacy, goal orientation, worry, self-regulated learning, and high stakes mathematics achievement.

Unpublished doctoral dissertation. Los Angeles: University of Southern California.

Newcomb, M. D. (1990). What structural modeling techniques can tell us about social support. In I.G. Sarason, B.R. Sarason & G.R. Pierce (Eds.), *Social Support: An Interactional View*, (26-63). New York: Wiley and Sons.

Newcomb, M.D. (1994). Drug use and intimate relationships among women and men: Separating specific from general effects in prospective data using structural equation models. *Journal of Consulting and Clinical Psychology*, 62(3), 463-476.

O'Neil, H.F., Jr., & Abedi, J. (1992). Japanese children's trait and state worry and emotionality in a high-stakes testing environment. *Anxiety, Stress, and Coping*, 5, 253-267.

O'Neil, H.F., Jr., & Abedi, J. (1996). Reliability and validity of the state metacognitive inventory: Potential for alternative assessment. *Journal of Educational Research*, 89, 234-245.

O'Neil, H.F., Jr., & Fukumura, T. (1992). Relationship of worry and emotionality to test performance in a Juku environment. *Anxiety, Stress, and Coping*, 5, 241-251.

O'Neil, H.F., Jr., & Herl, H.E. (1998, April). Reliability and validity of a trait measure of self-regulation. Paper presented at the *1998 Annual Meeting of the American Educational Research Association*, San Diego, CA.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66 (4), 543-578.

Pedhazur, E., & Schmelkin, L.P. (1991). *Measurement Design and Analysis: An Integrated Approach*. Hillsdale, NJ: Erlbaum.

Pintrich, P.R., & Schunk, D.H. (1996). *Motivation in Education: Theory, Research, and Applications*. New Jersey: Prentice Hall.

Pintrich, P.R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (Chap. 14). San Diego, CA: Academic Press.

Pintrich, P.R., & De Groot, E. (1990) Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.

Rigdon, E. (1996). CFI versus RMSEA: A comparison of two fix indexes for structural equation modeling. *Structural Equation Modeling*, 3, 369-379.

Rosna Awang Hashim (2000). Motivational Correlates as Mediating Factors in Academic Achievement, paper presented at the *7th International Workshop on Achievement Motivation (WATM): Motivation & Emotion, European Association for Research in Learning*, University of Leuven, Belgium, May 12-15, 2000.

Schunk, D.H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26, (3&4), 207-231.

Spielberger, C.D. (1975). Anxiety: State-trait process. In C.D. Spielberger & I.G. Sarason (Eds.), *Stress and anxiety*, 1, (115-143). Washington, DC: Hemisphere.

Spielberger, C.D. (1980). *Test Anxiety Inventory: Preliminary Professional Manual.* Palo Alto, CA: Consulting Psychologist Press.

Spielberger, C.D., & Vagg, P.R. (1995). Test anxiety: A transactional process. In C.D.

Spielberger C.D., & P. Vagg (Eds.). *Test Anxiety: Theory, Assessment and Treatment,* (3-14), Washington, DC: Taylor & Francis.

SPSS Inc. (1998). *SPSS for Windows V. 8.0. [Computer Software].* Chicago: SPSS Inc.

Ullman, J. (1996). Structural Equation Modeling. In B. G. Tabachnik & L. S. Fidell, (Eds.), *Using Multivariate Statistics* (3rd edn.), 709-811. New York: Harper Collins.

Weiner, B. (1994). Ability versus effort revisited: The moral determinants of achievement evaluation and achievement as a moral system. *Educational Psychologist,* 29,(3), 163-172.

Zeidner, M. (1998). *Test Anxiety: The State of the Art.* New York: Plenum Press.

Zimmerman, B.J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist,* 30(4), 217-221.

Zimmerman, B.J. (1996). Enhancing student academic and health functioning: A self-regulatory perspective. *School Psychology Quarterly,* 11 (1), 47-66.

Zimmerman, B.J. (1998). Academic studying and the development of personal skill: A self-regulatory perspective. *Educational Psychologist,* 33 (2/3), 73-86.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (Chap. 2). San Diego, CA: Academic Press.