

## Manuscript Title

The Effect of Conceptions of Assessment upon Reading Achievement: An Evaluation of the Influence of Self-efficacy and Interest

## Author(s)

Gavin T.L. Brown and K. Frank Walton, School of Learning, Development & Professional Practice, Faculty of Education & Social Work, The University of Auckland, New Zealand

## Summary

*The manuscript was received on June 01, 2017 and was peer reviewed by two reviewers and an editor.*

*The initial recommendation of Major Revision was made on June 29, 2017.*

*The first revision was submitted on July 08, 2017 and was re-evaluated by the editor.*

*The manuscript was accepted for publication on July 12, 2017.*

## Consolidated Peer Review Comments

The purpose of this manuscript was to examine the relationship between conceptions of assessment and reading achievement. Moreover, the relationship was examined across levels of self-efficacy and interest for reading. Overall, the results, indicate that the relationship between conceptions of assessment and achievement were weak and that there were no significant differences across self-efficacy and interest levels. The concept of this paper is interesting, however, there are several major concerns / suggestions. Particularly, the reviewer has major concerns regarding the methods for measuring self-efficacy, the use of two forms of a measurement, large variations in age across the sample, dated data, large proportions of missing data, and poor writing quality in the methods section.

First, the methods for assessing self-efficacy were concerning. Your measurement did not address the construct of self-efficacy. Two of the items regarding an individual's perceptions of others' perceptions is not self-efficacy. It may be more appropriate to use the single self-efficacy item than to use the aggregate of these three items.

Second, the authors' description of the use of two forms of assessment was confusing. It was difficult to follow and to determine whether these measures were in fact equivalent. It was also difficult to determine whether all items were administered or if the authors selected 11 items that matched across the two forms.

Third, the reviewer was surprised that the authors did not statistically examine or control for age in the model. There were some relatively large variations in age within your sample ranging from 13 to 17. It seems possible that there could be differences across these samples so the authors should have addressed whether it is appropriate to collapse these samples or whether some statistical control for age is needed in the model.

Fourth, the methods section in general was poorly written. It was difficult at times to understand what the authors actually did in the current study. The writing in this section need to be improved significantly so that a reader could better understand what actually occurred.

Fifth, the reviewer felt that 13-year-old data is fairly dated. The authors provided some justification that it was perceived that reading achievement had not changed over the course of this time frame but no data or evidence was cited regarding the stability of motivational variables.

It appears that approximately 20% of your data was missing. Why is such a large proportion of data missing following data recovery procedures? This seems problematic. Was data missing completely at random?

There is a need for a more substantial limitations section to address several limitations that may not be able to be remediated within this paper.

There were also a variety of minor concerns / suggestions that the reviewer noted. In particular, the abstract uses several statements that are somewhat vague given that reader may not actually know what "conceptions of assessment" means prior to reading the manuscript. I suggest using terminology that more specifically states what you examined. For example, "students' beliefs about the importance, usefulness, purpose of assessment"). Throughout the manuscript the authors write the name for some conceptions but the flow reads a little awkward. Rather than calling it "useless" why not saying "when students perceive assessment as useless they .....". The citation of other work is a bit too thin at times in the manuscript. When describing the literature it seemed somewhat awkward to say "when students agreed more that: (a) assessment makes students accountable." The "agreed more" part did not read well. What about "when students perceive that assessment makes them accountable....." Please define situational interest on page four given that some readers will not understand what that means. Identifying self-regulated learning (SRL) and using the acronym SRL throughout the paper might help reduce wordiness at times. The discussion section contained a few interpretations that seemed too bold, such as suggesting that conceptions of assessment may be more worthy of attention in an academic environment when compared to self-efficacy and interest. The vast literature supporting self-efficacy and interest shouldn't be shrugged off by your results, especially considering that the self-efficacy measurement was problematic. Finally, there is a need to clean up the writing throughout the manuscript. There were some grammatical errors, typos, or instances when the flow of writing could be improved. This was not an overwhelming problem, but would need some improvement before the paper is acceptable for publication.