Personal accountability versus excuse-making: The impact of secondary students' conceptions of assessment on academic performance mediated by self-efficacy and interest.

K. Frank Walton

University of Auckland

Gavin T. L. Brown

The Hong Kong Institute of Education

Paper presented at the 2009 Biannual Meeting of the European Association for Research in Learning & Instruction, Amsterdam, Netherlands, August 2009.

Correspondence concerning this paper should be addressed to:

Frank Walton, School of Teaching, Learning & Development, Faculty of Education, The University of Auckland, Private Bag 92019, Auckland, New Zealand or by email to f.walton@auckland.ac.nz

Acknowledgements.

The authors wish to thank Professor John Hattie for access to the asTTle database from which these data were drawn. The opinions expressed here are those of the authors and not of the Ministry of Education which funded the asTTle development. An earlier version of this paper was presented in Walton, K. F. (2009). *Secondary students' conceptions of assessment mediated by self-motivational attitudes: Effects on academic performance.* Unpublished masters thesis, University of Auckland, Auckland, NZ, supervised jointly by Dr Gavin T. L. Brown and Dr Susan Farruggia.

This study confirms and extends the findings of Brown and colleagues related to the impact of secondary students' conceptions of assessment on academic performance. One of their data sets (SCoA II) was reanalysed, including 15 previously unanalysed survey items on students' conceptions of assessment and six survey items on students' attitudes to reading. Analysis of the two SCoA II forms (Form 1, *N*=1774; Form 2, *N*=1623) via exploratory and confirmatory factor analyses, identified six conceptions of assessment. The resulting measurement models had acceptable fit with the sample data as did the structural models which linked the conceptions of assessment to students' asTTle reading scores. Reading performance was positively predicted by one conception ('assessment makes me accountable') and negatively predicted by the other five conceptions ('assessment makes schools accountable', 'assessment is helpful and enjoyable', 'assessment informs me', assessment is unfair and frustrating', 'assessment is useless and worthless'). Analysis of the self-motivational attitudes data produced two factors, self-efficacy and personal interest, both of which predicted small to medium positive effects on reading performance.
In order to explore the mediating effects of subject specific self-motivational attitudes on the predictive relationship between conceptions of assessment and reading achievement, multi-group invariance analysis was used. This analysis revealed that the high (top third) self-efficacy and interest groups were statistically different to the low (bottom third) self-efficacy and interest groups in their structural relations to reading achievement. It was concluded that subject motivational attitudes interact with conceptions of assessment to have a significantly differing effects on academic achievement. For students with high self-efficacy and interest, increased personal responsibility towards assessment predicts improved performance. For students with low self-efficacy and interest, increased personal responsibility towards assessment has a powerful positive effect on academic performance. These results have important implications for pedagogical practice around students' subject specific self-efficacy and interest, and their self-regulatory attitudes towards assessment.

Introduction

Self-regulation promises optimal academic advancement for students at all levels in both formal and informal contexts. Self-regulated learners take a significant degree of responsibility for managing their own learning and attribute "the consequences of their successes and failures to their own actions, rather than to forces and experiences beyond their control" . Self-regulation is a complex, multi-faceted construct and requires autonomous, agentic use of metacognitive and volitional strategies to manage cognitive and motivational resources in ways which are advantageous in achieving learning goals (Boekaerts, 1995; Boekaerts & Cascallar, 2006; Marsh, Hau, Artelt, Baumert, & Peschar, 2006; Schunk & Ertmer, 2000).

Boekaerts' Dual Processing Self-regulation Model

Boekaerts' (Boekaerts & Corno, 2005; Boekaerts & Niemivirta, 2000) model of self-regulation balances two fundamental types of goals that affect learning: growth and well-being. Growth goals strongly enhance and promote cognitive development, and are associated with mastery learning and motivations such as personal interest, self-efficacy and expected satisfaction in meeting challenges. Well-being goals are associated with maintaining emotional well-being and protecting one's self-image or ego. Both types of goals are legitimate, but the self-regulated learner places a priority on learning goals, while also maintaining their well-being goals. Environmental factors which produce stressors can threaten to 'derail' students from the growth/mastery learning track. Examples of stressors are difficult or tedious tasks, unfavourable or discouraging teacher feedback, pressure to perform, biological or developmental influences, and cultural constraints. Negative

emotional responses (e.g., feeling bored, isolated, coerced or threatened) can cause a shift in priority from learning goals to well-being goals.

Once on the well-being track, students' inability to deal with negative emotional responses can lead to behaviours that can be seriously detrimental or maladaptive to longer term learning, especially if they become entrenched. Examples of maladaptive behaviours are self-handicapping, for example 'I'm not going to try so I can have an excuse for poor performance' (Covington & Omelich, 1979), not asking for legitimate help from peers or the teacher, participating in inappropriate entertainment, refusing to comply with reasonable teacher directions, and not expressing one's own thoughts for fear of losing social acceptance. Boekaerts and Corno (2005) posited that students could be taught how to prevent themselves being derailed from the growth track, or how to switch back to the growth track by applying volitional strategies that help develop and maintain good work habits. By considering the different effects of adaptive and maladaptive beliefs, Boekaerts' two track model can be represented as a virtuous cycle and vicious cycle shown in Figure 1.

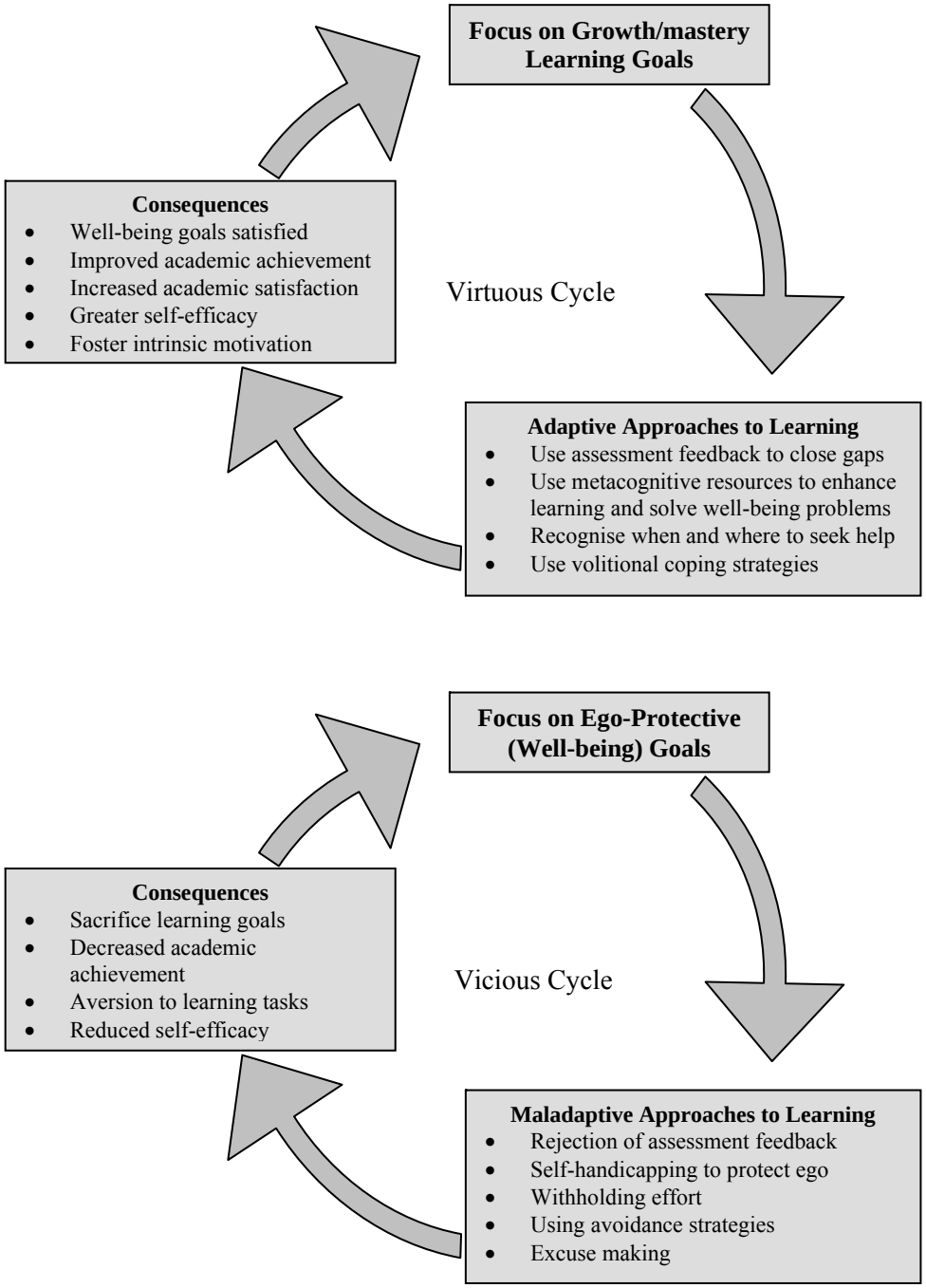


Figure 1: Virtuous Cycle (adaptive approaches) versus Vicious Cycle (maladaptive approaches)

Mediating effects of self-motivational attitudes

Self-regulation involves both knowledge of cognitive and metacognitive strategies and motivation to use the strategies (Pintrich & DeGroot, 1990). Motivation, or the willingness to learn, has been broadly categorised into intrinsic motivation and extrinsic motivation, depending on whether the motivation to learn is internally or externally sourced . Since the primary focus of this study is on personal self-regulatory attitudes, intrinsic motivation as indicated by levels of self-efficacy and personal interest are of relevance. Most studies have shown that both self-efficacy and personal interest have a small to medium positive effect on

academic performance (Hattie, 2004; Marsh et al., 2006; Otunuku & Brown, 2007; Schunk, 1983). Highly efficacious students respond positively to challenges and don't give up easily when difficulties arise. They have confidence in their own competency in a particular domain based on previous success and the affirmation from significant adults (e.g., a teacher or parent). Students who have a high level of personal interest in a subject area or task display high levels of engagement and enthusiasm, are willing to spend more time on a task and persist when facing difficult challenges (Hidi & Harackiewicz, 2000). Of particular interest in this study is the interaction between self-motivational attitudes and self-regulatory beliefs, and the effect of this interaction on academic achievement.

Research into conceptions of assessment

In a series of six studies into students' conceptions of assessment with the self-reported Students' Conceptions of Assessment (SCoA) inventory, Brown and colleagues (Brown, 2006; Brown & Hirschfeld, 2007, 2008; Brown, Irving, & Peterson, 2008, in press; Brown, Irving, Peterson, & Hirschfeld, 2009; Hirschfeld & Brown, 2009) have made a number of key findings related to attitudes to assessment, self-regulatory attitudes, and definitions of assessment, and the impact of each of these factors on academic achievement. They found that beliefs and attitudes towards assessment predicted 20-25% of the variability in academic achievement (Brown & Hirschfeld, 2008; Brown et al., 2008, in press).

Research into New Zealand secondary students' conceptions of assessment has shown that acceptance of the personal accountability role of assessment, consistent with an adaptive self-regulatory approach to assessment, was predictive of improved academic performance (Brown & Hirschfeld, 2008). It was presumed that students who viewed assessment as an opportunity to measure their progress against desired learning goals, valued the feedback information they gained from assessment events and sought to close any gaps between their goals and their present performance. These students were willing to submit themselves to the scrutiny of being assessed so that they could receive accurate information about their strengths and weaknesses and focus their efforts on improved performance. It was assumed that such students were better able to cope with the stresses of formal assessments since they accepted that, although testing was not an enjoyable experience, it had an important role in assisting their learning (Brown et al., 2009).

Conversely, Brown & Hirschfeld (2008) showed that some attitudes towards assessment (i.e., assessment could be ignored, assessment was enjoyable, and assessment evaluated schools) predicted decreased academic performance. These results were seen as indicative of maladaptive, ego-protective approaches to learning. Students who did not value the role of assessment as an adaptive process on the growth pathway were likely to make excuses for poor performances by adopting self-handicapping responses (i.e., it's about the school, it isn't serious). This pattern of attitudes could lead to intentional withholding of effort and thus reduced outcomes (Boekaerts & Corno, 2005; Cleary, 2006; Hattie, 2004). The pattern of score increasing and decreasing results were consistent with self-regulation theory (Zimmerman, 2001) and assessment-for-improvement pedagogical practices.

Data for this study is an expanded set of survey responses collected at the same time as the version 2 of the Students' Conceptions of Assessment inventory reported in Brown & Hirschfeld (2008). While a self-regulating framework has been successfully used to develop items for version 5 (Brown et al., 2009) and 6 (Brown et al., in press) of the SCoA inventory, "the explicit relationship between conceptions of assessment and personal motivational

beliefs has yet to be studied" (Brown et al., 2008, p. 3). Furthermore, Hirschfeld and Brown (2009), after analysing the effects of sex, ethnicity, and age upon the adapative and maladaptive conceptions of assessment, suggested that personal self-motivational beliefs may be of greater interest than demographic variables in understanding factors shaping the impact of student conceptions of assessment. Since the data sets used in Brown and Hirschfeld (2008) also contained self-reported self-motivational attitudes, it was possible to investigate the mediating effects of personal motivational attitudes on the adaptive and maladaptive conceptions of assessment. Furthermore, Brown and Hirschfeld (2008) only made use of 11 items from the SCoA inventory; a further 17 unanalysed items meant it was possible to extend our understanding of the structure of students' conceptions of assessment with responses from two very large data sets (i.e., N>1500 each).

The aim, then, of this study was to find out how subject specific self-efficacy beliefs and personal interest mediated the relationship between attitudes to assessment and academic achievement. In other words, how do the two categories of dispositional attitudes (self-regulatory attitudes towards assessment and subject self-motivational attitudes) interact with each other to influence academic performance? This was done by analysing three measures: students' conceptions of assessment, students' self-efficacy and interest in reading, and students' academic performance in reading comprehension as measured by a standardised test system. Factor analysis and structural equation modelling were used to model the relationship between attitudes to assessment and reading achievement, and multi-group invariance analysis was used to differentiate the effects of low and high self-efficacy and interest on the relationship between attitudes to assessment and reading performance.

It was hypothesised that the additional SCoA items would form conceptually meaningful factors consistent with previously conducted analyses. It was expected that the pathways from student conceptions of assessment to academic performance would not be invariant depending on the level of motivational attitudes students had. An inverse relationship was expected, such that students with high self-efficacy and high interest in reading would have greater positive weights on the adaptive pathways than students with low values. In contrast, students with low self-efficacy and low interest in reading would have stronger positive weights on the maladaptive or well-being oriented pathways.

Method

The study involved secondary analysis of self-report data on students' conceptions of assessment collected in 2004 in conjunction with a national survey trial of new test items for the Assessment Tools for Teaching and Learning (asTTle) software (Hattie et al., 2004).

Participants

Data were obtained from 3803 students (age range 13 – 17 years) from 58 different secondary schools in New Zealand (Table 1). For sex and ethnicity, the demographic break-down of students participating in the study approximately represented the proportions in the New Zealand secondary student population. The demographic breakdown of average age/year levels of participating students is somewhat weighted towards the year 9 and year 10 students.

Table 1. Student Participant Demographic Breakdown

Demographic category	*n*	%
Sex		
Boys	1759	46.3
Girls	2044	53.7
Ethnicity		
European	2110	56.5
Maori	523	14.0
Pasifika*	316	8.5
Asian	318	8.5
Others	468	12.5
(Not stated)	(68)	(1.8)
Average Age (year level)		
13 (year 9)	1622	42.7
14 (year 10)	1059	27.8
15 (year 11)	492	12.9
16 (year 12)	630	16.6

*Note: "Pasifika' is a term given to students of Pacific Island ethnicity, predominantly Samoan, Tongan and Cook Island

Instruments

The three instruments used to collect data were the Students' Conceptions of Assessment (SCoA II), the Students' Attitudes to Reading (SAR), and the standardised reading comprehension tests.

Students' Conceptions of Assessment (SCoA II) Instrument. The SCoA II instrument consisted of 29 items arranged into two forms to reduce the effects of response fatigue. The items were derived from an earlier study (Brown & Hirschfeld, 2007) and results for 11 were reported in Brown and Hirschfeld (2008). Form 1 contained 20 items and Form 2 contained 21 items; there were 12 items in common across the two forms. The item statements, labels and the expected factors are shown in Appendix A. The questionnaires used a positively packed six-point agreement response scale, with two negative options (strongly disagree, usually disagree) and four positive options (slightly agree, moderately agree, usually agree, strongly agree) (Brown, 2004). Since it was expected that students would tend to rate the various conceptions positively, positive packing was used to elicit increased variance in students' responses and thereby providing more precision in the analysis of the responses (Weekers, Brown, & Veldcamp, 2009).

Students' Attitudes to Reading (SAR) Instrument. Six items elicited motivational attitudes to reading. 'Otunuku and Brown (2007) reported that these items formed two scales (i.e., liking or interest and self-efficacy) with good fit properties. Students responded using a four point scale, identified by smiley face symbols indicating degree of affect (i.e., very happy face=4, happy face=3, sad face=2, very sad face=1). The items and expected factors are listed in Appendix B.

Academic Performance in Reading. Academic performance in reading comprehension was determined by performance on calibration test forms for the Assessment Tools for Teaching and Learning (asTTle) testing system (Hattie et al., 2004). The items were aligned to the New Zealand national English curriculum levels and objectives (Hattie, Brown, & Keegan, 2003; Ministry of Education, 2007) and scoring was done using single parameter

item response theory. This meant that regardless of test form completed, student performance was on a common transformed scale (Embretson & Reise, 2000). The asTTle scores used were the standardised, linear transformed scores derived from the IRT analysis (Year 6 M = 500, SD = 100) (Hattie et al., 2004).

Survey Procedures

For each of the four year levels (Year 9, 10, 11, 12), multiple asTTle reading tests were prepared, each containing items within an estimated appropriate range of difficulty. At the end of each test, either the Form 1 or Form 2 SCoA survey questionnaire was attached. It was intended that all test papers would have an equal chance of being assigned to any individual in any class so that any effect of the class or teacher on the distribution would be randomised. The teachers who supervised the tests were asked to remind students to complete the SCoA questionnaires when they had finished the one hour asTTle test. Student demographic information was gathered from the asTTle test included sex (male, female), ethnicity (European, Maori, Pasifika, Asian, Others) and year level (9, 10, 11, 12). Information about each school was also collected including school decile level (1-10), and school type (single sex or co-ed.).

Data analysis

The Form 1 and Form 2 SCoA data sets were cleaned by removing cases with 75% or more of the same response and cases with more than 10% missing responses. Remaining missing values were replaced with imputed values using the Expectation Maximisation procedure (Dempster, Laird, & Rubin, 1977). Comparison of the means and standard deviations for each item before and after the EM procedure was conducted to ensure no significant changes to the data had occurred.

Exploratory and confirmatory factor analyses of student responses to the two SCoA forms (Form 1 N=1774; Form 2 N=1623; nominally aged 13 to 16 years old) was conducted and the measurement models produced were tested for fit with the sample data. Structural equation modelling was used to produce models representing the relationship of the SCoA measurement models to performance on the asTTle reading tests. A range of indexes were used to test for model fit because of the different sensitivities to model features such as sample size, model complexity and model misspecification (Byrne, 2001; Cheung & Rensvold, 2002; Fan & Sivo, 2005, 2007; Hu & Bentler, 1998, 1999; Marsh, Hau, & Wen, 2004; Sivo, Fan, Witta, & Willse, 2006). Some of the indexes reported the degree of goodness-of-fit between the model and the sample, while others reported the degree of misfit (i.e., badness-of-fit). The cut-off values applied for each index to indicate acceptable fit were: p (χ^2/df) >.05, TLI ≥ .90, CFI ≥ .90, gamma hat >.90, SRMR ≤ .08 and RMSEA ≤ .08.

In order to examine the mediating effects of self-motivational attitudes towards reading, students were grouped into high-self-efficacy (top third, n = 550) and low-self-efficacy (bottom third, n = 550), and high-interest (top third, n = 550) and low-interest (bottom third, n = 550). These sized groupings provided sufficient differentiation between high and low attitudes as well as provide a sufficient sample size for stable analysis. Before invariance analysis of the SCoA models could be conducted, further cleaning of the Form1 and Form 2 SCoA data sets was required due to the combining of the SCoA and SAR data sets. Cases with missing values in the six attitudes to reading items were removed from the Form 1 and Form 2 data sets. This reduced the sample size for each data set to N = 1667 for Form 1 and

N = 1501 for Form 2. There were no cases left with missing values, so the expectation maximisation procedure was not required.

Multiple-group confirmatory factor analysis was used to test the invariance of the structural models linking SCoA to reading performance. The groups were the high and low self-efficacy and interest students and this studied the effect of self-motivational attitudes on adaptive and maladaptive pathways. If invariance cannot be demonstrated, then the different groups have different responses to the model being tested and must be considered to come from different populations (Cheung & Rensvold, 2002). Invariance analysis of covariance structures involves using a series of increasingly stringent tests which are carried out using a set of nested multi-group models in which selected model parameters are constrained to be equivalent (Byrne, 2004; Wu, Li, & Zumbo, 2007). The fit characteristics of the more constrained model and less constrained model are compared and if the changes are not statistically significant, the invariance assumption can be accepted at each level of testing.

The difference in chi-square ($\Delta\chi^2$) between two nested models, while conventionally applied, is impractical on its own because of its sensitivity to large sample sizes (Cheung & Rensvold, 2002; Wu et al., 2007). A change in the comparative fit index (ΔCFI $\leq$ .01) or gamma hat (Δgamma hat $\leq$ .001) are preferred criteria for accepting invariance between nested models since they are independent of both sample size and model complexity (Cheung & Rensvold, 2002). Stringent criteria were set to accept the invariance assumption. Invariance was accepted if p ($\Delta\chi^2/\Delta df$) $\geq$ .05 *or* if both ΔCFI $\leq$ .01 *and* ΔGamma hat $\leq$ .001. In other words, if the $\Delta\chi^2$ test failed, both the ΔCFI and Δgamma hat criteria had to be met.

Four invariance tests were carried out (Cheung & Rensvold, 2002; Wu, Li, & Zumbo, 2007).

1. Test for configural invariance (unconstrained model). The number of factors and the loading pattern (model specification) is assumed to be equivalent across groups. There are no additional constraints placed on the model parameters. The model configuration, which includes each group, is estimated simultaneously. Configural invariance is accepted if RMSEA $\leq$ .05.

2. Test for metric (weak) invariance. Using the same model configuration the regression weights between the latent factors and the observed variables are constrained to be equivalent.

3. Test for scalar (strong) invariance. In addition to the metric invariance constraints, the regression line intercepts (item scores when the factor score is zero) are constrained to be equivalent. The criteria used for accepting the invariance assumption for a model across population sub-groups was that scalar invariance needed to be demonstrated (Wu et al., 2007).

4. Test for residual (strict) invariance. In addition to scalar invariance constraints, the regression residual variances are constrained to be equal. However, Wu and colleagues (2007) comment that if the item communalities are high (i.e., > .50) the strict invariance test is likely to give the same result as the strong invariance test. Hence, strict invariance is not required to accept that parameter values are equivalent between groups.

Results

Student Conceptions of Assessment

Measurement Models. For each form, exploratory and confirmatory factor analysis produced a measurement model which consisted of six 1st-order conceptions of assessment. Three factors aggregated into a positive benefit 2nd-order factor and two factors aggregated into a negative 2nd-order factor. These two 2nd-order factors inter-correlated with a sixth 1st-order factor. The models exhibited goo fit characteristics (**Form 1** [Figure 2]: n=1774; χ^2= 484.48; df=111; $\chi^2/df=4.37$, $p=<.01$; gamma hat=.98; RMSEA=.044; SRMR=.038; **Form 2** [Figure 3]: n=1623; χ^2=1050.79; df=183; $\chi^2/df=6.45$; $p<.01$; gamma hat=.95; RMSEA=.058; SRMR=.065).

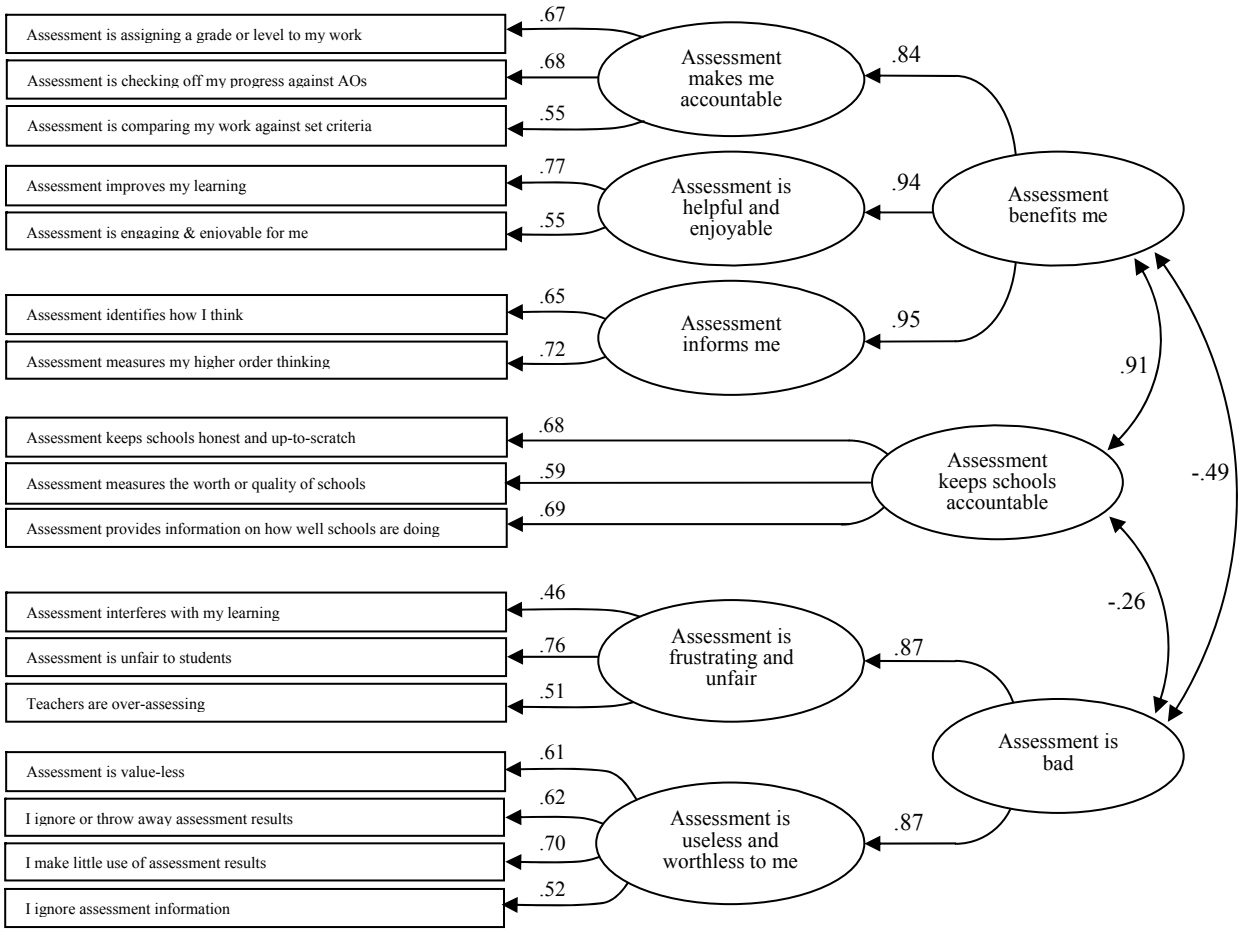


Figure 2. Form 1 SCoA Measurement Model

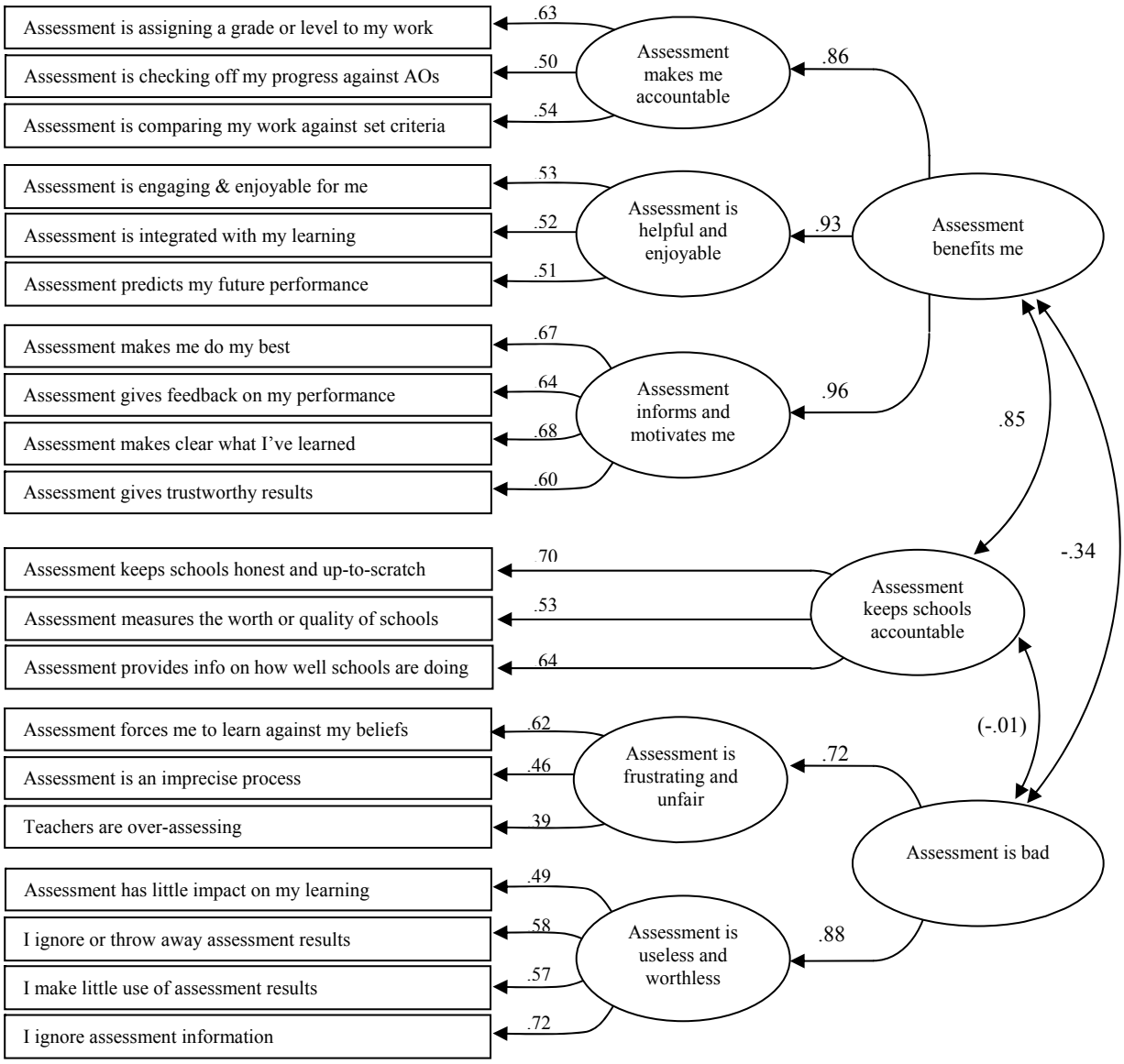


Figure 3. Form 2 SCoA Measurement Model

Structural Models. The six 1st-order factors were allowed to regress simultaneously onto the overall reading score. Each model kept the inter-correlated and 2nd-order structure demonstrated in the measurement models. The structural models had acceptable fit (**Form 1** [Figure 3]: n=1774; χ^2=520.99; df=122; χ^2/df *4.27; p=<.01*; gamma hat=.97; RMSEA=.043; SRMR=.038; **Form 2** [Figure 4]: n=1623; χ^2=1160.70; df=177; χ^2/df*=6.56; p=<.01*; gamma hat=.94; RMSEA=.059; SRMR=.065).

For both forms, reading performance was positively predicted by one conception (i.e., assessment makes me accountable), while five conceptions predicted decreased performance (i.e., assessment is helpful and enjoyable, assessment informs me, assessment keeps schools accountable, assessment is frustrating and unfair, assessment is useless and worthless). These results were consistent with Brown and Hirschfeld (2008). Only two paths were equally

statistically significant in both models (i.e., assessment makes me accountable and assessment keeps schools accountable).

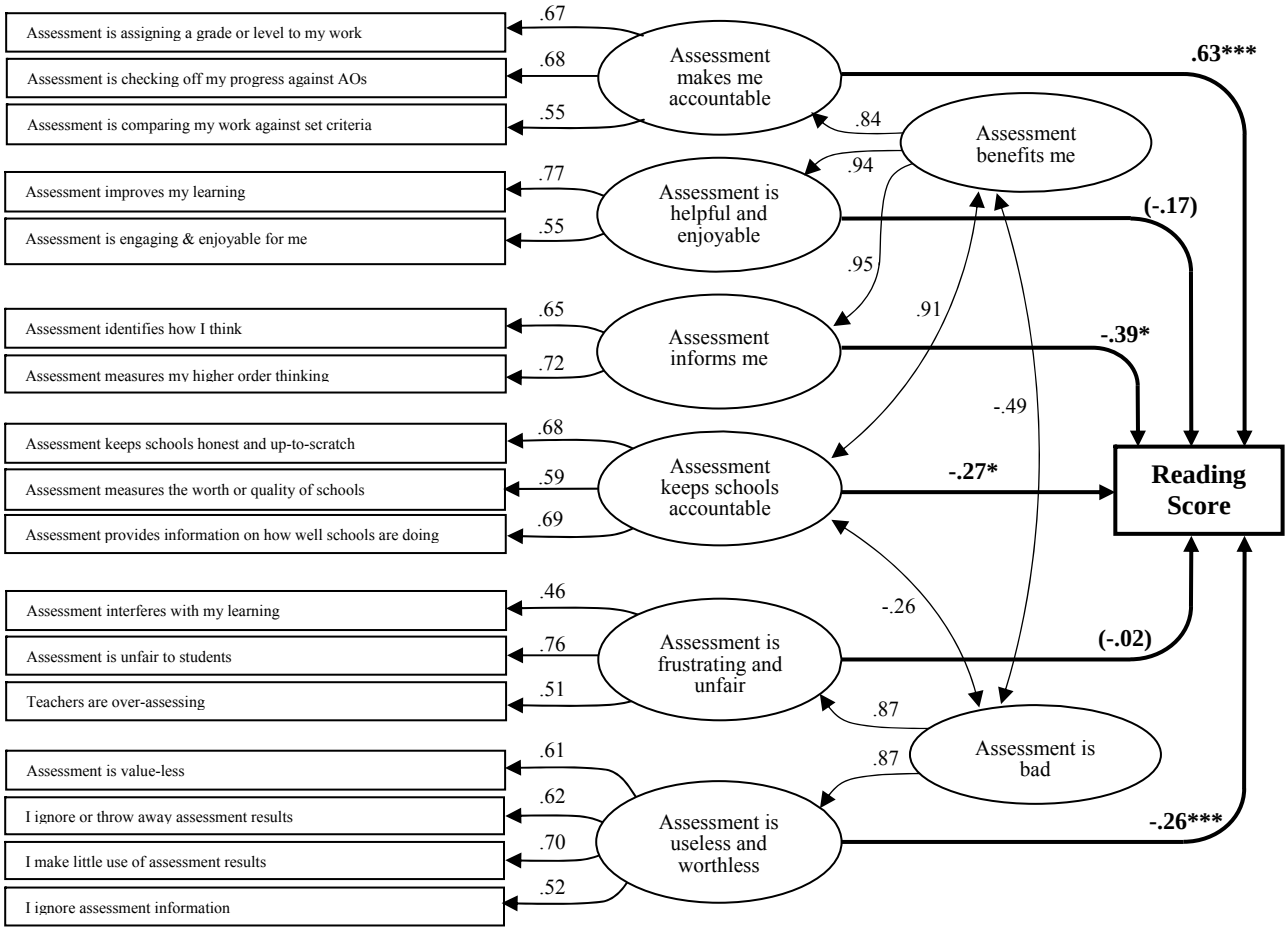


Figure 4. Form 1 SCoA Structural Model

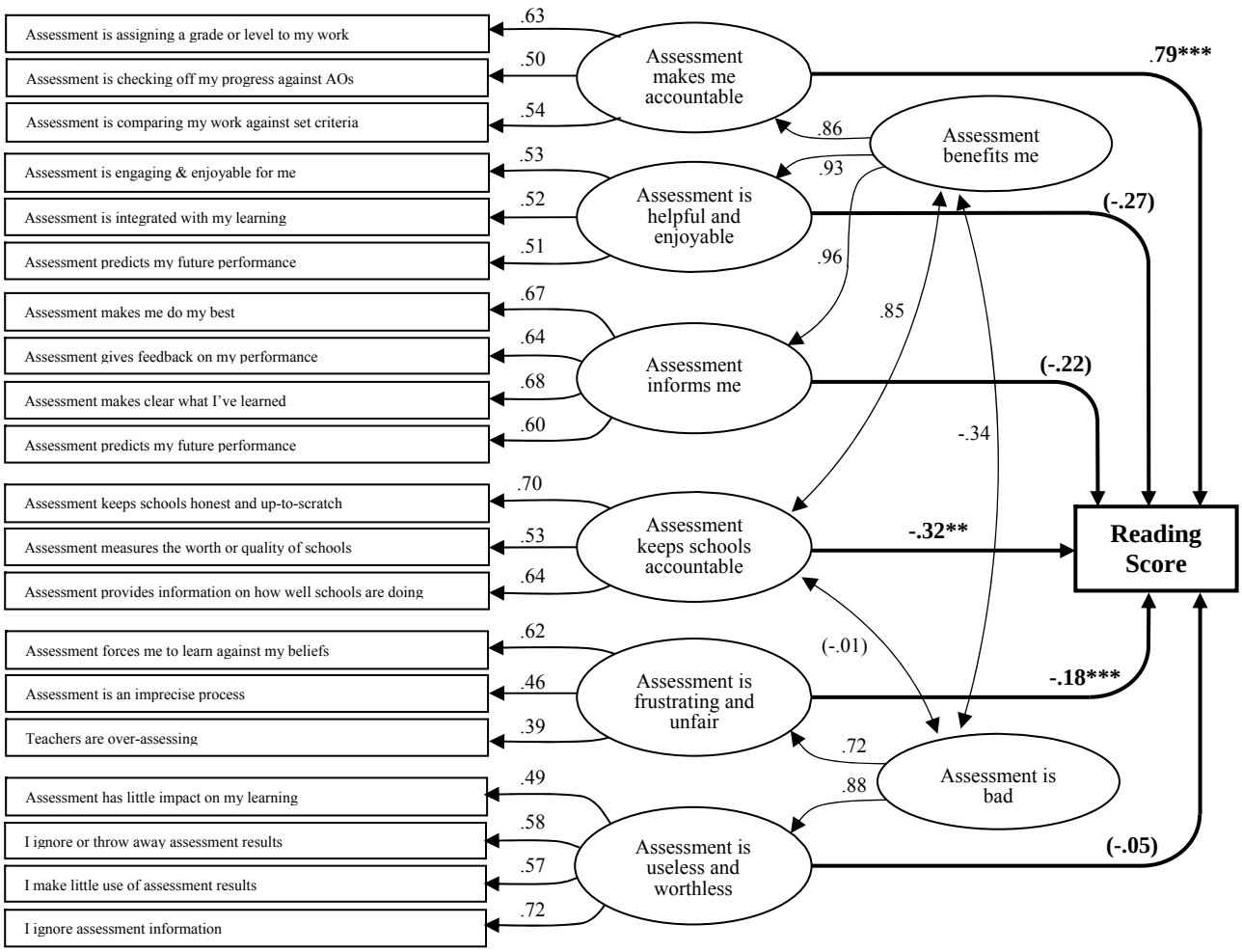


Figure 5. Form 1 SCoA Structural Model

Invariance Analysis of SCoA Structural Models across Low/high self-motivational attitudes

Tables 2 and 3 report the invariance test statistics for the self-efficacy and interest group comparisons respectively. Under the criteria set, configural invariance was demonstrated, whereas only metric invariance was demonstrated in Form 1 for both self-efficacy and interest groupings. While the ΔGamma hat cut-off criterion was generally more stringent than the ΔCFI criterion, we concluded that the groups, which differed by degree of self-efficacy and interest in reading, came from different populations which, in turn, had different structural relations in their conceptions of assessment to reading performance. This provided statistical evidence that student responses to assessment were mediated by their subject attitudes.

Table 2: SCoA Invariance Analysis across Low/High 'Self-efficacy in Reading' Groups

(*N* = 550)	RMSEA	$\Delta\chi^2$	Δ*df*	*p*	CFI	ΔCFI	ΔGh
Form 1 Measurement Model							
Test 1 Configural	.035	-	-	-	.937	-	-
Test 2 Metric (weak)	.035	27.34	11	<.01*	.934	.003	.0008
Test 3 Scalar (strong)	**.039**	**116.23**	**17**	**<.01***	**.913**	**.021***	**.0054***
Form 1 Structural Model							
Test 1 Configural	.034	-	-	-	.935	-	-
Test 2 Metric (weak)	**.034**	**41.55**	**20**	**<.01***	**.931**	**.004**	**.0012***
Form 2 Measurement Model							
Test 1 Configural	.043	-	-	-	.871	-	-
Test 2 Metric (weak)	**.043**	**42.20**	**14**	**<.01***	**.865**	**.006**	**.0012***
Form 2 Structural Model							
Test 1 Configural	.041	-	-	-	.862	-	-
Test 2 Metric (weak)	**.041**	**52.21**	**21**	**<.01***	**.856**	**.006**	**.0016***

Note. RMSEA = root mean square error of approximation; CFI = comparative fit index; Δgh = Δgamma hat; *=invariance cut-off NOT met; Bold = invariance test rejected

Table 3: SCoA Measurement Invariance across Low/High 'Interest in Reading' groups

(*N* = 550)	RMSEA	$\Delta\chi^2$	Δ*df*	*p*	CFI	ΔCFI	Δgh
Form 1 Measurement Model							
Test 1 Configural	.039	-	-	-	.927	-	-
Test 2 Metric (weak)	.038	21.42	11	.03*	.925	.002	.0001
Test 3 Scalar (strong)	**.045**	**202.22**	**17**	**<.01***	**.888**	**.037***	**.0095***
Form 1 Structural Model							
Test 1 Configural	.037	-	-	-	.927	-	-
Test 2 Metric (weak)	**.037**	**38.56**	**20**	**<.01***	**.923**	**.004**	**.0015***
Form 2 Measurement Model							
Test 1 Configural	.041	-	-	-	.875	-	-
Test 2 Metric (weak)	**.041**	**36.50**	**14**	**<.01***	**.871**	**.004**	**.0012***
Form 2 Structural Model							
Test 1 Configural	.041	-	-	-	.867	-	
Test 2 Metric (weak)	**.041**	**50.39**	**21**	**<.01***	**.861**	**.006**	**.0016***

Note. RMSEA = root mean square error of approximation; CFI = comparative fit index; Δgh = Δgamma hat; *=invariance cut-off NOT met; Bold = invariance test rejected

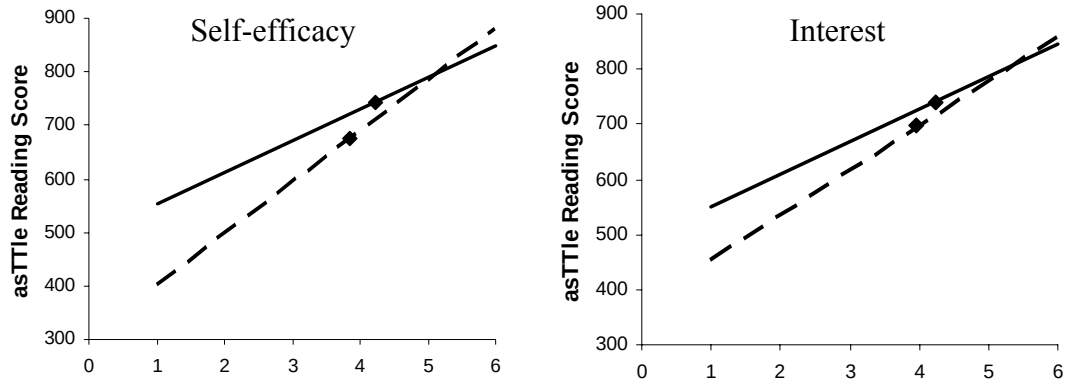
Having demonstrated that self-motivational attitudes towards reading had a statistically significant effect on the regression from conceptions of assessment to reading performance, the structural model regression weights were examined for each group to refine in what way the groups differed (Table 4).

Table 4: SCoA Structural Models Regression Weights for Factors predicting Reading Achievement by Attitude and Form

	Conceptions of Assessment and Attitude Level											
	Makes me accountable		Keeps schools accountable		Helpful & enjoyable		Informs & motivates me		Unfair & frustrating		Useless & worthless	
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Self-efficacy												
Form 1	.61	.95	.08	-.15	-.15	-.27	-.83	-.70	.03	-.03	-.46	-.20
Form 2	.64	.86	-.05	-.58	-.64	-.08	-.04	-.21	-.28	-.12	-.05	-.02
average	**.63**	**.91**	**.02**	**-.37**	**-.40**	**-.18**	**-.44**	**-.46**	**-.13**	**-.08**	**-.26**	**-.11**
Interest												
Form 1	.53	.60	-.08	-.19	-.01	-.09	-.79	-.43	-.03	-.04	-.38	-.11
Form 2	.65	.98	-.08	-.80	-.41	-.15	-.35	.01	-.36**	.04	.01	-.04
average	**.59**	**.79**	**-.08**	**-.50**	**-.21**	**-.12**	**-.57**	**-.22**	**-.20**	**.00**	**-.19**	**-.08**

A number of interesting general patterns were detected. Only the 'makes me accountable' factor had large regressions across groups and forms, with a noticeably larger effect for the low efficacy/interest group. Three factors generally had very similar effects across group, form, and attitude (i.e., informs and motivates me, unfair and frustrating, and useless and worthless). The 'keeps schools accountable' had a larger negative effect for the low self-efficacy/interest groups, while the 'helpful and enjoyable' factor had a more pronounced negative effect for the high self-efficacy/interest groups.

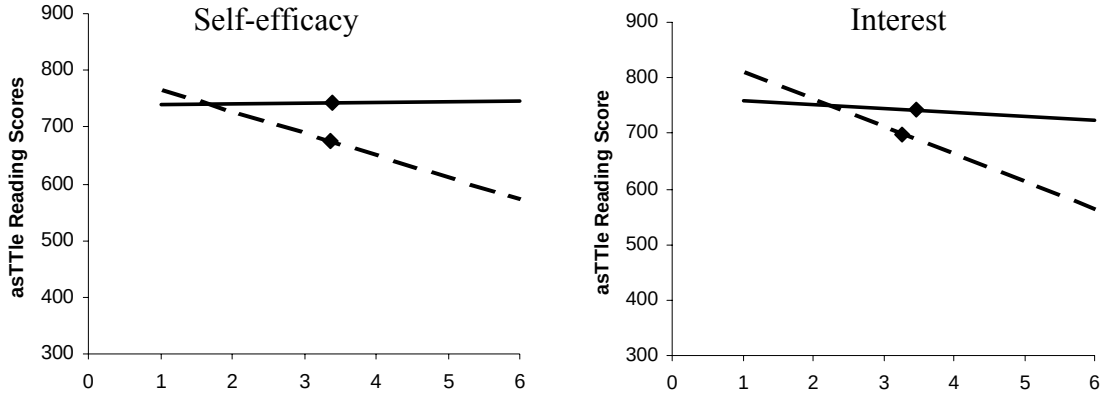
The regression weights were used to construct plots of predicted reading comprehension score by strength of conception for the high and low self-efficacy and interest groups. (Figure 6 to Figure 11). The gradient of each plot was calculated by multiplying the standard deviation of the reading score by the regression weight and dividing by the standard deviation of the conception factor score.



Note. Solid Line=high group; dashed line=low group

Figure 6. Predicted asTTle Reading Scores vs Strength of 'Assessment makes me accountable' Conception by Self-efficacy and Interest and Grouping

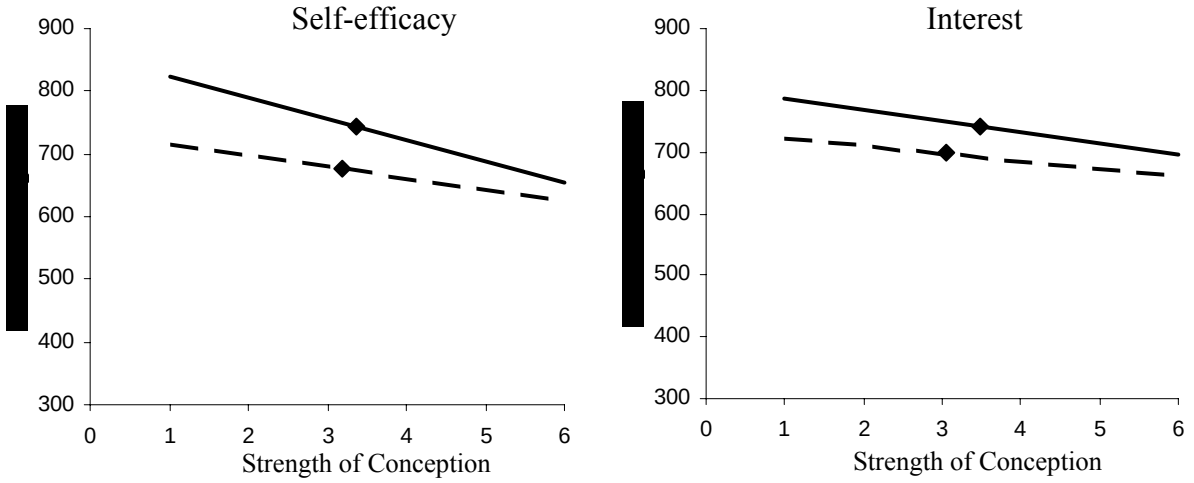
Figure 6 showed that in general, while students with higher self-motivational attitude levels performed better academically, the conception that 'assessment makes me accountable' had a stronger positive effect for the low attitude groups. This indicates the significant adaptive role that personal accountability attitudes and approaches have in improving student achievement, most especially for those with the least self-efficacy or interest. Taking assessment seriously, helped the weaker students more than the academically stronger students. This indicates that a virtuous cycle of increased academic performance can begin with increased endorsement of the conception that assessment evaluates the learner. This should in turn lead to greater self-efficacy beliefs and interest. Appropriate extrinsic motivation such as the pressure of a formal assessment event, if mediated appropriately by the teacher, may provide these students with the motivational push needed to raise their performance (Hidi & Harackiewicz, 2000). It is worth noting that for students with highly positive subject attitudes, academic performance also increases with the same belief. Hence, even for students with the most positive attitudes, self-regulation requires acceptance of being evaluated.



Note. Solid Line=high group; dashed line=low group

Figure 7. Predicted Reading Scores vs Strength of 'Assessment makes schools accountable' Conception by Self-efficacy and Interest and Grouping

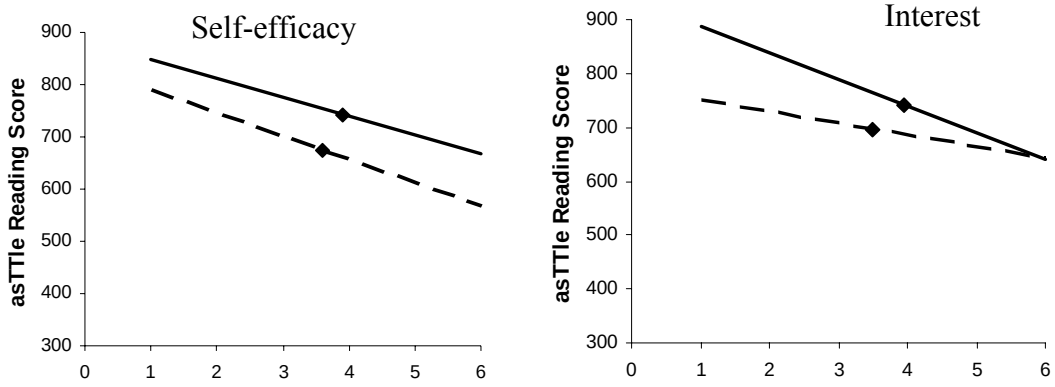
Figure 7 shows the strongly negative impact on students with low self-efficacy/interest with greater endorsement of the conception that 'assessment makes schools accountable'. Students with low self-motivation attitudes are more likely to shun personal responsibility in their learning by blaming external agencies for their underachievement. Thus, students with low efficacy or low interest in a subject are more likely to take a maladaptive approach towards assessment by making ego-protective excuses for poor results, with deleterious effects on their academic performance. It is likely that this effect is driven by attributions and locus of control processes that place responsibility for educational outcomes on the learner's environment. Such beliefs are known to be maladaptive and this result shows that it is the least motivated students who are hurt by this emphasis. In contrast, this conception has nearly no effect on students with high self-efficacy/interest, indicating the irrelevance of placing responsibility on an external agency for this group of students. This suggests that focusing on evaluating the school as a motivation for participation in an assessment task will reduce the performance of those with the least interest and self-confidence. There are strong implications for all low-stakes, system monitoring assessments; the least motivated students will do considerably worse, providing a very poor index of system performance. Clearly, helping low-motivation students to see assessment as a personal accountability event has the opposite and adaptive effect.



Note. Solid Line=high group; dashed line=low group

Figure 8. Predicted Reading Scores vs Strength of 'Assessment is helpful and enjoyable' Conception by Self-efficacy and Interest and Grouping

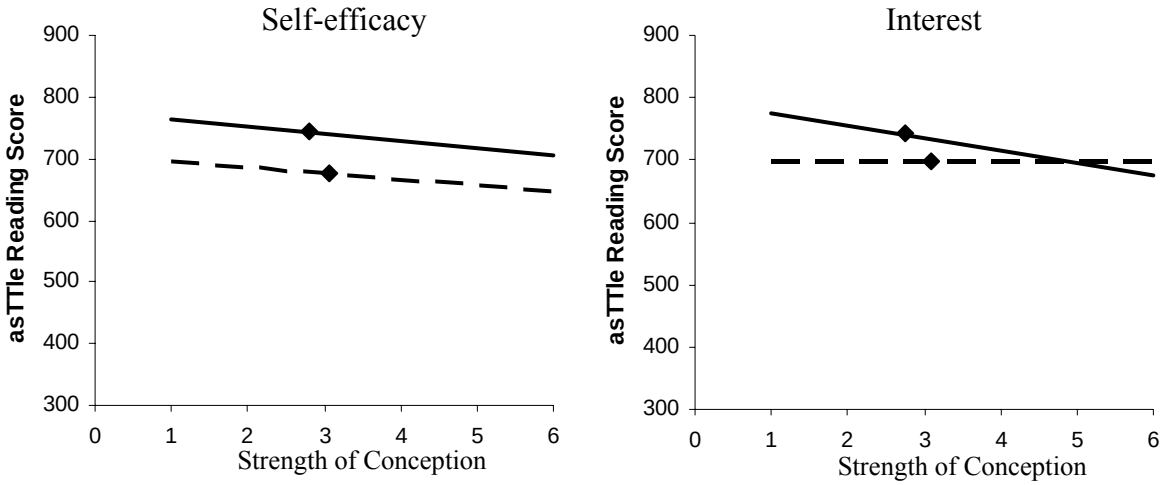
Agreement with the conception 'assessment is helpful and enjoyable' negatively affected both high and low self-motivation groups in very similar ways (Figure 8). Students who expect assessment to be a pleasant experience, could be placing well-being goals above learning goals (Boekaerts & Corno, 2005), and are likely to perform less well compared to their peers who take assessment more seriously and accept the reality of the unpleasant stresses that typically accompany rigorous assessment. Students who are highly efficacious are somewhat more at risk of underachieving if they take the attitude 'I'm good at this, so I don't need to take assessment too seriously'. While teachers have an important role in encouraging and motivating students, an over-emphasis on trying to make assessment pleasant by using ego-enhancing praise and other 'feel good' strategies is unlikely to help students develop a more adaptive attitude and approach to assessment. Since these students are already interested and motivated, the focus for these students should not be on 'feel good' factors such as making learning fun, easy and entertaining, but on making it challenging, demanding, and intrinsically rewarding.



Note. Solid Line=high group; dashed line=low group

Figure 9. Reading Scores vs Strength of 'Assessment informs and motivates me' Conception by Self-efficacy and Interest and Grouping

Agreement with the conception 'assessment informs and motivates me' also has a serious detrimental effect on achievement for all students irrespective of efficacy and interest levels (Figure 9). It suggests that for some students the formative emphasis for assessment is not having the desired or anticipated effect. It might be that students see this as a 'soft option', in that, assessment will do the learning for me. If this is the case, then students may see themselves as being let off the hook of having to take personal responsibility for their progress. It is interesting that this current educational emphasis on student-engaged and motivated formative assessment has the greatest negative impact on the most positively motivated of students. Clearly, such students do not need this type of message—motivation for performance and learning appears not to come from assessment. Perhaps for self-regulating students especially, motivation comes from internal sources.



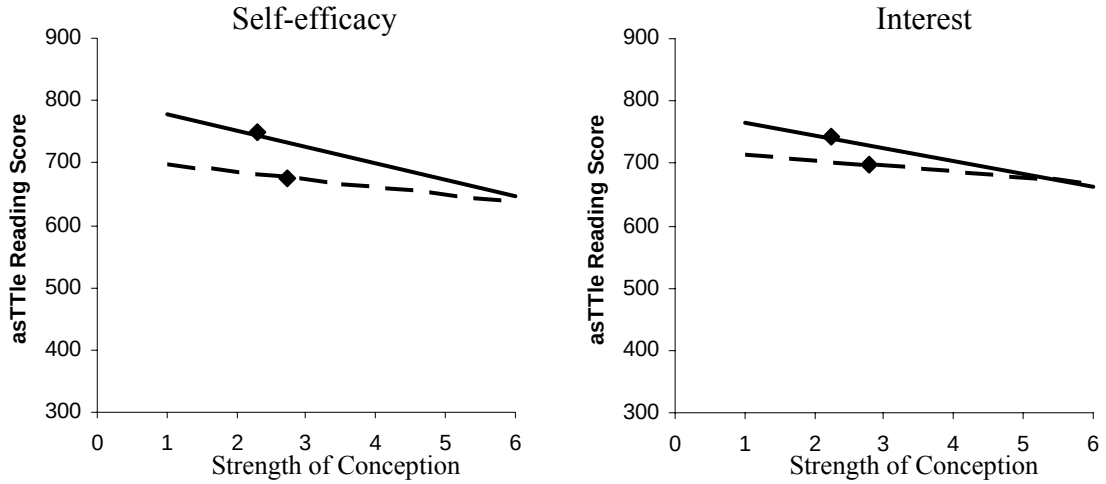
Note. Solid Line=high group; dashed line=low group

Figure 10. Predicted Reading Scores vs Strength of 'Assessment is unfair and frustrating' Conception by Self-efficacy and Interest and Grouping

Figure 10 demonstrates that stronger agreement with the conception 'assessment is unfair and frustrating' predicted a small detrimental effect on reading scores. While the negative effect is similar regardless of level of self-efficacy, there appeared to be a greater negative effect in believing that assessment was at fault among students with the highest interest in reading. It might be that those who like reading more find assessment much more unfair; it's as if students are indicating "I'd rather read than do a test about what I'm reading". Nonetheless, across all groups the greater the belief in the unfairness of testing, the worse the performance. This is a fundamentally maladaptive belief, perhaps centred on ego-protection and well-being emphasis; avoiding of discomfort leads to lowered performance. Some performance anxiety and discomfort appears necessary to greater performance. Students who accept this outperform those who reject it.

Similar to Figure 10, Figure 11 shows that agreement with the conception 'assessment is useless and worthless' predicted a small reduction in reading performance. Again, the negative impact of this conception of assessment is somewhat greater for the high self-efficacy/ interest students. Totally rejecting the usefulness of assessment is a fundamentally maladaptive belief. It probably originates along with the conception that assessment is unfair in ego-protection and well-being emphasis. Rejecting the act of assessment leads to lowered

performance. Students who accept either the inevitability or usefulness of assessment are more self-regulating than those who reject the processes.



Note. Solid Line=high group; dashed line=low group

Figure 11. Predicted Reading Scores vs Strength of assessment is useless and worthless Conception by Self-efficacy and Interest and Grouping

Discussion

This study has shown that the expanded SCoA inventory analysed here generated conceptually similar factors to those reported previously (Brown & Hirschfeld, 2008; Brown et al. 2009) and that the same factors had similar regression weights to previous studies (Brown & Hirschfeld, 2008; Brown et al. 2008). The findings of this study demonstrate that students' academic achievement is significantly advanced by their viewing of assessment as a regulatory tool to help them monitor their progress towards mastery learning goals. By taking an adaptive, self-regulatory approach to assessment, students can take full advantage of assessment information to evaluate their subject competency, and plan where next to focus their energy and effort. As their efforts are rewarded with greater academic competency, their self-efficacy improves and this feeds positively into their self-motivation producing a 'virtuous cycle' of progress.

Students who do not view the role of assessment as primarily keeping them personally accountable, but place a higher priority on protecting their self-image, tend to make excuses, adopt avoidance strategies or blame others for poor performances. Their extended focus on well-being goals causes them to look for temporary 'feel good' experiences rather than accepting the necessity of being put under pressure with suitably challenging assessment tasks. If students remain in this mode of thinking for any length of time, there is a risk these maladaptive attitudes will become more entrenched, causing learning goals to be abandoned thereby producing poor work habits. The resulting academic under-achievement is likely to have a further de-motivating effect producing a 'vicious cycle'. Whether these maladaptive attitudes are expressed negatively (assessment is unfair, frustrating and worthless) or masked by being expressed positively (assessment is meant to make others accountable, be pleasant, and help me learn), there is likely to be a detrimental effect on their academic achievement.

The finer-grained mediating effects of subject specific self-motivational attitudes give further clues as to how students' academic achievement can be optimised. Students who have positive self-motivational attitudes towards a subject and also view assessment as a way to check their progress, develop a positive, realistic self-efficacy based on accurate feedback information. They are willing to tackle challenging problems and persist in the face of difficulties (Dweck, 2000). However, students who believe they are good at a subject and are interested in it cannot afford to become complacent about assessment. They are at risk of serious underachievement if they begin to adopt attitudes such as 'I'm doing okay, so I don't need to take assessment too seriously', or give up on valuing assessment feedback due to what is perceived as an unfair or frustrating experience. These students need to continue giving priority to mastery learning goals and not let themselves be side-tracked into an ego-protective, excuse-making way of thinking. Teachers have an important contribution to make by ensuring assessment events are conducted in a fair manner with minimal frustration, and by giving quality feedback to students. They also need to attend to the affective domain by providing a positive classroom climate where students' contributions are welcomed, by expressing high expectations of their students, by giving positive encouragement for persistent effort to maintain students' sense of well-being and belonging. This will prevent students from experiencing unnecessary stressors which could push them towards a more ego-protective way of thinking (Boekaerts & Corno, 2005).

Turning around students with low self-motivational subject attitudes poses a particular challenge for teachers. These students are more likely to make excuses for poor performance, blame others and retain entrenched negative attitudes towards assessment. However, if they can be helped to adopt realistic learning goals, see assessment as a useful tool in helping them achieve those goals, and stop making excuses for poor achievement there are significant pay offs in terms of their academic progress. For these students changing both their motivational attitudes and their attitudes towards assessment appears to be the most effective strategy.

References

Biggs, J., & Moore, P. (1993). *The process of learning*. New York: Prentice Hall.

Boekaerts, M. (1995). Self-regulated learning: Bridging the gap between metacognitive and metamotivation theories. *Educational Psychologist, 30*(4), 195-200.

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review, 18*, 199-210.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: a perspective on assessment and intervention. *Applied Psychology: An International Review, 54*(2), 199-231.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 416-450). San Diego: Academic Press.

Brophy, J. (1998). *Motivating students to learn*. Boston: McGraw-Hill.

Brown, G. T. L. (2002). *Teachers' conceptions of assessment.* University of Auckland, Auckland.

Brown, G. T. L. (2004). Measuring attitudes with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports, 94*, 1015-1024.

Brown, G. T. L. (2006). Secondary school students' conceptions of assessment: A survey of four schools *Conceptions of Assessment and Feedback Project Report #5*. Auckland: University of Auckland.

Brown, G. T. L., & Hirschfeld, G. H. F. (2007). Students' conceptions of assessment and mathematics: Self-regulation raises achievement. *Australian Journal of Educational and Developmental Psychology, 7*(63-74).

Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice, 15*(1), 3-17.

Brown, G. T. L., Irving, S. E., & Peterson, S. R. (2008). *Beliefs that make a difference: Students' conceptions of assessment and mathematics achievement.* Paper presented at the International Test Commission Biannual Conference, Liverpool, UK.

Brown, G. T. L., Irving, S. E., Peterson, S. R., & Hirschfeld, G. H. F. (2009). Use of informal-interactive assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction, 19*(2), 97-111.

Brown, G. T. L., Peterson, E. R., & Irving, S. E. (in press). Self-regulatory beliefs about assessment predict mathematics achievement. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.) *Student perspectives on assessment: What students can tell us about assessment for learning*. Charlotte, NC: Information Age Publishing.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Byrne, B. M. (2004). Testing for multigroup invariance using AMOS graphics: A road less travelled. *Structural Equation Modeling, 11*(2), 272-300.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.

Cleary, T. C. (2006). The development and validation of the self-regulation strategy inventory -- self-report. *Journal of School Psychology, 44*, 307-322.

Covington, M. V., & Omelich, C. L. (1979). Effort: the double-edged sword in school achievement. *Journal of Educational Psychology, 71*, 169-182.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah N.J.: Lawrence Erlbaum Associates.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Models, 12*, 343-367.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research, 42*(3), 509-529.

Hattie, J. A. (2004). *Models of self-concept that are neither top-down or bottom-up: The rope model of self-concept.* Paper presented at the 3rd International Biennial Self Research, Berlin.

Hattie, J. A., Brown, G. T. L., & Keegan, P. J. (2003). A national teacher-managed, curriculum-based assessment system: Assessment tools for teaching and learning (asTTle). *International Journal of Learning, 10*, 771-778.

Hattie, J. A., Brown, G. T. L., Keegan, P. J., MacKay, A. J., Irving, S. E., Cutforth, S., et al. (2004). *Assessment Tools for Teaching and Learning (asTTle) Version 4, 2005: Manual*. Wellington: University of Auckland/ Ministry of Education/ Learning Media.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*(2), 151-179.

Hirschfeld, G. H. F., & Brown, G. T. L. (2009). Factorial and structural invariance of the SCoA across sex, age and ethnicity. *European Journal of Psychological Assessment, 25*(1), 30-38.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modelling: Sensitivity to underparametized model misspecification. *Psychological Methods, 3*, 424-453.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55.

Marsh, H., Hau, K., Artelt, C., Baumert, J., & Peschar, J. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing, 6*(4), 311-360.

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers of overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320-341.

Ministry of Education. (2007). *The New Zealand curriculum for English-medium teaching and learning in years 1-13*. Wellington: Learning Media Ltd.

Otunuku, M., & Brown, G. T. L. (2007). Tongan students' attitudes towards their subjects in New Zealand relative to their academic achievement. *Asia Pacific Education Review, 8*(1), 117-128.

Paris, S. G., & Turner, J. C. (1994). Situated motivation. In P. R. Pintrich, D. R. Brown & C. E. Weinstein (Eds.), *Student motivation, cognition and learning: Essays in honor of Wlibert J. McKeachie* (pp. 213-237). Hillsdale, NJ: Lawrence Erlbaum.

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Schunk, D. H. (1983). Progress self-monitoring: Effects on children's self-efficacy and achievement. *Journal of Experimental Education, 51*, 89-93.

Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631-649). San Diego: Academic Press.

Sivo, S. A., Fan, X., Witta, E. L., & Willse, J. T. (2006). The search for "optimal" cutoff properties: Fit index criteria in structural equation modelling. *The Journal of Experimental Education, 74*(3), 267-288.

Weekers, A. M., Brown, G. T. L., & Veldcamp, B. P. (2009). Analyzing the dimensionality of the Students' Conceptions of Assessment. In D. M. McInerney, G. T. L. Brown & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about improving school outcomes*. Greenwich, CT: Information Age Press.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. . *Practical Assessment, Research & Evaluation, 12*(3), 1-26.

Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix A. SCoA II Form 1 and Form 2 Items and Expected Factors

FORM 1 (20 items)		
Label	Item Statement	Expected Factor
COAac4*	Assessment is assigning a grade or level to my work	accountability (me)
COAac5*	Assessment is checking off my progress against achievement objectives	accountability (me)
COAac6*	Assessment is comparing my work against set criteria	accountability (me)
COAac8*	Assessment keeps schools honest and up-to-scratch	accountability
COAac9*	Assessment measures the worth or quality of schools	accountability
COAac11*	Assessment provides information on how well schools are doing	accountability
COAimp4	Assessment helps me improve my learning	improvement
COAimp8*	Assessment is a positive force for improving the social climate	improvement
COAimp9*	Assessment is an engaging and enjoyable experience for me	improvement
COAimp10	Assessment is appropriate and beneficial for me	improvement
COAval3	Assessment identifies how I think	valid
COAval5	Assessment makes clear and definite what I have learned	valid
COAval7	Assessment measures my higher order thinking	valid
COAir3	Assessment interferes with my learning	irrelevant
COAir5	Assessment is unfair to students	irrelevant
COAir6	Assessment is value-less	irrelevant
COAir8*	I ignore or throw away assessment results	irrelevant
COAir9*	I make little use of assessment results	irrelevant
COAir10*	I ignore assessment information	irrelevant
COAir13*	Teachers are over-assessing	irrelevant
FORM 2 (21 items)		
Label	Item Statement	Expected Factor
COAac4*	Assessment is assigning a grade or level to my work	accountability (me)
COAac5*	Assessment is checking off my progress against achievement objectives	accountability (me)
COAac6*	Assessment is comparing my work against set criteria	accountability (me)
COAac8*	Assessment keeps schools honest and up-to-scratch	accountability
COAac9*	Assessment measures the worth or quality of schools	accountability
COAac11*	Assessment provides information on how well schools are doing	accountability
COAimp8*	Assessment is a positive force for improving the social climate	improvement
COAimp9*	Assessment is an engaging and enjoyable experience for me	improvement
COAimp11	Assessment is integrated with my learning	improvement
COAimp12	Assessment makes me do my best	improvement
COAimp13	Assessment provides feedback to me about my performance	improvement
COAval2	Assessment makes clear and definite what I have learned	valid
COAval9	Assessment results are trustworthy	valid
COAval11	Assessment results predict my future performance	valid
COAir1	Assessment forces me to learn in a way against beliefs about learning	irrelevant
COAir2	Assessment has little impact on my learning	irrelevant
COAir4	Assessment is an imprecise process	irrelevant
COAir8*	I ignore or throw away assessment results	irrelevant
COAir9*	I make little use of assessment results	irrelevant
COAir10*	I ignore assessment information	irrelevant
COAir13*	Teachers are over-assessing	irrelevant

* items common to both Forms

Note: In the actual questionnaires given to students, the items appeared ungrouped in random sequence to spread any fatigue effects evenly among the factors.

Appendix B. SAR Items and Expected Factors

Label	Item Statement	Expected Factor
Att 1	How good do you think you are at reading?	Self-efficacy
Att 5	How good does your teacher think you at reading?	Self-efficacy
Att 6	How good does your mum or dad think you are at reading?	Self-efficacy
Att 2	How much do you like doing reading at school?	Liking/interest
Att 3	How much do you like doing reading in your own time (not at school)?	Liking/interest
Att 4	How do you feel about going to the library to get something to read?	Liking/interest

The numbering of the labels indicates the sequence in which the items appeared in the asTTle tests.