

Modeling Student Motivation and Students' Ability Estimates From a Large-Scale Assessment of Mathematics

Carlos Zepa, Krystal Hachey, Christina van Barneveld and Marielle Simon

SAGE Open 2011 1: originally published online 15 September 2011

DOI: 10.1177/2158244011421803

The online version of this article can be found at:
<http://sgo.sagepub.com/content/1/2/2158244011421803>

Published by:



<http://www.sagepublications.com>



Additional services and information for *SAGE Open* can be found at:

Email Alerts: <http://sgo.sagepub.com/cgi/alerts>

Subscriptions: <http://sgo.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Modeling Student Motivation and Students' Ability Estimates From a Large-Scale Assessment of Mathematics

SAGE Open
1-9
© The Author(s) 2011
DOI: 10.1177/2158244011421803
<http://sgo.sagepub.com>


Carlos Zerpa¹, Krystal Hachey², Christina van Barneveld¹,
and Marielle Simon²

Abstract

When large-scale assessments (LSA) do not hold personal stakes for students, students may not put forth their best effort. Low-effort examinee behaviors (e.g., guessing, omitting items) result in an underestimate of examinee abilities, which is a concern when using results of LSA to inform educational policy and planning. The purpose of this study was to explore the relationship between examinee motivation as defined by expectancy-value theory, student effort, and examinee mathematics abilities. A principal components analysis was used to examine the data from Grade 9 students ($n = 43,562$) who responded to a self-report questionnaire on their attitudes and practices related to mathematics. The results suggested a two-component model where the components were interpreted as task-values in mathematics and student effort. Next, a hierarchical linear model was implemented to examine the relationship between examinee component scores and their estimated ability on a LSA. The results of this study provide evidence that motivation, as defined by the expectancy-value theory and student effort, partially explains student ability estimates and may have implications in the information that get transferred to testing organizations, school boards, and teachers while assessing students' Grade 9 mathematics learning.

Keywords

achievement, educational testing, large-scale assessments, motivation, self-report data

When test scores from large-scale assessments (LSA) are reported and interpreted for individual students, classes, schools, states, or nations, there is often an implicit assumption that the scores represent the best effort of the student (Wolf & Smith, 1995). Researchers in the field of educational measurement have questioned this assumption by stating that if the test score is not consequential or important to the student, then one cannot be sure how much the observed score is influenced by the lack of effort (DeMars, 2000; Wolf & Smith, 1995). This leads to the argument that test consequences influence motivation, and motivation influences students' effort, test performance, and estimates of academic achievement (DeMars, 2000; Sundre & Moore, 2002; Wolf, Smith, & Birnbaum, 1995).

Harlen and Crick (2003) have suggested that motivation is related to effort and the assessment learning context. Other researchers have proposed that motivation is closely aligned with the will to learn. Motivation encompasses self-esteem, self-efficacy, values, and students' perception of their abilities to accomplish a particular task. All these components of motivation affect effort and ultimately achievement (Brookhart & Durkin, 2003; Eccles & Wigfield, 2002; Salomon, 1983, 1984; Scheifele, 1991; Wigfield & Eccles, 2000).

In Ontario, Canada, LSA of mathematics do not explain the effect of students' motivation and effort on their estimates of academic achievement. This is a concern because the results from these LSA are used to provide accountability in the educational system for the allocation of funds, justify changes to the mathematics curriculum, inform parents about their children's progress, and help children adapt to changes in today's world (Education Quality Accountability Office [EQAO], 2011). This is one of the reasons why the current study focused on examining student motivation and student effort in relation to their estimates of academic achievement by using Grade 9 LSA of mathematics based on self-report data and test scores. The mathematics self-report questionnaires used in the current study to measure student motivation were administered by the EQAO in 2007. Because EQAO self-report questionnaires were not designed as a measure of student motivation, we decided to use the research

¹University of Lakehead, Thunder Bay, Ontario, Canada

²University of Ottawa, Ontario, Canada

Corresponding Author:

Carlos Zerpa, University of Lakehead, Faculty of Education, 955 Oliver Road, Thunder Bay, Ontario P7B 5E1, Canada
Email: czerpa@lakeheadu.ca

work done by Wigfield and Cambria (2010) to guide us in identifying motivation items from EQAO self-report questionnaires related to students' values and effort. Wigfield and Cambria's research provided us with an extensive review of students' motivation constructs related to achievement values, goal orientations, and interest that could be used in educational research to measure student motivation. For the current study, we used these motivation constructs (achievement values, goal orientations, and interest) and related them to EQAO self-report items to address the research question. For example, EQAO self-report items such as "I like math" was classified as intrinsic or interest value because it related to students' enjoyment from doing mathematics; "Math is boring" was classified as attainment value because it is related to the importance that the student placed on the mathematics tasks; "The math I learn now is very useful for everyday life" and "I need to keep taking math for the kind of job I want after school" were classified as utility values because they reflected the importance of mathematics for student future plans; "I am good in math" and "Mathematics is an easy subject" were classified as achievement values because they related to student goal orientations (Wigfield & Cambria, 2010, p. 10). Finally, items such as "How often do you complete your math homework" and "How much time do you usually spend in math homework" were classified as student effort because these items related to students' well-developed interest to engage in mathematics tasks frequently (Wigfield & Cambria, 2010, p. 11). The outcome of the current study sheds light and provides an avenue for researchers, teachers, and educational agencies to better explain the impact of motivation on the estimates of students' academic achievement, while using EQAO's LSA of Grade 9 mathematics.

Context

In the context of education, when researchers address students' motivation, they focus on the theory of motivation related to students' beliefs, values, and goals to better assess students' academic performance and achievement (Bishop, Clark, Corrigan, & Gunstone, 2006; Eklof, 2006; Stipek, Givven, Salmon, & MacGyvers, 1998; Sundre & Moore, 2002; Wolf et al., 1995; Wolf & Smith, 1995). As Eccles and Wigfield (2002) stated, these constructs (beliefs, values, and goals) are the most immediate and direct predictors of academic achievement, performance, and choice, and are themselves influenced by a variety of psychological, social, and cultural determinants. One theory that encapsulates these constructs is the expectancy-value theory of motivation (Atkinson, 1964; Cole, Bergin, & Whittaker, 2008; Eccles & Wigfield, 2002; Putwain, 2008). Expectancy-value theory links achievement performance, persistence, and choice directly to individuals' expectancy-related and task-value beliefs. Expectancy-related beliefs refer to individuals' beliefs about how well they will do on an

upcoming task, either in the immediate or upcoming future (Eccles & Wigfield, 2002). Task-value beliefs are defined by four components: (a) attainment value—the personal importance of doing well on a task, (b) intrinsic value—the enjoyment the individual gets from performing the task, (c) utility value—how well the task relates to current and future goals, such as career goals, and (d) cost—negative aspects of engaging in the task, such as fear of failure (Eccles & Wigfield, 2002).

As defined by expectancy-value theory, students' motivation in relation to academic achievement depends on students' general-ability beliefs and task-value beliefs (Eccles & Wigfield, 2002; Eklof, 2006; McMillan, Simonetta, & Singh, 1994). Applied to LSA, general-ability beliefs relate to a student expectancy-related belief about his or her ability to be successful on LSA. Task-value beliefs relate to the importance the student places on a successful performance on LSA. In effect, if the student values the outcome of the LSA, then there are more chances that the student will be motivated, make an effort on tasks, and engage with the tasks to the best of his or her ability (Dweck & Elliot, 1983; Eccles & Wigfield, 2002; Ryan, Ryan, Arbuthnot, & Samuels, 2007; Wigfield & Eccles, 2000).

Conditions of Testing and Low-Effort Student Behaviors on LSA

During the administration of LSA, students' low motivation and the conditions of testing may influence their effort in responding to mathematical test items (DeMars, 2000; Putwain, 2008; Wolf et al., 1995). For instance, students' low motivation as a result of not valuing the outcome of the test may trigger certain low-effort test-taking behaviors such as guessing, omitting items, or quitting entirely on the large-scale examination, and these behaviors may cause an underestimation of students' abilities (De Ayala, Plake, & Impara, 2001; Meijer, 1996; Meijer & Sijtsma, 1995, 2001). The conditions of the test such as item difficulty, mental taxation, and item position may also affect students' motivation and effort, and may affect their ability estimates, especially in situations when the students do not value the outcome of the LSA (Wolf & Smith, 1995).

Wolf et al. (1995) studied students' motivation and effort during a large-scale examination by examining item-by-item differences in performance between two groups of students taking the same test under different conditions. Participants included 168 Grade 10 and 133 Grade 11 students from the same high school. The researchers found that the conditions of testing influenced test performance, and this influence varied for different kinds of items. For instance, if there were no consequences linked to assessment results, items related to nonconsequential conditions appeared unnaturally difficult because they did not motivate and capture the complete effort of students. The researchers concluded that this lack of student motivation and effort due to the conditions of testing

poses a threat to the validity of interpretation of test results when assessing students' test performance.

Impact of Low-Motivation Behaviors on the Validity of Interpretation of Test Scores

Low motivation and the conditions of testing may affect students' effort and the estimates of students' abilities during LSA (DeMars, 2000; Kane, 2006; Putwain, 2008; Wolf et al., 1995) and may also affect the validity of the interpretation of test scores when assessing student academic performance and achievement (Kane, 2006; Messick, 1989). For instance, Schmitt, Chan, Sacco, McFarland, and Jennings (1999) found that low-motivation behaviors (e.g., guessing, omitting items, or quitting entirely on a test) affect the validity of interpretation of test results during LSA and can either artificially inflate or deflate estimates of students' abilities. The researchers stated that inaccurate estimates of students' abilities negatively affect individuals and test organizations when assessing student academic performance and achievement. For example, an inflated estimate of ability, as a result of guessing the correct answer (an examinee of low ability guesses the correct answer on medium difficulty items and on more difficult items), may cause educational authorities to think that students are able to perform at the expected level. On the contrary, a deflated estimate of ability, as a result of omitting items on a test, may deprive students of opportunities where they can be exposed to higher levels of knowledge. These are examples of concerns to highlight the effect of low-motivation behaviors on the validity of LSA data (Linn & Baker, 1996; Meijer, 1996; Meijer & Sijtsma, 2001; van Barneveld, 2007). As Kane (2006) and Messick (1989) stated, to validate a proposed interpretation or use of test scores, is to evaluate the rationale used for the interpretations of the test scores. Valid interpretations of test results such as those obtained from LSA may lead to a more rational argument of why certain changes need to be implemented in the educational system (Kane, 2006). These changes may include curriculum modifications and allocation of funding by administrators and policy makers.

Techniques Used to Measure Motivation

There are models and indices used to identify low motivation that affects students' performance on tests. These models and indices provide an avenue to better understand students' academic performance and make valid interpretations of LSA results (Fraire, Tideman, & Watts, 1997; Karabatsos, 2003; Meijer, 1996; Meijer & Sijtsma, 2001; Putwain, 2008; Sotaridona, Linden, & Meijer, 2006; Sotaridona & Meijer, 2003; L. Wise, 1996). One statistical modeling index that researchers have used quite extensively

to identify low motivation is the *lz* index statistic, which was developed by Drasgow, Levine, and Williams (1985). The *lz* is a standardized statistical estimate used to detect the percentage of low motivation test-taking behaviors manifested in the data for low- and high-stake examinations (Drasgow et al., 1985; Karabatsos, 2003; Meijer, 1996; Nering & Meijer, 1998; L. Wise, 1996; Wolf et al., 1995). The limitation of the *lz* statistical technique, however, is that it is only based on test scores and does not measure students' effort (DeMars, 2000). Based on this concern, researchers have developed statistical indices to measure low motivation by using response-time effort (S. Wise & Kong, 2005). Response-time effort is based on the hypothesis that when an item is administered, unmotivated students have the tendency to answer the item too quickly (S. Wise & Kong, 2005). This means that students lack performance effort in their responses on the test due to low motivation. Although the response-time effort technique seems to be more promising than the *lz* statistical index in detecting low motivation because it takes into consideration student effort, the challenges with the response-time effort technique is that it requires the use of computer-based technology. LSA and instruction in the classroom, however, are usually conducted using pencil-and-paper methods (Camara, 2009; S. Wise & DeMars, 2006). This is one of the reasons why some researchers choose to use student self-reported questionnaire measures as another method to examine the relationship of motivation and effort with student estimates of academic achievement during LSA (Eklof, 2006; Pintrich, Smith, Garcia, & McKeachie, 1993; Wolf & Smith, 1995).

A number of self-report questionnaires have been developed, validated, and used to measure motivation. For instance, the Motivated Strategies Learning Questionnaire (MSLQ; Pintrich et al., 1993) and the Student Opinion Survey (Sundre & Kitsantas, 2004; Wolf & Smith, 1995) have been used by researchers to measure students' motivational beliefs and values ranging in age from late elementary to university. Marsh, Koller, Trautwein, Ludtke, and Baumert (2005) developed a learning survey to measure how much students look forward to learning mathematics, how important mathematics is to them, the importance of being a good mathematician and the enjoyment of learning mathematics by drawing on the expectancy-value theory of motivation. O'Neil, Abedi, Miyoshi, and Mastergeorge (2005) used an adaptation of the State Thinking Questionnaire (O'Neil, Sugrue, Abedi, Baker, & Golan, 1997) to measure motivation using a monetary incentive as a way to increase student effort and performance. Roderick and Engel (2001) used the Reynolds Adolescent Depression Scale (Reynolds, 1984) to cross check interview data of students' descriptions of their motivation.

One of the strengths of using self-reported questionnaires to measure motivation is that they can be easily implemented using a pencil-and-paper method as opposed to other measurement techniques that may require the use of computer-based technology (Camara, 2000). In addition, self-reported

questionnaires' variables and constructs can be grounded on the expectancy-value theory of motivation or other motivation theories (i.e., attribution theory, achievement goal theory, and self-efficacy) to assess students' motivation and effort in relation to their academic achievement (Eccles & Wigfield, 2002; Pintrich, 2004; Pintrich et al., 1993; Pintrich & Schunk, 1996; Wigfield & Cambria, 2010). One of the challenges, however, is to develop and use self-reported questionnaires that have a clear structure, high internal consistency, and a strong evidence of validity measures. Another challenge is to create a motivation self-report questionnaire that clearly addresses factors and variables as they relate to test effort and performance using a motivation theory (Cole et al., 2008; Eccles & Wigfield, 2002; Harlen & Crick, 2003; Putwain, 2007, 2008; Wigfield & Cambria, 2010).

In Ontario, Canada, Grade 9 EQAO self-reported questionnaires are used to obtain students' background information, which can be linked to their achievement, interest, values, effort, and goals in different mathematics strands such as numeracy, algebra, and geometry. For instance, how often do you complete all your mathematics homework? I like math and I am good in math are questions that can be related to student effort and task-value beliefs as defined by the expectancy-value theory of motivation (Brookhart, Walsh, & Zientarski, 2006; Kloosterman, 1996; Wigfield & Cambria, 2010).

The intention of the self-reported questionnaires and tests are to monitor how well students are meeting the expectations of the mathematics curriculum. The information obtained from these assessment tools is used to better inform schools, teachers, and parents about students' mathematics achievement in relation to a provincial standard (Volante, 2006). These provincial self-reported questionnaires and mathematics tests are administered each year by teachers in the schools and then returned to EQAO for marking and reporting. The first administration takes place in the winter semester and the second one in the spring.

A portion of these EQAO Grade 9 mathematics tests (0%-30%) may or may not count toward students' final grades. If students know that assessment results do not count, their task-values and effort may change as they may not place too much importance on a successful performance for the large-scale examination (DeMars, 2000; Eccles & Wigfield, 2002). These variations in test stakes as well as the results from EQAO tests not explaining if students' motivation affects their academic achievement led us to the development of the current study. We addressed a portion of the motivation problem in the current study by using students' self-reported questionnaires data from the EQAO Grade 9 LSA of mathematics. This approach allowed us to identify motivation components (task-values and effort) as defined by the expectancy-value theory and examine if these motivation components were significant predictors of students' academic achievement. The intention of this research was to provide an avenue for researchers, teachers, and the EQAO officials

to better explain the effect of student motivation on his or her academic performance during LSA of mathematics.

In summary, the research work presented in this article is mostly concerned with theoretical constructs, predictions, and relationships among motivational variables such as expectancy value, achievement goals, effort, and interest with students' academic achievement during LSA. The study also builds on work done by DeMars (2000), Maehr and Meyer (1997), and Wolf and Smith (1995). The question that guided this study was as follows:

1. To what extent does the expectancy-value theory of motivation and student effort relate to students' academic achievement on a large-scale assessment of mathematics?

Method

Instrument and Participants

EQAO Grade 9 assessment of mathematics, 2007, was used as the source of data. The EQAO data included students' self-reported questionnaires and test scores. The EQAO tests were administered twice during the year. The first administration took place in the winter semester and the second administration in the spring semester. For the current study, the results from students in the academic program (students who will be attending university) who wrote the test in the spring were used. We chose these results because they represented the largest sample of data from the EQAO student self-reported questionnaires and test scores. This provided us with a database sample of 43,562 students.

Procedure and Analysis

First, 11 items related to expectancy-value theory of motivation and student effort were selected from Grade 9 LSA of mathematics using EQAO student self-report questionnaires. This selection was based on Wigfield and Cambria's (2010) study as previously stated. See Formula 1 and 2 for a list of items. Once the items were selected from the student self-reported questionnaire data, a principal components analysis (PCA) for categorical data was conducted. From the PCA, the selected items from the student self-reported questionnaire data were reduced into two components, one component representing task-values and the other representing student effort (Wigfield & Cambria, 2010). After the motivation components were identified based on the literature and the PCA, component scores were computed for each case using the modeling Equations 1 and 2.

$$\begin{aligned} \text{Task Value} = & \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5 \\ & + \alpha_6 X_6 + \alpha_7 X_7 + \alpha_9 X_9 + \alpha_{10} X_{10} \\ & + \alpha_{11} X_{11}, \end{aligned} \quad (1)$$

Table 1. Results of the PCA

Variables	Components	
	Task-values	Effort
I like math	.83	-.09
I am good in math	.79	-.31
I understand most of the mathematics I am taught	.77	-.24
The mathematics I learn now is very useful for everyday life	.56	.18
I need to keep taking mathematics for the kind of job I want after I leave school	.50	.07
Mathematics is boring	-.61	-.18
Mathematics is an easy subject	.71	-.40
How much time do you usually spend on mathematics homework (in or out of school) on any given day?	.11	.66
How often do you complete all of your mathematics homework?	-.45	-.54
How often have you been absent from your Grade 9 mathematics class this year?	.22	.41
How often have you been late for your Grade 9 mathematics class this year?	.27	.50

Note: PCA = principal components analysis. Variables with a component loading equal or higher than .40 were considered to have high loadings.

$$\text{Effort} = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \alpha_{11} X_{11}, \quad (2)$$

where X_1 = I like Math, X_2 = I am good in math, X_3 = I understand most of the mathematics I am taught, X_4 = The mathematics I learn now is very useful for everyday life, X_5 = I need to keep taking mathematics for the kind of job I want after I leave school, X_6 = Mathematics is boring, X_7 = Mathematics is an easy subject, X_8 = How much time do you usually spend on mathematics homework (in or out of school) on any given day? X_9 = How often do you complete all of your mathematics homework? X_{10} = How often have you been absent from your Grade 9 mathematics class this year? X_{11} = How often have you been late for your Grade 9 mathematics class this year; $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9, \beta_{10}, \beta_{11}$ are the coefficients for the variables in the task-value component; $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8, \alpha_9, \alpha_{10}, \alpha_{11}$ are the coefficients for the variables in the expectancy-performance component.

Next, a two-level hierarchical linear model (HLM)—students nested in schools—was used to determine the significance of the relationship between students’ motivation and their mathematics achievement, as measured by the EQAO test. The first level of the HLM contained a fixed-model effect, which was used to determine how significant the extracted components (task-values and effort) were in relation to students’ academic achievement. The second level contained a random-model effect that was used to determine the impact of different schools on students’ academic achievement at random. The modeling Equations 3, 4, and 5 were used for the two-level HLMs.

Level 1 or fixed-model effect:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2, \quad (3)$$

Level 2 or random-model effect:

$$\alpha_0 = \beta_0 + \beta_1 z_1, \quad (4)$$

Combined:

$$y = \beta_0 + \beta_1 z_1 + \alpha_1 x_1 + \alpha_2 x_2 + r \quad (5)$$

where y = students’ academic achievement, x_1 = task-values, x_2 = expectancy performance, r = fixed-model effect residual variance, α_0 = mean student academic achievement for a given school, β_0 = intercept with represents the grand mean due to schools, β_1 = variance of the intercept due to schools, α_1, α_2 = coefficient for the fixed factors of Level 1.

Results

The results of the PCA suggested that the students’ self-reported EQAO questionnaire variables clustered into two components. These two components were interpreted by the research team as task-values and effort in mathematics based on Wigfield and Cambria’s (2010) study. See Table 1 for results of the PCA.

The results of the HLM first-level analysis (fixed analysis effect model) depicted in Table 2 suggested that the task-values and effort components were significant predictors of students’ academic achievement. Although significant in the

Table 2. HLM Fixed Analysis Effect Model

Source	Numerator <i>df</i>	Denominator <i>df</i>	<i>F</i>	<i>p</i>	<i>R</i> ² %
Intercept	1	1638.91	127661.10	.00	
Task-values	1	41491.13	8790.70	.00	17.90
Effort	1	41549.10	256.85	.00	

Note: HLM = hierarchical linear model. Dependent variable-academic achievement.

Table 3. HLM Random Analysis Effect Model

Parameter	Estimate	Std. Error %	Wald <i>Z</i>	<i>p</i>
Residual	.36	.25	142.14	.00
Intercept/variance subject = school	.08	.34	24.05	.00

Note: HLM = hierarchical linear model. Dependent variable-academic achievement.

first-level analysis of the HLM, these predictors only accounted for 17.90% of the variance, which left 82.10% of the variance unexplained in Level 1. The second level of the HLM analysis, which was the intraclass correlation coefficient computations showed that from the total variance accounted for at this level, there was 18.18% of between-school variance and 81.81% of within-school variance that affected students' academic achievement as shown in Table 3. In addition, the total coefficient of determination ($R^2 = 34.69\%$) between the observed and predicted students' achievement data computed for the entire HLM model suggested that the motivation components (task-values, student effort, and score variability within and between schools) accounted for 34.69% of the variance in relation to students' academic achievement. This means that 65.31% of the variance was unaccounted by the HLM model and might be related to other factors besides motivation and student score variability within and between schools that affected students' academic achievement.

Discussion

The results of the PCA using students' self-reported EQAO questionnaire data suggest a two-component model, which was interpreted as task-values and effort using the research framework from Wigfield and Cambria (2010). The results of the PCA support the research literature in that it is possible to identify motivation components related to student task-values as defined by the expectancy-value theory and student effort, using students' self-reported data (Cole et al., 2008; Eccles & Wigfield, 2002; Putwain, 2008).

The results of the first level (fixed-model effect) of the HLM statistical analysis conducted on the Grade 9 students' self-reported questionnaire EQAO data suggest that students' task-values and effort are significant predictors of their academic achievement on the EQAO test. These findings may provide relevant information for teachers and educational

agencies to help them make more valid interpretations of EQAO data when assessing student academic performance (Kane, 2006; Linn & Baker, 1996; Meijer, 1996; Meijer & Sijtsma, 2001; Messick, 1989; Schmitt et al., 1999).

The second level of the HLM analysis suggests that the variance within and between schools is also a significant predictor of student academic achievement for EQAO Grade 9 mathematics assessments. There is 81.81% of score variability within schools and 18.18% between schools based on 16.79% of the total variance accounted at this level. The results from the HLM statistical analysis indicated that student task-values, effort, and students nested in schools are significant predictors of achievement for EQAO data. These findings support the literature as student task-values and effort are considered among the most immediate predictors of academic achievement and performance (Eccles & Wigfield, 2002).

One of the limitations of the current study is that it only takes into consideration students' overall score, and it does not explain the motivation effect per item. This is because item position can have an effect on student motivation in relation to the effort that the student puts forth in answering the item correctly on the LSA (Wolf et al., 1995). Another limitation is that the current study does not explore the impact of low motivation on student academic achievement per item or question but rather the impact of motivation as a whole. There is, however, a need to use other statistical techniques to explore the relationship of low motivation with items that count and do not count on EQAO LSA toward the student grade. There is also a need to develop item response theory models using pencil-and-paper tests that include motivation as a parameter estimate in the model. This will permit a better design of EQAO tests to more accurately estimate student abilities. As a result of addressing these needs, more valid interpretation of test results can be made when assessing student academic achievement in relation to the Grade 9 mathematics curriculum.

Conclusion

The research question that guided the current study was “to what extent does the expectancy-value theory of motivation and student effort relate to students’ academic achievement on a LSA of mathematics?” The outcome of the study suggests that student effort and task-values as defined by the expectancy-value theory are related to students’ academic achievement on EQAO Grade 9 mathematics assessments.

In the context of educational assessment, the current study provides important evidence against the common assumption that the impact of student motivation on LSA results may be negligible, as the outcome of this study reveals that these motivation components (task-values and effort) are significant predictors of students’ Grade 9 EQAO academic achievement in mathematics. The results of this study, however, can have implications for teachers and test organizations to help them better understand and assess students’ academic performance in relation to motivation while using LSA. It also may help them make better decisions about educational policies and curriculum changes, which are sometimes implemented based on EQAO test scores. The results of this study can also have implication for research because it builds on the work by DeMars (2000), Maehr and Meyer (1997), and Wolf and Smith (1995) by providing another avenue to examine student motivation in relation to LSA.

Declaration of Conflicting Interests

The opinions presented in this paper are solely those of the authors and do not necessarily reflect the opinions of SSHRC and EQAO

Funding

The authors received the following financial support for the research and/or authorship of this article.

This research was funded by the Social Science and Humanities Research Council (SSHRC) and the data were provided by the Education Quality Accountability Office (EQAO). The authors would also like to acknowledge the contribution of Kariann Brinson for her comments and edits related to this manuscript.

References

- Atkinson, J. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.
- Bishop, A., Clark, B., Corrigan, D., & Gunstone, D. (2006). Values in mathematics and science education: Researchers’ and teachers’ views on the similarities and differences. *International Journal of Mathematics Education*, 26, 7-11.
- Brookhart, S. M., & Durkin, D. T. (2003). Classroom assessment, student motivation, and achievement in high school social studies classes. *Applied Measurement in Education*, 16, 27-54.
- Brookhart, S. M., Walsh, J., & Zientarski, W. (2006). The dynamics of motivation and effort for classroom assessments in middle school science and social studies. *Applied Measurement in Education*, 19, 151-184.
- Camara, W. (2009). Moving large scale assessments to computer platform. *Large-Scale Assessment SIG*, 2(1), 2-4.
- Cole, J., Bergin, D., & Whittaker, T. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology*, 33, 609-624.
- De Ayala, R., Plake, B., & Impara, J. (2001). The impact of omitted responses on the accuracy of ability estimation in item response theory. *Journal of Educational Measurement*, 38, 212-234.
- DeMars, C. (2000). Test stakes and item format interactions. *Applied Measurement in Education*, 13, 55-77.
- Drasgow, F., Levine, M., & Williams, E. (1985). Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, 38, 67-86.
- Dweck, C. S., & Elliot, E. S. (1983). Achievement motivation. In E. M. Hetherington (Ed.) & P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 4. Social and personality development* (pp. 643-691). New York, NY: John Wiley.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values and goals. *Annual Reviews*, 53, 109-132.
- Education Quality Accountability Office. (2011, May 12). *EQAO*. Available from <http://www.eqao.com>
- Eklöf, H. (2006). Development and validation of scores from an instrument measuring student test-taking motivation. *Educational and Psychological Measurement*, 66, 1-15.
- Fraire, R., Tideman, T., & Watts, T. (1997). Indices of cheating on multiple-choice tests. *Journal of Educational Statistics*, 6, 152-165.
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education*, 10, 169-207.
- Kane, M. (2006). Current concerns in validity theory. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17-64). National Council on Measurement in Education and American Council on Education. Westport, CT: Praeger.
- Karabatsos, G. (2003). Comparing the aberrant response detection performance of thirty six person-fit statistics. *Applied Measurement in Education*, 16, 277-298.
- Kloosterman, P. (1996). Students’ beliefs about knowing and learning mathematics: Implications for motivation. In M. Carr (Ed.), *Motivation in mathematics* (pp. 131-156). Cresskill, NJ: Hampton Press.
- Linn, R., & Baker, E. (1996). *Assessing the validity of the National Assessment of Educational Progress: NAEP technical review panel white paper*. Washington, DC: National Center for Research on Evaluation, Standards, and Student Testing/ National Center for Education Statistics.
- Maehr, M. L., & Meyer, H. (1997). Understanding motivation and schooling. Where we’ve been, where we are, and where we need to go. *Educational Psychology Review*, 9, 371-409.
- Marsh, W., Koller, O., Trautwein, U., Ludtke, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effect models of casual ordering. *Child Development*, 76, 376-416.

- McMillan, J., Simonetta, L., & Singh, J. (1994). Student opinion survey: Development of measures of student motivation. *Educational and Psychological Measurement, 54*, 498-505.
- Meijer, R. (1996). Person-fit research: An introduction. *Applied Measurement in Education, 9*, 3-8.
- Meijer, R., & Sijtsma, K. (1995). Detection of aberrant item score patterns: A review of recent developments. *Applied Measurement in Education, 8*, 261-272.
- Meijer, R., & Sijtsma, K. (2001). Methodology review: Evaluating person fit. *Applied Psychological Measurement, 25*, 107-135.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York, NY: American Council of Education.
- Nering, M., & Meijer, R. (1998). A comparison of the person response function and the Iz person fit statistics. *Applied Psychological Measurement, 22*, 53-69.
- O'Neil, H. F., Abedi, J., Miyoshi, J., & Mastergeorge, A. (2005). Monetary incentives for low-stakes tests. *Educational Assessment, 10*, 185-208.
- O'Neil, H. F., Jr., Sugrue, B., Abedi, J., Baker, E. L., & Golan, S. (1997). *Final report of experimental studies on motivation and NAEP test performance* (CSE Tech. Report No. 427). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385-407.
- Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research, and applications*. Englewood Cliffs, NJ: Merrill-Prentice Hall.
- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement, 53*, 801-813.
- Putwain, D. (2007). Test anxiety in UK school children: Prevalence and demographic patterns. *British Psychological Society, 77*, 579-593.
- Putwain, D. (2008). Do examinations stakes moderate the test anxiety-examination performance relationship? *Educational Psychology, 28*, 109-118.
- Reynolds, W. M. (1984). Depression in children and adolescents: Phenomenology, evaluation and treatment. *School Psychology Review, 13*, 171-182.
- Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high-stakes testing. *Educational Evaluation and Policy Analysis, 23*, 197-227.
- Ryan, K., Ryan, A., Arbuthnot, K., & Samuels, M. (2007). Students' motivation for standardized math exams. *Educational Researcher, 36*, 5-13.
- Salomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychologist, 18*, 42-50.
- Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort as a function of perceptions and attributions. *Journal of Educational Psychology, 76*, 647-658.
- Scheifele, U. (1991). Interest, learning, and motivation. *Educational Psychologist, 26*, 299-323.
- Schmitt, N., Chan, D., Sacco, J., McFarland, L., & Jennings, D. (1999). Correlates of person fit and effect of person fit on test validity. *Applied Psychological Measurement, 23*, 41-53.
- Sotaridona, L., Linden, W., & Meijer, R. (2006). Detecting answer copying using kappa statistics. *Applied Psychological Measurement, 30*, 412-431.
- Sotaridona, L., & Meijer, R. (2003). Two new statistics to detect answer copying. *Journal of Educational Measurement, 40*, 53-69.
- Stipek, D., Givven, K. B., Salmon, J. M., & MacGyvers, V. L. (1998). The value (and convergence) of practices suggested by motivation research and prompted by mathematics education reformers. *Journal for Research in Mathematics Education, 29*, 465-488.
- Sundre, D., & Kitsantas, A. (2004). An exploration of the psychology of the examinee: Can examinee self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemporary Educational Psychology, 29*, 6-26.
- Sundre, D., & Moore, D. (2002). *The student opinion scale: A measure of examinee motivation*. Retrieved from http://www.jmu.edu/assessment/resources/resource_files/sos_assessment_update_article.pdf
- van Barneveld, C. (2007). The effect of examinee motivation on test construction within an IRT framework. *Applied Psychological Measurement, 31*, 31-46.
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*, 1-35.
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81.
- Wise, L. (1996). *A persistent model of motivation and test performance*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.
- Wise, S., & DeMars, C. (2006). An application of item response time: The effort-moderated IRT model. *Journal of Educational Measurements, 43*, 19-38.
- Wise, S., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education, 18*, 163-183.
- Wolf, L., & Smith, J. (1995). The consequence of consequence: Motivation, anxiety and test performance. *Applied Measurement in Education, 8*, 227-242.
- Wolf, L., Smith, J., & Birnbaum, M. (1995). Consequences of performance, test motivation, and mentally taxing items. *Applied Measurement in Education, 8*, 341-351.
- Volante, L. (2006). An alternative vision for large-scale assessment in Canada. *Journal of Teaching and Learning, 4*, 1-14.

Bios

Carlos Zerpa is a PhD candidate in the Faculty of Education at Lakehead University. His research interest is based on measurement and evaluation of human physical and cognitive performance by implementing mathematical modeling techniques.

Krystal K. Hachey is a full-time PhD student at the University of Ottawa in Education. Her research interests are focused on measurement and evaluation including Item Response Theory, Structural Equation Modeling, Validation, and Survey Development.

Dr. Christina van Barneveld is an Associate Professor in the Faculty of Education at Lakehead University. Her research interests include educational measurement and program evaluation.

Dr. Marielle Simon is a full professor at the Faculty of Education, University of Ottawa. She teaches and conducts research in the areas of classroom assessment, large-scale assessments and research methods.