The more I enjoy it the less I achieve: The negative impact of socio-emotional purposes of assessment and feedback on academic performance

Gavin T. L. Brown[a]
The Hong Kong Institute of Education
S. Earl Irving[b]
Elizabeth R. Peterson [c]
The University of Auckland

[b] Corresponding Author:
Faculty of Education, The University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, NEW ZEALAND

[a] email: gtlbrown@ied.edu.hk
[b] email: e.irving@auckland.ac.nz
[c]email: e.peterson@auckland.ac.nz

Citation:
Brown, G. T. L., Irving, S. E., & Peterson, E. R. (2009, August). *The more I enjoy it the less I achieve: The negative impact of socio-emotional purposes of assessment and feedback on academic performance*. Paper presented at EARLI 2009 biennial conference, Amsterdam, NL.

Abstract

How students understand the role and nature of feedback in relation to assessment is not well understood, but it is assumed that their belief systems can add to or detract from their academic performance. Assessment reforms have tried to increase performance through 'formative' assessment practices (e.g., greater engagement and motivation through alternative assessment practices such as self- and peer-assessment), but such processes are mediated by students' thinking about them. Within a self-regulation framework, relative to growth in learning some beliefs are adaptive (i.e., increase performance), while others are maladaptive (i.e., decrease performance). In contrast, other beliefs increase sense of personal well-being which may or may not be adaptive. It was anticipated that (a) conceptions of assessment and feedback would be complementary and (b) conceptions consistent with self-regulation theory would predict greater learning outcomes. The degree to which students' conceptions of assessment and feedback could predict their mathematics achievement was investigated in a study of 499 Year 10 New Zealand secondary school students. The students' conceptions were measured using the Students' Conceptions of Assessment –V (SCoA-V) Inventory (Brown, 2006) and the Conceptions of Feedback-III (COF-III) Inventory (Irving, Peterson, & Brown 2008). A standardised achievement test (asTTle) was used to measure mathematics achievement. Measurement models for the SCoA-V and the COF-III were acceptable and the structural relations between the conceptions and achievement found conceptually and practically meaningful results. The relations between conceptions of assessment and feedback formed four meta-conceptions (i.e., Assessment and Feedback are for Improvement; Assessment and Feedback reflect Extrinsic Attributions; Feedback and Assessment are Irrelevant; and Feedback and Assessment have Socio-Emotional Effects), which in turn explained 13% of variance in mathematics achievement. The latent concept Assessment and Feedback for Improvement (β=.30) was positive, while the remaining concepts had negative regressions to performance (i.e., Extrinsic Attributions (β=-.17); Feedback and Assessment are Irrelevant (β=-.26); and Feedback and Assessment have Socio-Emotional Effects (β=-.32). This suggests that NZ secondary students deem emphasis on socio-emotional purposes of feedback and assessment as counter-productive on a self-regulation growth pathway. From the student point of view, assessment reforms which maximise well-being are maladaptive. Greater emphasis on adaptive growth pathway beliefs about assessment and feedback are required to enhance learning outcomes.

Introduction

Assessment and feedback are integral parts of a learning system. Reform pressures in educational assessment suggest that students who receive feedback from assessment processes that they enjoy will achieve more (Harlen, 2007; Weeden, Winter, & Broadfoot, 2002). At the heart of these reforms is **formative** feedback in a context of assessment **for** learning – the assessment provides an opportunity to provide feedback (from a variety of sources such as parents, peers, self and teachers) and this feedback is focused on **what** to improve and **how** to improve.

In this paper we seek to explain how student conceptions of assessment and feedback relate to each other and impact on their learning outcomes. We do so by examining the purposes of assessment and feedback, the emotional effects of assessment and feedback on students, and how assessment and feedback impact on student motivation.

The purposes of assessment and feedback

Feedback is normally provided to students after an assessment event. Assessment and feedback have a number of purposes depending on the audience or stakeholder. For example, Hornby (2003) argues that assessment can be: i) summative, providing information about current achievement, ii) formative, providing guidance to improve student learning, iii) certificating, assisting with identifying individuals with appropriate qualifications and iv) evaluative, providing information to stakeholders (e.g., parents, teachers, schools) to enable them to evaluate the success of the educational system. Brown (2008) has argued that all variations as to the purpose of assessment reduce to one of four uses: a) improving teaching and learning, b) evaluating students, c) evaluating schools and teachers, or d) rejecting the use of assessment. The true educational purpose of assessment and feedback is to improve student academic performance (Popham, 2000).

While feedback may most typically have its effects after an assessment event, assessment itself seems to have an impact on students before it is conducted (i.e., it shapes learning and studying), after the event (i.e., feedback from performance), and even during the event (i.e., students learn while being assessed) (Segers, Dochy, Gijbels, Struyven, in press).

Emotional affects of assessment and feedback

While assessment and feedback is argued to have numerous formal roles within the education system, there may be unintended consequences and effects beyond influencing and monitoring student learning outcomes, and motivating students to improve. In particular, an individuals' perception and interpretations of assessment and feedback has been found to influence students' sense of self and well-being (Black & Wiliam, 1998). A number of researchers have found that assessment practices have strong emotional impacts on students both in terms of their self-concept and in their relations with other students (Cowie, in press; Harris, Harnett, & Brown, in press; Remesal, in press).

Models of conceptions of assessment and feedback.

The importance of involving students in feedback and assessment processes especially in a framework of assessment for learning has led to research that attempts to model how it is students understand, experience, and respond to assessment and feedback. Our research program in New Zealand has focused on developing inventories and measurement models that explicitly identify these conceptions and how they relate to each other and academic performance.

Conceptions of assessment

Earlier research in our programme of work has identified high school students' conceptions of assessment (Brown & Hirschfeld, 2008; Brown, Irving, Peterson, & Hirschfeld, 2009) and conceptions of feedback (Irving, Peterson, & Brown, 2007, 2008; Peterson & Irving, 2008). The interpretation of student conceptions of assessment has gravitated towards a self-regulation framework with explicitly adaptive and maladaptive responses on a growth pathway that increase and decrease respectively academic performance and on a well-being pathway which was independent of academic performance (Brown, Peterson, & Irving, in press). However, the inter-relationship of these two constructs has not been systematically studied.

For example, it is likely that students who are on the growth pathway or take an adaptive self-regulating approach will agree that (1) assessment can identify gaps in their learning, (2) feedback helps them to improve, and (3) studying is important for their future. In contrast, students who think that feedback does not tell them the truth or that the assessment and feedback are irrelevant are much more likely to do poorly when assessed (Peterson & Irving, 2008). Being able to identify unproductive conceptions of assessment and feedback and students with those conceptions should enable action that addresses misguided conceptions and subsequently help such students overcome barriers to educational achievement (Fransson, 1977; Elen & Lowyck, 1999).

From the student perspective, using factor-analysed modeling, four major conceptions about the purposes of and nature of assessment have been found (Table 1; Weekers, Brown, & Veldkamp, in press). It has been shown (Brown & Hirschfeld, 2008; Brown, Peterson, & Irving, in press) that the improvement conceptions lead to increased academic performance in mathematics (hence, are adaptive on the growth pathway). At the same time they found that the external pressures of the future and school accountability were maladaptive on the growth pathway (i.e., scores decreased as agreement increased). The consequence of the affective impact was negative towards teacher testing, but had a zero relationship with academic performance.

Table 1. Relationships between Conceptions of Assessment and Academic Performance

Conceptions of Assessment Factors with sub-factors	Relationship to Academic Performance
Assessment Improves Teaching and Learning	
Assessment is used by teachers to improve teaching	Adaptive
Assessment is used by students to guide learning	
Assessment makes me accountable	
Assessment Relates to External Factors	
Assessment measures school quality	Maladaptive
Assessment predicts my future	
Assessment has Affective Impact or Benefit	
Assessment is enjoyable	Neutral
Assessment improves class social climate	
Assessment is Irrelevant.	
Assessment is unfair	Maladaptive
Assessment is ignored	

Conceptions of feedback

In a series of studies (Peterson & Irving, 2008; Irving, Peterson, & Brown, 2007) a student-self report Students' Conceptions of Feedback inventory has been developed. The third version (SCoF-III; Irving, Peterson & Brown, 2008) found seven dimensions –

- Parental feedback,
- Feedback is irrelevant/ignored,
- Feedback is enjoyable,
- Feedback provides information about standards,
- Teachers give trustworthy feedback,
- Feedback is motivating, and
- Peer feedback.

In relating these factors to observed scores on a standardised mathematics test, only two of the dimensions were significant predictors of achievement – parental feedback, and feedback is motivating. Interestingly, parental feedback positively predicted achievement, while the motivating aspects of feedback negatively predicted

achievement. This latter result was curious, as it suggested that where students regarded feedback as a motivating influence, this did not have a concomitant effect on their achievement. Our interpretation of this result is that low ability students held the view that feedback was motivating, but higher ability students did not. Indeed, it has been found that high ability students do not need feedback in mathematics as they already know what they need to improve and how to do so (Kulhavy, Stock, Hancock, Swindell, & Hammrich, 1990). Thus, the very things that have been at the heart of the assessment reforms were not shown to impact on student achievement in mathematics for this group of students.

It is important to note that whereas meta-analyses such as Shute (2008) have shown which features, functions, and interactions of feedback are linked to improved learning outcomes, this study concerns the way in which students conceive of feedback, and how those conceptions are related to greater learning outcomes. Thus on the one hand we have what the student received by way of feedback and its impact on learning, and on the other hand, what the students perceived (i.e., the meanings that students attach to feedback) that may lead to further action (or inaction) and thus on their learning.

The current study

This study explores the relationship of secondary students' conceptions of assessment with their conceptions of feedback and how these structures interact to predict mathematics performance. Additionally, further validation evidence for the Student Conceptions of Assessment and Student Conceptions of Feedback inventories is generated. Because feedback is a consequence of assessment, it seems natural that the seven conceptions of feedback should be predicted by how students conceive of assessment. Hence, we hypothesised that:

1. The irrelevance of assessment would be strongly connected to the irrelevance of feedback factor
2. The affective benefit of assessment would be strongly connected to the enjoyable and peer feedback factors

3. The improvement conception of assessment factors would be strongly connected to the feedback factors that it is motivating, provides information, and teacher giving trustworthy feedback factors
4. The external attribution factor of assessment would be strongly connected to the parental feedback factor.

Based on our self-regulating interpretive framework, we further hypothesised that:

5. The joint irrelevance factors would be maladaptive to academic performance
6. The joint affective factors would be neutral towards academic performance
7. The joint improvement factors would be adaptive towards academic performance
8. The external factors would be maladaptive towards academic performance.

Method

Instruments

In this study, student were administered three instruments - the Student Conceptions of Assessment – version V (SCoA-V, Brown, 2006), the Student Conceptions of Feedback – version III (SCoF-III, Irving & Peterson, 2008), and a nationally standardised mathematics assessment, asTTle.

Student Conceptions of Assessment Questionnaire - V (SCoA-V). The SCoA-V inventory has been shown to have four domains – assessment improves learning, assessment makes students accountable, assessment is enjoyable, and assessment is irrelevant (Brown et al 2009). The inventory consists of 33 statements about the purposes of assessment, and 12 assessment practices. The students responded to the 33 statements on a six-point positively packed agreement scale, which has two negative responses (Strongly Disagree, and Mostly Disagree), and four positive responses (Slightly Agree, Moderately Agree, Mostly Agree, and Strongly Agree) (see Brown, 2004; Klockars & Yamagishi, 1988; Lam & Klockars, 1982 for a discussion of this type of scale). For the 12 assessment practices (such as traditional testing, grading of homework, alternative assessment practices and peer assessment),

the students endorsed those that "come to mind when you think of the word assessment".

Student Conceptions of Feedback Questionnaire -III (SCoF-III). A revised Student Conceptions of Feedback inventory (SCoF-III) was administered. The SCoF-III inventory contained 42 statements and 16 practices associated with feedback. These statements and practices have been identified by students as being associated with feedback (Peterson and Irving, 2007). Students rated the statements using the same rating scale described above, and also endorsed "the kinds and types of feedback that come to mind".

Academic Performance Measure: Mathematics. Academic performance was measured using a mathematics assessment generated by the Assessment Tools for Teaching and Learning (asTTle) test system (Hattie et al., 2004). asTTle is a computer-assisted, school-based assessment system for measuring learning in reading, writing, and mathematics,[1] and provides teachers, students and parents with standardised information about individual and group performance against national norms and national curriculum standards. Schools may choose whether to assess using the asTTle programme, and national reporting of results is not mandated (Brown and Irving, 2008). Schools can create their own assessment from a large item bank in each of the test domains, and receive a variety of reports on individual and group performance. The asTTle programme uses an item response theory scoring procedure (Embretson & Reise, 2000) so that, irrespective of which test is sat by the students, their total scores can be compared across classes, years, and schools. In this study, the schools administered a reading and/or a mathematics assessment of their own design, and provided at least one academic performance score for each student who completed the two conceptions inventories.

Participants

The participants were drawn from three secondary schools in the greater Auckland region. These schools were a small, convenience sample of all the schools in Auckland (N=75). In two of the schools, the full Year 9 (equivalent to about Grade 8) cohort participated, while the third school provided almost equal number of Year 9

[1] For further details about the software see www.asttle.org.nz

and Year 10 (equivalent to about Grade 9) students. Schools administered the two inventories, and the students responses were captured in Excel; then, the data were transferred to SPSS and AMOS for further analysis. Schools also provided the asTTle mathematics scores of the participating students for analysis.

A total of 721 completed inventories were received. The responses of participants were deleted if they had not completed at least 90% of the responses (*n*=45, 6.2%) (Kline, 1994), or if the school did not have asTTle mathematics scores for the student (n=177, 24.5%)[2]. In all, 222 students (30.8% of all participants) were deleted, which left a total of 499 cases for analysis. As full data was required for analysis, where a student had less than 10% missing data, missing value analysis using the expectation-maximisation procedure (Dempster, Laird, & Rubin, 1977) was employed to impute any remaining missing values.

A total of 499 students provided data for all three measures. Sex distribution was nearly equal (female = 259; male =239; 1 not specified). The mean age was 13.6 (SD = .59). The vast majority were in Year 9, the first year of secondary school (N = 441, 88.4%) and the balance were in Year 10. Just over a half (54.7%) of the participants were New Zealand European/Pakeha (N = 273), 40 (8.0%) were Māori, 31 (6.2%) were Pasifika, 60 (12.0%) were Asian and the remainder were classified as other, or not given (N = 93; 18.6%)[3]. The largest number of participants from any one school was 203 (40.7%), and the smallest number 110 (22.0%). The remaining 186 participants (37.3%) came from the third school.

Data analysis

Since existing inventories with reported data structures were used, a confirmatory approach was taken to ascertaining the validity of the pre-existing models for use in this study. The existing measurement models were tested with confirmatory factor analysis using AMOS version 7 (Arbuckle, 2006). Multivariate regression analyses were used to determine the most likely predictive relationship of SCoA factor to SCoF factors. The fit of the proposed paths between inventories and towards

[2] Reading scores were provided for these students, but these are not used in this study.
[3] The ethnicity distribution does not reflect the composition of students in New Zealand high schools. However, not attempt is made to analyse the data using any of the demographic details. These are provided for completeness.

mathematics performance were tested in structural equation models (again with AMOS v7). The proportion of variance explained in mathematics performance was determined by examination of the squared multiple correlations.

Identifying the most suitable measures of fit for measurement and structural models is an area of considerable contention (Barrett, 2007). However, there is consensus that multiple measures of goodness-of-fit (e.g., χ^2, gamma hat) and badness-of-fit (e.g., SRMR, RMSEA) should be reported (Fan & Sivo, 2007; Hu & Bentler, 1999). Increasingly there is a preference for reporting model fit statistics that are more robust against differences in model complexity, sample size and model misspecification. For example, the standardised root mean square residual (SRMR), the root mean square error of approximation (RMSEA), and gamma hat are argued to be more resistant to the impact of large samples, complex models, and model misspecification (Fan & Sivo, 2007; Hu & Bentler, 1999; Vandenberg & Lance, 2000). In contrast, the $\chi 2$ statistic is sensitive to large sample sizes and the Tucker-Lewis Index and comparative fit index are sensitive to models that have more than three factors or a hierarchical structure (Cheung & Rensvold, 2002). Choosing the most suitable standard for each index is also not straightforward. Hu and Bentler (1999) suggest that gamma hat should be around (or greater) than .95, SRMR should be around or less than .08, and RMSEA should be around or less than .06. In contrast, Marsh, Hau, and Wen (2004) proposed that indices such as statistically non-significant χ^2/df (α>.05) and goodness-of-fit values >.90 be used. Steiger (2000) has suggested RMSEA values as high as .10 point to theoretically interesting models. Therefore, it seems appropriate to interpret models with statistically non-significant χ^2 per *df*, gamma hat >.90, and RMSEA and SRMR <.08 as not being rejected.

Results

Testing the COA and COF measurement models

In keeping with Marsh and Hocevar (1985), we tested our measurement models of the SCoA V and SCoF-III first before identifying a structural equation model of how conceptions of assessment and feedback related to each other and then to academic performance in Mathematics.

SCoA-V. A four 2nd-order hierarchical factor model derived from Brown et al. (2009) was tested. In order to avoid negative error variances, the sub-factors ignore, bad, and personal future were removed and items were directly predicted by the appropriate higher-order factor. The model (Figure 1) has acceptable fit ($\chi^2_{484, 499}$=1460.45, p<.001; χ^2/df=3.02, p=.08; gamma hat=.89; RMSEA=.064, SRMR=.072).

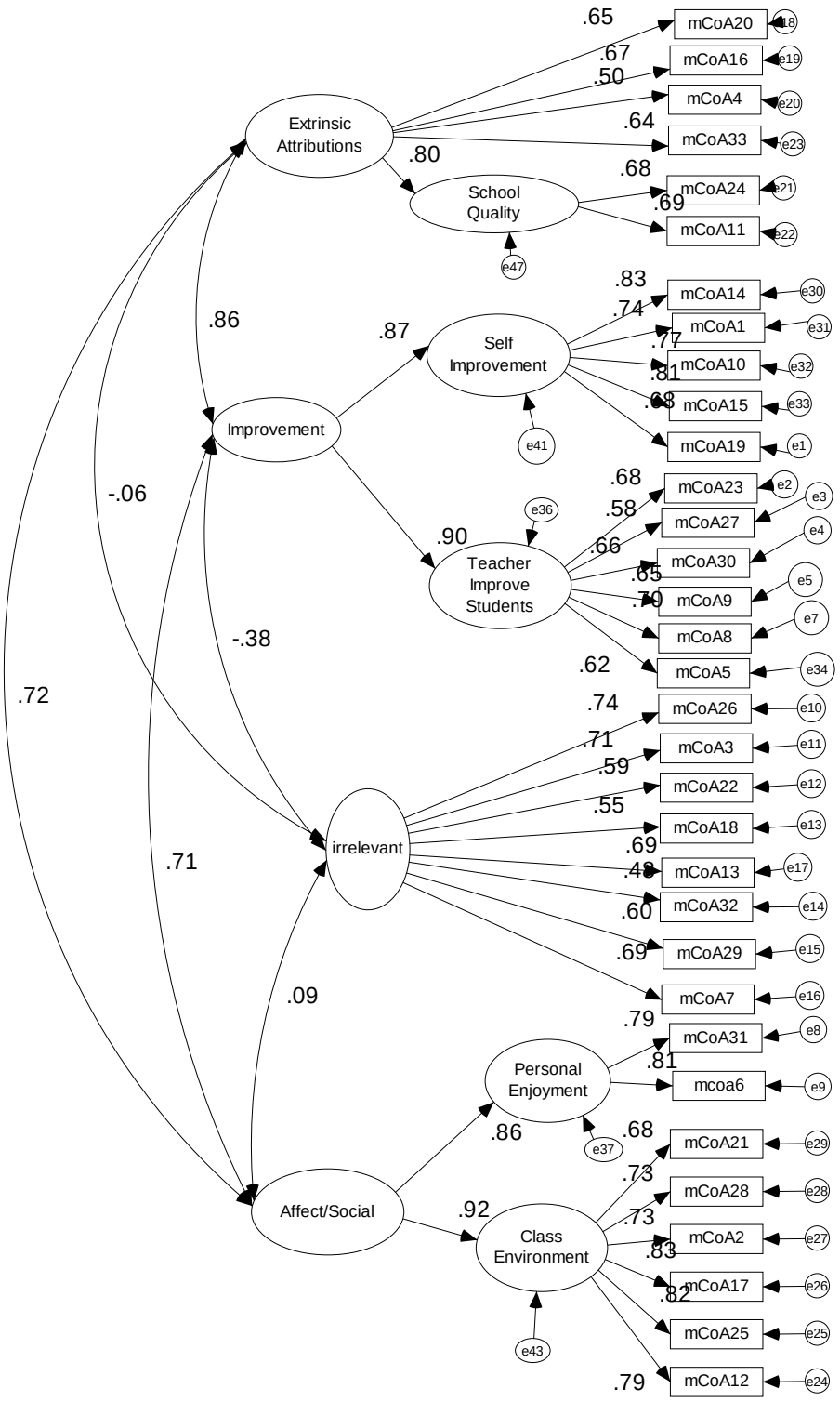

Figure 1. Simplified Students' Conceptions of Assessment Measurement Model

SCoF-III. A simple seven inter-correlated factors was tested. The model (Figure 2) has acceptable fit ($\chi^2_{475, 499}$=1166.29, p<.001; χ^2/df=2.46, p=.12; gamma hat=.92; RMSEA=.054, SRMR=.054).

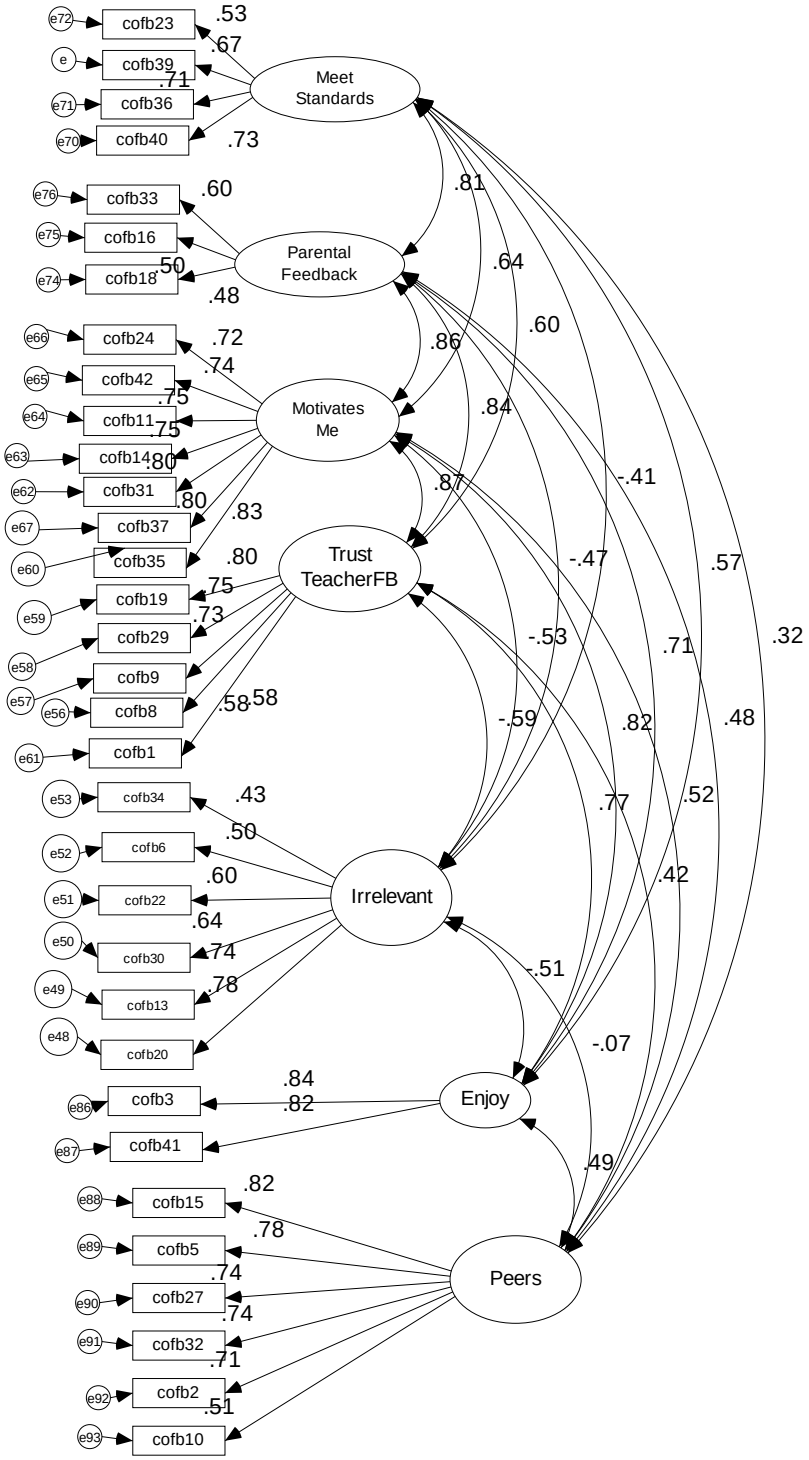

Figure 2. Simplified Students' Conceptions of Feedback Measurement Model

Testing the structural model COA and COF and mathematics achievement

Our next analysis tested the relationships between the SCoA-V and the SCoF-III and student asTTle mathematics achievement. In the first instance we had the four conceptions of assessment regressed onto their respective seven conceptions of feedback as per our hypotheses. In order to overcome negative error variances, various sub-factors were removed (i.e., enjoy feedback and peer feedback); those items were given freely estimated paths to a common personal enjoyment and peers factors consisting of both feedback and assessment items. In a similar vein all the irrelevant assessment and feedback items were allowed to be predicted by one common latent trait of irrelevance. Similarly, the improvement oriented assessment and feedback factors were allowed to be predicted by one common latent trait, rather than have their own separate improvement factors. In contrast to our hypotheses 3 and 4, we found that the parental feedback was best predicted by the improvement metafactor; while the standards factor was best predicted by the external attributions factor. These paths are shown in Figure 3 and psychometric properties are outlined in Table 1.

Assessment and feedback shared four super factors (i.e., Improvement, Irrelevant, Affect/Social, and Extrinsic Attributions). The improvement conception of feedback and assessment consisted of five sub-factors (i.e., students use assessment to improve their learning, teachers use assessment to improve their learning, parental feedback improves learning, feedback exists to improve learning, and teacher feedback is trusted). The extrinsic attribution conceptions of assessment and feedback consisted of items drawn from three factors (i.e., assessment predicts my personal future, assessment measures school quality, and feedback shows if I have met expected standards). The irrelevance factor consists of all the items having to do with assessment being bad and with ignoring assessment and feedback. The affect/social factor consisted of the all the items related to peers or classmates and personal enjoyment drawn from both assessment and feedback.

The students agreed equally with the improvement and extrinsic attribution factors (M=4.24, 4.38 respectively), agreed weakly with the affect/social factor (M=3.40), and rejected the irrelevance factor (M=2.67). The inter-correlations among three of the super factors was reasonably robust (range of r .75 to .88) indicating that

Improvement, Extrinsic Attributions, and Affect/Social were positively related even though they logically represented quite different facets of assessment and feedback. Unsurprisingly, the Irrelevance conception was negatively correlated with the other three super factors (*r* range -.22 to -.48).

Table 1. Descriptive Statistics for Four Super Factors of Assessment and Feedback Conceptions

Factor	# variables	Cronbach's alpha[1]	*M* (*SD*)[2]	Inter-correlations I[3]	II[3]	III[3]	β to Academic performance[3]
I. Improvement	11 SCoA + 15 SCoFB	.94	4.28 (.87)				.30
II. Extrinsic Attributions	6 SCoA + 4 SCoFB	.82	4.34 (.87)	.82			-.17
III. Affect-Social	8 SCoA + 8 SCoFB	.90	3.40 (.94)	.88	.75		-.32
IV. Irrelevant	8 SCoA + 6 ScoFB	.88	2.67 (.90)	-.48	-.22	-.24	-.26

Note.[1] *n*=600; [2] *n*=666; [3] *n*=499

These four metafactors were then regressed onto mathematics performance in a structural model. The model had acceptable fit ($\chi^2_{2126, 499}$=6012.18, p<.001; χ^2/df=2.83, p=.09; gamma hat=.81; RMSEA=.061, SRMR=.079). The model explains 13% of the variance in mathematics scores.

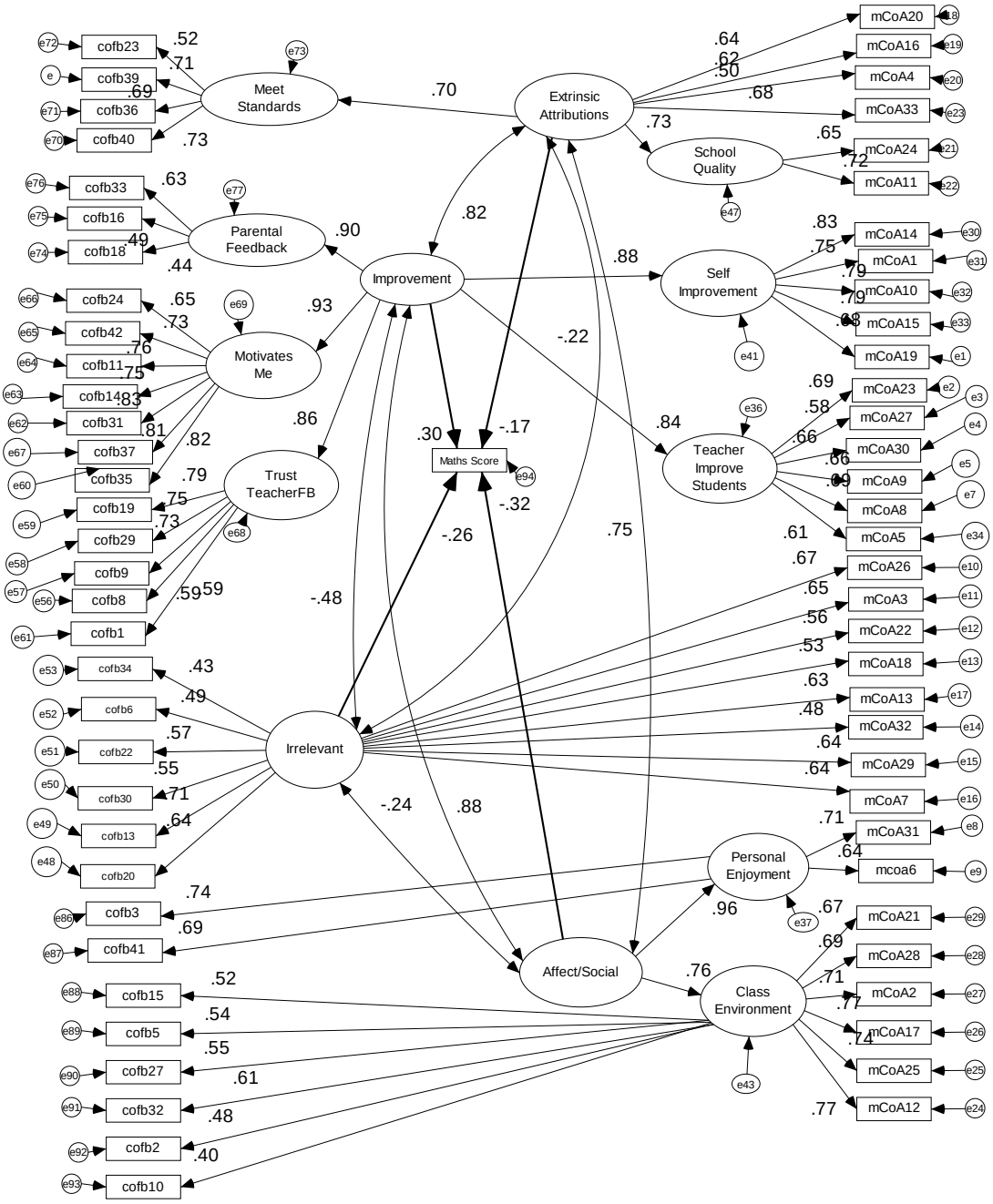

Figure 3. Structural model of student conceptions to mathematics performance

Of our four hypotheses concerning adaptive relations to academic performance, three were supported. Improvement was adaptive (β=.30), while extrinsic attributions (β=-.17) and irrelevance (β=-.26) were maladaptive. Unlike Brown et al. (in press) the social/affect priorities in assessment and feedback were not neutral relative to academic performance; instead they were maladaptive (β=-.32). This gives further evidence to the possibility that students who emphasise well-being over growth suffer

in their academic performance. It is worth noting that only the path from the irrelevance factor to academic performance was statistically significant at α=.05.

Discussion

In terms of our starting hypotheses, this study has provided corroborating evidence for three of the hypotheses concerning a strong relationship between the common factors in the assessment and feedback inventories. That is, the two irrelevance factors were linked; the two sets of affective and peer factors were linked; and the improvement oriented factors were linked. In contrast, the external attributions assessment factor did not link to the parental feedback factor (this was instead linked to improvement) and was linked instead to feedback being linked to standards. This last result strongly suggests students see parental feedback as having an improvement effect, while doing assessment for parents instead has a maladaptive external attribution effect on performance. Likewise, linking feedback to qualifications, which are in the future, also taps into an external maladaptive attribution.

Based on our self-regulating interpretive framework, we found support for three of our hypotheses. Improvement was adaptive and irrelevance and external attributions were maladaptive. Students who pay attention to assessment results in order to improve their learning and take responsibility for their learning achieved higher mathematics grades. This is in keeping with the self-regulation models of both Zimmerman (2008) and Boekaerts and colleagues (Boekaerts & Cascallar, 2006; Boekaerts & Corno, 2005) in which student beliefs and actions which involve taking responsibility for learning outcomes leads to greater academic performance. As might be expected, students who reject assessment and feedback as irrelevant or something that they can safely ignore, have lower mathematics scores as a consequence.

In contrast the affective, social factor was not neutral as we had anticipated—instead, this pathway was found to be maladaptive. This key aspect of the assessment *for* learning reform literature was not endorsed by these findings. Where students believed that assessment and feedback had a positive effect on the classroom environment, and on their personal enjoyment, the results were maladaptive. There is a clear indication that emphasis on the well-being pathway is not just neutral towards academic outcomes but is rather inimical to them. It would seem that students are

more instrumental about the role of assessment and feedback in their learning, and want to be assessed and given feedback that provides them with information that will improve their academic results, not just make them feel good (Pajares & Graham, 1998). There is strong reason to believe that practices which shield students from the awkward information of academic failure in order to preserve their well-being are not wanted by students and, when sought out leads to depressed academic results.

Our findings indicated that conceptions of assessment and conceptions are linked in ways which impact on academic performance in mathematics and together these conceptions explain 13% of the variance in mathematics achievement.

Teachers would be well advised to continue to emphasise the improvement purposes of assessment and feedback, and to encourage their students to view assessment and feedback in this way. On the evidence of this study, students receptive to this message could expect to see their mathematics scores increase. While it may be easy for teachers to students with such negative conceptions of assessment and feedback, the challenge is how to confront these maladaptive conceptions and turn them into more positive ones.

Both teachers and students need to understand that assessment might not make students feel good, and that it is better to know the truth so as to lead to improvement rather than avoid this information. Students who see improvement as a vital part of their beliefs about assessment and feedback, are less likely to accept positive affective statements about their performance that contain little or no information directly related to the learning/assessment task, and less likely to view assessment and feedback as something of little relevance to them.

Items relating to the more distal purposes of assessment and feedback such as (future job opportunities and the quality of the school, or feedback that communicates standards) negatively predicted mathematics achievement (β=.-17). This is in keeping with the motivational literature which suggests that events that are perceived as having only a distant relationship to the immediate performance have weaker impact than the more proximal factors. Furthermore, the extrinsic nature of these beliefs, even though the students agree with them, contributes to maladaptive outcomes. Students have little control over their schools, standards, or their futures. We suggest

that the lack of control probably feeds into a fatalistic attitude which converts to inappropriate study and learning approaches and thus generates maladaptive consequences for learning outcomes.

Perhaps the most controversial finding was that the socio-emotional nature of feedback (i.e., enjoying feedback, and receiving feedback from your peers) and of assessment (i.e., enjoying assessment, and believing that my class benefits from assessment') negatively predicted mathematics (β=.-26). This suggests that NZ secondary students deem the affective and beneficial purposes of feedback and assessment as having a negative impact on performance. From the student point of view, the reform emphases appear to be wrong. While we can only speculate as to the root causes of this maladaptive pathway (see discussion in Brown et al. 2009), the persistence of either a neutral or negative pathway from emotive and social responses to assessment needs to be taken seriously by assessment reform advocates. As Cowie (in press) and Segers et al. (in press) have shown assessment innovations have mixed and complex impacts on students' emotions and social relations. Hence, whatever is currently happening is clearly not intended and further revision either to the assessment for learning rhetoric or practices is needed.

An additional implication of this study is that the two inventories could be merged to provide a new, more complete single inventory of assessment and feedback, and the beliefs that students have about them.

Overall this study demonstrates that students' conceptions of assessment and feedback are meaningfully related to each other and to mathematics performance. The results raise serious doubts as to the effectiveness of prioritising the socio-emotional aspects of assessment and feedback and draw attention to the positive effect of a systematic commitment to improvement on the part of the teachers/school and the students themselves.

References

Arbuckle, J. L. (2006). AMOS (Version 7.0.0). Spring House, PA: Amos Development Corporation.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42*, 815–824.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Boekaerts, M., & Cascallar, E. (2006). How far have we moved towards the integration of theory and practice in self regulation? *Educational Psychology Review, 18*(3), 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.

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

Brown, G. T. L. (2006). Students' conceptions of assessment (SCoA) inventory (Version 5). Unpublished test. University of Auckland.

Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York: Nova Science Publishers.

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

Brown, G. T. L., Irving, S. E., & Peterson, E. R. (2008, July). *Beliefs that make a difference: Students' conceptions of assessment and academic performance.*

Paper presented at the Biannual Conference of the International Test Commission, Liverpool, UK.

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

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

Cowie, B. (in press). My teacher and my friends helped me learn: Student perceptions and experiences of classroom assessment. 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.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B, 39*, 1-38.

Elen, J., & Lowyck, J. (1999). Metacognitive instructional knowledge: Cognitive mediation and instructional design. *Journal of Structural Learning and Intelligent Systems, 13*, 145-169.

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

Fransson, A. (1977). On qualitative differences in learning. IV - Effects of intrinsic motivation and extrinsic test anxiety on process and outcome. *British Journal of Educational Psychology, 47*, 244-257.

Harlen, W. (2007). *Assessment of learning*. Los Angeles: Sage.

Harris, L. R., Harnett, J., & Brown, G. T. L. (2009, accepted). Assessment from students' perspectives: Using pupil drawings to examine their 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 assessment for learning*. Charlotte, NC: Information Age Publishing.

Hattie, J. A. C., 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) Manual* (Version 4, 2005 ed.). Wellington, NZ: University of Auckland / Ministry of Education / Learning Media.

Hornby, W. (2003). Assessing using grade-related criteria: a single currency for universities? *Assessment and Evaluation in Higher Education, 28*, 435-454.

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

Irving, S. E., Peterson, E. R., & Brown, G. T. L (2008, July). *Feedback and Academic Achievement: The Relationship between Students' Conceptions of Feedback and Achievement.* Paper presented at the 6th Biennial Conference of the International Test Commission, Liverpool, UK.

Irving, S. E., Peterson, E. R., & Brown, G. T. L. (2007, August). *Student Conceptions of Feedback: A study of New Zealand secondary students*. Paper presented to the

Biennial Conference of the European Association for Research in Learning and Instruction (EARLI), Budapest, Hungary.

Irving, S. E., Peterson, E. R., & Brown, G. T. L. (2008). Conceptions of feedback (CoF) inventory (Version 3). Unpublished test. University of Auckland.

Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.

Klockars, A. J., & Yamagishi, M. (1988). The influence of labels and positions in rating scales. *Journal of Educational Measurement, 25*(2), 85-96.

Kulhavy, R. W., Stock, W. A., Hancock, T. E., Swindell, L. K., & Hammrich, P. L. (1990). Written feedback: Response certitude and durability. *Contemporary Educational Psychology 15*(4), 319-332.

Lam, T. C. M., & Klockars, A. J. (1982). Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement, 19*(4), 317-322.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin, 97*(3), 562-582.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320-341.

Pajares, M. F., & Graham, L. (1998). Formalist thinking and language arts instruction: Teachers' and students' beliefs about truth and caring in the teaching conversation. *Teaching & Teacher Education, 14*(8), 855-870.

Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*(3), 238-250.

Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (6th ed.). Boston: Allyn & Bacon.

Remesal, A. (in press). Accessing primary pupils' conceptions of daily classroom assessment practices. 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.

Segers, M., Dochy, F., Gijbels, D., & Struyven, K. (in press). Changing insights in the domain of assessment in higher education: Novel assessments and their pre-, post- and pure effects on student learning. 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.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Steiger, J. H. (2000). Point estimation, hypothesis testing, and interval estimation using the RMSEA: Some comments and a reply to Hayduk and Glaser. *Structural Equation Modeling, 7*(2), 149-162.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(4), 4-70.

Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment: What's in it for schools?* London: RoutledgeFalmer.

Weekers, A. M., Brown, G. T. L., & Veldkamp, B. P. (in press). Analyzing the dimensionality of the Students' Conceptions of Assessment inventory. 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.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166-183.