Evaluation of Student Skills and Beliefs Survey:

Item Analysis, Factor Structure, and Measurement Invariance Analysis

Anna Sytnik

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Education

University of Washington

2014

Committee:

Min Li

Catherine Taylor

Program Authorized to Offer Degree:

College of Education

University of Washington

Abstract

Evaluation of Student Skills and Beliefs Survey:

Item Analysis, Factor Structure, and Measurement Invariance Analysis

Anna Sytnik

Chair of the Supervisory Committee:

Associate Professor Min Li

Educational Psychology

The present study evaluated the psychometric properties and dimensionality of the Student Skills and Beliefs Survey administered to N = 5,981 students in a local school district in South King County, WA. Item analysis revealed that the overall internal consistency of the survey is excellent, as measured by Cronbach's α = .91. Factor analysis revealed three stable factors: academic responsibility (8 items, α = .85), social-cognitive skills (12 items, α = .85), and academic competence (6 items, α = .79). Configural (factorial) invariance is stable across gender and ELL-status, but not for different grade bands. Metric invariance did not hold for gender and ELL-status, but after identifying and removing three non-invariant items, partial metric invariance was achieved. The study concludes with recommendations for further development of the Student Skills and Beliefs Survey, as well as possible future analyses.

Keywords: factor analysis, measurement invariance, motivation, school engagement, ELL

Evaluation of Student Skills and Beliefs Survey:
Item Analysis, Factor Structure, and Measurement Invariance Analysis

A recent implementation of Common Core State Standards Initiative is a national effort to raise mathematics and English language arts (ELA) academic standards in the United States. The goal of the initiative is to provide high-quality education to ensure that students graduate with skills and knowledge necessary to succeed in college and careers, as well as to remain competitive in the global economy (Porter, McMaken, Hwang, & Yang, 2011). However, raising standards alone is not likely to result in increased academic achievement (National Research Council and Institute of Medicine, 2004). Succeeding at school requires that students remain actively engaged in the learning process, are highly motivated, and view education as valuable for their future. Research on student engagement has been gaining momentum in recent years due to evidence that school engagement is linked to academic achievement (Skinner, Wellborn, & Connell, 1990; Finn & Rock, 1997; Wang & Holcombe, 2010; Sedaghat, Abedin, Hejazi, & Hassanabadi, 2011). Increasing student motivation and engagement is the focus of many modern school reform efforts, as it is believed that students who are actively engaged in the learning process generally have better grades and better performance on standardized tests (Marks, 2000; National Research Council and Institute of Medicine, 2004). School engagement is viewed as a possible solution to combat educational problems such as student alienation, boredom, lack of interest, and high dropout rates (Fredricks, Blumenfeld, & Paris, 2004).

According to Fredricks et al. (2004), school engagement is a multidimensional construct that incorporates student behavior, emotion, and cognition. Students' actions, feelings and thoughts are all dynamically interconnected within the individual, making it important to look at these factors simultaneously in order to get a richer description of students' engagement

(Fredricks et al., 2004). School engagement has also been shown to be malleable, it can be shaped and transformed with external forces; in other words, it is not a "fixed" internal trait of each student (Connell, 1990). The multifaceted nature of school engagement, along with its malleability, makes it an attractive area of educational research.

Schools and districts that turn their focus to promoting student engagement need ways to measure engagement to determine whether their strategies and intervention programs are working. According to Fredricks et al. (2004), there is a perpetual tension between conceptual precision and practical reality when it comes to measuring engagement. There is a wide net of student emotions, actions and thoughts that fall under the umbrella of school engagement, and carefully defining and measuring each one is not practical, as it would take an immense amount of time and resources. The vast majority of existing measurements focus on just one or at most two types of engagement. However, limiting the measurement to just one or two components of engagement (e.g. student behavior) will likely result in an incomplete or even deceiving picture of student engagement (Fredricks, et al., 2004; National Research Council and Institute of Medicine, 2004). Thus, there is a need for instruments that can measure behavioral, emotional, and cognitive engagement simultaneously, are straightforward in use and administration, and are available to schools that are interested in measuring school engagement.

A non-profit organization called the Youth Development Executives of King County (YDEKC) has developed a survey designed to measure student motivation and engagement. They used a mix-and-match approach suggested by Fredricks et al. (2004), whereby they selected a few constructs that collectively tap into each type of student engagement. Student Skills and Beliefs Survey consists of six different subscales designed to measure various constructs that contribute to school engagement: academic behavior, future orientation, mindsets,

sense of belonging, interpersonal skills, and thinking and learning. The ultimate goal of designing this survey is to provide schools with a free, easy-to-use, and high-quality instrument that will provide a practical measure of student engagement and motivation.

The purpose of this study is to help advance further development of Student Skills and Beliefs Survey instrument by (a) examining the psychometric properties of individual items, of the six subscales, and of the survey instrument as a whole, (b) validating the multidimensionality of the survey by examining its factor structure, and (c) investigating whether the measurement model functions similarly across different subgroups of students.

What is Engagement?

Skinner et al. (1990) define the term engagement as "initiation of action, effort, and persistence on schoolwork, as well as their ambient emotional states during learning activities" (p. 24). Lamborn, Newmann, and Wehlage (1992) define engagement as "the student's psychological investment in and effort directed toward learning, understanding, and mastering the knowledge, skills, or crafts that academic work is intended to promote" (p. 12). Similarly, Marks (2000) defines engagement as "a psychological process, specifically, the attention, interest, and investment and effort students expend in the work of learning." Engagement is typically viewed as either a two- or a three-dimensional construct. Earlier studies adopt a two-dimensional model of engagement that consists of behavioral and emotional components (Finn, 1989; Skinner et al., 1990; Marks, 2000). Latest reviews of the literature have also included a third, cognitive component (Jimerson, Campos, & Greif, 2003; Fredricks et al., 2004).

Behavioral engagement denotes actions and practices that students direct towards their learning and school (Wang, Willett, & Eccles, 2010). These include positive conduct such as completing homework and attending class, following classroom and school rules, participating in

extracurricular activities, staying on task, and initiating action when given the opportunity, as well as the lack of disruptive behaviors such as tardiness, unexcused absences, and fighting (Connell, 1990; Finn, 1989; Klem & Connell, 2004).

Emotional engagement refers to students' sense of connectedness to their school, interest in learning and school activities, valuing school, and having a positive attitude towards school and learning (Voelkl, 1997; Finn, 1989; Wang et al., 2010). Concepts included in this construct overlap heavily with constructs used in motivational research, and for practical purposes some researchers use the terms motivation and engagement interchangeably (National Research Council and Institute of Medicine, 2004). However, concepts and their definitions in engagement literature are more general and broad than those found in motivational literature (Fredricks et al., 2004). One exception is the concept of flow introduced by Csikszentmihalyi (1988), which is defined as a state of complete involvement, whereby individuals are so engrossed in an activity that they lose awareness of everything else, including time and space. Flow represents the highest level of emotional engagement (Fredricks et al., 2004).

Cognitive engagement refers to students' cognitive investment in learning, including use of self-regulated strategies, and application of mental efforts necessary for comprehension of complex ideas (Pintrich & De Groot, 1990; Zimmerman, 1990; Wang & Eccles, 2012). Connell & Wellborn (1991) define cognitive engagement as flexibility in problem solving, preference for challenging academic work, and employing positive coping strategies in case of failure. Students who are deeply cognitively engaged tend to have better control over their learning, for example they can suppress distractions, use learning strategies like rehearsal and memorization, can better organize their thoughts and materials, and as a result can achieve a deeper level of understanding (Fredricks et al., 2004).

As mentioned in the introduction, the three types of engagement do not exist in isolation; they are dynamically interconnected within each individual student. Finn (1989) proposed a participation-identification model, in which participation in class (behavioral engagement) increases identification with school (emotional engagement), which in turn increases participation. In addition, it is likely that emotional engagement may result in increased cognitive engagement (Fredricks et al., 2004). If students feel genuinely interested and excited about a topic (emotional engagement), they are more likely apply themselves to understanding of complex ideas related to that topic (cognitive engagement). However, it is important to note that, while we can make educated guesses about how the different types of engagement interact, we currently do not have empirical studies that focus specifically on these relationships (Guthrie & Wigfield, 2000; Fredricks et al., 2004).

Mediators of Engagement

Connell and Wellborn (1991) propose a self-system process model, in which they postulate that, in order for school engagement to be optimized, students' individual needs for competence, autonomy, and relatedness need to be addressed. If students perceive that their individual needs are being met by the school context, they will be more engaged, whereas if these needs are not met then students are more likely to become disaffected. Thus, these individual needs are considered to be psychological mediators of engagement (Fredricks et al., 2004; National Research Council and Institute of Medicine, 2004; Wang & Eccles, 2012).

Competence refers to students' beliefs about their level of control over their own success, understanding of what it takes to be successful, and beliefs that they have what it takes to succeed (Connell & Wellborn, 1991). Students who are convinced that success is outside of their potential, or that they have no control over academic outcomes are less likely to exert the

necessary effort required to complete academic work (National Research Council and Institute of Medicine, 2004). Psychological concept of self-efficacy closely overlaps with competence. First introduced by Bandura (1986), self-efficacy refers to individuals' perception of whether or not they will be able to succeed at something. According to Bandura (1986), people tend to engage in activities when they are confident in their ability to complete them successfully. On the other hand, individuals are more likely to avoid activities when that confidence is lacking. Another important notion for school engagement is that students who lack academic competence tend to attribute failures to their innate lack of ability, rather than a lack of effort (Skinner et al., 1990; Dweck, 2000). In other words, they view their intelligence as a "fixed entity" that will not change with time and effort (Bandura & Dweck, 1985). Students who lack confidence in their academic abilities and believe that there is nothing they can do to change are less likely to raise their hands in class for fear that they have the wrong answer, and are also less likely to ask for help from their teachers or peers for fear of appearing "dumb" (Dweck, 2000).

Autonomy refers to students' desire and ability to do things for personal reasons, rather than external pressures (Ryan & Connell, 1989; Fredricks et al., 2004). Students whose reasons for attending school and completing schoolwork revolve around avoiding bad grades and punishment are less likely to become actively engaged in their own learning beyond a superficial level (i.e. doing the bare minimum). On the other hand, students who feel genuinely interested in the content area they are learning, experience a sense of pride in their academic accomplishments, and see school as valuable to their current and future goals are more likely to become deeply engaged and reap all the benefits that school has to offer (National Research Council and Institute of Medicine, 2004).

Relatedness encompasses students' desire to be in an environment in which they feel cared for and supported by their peers and teachers (Connell & Wellborn, 1991). Belonging is a similar, albeit a little bit broader concept that involves identifying with the values and goals of the school, as well as feeling connected to other individuals, such as teachers and peers (Finn, 1989; National Research Council and Institute of Medicine, 2004). When students feel that they are part of school or classroom community, they experience a wide range of psychological benefits. Peer acceptance has been associated with school satisfaction and enjoyment of attending school, which are both aspects of emotional engagement (Wentzel, 1994). Students who feel accepted and supported by peers who value school are also more likely to exhibit positive social conduct (Osterman, 2000), which increases levels of behavioral engagement. Cognitive engagement can also be increased with greater peer support, particularly when students engage in active discussions of ideas and class material, debate different points of view, and give each other constructive criticism (Newmann, 1992). Teacher support has been linked to levels of behavioral, emotional, and cognitive engagement (Connell & Wellborn, 1991; Fredricks et al., 2004). According to Marks (2000), students who feel support from both teachers and peers have greater levels of school engagement across all grade levels.

Engagement and Academic Achievement

A number of empirical studies link school engagement to academic achievement. Skinner et al. (1990) examined the relationships between elementary students' perceived level of control, capacity beliefs, school engagement, and academic achievement variables. Perceived level of control and capacity beliefs were measured with self-report questionnaires administered to students; school engagement was measured using teacher ratings; and academic achievement was measured by student grades and scores on math and reading subtests of Stanford

Achievement Test (Psychological Corporation, 1970). Path analysis results showed that perceived level of control promotes or undermines school engagement, which in turn has a direct, positive relationship with academic achievement (Skinner et al., 1990).

Finn and Rock (1997) looked at the relationships between school completion, academic achievement, and engagement levels among low-income, minority youths (n = 1803). They have classified "at risk" high school students into three groups: resilient students (defined by school completion and high academic achievement), nonresilient students (school completion and low academic achievement), and noncompleters (students who have dropped out). The concept of resilience is defined as "successful adaptation to life tasks in the face of social disadvantage or highly adverse conditions" (p. 222). Student engagement was measured using a composite score of teacher ratings (the extent to which student completes homework, comes prepared, attentive in class, and not disruptive) and self-report questionnaire administered to students (questions related to in-school and out-of-school initiative taking). The results showed large, significant differences in behavioral engagement among all three groups, with resilient students having the highest level of behavioral engagement. The findings suggest that school engagement is a vital element of academic resilience (Finn & Rock, 1997).

In a short-term longitudinal study, Wang and Holcombre (2010) examined the relationships between student perceptions of school environment, school engagement, and academic achievement among 1,046 urban, middle school students. Specifically, the focus of the study was to determine whether school engagement mediates the relationship between students' perceptions of school environment and academic achievement using structural equation modeling (SEM). Academic achievement was measured by student grade point average (GPA), school engagement was measured using a 14 item school engagement index (Eccles, Midgley,

Wigfield, Buchanan, Reuman, Flanagan, & Mac Iver, 1993), and perceptions of school environment were measured by self-report questionnaire called the School Climate Perception Measure (Eccles & Midgley, 1989). Results indicated that student perceptions of school environment contribute to all three types of school engagement, and school engagement has an effect on academic performance, even after accounting for gender, race/ethnicity, and SES status (Wang & Holcombre, 2010).

Similar results were found in an international study by Sedaghat et al. (2011), which examined the effects of perceived ability, perceived instrumentality, and achievement goals on cognitive engagement and academic achievement among 1,371 junior high school students in Tehran, Iran. Perceived ability, instrumentality, and achievement goals were measured by Approaches to Learning (ATL) scale comprised of 32 survey items (Raymond, DeBacker, & Green, 1999). Cognitive engagement was measured by Motivated Strategies Learning Questionnaire (MSLQ), which comprised of 19 items designed to assess shallow and deep cognitive learning strategies (Pintrich, Smith, Garcia, & McKeachie, 1993). Structural equation modeling was used to test hypothesized relationships among the variables of interest. Results showed that academic achievement was significantly and positively predicted by perceived ability and deep cognitive strategy use (Sedaghat et al., 2011).

Summary

Several empirical studies have examined the link between school engagement and academic achievement. All three types of school engagement have been shown to have a direct, positive relationship with academic achievement. Studies have also shown that school engagement often acts as a mediator between various aspects of school context (e.g., perceptions of school environment, teacher support) and academic achievement (e.g., grades). It is possible

that the causal relationship is bidirectional; with higher levels of school engagement resulting in higher levels of academic achievement, which in turn result in increased engagement (Fredricks et al., 2004). It is important to note that most research precludes causal inference because the independent and dependent variables were measured simultaneously. Nevertheless, research supports a strong link between various components of school engagement and academic achievement.

Measuring Engagement

There is little consistency in existing literature when it comes to measuring school engagement. The majority of studies choose to either use a mix-and-match approach by combining scales from different existing measures (e.g., Perry, Liu, & Pabian, 2010), appropriate existing scales from other areas of research such as motivation (e.g., Sedaghat et al., 2011), or researchers create their own instruments to measure various aspects of school engagement (e.g., Finn & Rock, 1997). In addition to measurements of engagement described in the previous section, a number of instruments have been developed specifically to measure school engagement.

The purpose of this section is to take a closer look at some of these instruments to get a sense of how various aspects of school engagement and psychological mediators of engagement have been measured in the past, as well as to see whether Student Skills and Beliefs Survey can offer any advantages when compared to existing body of measurement work. While there are some measurements that use teacher or parent ratings of student engagement, the focus of this section will be on reviewing the instruments that measure engagement from the student's point of view. Using teacher and parent reports to measure student engagement is problematic because only behavioral (i.e. observable) student engagement can be assessed this way, which can lead to

erroneous conclusions because students sometimes "go through the motions" of what is required in class without actually being deeply engaged in their learning (Lamborn et al., 1992). Table 1 provides a summary of the instruments reviewed in this section.

Table 1

Summary of Measures of School Engagement

Instrument Name	Subscales	Type of Engagement	Cronbach's α
Rochester Assessment Package for Schools (RAPS)	Ongoing engagement Reaction to change Teacher support	Behavioral Emotional	.68 .69 .82
School Engagement Questionnaire (SEQ)	Homework Classroom attention Attendance Classroom concentration	Behavioral	.74-.86
Identification With School Questionnaire (IWS)	Belongingness Valuing school	Emotional	.84
Student Engagement Instrument (SEI)	Teacher-student relationships Control and relevance of school work Peer support for learning Future aspirations and goals Family support for learning Extrinsic motivation	Emotional Cognitive	.88 .80 .82 .78 .76 .72
High School Survey of Student Engagement (HSSSE)	Cognitive/Intellectual/Academic engagement Social/Behavioral/Participatory engagement Emotional engagement	Behavioral Emotional Cognitive	--

Rochester Assessment Package for Schools (RAPS) was developed by Wellborn and Connell (1987) to measure behavioral and emotional components of engagement. RAPS measures school engagement in three ways: ongoing engagement, reaction to challenge, and teacher support. Ongoing engagement (α = .68) is measured by the degree to which students exert effort on schoolwork, pay attention and prepare for class activities, and believe that school is important (e.g., "When I am in class, I just act like I am working"). Using previously defined framework, ongoing engagement can be viewed as measuring both behavioral engagement as well as the need for autonomy (i.e., valuing school). Reaction to challenge (α = .69) measures how well students cope with negative school or learning events (emotional engagement). Students also evaluated the degree to which they feel supported by their teachers (α = .82), which is a measure of relatedness (e.g., "My teacher is fair to me").

School Engagement Questionnaire (SEQ) is a self-report survey consisting of 12 Likert-type items designed to measure behavioral engagement (Dornbusch & Steinberg, 1990). Internal consistency as represented by Cronbach's alpha for this scale ranges from .74 to .86 for different studies (Fredricks & McColskey, 2012). The scale is broken down into four factors: homework (e.g. "How much time do you put into homework each week?"), classroom attention (e.g., "How often do you really pay attention to the class work?"), attendance (e.g., "How often do you cut each of these classes?"), and classroom concentration (e.g., "How often does your mind wander in each of these classes?"). Higher total score on SEQ indicates a stronger level of behavioral engagement (Perry et al., 2010).

Identification With School Questionnaire (IWS) is a self-report questionnaire designed to measure the degree to which students identify with school (Voelkl, 1996). This instrument consists of 16 Likert-type items scored on a four-point scale (1 = strongly agree to 4 = strongly

disagree) and has two dimensions: belongingness (e.g., “I feel proud of being a part of my school”), and valuing school/attitudes toward school (e.g., “School is often a waste of time”). The two dimensions have been examined using confirmatory factor analysis, and it was determined that a single factor identification model is a better fit than the two-factor model (Voelkl, 1997). Item scores are combined into one total score, with higher scores indicating higher levels of school identification. Cronbach’s alpha reliability coefficient for this scale is .84. Although Voelkl did not use the same terminology, IWS can be considered a measure of student emotional engagement.

Student Engagement Instrument (SEI) is an instrument designed to measure student cognitive and psychological engagement from student perspective (Appleton, Christenson, Kim, & Reschly, 2006). The survey consists of 56 Likert type items, with 30 items designed to measure levels of cognitive engagement (e.g., “After finishing my schoolwork I check it over to see if it’s correct”) and 26 items designed to measure psychological engagement (e.g., “Other students at school care about me”). Items were scored on a four-point scale, ranging from strongly disagree to strongly agree (Appleton et al., 2006). SEI was orally administered to n = 1931 eighth grade students in an urban setting. The sample was 40.4% African American, 35.1% White, 10.8% Asian, 10.3% Hispanic, and 3.5% American Indian. Of all the participants, 22.9% reported that their home language is not English. Using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), the researchers have determined that a six-factor model is the best fit for the data. The following six factors have been identified: Teacher–Student Relationships (α = .88), Control and Relevance of School Work (α = .80), Peer Support for Learning (α = .82), Future Aspirations and Goals (α = .78), Family Support for Learning (α = .76), and Extrinsic Motivation (α = .72).

High School Survey of Student Engagement (HSSSE) is a survey instrument designed to measure three dimensions of engagement among high school students (Yazzie-Mintz, 2007). Cognitive/Intellectual/Academic Engagement dimension includes questions about homework, classwork, preparation for class, and level of academic challenge (e.g., "I enjoy working on tasks that require a lot of thinking and mental effort"). Social/Behavioral/Participatory Engagement dimension focuses on student actions, interactions, and participation within the school community (e.g., "During this school year, about how often have you asked or answered questions in class?"). Emotional Engagement includes feelings of connectedness to school, general feelings about school, and feelings about where they are in school (e.g., "My opinions are respected in this school"). This survey is quite extensive, the current version is four pages of densely packed items, a total of 31 broad questions (e.g., "How much does each of the following classroom activities and assignments interest you?") with over 120 individual statements designed to measure various aspects of school engagement. Most items are of Likert-type with four answer options (e.g., 1 = strongly disagree to 4 = strongly agree; 1 = never to 4 = often). Unfortunately, the reliability information is not available for this measure.

Summary

Of all the instruments reviewed, the HSSSE instrument is the most current, comprehensive measure of school engagement; however, it has certain limitations. First, the absence of reliability information is alarming. Second, the distinctions between different dimensions are not very clear, for example the social component is included in both Social/Behavioral/Participatory and Emotional Engagement dimensions. Third, the inclusiveness of so many different aspects of engagement comes with a price of lengthiness, which can has been known to cause fatigue among the participants (Dillman, Smyth, and Christian, 2009). Last

but not least, this instrument can only be used at a high school level. It is important to keep track of student engagement during elementary and middle grades as well as in high school, as research has shown that levels of school engagement tends to decline over time (National Research Council and Institute of Medicine, 2004).

The main limitation of the other instruments reviewed is that they do not measure all three types of school engagement simultaneously; at most one or two types of engagement are measured. Measuring only one type of engagement (e.g. behavior) can result in an incomplete, and sometimes even erroneous picture of student engagement (Fredricks et al., 2004). Another limitation of the instruments reviewed in this section is limited attention is being paid to the psychological mediators of school engagement (need for autonomy, relatedness, and competence). Also, some instruments do not seem to have an adequate number of items when compared to the number of factors they are supposed to measure (e.g., SEQ consists of only 12 items that are measuring four different factors).

While recent research on engagement proposes a three-dimensional model that consists of behavioral, emotional, and cognitive components (Fredricks et al., 2004), few existing instruments sufficiently address this model. Student Skills and Beliefs Survey was constructed using a mix-and-match approach advocated by Fredricks et al. (2004), whereby different factors, which contribute to school engagement and psychological mediators of engagement, have been selected to tap into each component of school engagement. When finalized, it has the potential to become a valuable tool for measuring different dimensions of school engagement, as well as the psychological mediators of engagement for students at elementary, middle, and high school grade levels.

Method

Participants

The survey was administered to students in 5th through 12th grades, in one of the local school districts in South King County. A total of *N* = 5,981 students completed the survey. The sample is equally distributed by gender, race/ethnicity, and grade level. The demographic characteristics of the data are displayed in Table 2.

Table 2

Descriptive statistics of Student Skills and Beliefs survey data

Characteristics	*N*	*%*
Gender		
Males	2845	47.6%
Females	2902	48.5%
Other	155	2.6%
Grade		
5th	931	15.6%
6th	672	11.2%
7th	645	10.8%
8th	598	10.0%
9th	818	13.7%
10th	893	14.9%
11th	722	12.1%
12th	575	9.6%
Ethnicity		
American Indian/Alaskan Native	85	1.4%
Asian	1305	21.8%
Pacific Islander/Native Hawaiian	183	3.1%
Black	686	11.5%
Hispanic	833	13.9%
White	1507	25.2%
Two or more races	962	16.1%
English Spoken at Home		
Always	2974	49.7%
Most of the time	709	23.6%
Sometimes	595	18.6%
Rarely or Never	166	6.3%

Note. Total *N* = 5,983

For practical purposes, “English Spoken at Home” survey question will be used to identify English language learners (ELL) students: “Sometimes” and “Rarely or never” answer categories will be combined as a proxy for ELL status (25%), whereas students who answered “Always” or “Most of the time” will be classified as non-ELL students (73%).

Instrument

The Student Skills and Beliefs Survey consists of 34 items (excluding demographic questions) that fall into one of the six subscales: academic behavior, future orientation, mindsets, sense of belonging, interpersonal skills, and thinking and learning. The items are of 5 point Likert-type, with the following answer options: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree. The stems are “I” statements describing a particular behavior, strategy or disposition. YDEKC acknowledge the use of items from the following measurement tools in the construction of the survey:

- School Attitudes Assessment Survey (McCoach, 2002; Cronbach’s α = 0.87). The following items were used for the academic behavior subscale:
    - I concentrate on my schoolwork
    - I am a responsible student
    - I complete my schoolwork regularly
- Patterns of Adaptive Learning Scales (PALS), Academic Efficacy subscale (Midgley, Maehr, Hruda, Anderman, Anderman, Freeman, ... & Urdan, 2000; α = 0.78). The following items were used for the mindsets subscale:
    - I can do even the hardest school work if I try
    - Even if the work is hard, I can learn it
    - I'm certain I can master the skills taught in class this year

    - I'm certain I can figure out how to do the most difficult class work
    - I can do almost all the work in class if I don't give up

- Theory of Intelligence Scale (Blackwell, Trzesniewski, & Dweck, 2007; Cronbach's α = 0.78). The following items were used as part of the mindsets subscale:
    - I really can't do much to change how smart I am
    - I can learn new things but I can't really change how smart I am
    - How smart I am is something that I can change

Table 3 displays the subscales, content summary, as well as type of engagement and/or psychological need the subscale is intended to measure (see Appendix B for entire survey instrument as presented to participants).

Analysis

This section describes the analysis techniques that will be performed to address the goals of this study. As mentioned in the introduction, these goals are to (a) examine the psychometric properties of individual items, of the six subscales, and of the survey instrument as a whole, (b) validate the multidimensionality of the survey by examining its factor structure, and (c) investigate whether the measurement model functions similarly across different subgroups of students. Part (a) will be addressed by conducting item analysis and examining the reliabilities of the six subscales, as well as the total score; part (b) will be performed using exploratory and confirmatory factor analyses; and part (c) will be performed by conducting multi-group confirmatory factor analyses to test for measurement invariance.

Table 3

Subscales, Summary of Content of Items, Type of Engagement Intended and/or Psychological Needs Measured

Subscale	Summary of Content	Type of Engagement	Psychological Need
Academic Behavior	Three items assessing student concentration on schoolwork, work completion regularity, and being a responsible student	Behavioral	
Future Orientation	Five items assessing goals management, emotional control, hope for the future, and plans for graduation	Behavioral Emotional	Autonomy
Interpersonal Skills	Five items assessing listening skills, ability to work in a group, sharing thoughts and ideas with others, and social problem resolution skills	Emotional	Relatedness
Sense of Belonging	Three items assessing feelings of belonging to a group of peers, level of comfort asking teachers for help, as well as involvement in extracurricular activities	Emotional	Relatedness
Mindsets	Seven items assessing student competence, self-efficacy beliefs, family culture and values, and sense of control over their learning	Behavioral Emotional	Competence
Thinking and Learning	Nine items assessing decision making skills, problem-solving strategies, ability to come up with new ideas, and metacognitive skills	Cognitive	

Note. A total of $N = 34$ Likert-scale items

Item Analysis. Item analysis includes a variety of statistical procedures designed to examine the quality of the items. It provides information about how participants responded to each item, as well as well as how each item relates to the overall performance on the survey. SPSS version 19 was used to conduct all item analyses.

First, means and standard deviations for each item were calculated. Normality will be examined by looking at skewness and kurtosis for each item. The criteria for acceptable skewness and kurtosis are +/- 2 for psychometric purposes (Bentler & Wu, 2003). Next, response rate for each answer option for each item, as well as non-response rate will be examined to determine whether all answer options were appropriate choices (e.g., if all students chose only Agree or Strongly Agree options and omitted all the others, that could indicate a potential problem with an item stem). Non-response rates above 5% may indicate among other things potential confusion among the participants, sensitivity of information requested, or difficulty recalling the information asked for (Eftekhari-Sanjani, 2008).

The next step involves looking at the relationship of individual item scores to the overall performance on the survey. Corrected item-total correlation measures the correlation between an individual item and the adjusted total score. The term corrected is used because the total score is adjusted to omit scores from the item in question. Low corrected item-total correlation values indicate that items do not measure the same construct(s) as the rest of the items in the scale. A general rule of thumb is to consider item-total correlations below .30 as low (Traub, 1994; Field, 2005), although some suggest using .40 as the cutoff point (Leong & Austin, Eds. 2006).

Reliability. Internal consistency is one common way to estimate the reliability of survey scores. It refers to a general agreement of multiple survey items that are designed to measure the same underlying construct (in this case, the construct is school engagement), in other words,

internal consistency measures how well the scores from the instrument "hold together". Internal consistency was calculated using SPSS version 19 software by obtaining Cronbach's alpha (α) reliability coefficient (Cronbach, 1951) for the total score, as well as for scores from different subscales (i.e. different factors). Cronbach's α values range from 0 to 1, with values above .70 generally considered acceptable (George & Mallery, 2006).

In addition, "Cronbach's alpha if item deleted" (a subroutine available in the SPSS reliability estimation routine) was calculated for each subscale. These values are indicators of what the Cronbach's alpha for each subscale will be if we remove a particular item from the survey. Typically, if the alpha increases by .05 through omission of an item, it is a good idea to drop the item.

Factor Analysis. Factor analysis is defined as a "set of statistical procedures designed to determine the number of distinct constructs needed to account for the pattern of correlations among a set of measures" (Fabrigar & Wegener, 2011). In this study, the measures are the items in the Student Skills and Beliefs Survey, and latent (i.e. unobservable) constructs (also called factors) are different components of school engagement. Two different types of factor analyses often used to determine the underlying structure of survey instruments are referred to as exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The key difference between EFA and CFA is that EFA analysis is that it is unrestricted, in other words there are no clear expectations about the number of underlying factors, nor which specific items each latent factor will influence (Fabrigar & Wegener, 2011). On the contrary, CFA is a type of restricted analysis, in other words the number of factors and which items are influenced by which factor are a priori defined by the researcher.

Factor analysis is widely used in the development of measurement instruments (e.g., Appleton et al., 2006). It provides important information about the dimensionality of the scale, as well as additional information about psychometric properties of each item (i.e. the strength of factor loadings). Following a suggestion by Fabrigar & Wegener (2011), for this part of analysis, data has been randomly divided into two equal parts (see Table 4 for demographic descriptions of the two parts). This is also a common approach found in the literature. Part 1 will be used to conduct the EFA, whereas Part 2 will be used to conduct the CFA.

Table 4

Descriptives of the two parts of the data used for factor analyses

Characteristics	Part 1 (*N* = 2991)		Part 2 (*N* = 2992)	
	N	*%*	*N*	*%*
Gender				
Male	1405	47.0%	1440	48.1%
Female	1455	48.6%	1447	48.4%
Other	79	2.6%	76	2.5%
Grade				
5th	459	15.3%	472	15.8%
6th	325	10.9%	347	11.6%
7th	330	11.0%	315	10.5%
8th	300	10.0%	298	10.0%
9th	423	14.1%	395	13.2%
10th	430	14.4%	463	15.5%
11th	362	12.1%	360	12.0%
12th	294	9.8%	281	9.4%
Race/Ethnicity				
American Indian/Alaskan Native	47	1.6%	38	1.3%
Asian	649	21.7%	656	21.9%
Pacific Islander/Native Hawaiian	80	2.7%	103	3.4%
Black	324	10.8%	362	12.1%
Hispanic	420	14.0%	413	13.8%
White	759	25.4%	748	25.0%
Two or more races	509	17.0%	453	15.1%
English Spoken at Home				
Always	1462	48.9%	1512	50.5%
Most of the time	709	23.7%	701	23.4%
Sometimes	595	19.9%	515	17.2%
Rarely or Never	166	5.5%	213	7.1%

Note. Total *N* = 5,983.

Exploratory Factor Analysis (EFA). Although the makers of the survey believe that there are six distinct subscales (factors) present in the survey, qualitative examination of the items along with uncertainty and overlap of concepts related to school engagement in the literature leaves us with incomplete expectations about the nature of the true underlying structure of the Student Skills and Beliefs Survey instrument. Therefore, it was determined that the best course of action is to first conduct an unrestricted factor analysis (EFA). EFA identifies factors by providing estimates (referred to as factor loadings) of the strength and direction of influence between the factors and each item. EFA will be performed using SPSS version 19 on Part 1 of the data.

Conducting EFA involves several steps. First, factorability of the data needs to be examined. Tabachnic and Fidell (2007) suggest that item correlation matrix should be examined to determine factorability - correlations above .30 among multiple items indicate that the items share at least 10% of the variance, suggesting reasonable factorability. SPSS output provides two additional tests of factorability: Keiser-Meyer-Olkin (KMO) measure of sampling adequacy (values above .90 are desirable), and the Bartlett's Test of Sphericity, which needs to be significant (Tabachnic & Fidell, 2007; Williams, Brown, & Onsman, 2010). If these criteria are not met, factor analysis should not be conducted.

The second step involves choosing factor extraction method and number of factors to extract. There are numerous options for factor extraction available in SPSS, and literature is scarce on the advantages of different methods. According to Williams et al. (2010), Principal Axis Factoring (PAF) is among one of the most popular methods used in research, therefore this method was chosen for this study. Several criteria were used to identify the optimal number of

factors to extract: 1) “percent of cumulative variance accounted for by the factors” and “eigenvalues-above-one-rule”, 2) scree test, and 3) parallel analysis.

Although there is no strict rule regarding cumulative variance accounted for by the factors, Williams et al. (2010) suggest that in the social sciences the percentages typically range from 50-60%. There is a general consensus among the researchers that strictly applying the “eigenvalues > 1.0” rule is not a good practice and often results in retaining too many factors, therefore, the use of more than one criteria for determining the number of factors to extract is recommended as best practice (Velicer & Jackson, 1990; Tabachnic & Fidell, 2007; Costello & Osborne, 2005; Williams et al., 2010).

Scree test is a popular technique for identifying the number of potential factors. It involves the examination of the plot of the eigenvalues in order to find a natural breaking point or bend at which the curve levels off and becomes flat. The number of points above the “break” is the number of factors that should be retained (Costello & Osborne, 2005). Scree test is often recommended in the literature (Tabachnic & Fidell, 2007, Costello & Osborne, 2005; Fabrigar & Wegener, 2011) as a good “eyeball” method for identifying the potential number of factors to extract, although most concede that this technique is somewhat subjective. Sometimes with real data it is not clear at which point the line becomes flat. Nevertheless, it is a useful procedure and when used in conjunction with other methods provides a good estimate of the number of factors to retain.

Parallel analysis is a procedure that involves comparing the eigenvalues from the sample data to the eigenvalues that would result from a completely random data of the same size, with the same number of variables as specified by the researcher (Fabrigar & Wegener, 2011). The factors are retained if the eigenvalues from the actual data exceed the eigenvalues of the

corresponding factors generated by the random data (see O'Connor, 2000, for a more detailed description of this procedure). According to the literature, parallel analysis is among the best, albeit under-used, methods for determining the number of factors to extract (O'Connor, 2000; Fabrigar & Wegener, 2011). If more than one model seems plausible, multiple EFAs are conducted to determine which model is best. The decision on which number of factors is the "best" fit for the data is often up to the researcher and depends on results of the EFAs as well as on the theoretical basis.

The third step involves choosing a rotation method. Factors exist in multidimensional space and the purpose of the rotation is to simplify and clarify the structure of the data. Rotation produces more interpretable solutions by maximizing high item loadings and minimizing low item loadings (Costello & Osborne, 2005; Williams et al., 2012). There are two kinds of rotation: orthogonal and oblique. Orthogonal rotation (e.g. Varimax, Equamax) produces factors that are not correlated, whereas oblique rotation (e.g. Promax, Direct Oblimin) produces correlated factors. Promax rotation method will used in this study due to the expectation that different factors related to school engagement will be correlated (Fredricks et al., 2004).

After rotation, communalities and rotated factor loadings of all items are examined. According to Fabrigar and Wegener (2011), communality is the "variance in that variable explained by all the common factors in the model" (p. 134). Items that have low communalities (less than .20) are either not related to the other items, or they may be part of an additional domain that should be explored by adding more items in the future. Tabachnic and Fidell (2007) recommend minimum factor loadings of .32 of three to five variables for each factor. According to Comrey and Lee (1992), factor loadings above .71 are excellent, around .45 are fair, and

loadings below .32 are poor. During this part of EFA, certain items may be removed due to low communalities or factor loadings.

The final step of the EFA is interpreting the results. Examining the factors and item loadings may lead to the need to rename the factors (e.g. if items from two different subscales load on to the same factor). This is not an exact science and requires considerable amount of input from the researcher, but nevertheless a very important step, especially if the results of the EFA differ from the preconceived beliefs about the structure of the instrument. Initial subscale names may not be a good representation of the factors revealed by the EFA.

Confirmatory Factor Analysis (CFA). Confirmatory Factor Analysis (CFA) is a Structural Equation Modeling (SEM) procedure commonly used to identify nondirectional relationships (e.g. correlations) between the indicators and latent factors (Geiser, 2012). Relevant to the present study, SEM has been a popular technique to examine various hypothesized relationships between school engagement, psychological factors, context variables, and academic achievement. As mentioned previously, factors and their indicators (i.e. items) are specified a priori. CFA were performed on Part2 of the data using Mplus version 7. Several CFA models were tested, including a unidimentional model in which all the items load on to one factor, a six-factor model which represents the initial six subscales of the survey instrument, and a model with the number of factors suggested by EFA.

One of the benefits of conducting CFA is the ability to compare model fit of different models tested. Mplus software provides several model fit indices in the output. These include Chi-square goodness-of-fit test (χ^2), comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Ideally, the chi-square goodness of fit test should not be significant, however, it is very

sensitive to sample size; therefore, researchers caution against strict application of rigid cutoff values for this fit statistic (Fornell & Larcker, 1981). Values above .90 for CFI and TLI indicate adequate model fit, according to Kline (2011). For RMSEA, values below .05 indicate good model fit (Browne & Cudeck, 1993). SRMR values at or below .08 indicate an acceptable model fit (Hu & Bentler, 1999).

If the model fit statistics indicate that there is a problem with model fit, modifications to the best fitting model will be considered. Modifications include possible removal of items from analysis and/or moving items to load on to different factors. Both theoretically and statistically justifiable modifications can be used. Mplus output provides modification indices, which are values that display the potential improvement in model fit if a certain item was to load onto a different factor. These were used along with residuals and parameter estimates (i.e. factor loadings) to make modification decisions, along with theoretical justifications. Once the final measurement model was identified, Cronbach's α internal consistency coefficient was recalculated for the new factors.

Measurement Invariance

Measurement invariance occurs when the measurement model remains consistent across different groups (Meredith, 1993; Van de Schoot, Lugtig, & Hox, 2012). In other words, the individuals in different groups (e.g. males and females) interpret the items the same way. In order to be able to compare different groups (e.g. males and females) on their levels of school engagement as measured by the Student Skills and Beliefs survey, measurement invariance was tested. If the measurement tool is non-invariant, it means that different groups do not respond the same way to the questions, and any comparisons made with scores from such an instrument would be misleading.

Two different invariances will be tested: configural invariance and metric invariance. Configural invariance occurs when the factor structure of the instrument is stable across different groups. According to Wu, Li, & Zumbo (2007), if configural invariance does not hold, it is an indication that different constructs were measured for different groups. Configural invariance is a prerequisite for metric invariance (sometimes referred to as weak factorial invariance), which exists when the scales of latent factors are identical across groups. In other words, "one unit change in the item score is scaled to an equal unit change in the factor score across groups" (Wu et al., 2007).

Measurement invariance is tested by conducting multiple-group CFAs. First, the configural model is tested by specifying comparison groups and leaving all the parameters unrestricted (i.e. free-to-vary). This is also known as the baseline model. The same model fit indices are used to evaluate model fit as with regular CFA described above. If configural invariance is established, metric invariance is tested by constricting factor loadings to be equal among the specified groups. The metric invariance model is then compared to the baseline (configural) model. If the differences in the fit statistics between the models are not significant, metric invariance is established. While many researchers use the Chi-squared difference test ($\Delta\chi^2$) to assess model fit of metric invariance models, Wu et al. (2007) warn that just like χ^2, $\Delta\chi^2$ is also vulnerable to sample size. In addition to $\Delta\chi^2$ test, CFI difference (ΔCFI) of less than -.01 was used to evaluate model fit (Wu et al., 2007), along with along with RMSEA $\leq$. 05, and SRMR < .08 criteria.

In order to make meaningful comparisons across different groups, both configural and metric invariance are necessary. If metric invariance does not hold, partial invariance can be established. This is done in an iterative process, whereby factor loading constraints on one

variable are relaxed, while keeping the rest constricted. Changes in model fit indices are evaluated after each multiple-group CFA is run in order to determine which items are non-invariant (Millsap, 2007). Van de Schoot et al. (2012) suggest examining factor loadings and modification indices in unrestricted models to determine which items are potentially non-invariant.

Summary

Various statistical analyses were performed to address the goals of this study. SPSS software was used to conduct item analysis, reliability analysis, and EFA. Mplus software was used to conduct CFA and multiple-group CFA to test for measurement invariance. For EFA and CFA parts of analysis, data was randomly divided into two equal parts.

Results

Item Analysis

Means and standard deviations for each item are displayed in Table 5. It is important to note that for, clarity purposes, items have been rearranged during analysis to keep the items that belong to the same subscales together. All item numbers mentioned in the results and discussion sections refer to Table 5. Item means range from 3.01 to 4.35, on a 5-point scale. Standard deviations range from .84 to 1.23. All of the items are slightly negatively skewed, with skewness values ranging from -0.34 to -1.45. Most items are mesokurtic, with values ranging from -0.12 to +1.82. Analysis of response rates showed that participants chose all options for each item to some degree, indicating no problems with appropriateness of answer options for all item stems. In addition, non-response rate for items was quite low, varying from .6% to 2%, not raising any concerns (see Table 6). Corrected item-total correlations for each item are displayed in Table 5.

Table 5

Survey items from each subscale and psychometric properties of subscales and individual items.

Subscales (Cronbach's α)/Items	Mean (SD)	Corrected Item-Total Correlation	Cronbach's α if Item Deleted
Academic Behaviors (α = 0.78)			
1. I concentrate on my schoolwork	3.93 (0.84)	0.65	0.67
2. I am a responsible student	3.98 (0.86)	0.62	0.69
3. I complete my schoolwork regularly	3.81 (1.03)	0.59	0.74
Future Orientation (α = 0.72)			
4. I take responsibility for working on my goals	4.02 (0.84)	0.56	0.64
5. I am hopeful about my future	4.34 (0.87)	0.50	0.66
6. I know I will graduate from high school	4.35 (0.90)	0.44	0.69
7. I am able to control my emotions	3.81 (1.04)	0.36	0.72
8. I am good at staying focused on my goals	3.67 (0.92)	0.56	0.64
Interpersonal Skills (α = 0.69)			
9. I listen to other ideas even if those ideas are different from mine	3.94 (0.89)	0.43	0.65
10. I work well in a group or team	3.82 (1.04)	0.42	0.65
11. I am good at sharing ideas in front of other people	3.40 (1.16)	0.48	0.63
12. I am good at writing thoughts and ideas	3.60 (1.04)	0.47	0.63
13. I am able to handle problems with people without fighting	3.80 (1.01)	0.43	0.65
Sense of Belonging (α = 0.51)			
14. I feel that I belong to a group of friends	3.85 (1.11)	0.32	0.43
15. I am comfortable asking my teacher(s) for help	3.75 (1.05)	0.37	0.35
16. I am involved in activities to support my school or community	3.39 (1.23)	0.31	0.46
Mindsets (α = 0.73)			
17. I can do even the hardest school work if I try	3.89 (0.96)	0.47	0.70
18. I'm certain I can master the skills taught in class this year	3.79 (0.94)	0.49	0.69
19. I really can't do much to change how smart I am*	3.29 (1.20)	0.24	0.74
20. I can do almost all the work in class if I don't give up	4.13 (0.88)	0.49	0.69
21. I can learn new things but I can't really change how smart I am*	3.01 (1.20)	0.22	0.75
22. I understand my family culture and values	4.15 (0.94)	0.28	0.73
23. How smart I am is something that I can change	3.73 (1.06)	0.50	0.69
24. Even if the work is hard, I can learn it	3.94 (0.93)	0.56	0.68
25. I'm certain I can figure out how to do the most difficult class work	3.68 (0.97)	0.53	0.68
Thinking and Learning (α = 0.84)			
26. I make good decisions in my school life	3.83 (0.93)	0.56	0.82
27. I solve problems by first breaking them into smaller steps	3.43 (0.97)	0.48	0.83
28. I am able to think of new ideas or other ways to solve a problem	3.86 (0.87)	0.53	0.83
29. I come up with creative ideas that other people find useful	3.73 (0.95)	0.51	0.83
30. I think about my options before I make a decision	3.84 (0.94)	0.57	0.82
31. I make good decisions in my personal life	3.86 (0.96)	0.59	0.82
32. I have ways to help me remember or help me learn	3.80 (0.91)	0.63	0.82
33. I think about how I am growing or progressing	3.89 (0.96)	0.58	0.82
34. To overcome challenges, sometimes I need to change what I do or try a different approach	3.96 (0.90)	0.53	0.83

*Note.** Negatively worded items have been reverse-coded. Items have been rearranged for analysis purposes to keep the items that belong to the same subscales together.

Table 6

Option response rates and nonresponse rates for each item

Survey Items	%					
	SD	D	N	A	SA	NR
Item1. I concentrate on my schoolwork	1.3	2.9	22.5	47.3	25.4	.6
Item2. I am a responsible student	1.1	3.4	21.0	44.5	29.3	.7
Item3. I complete my schoolwork regularly	2.8	7.9	23.2	36.3	28.4	1.3
Item4. I take responsibility for working on my goals	1.1	2.7	19.5	45.4	30.0	1.4
Item5. I am hopeful about my future	1.4	2.1	10.9	31.2	52.9	1.5
Item6. I know I will graduate from high school	1.5	2.1	13.2	24.6	56.7	1.8
Item7. I am able to control my emotions	3.5	6.5	23.6	36.3	28.4	1.8
Item8. I am good at staying focused on my goals	1.7	7.3	31.0	40.6	18.1	1.3
Item9. I listen to other ideas even if those ideas are different from mine	1.9	3.8	19.8	46.4	26.8	1.3
Item10. I work well in a group or team	3.6	6.1	22.8	37.1	28.5	1.9
Item11. I am good at sharing ideas in front of other people	7.3	13.3	28.7	30.5	18.6	1.7
Item12. I am good at writing thoughts and ideas	3.7	9.9	29.0	35.1	20.4	1.9
Item13. I am able to handle problems with people without fighting	3.4	5.4	25.5	37.1	26.8	1.8
Item14. I feel that I belong to a group of friends	4.5	6.9	20.6	32.7	33.5	1.9
Item15. I am comfortable asking my teacher(s) for help	3.6	8.1	23.8	36.4	26.6	1.5
Item16. I am involved in activities to support my school or community	8.5	15.0	25.8	26.9	21.8	2.0
Item17. I can do even the hardest school work if I try	1.8	6.0	22.8	38.6	29.5	1.3
Item18. I'm certain I can master the skills taught in class this year	2.1	6.0	26.9	40.8	23.1	1.2
Item19. I really can't do much to change how smart I am	8.1	18.0	27.4	25.4	18.7	2.3
Item20. I can do almost all the work in class if I don't give up	1.6	3.4	15.1	40.9	37.7	1.4
Item21. I can learn new things but I can't really change how smart I am	11.6	23.2	29.2	21.3	12.9	1.8
Item22. I understand my family culture and values	2.1	3.8	14.6	35.6	42.5	1.5
Item23. How smart I am is something that I can change	4.2	7.1	26.5	33.3	26.9	2.1
Item24. Even if the work is hard, I can learn it	1.9	4.4	21.7	40.3	30.2	1.6
Item25. I'm certain I can figure out how to do the most difficult class work	2.5	7.6	30.8	36.5	20.9	1.7
Item26. I make good decisions in my school life	1.8	5.4	25.2	41.4	24.5	1.7
Item27. I solve problems by first breaking them into smaller steps	3.7	10.9	35.7	35.5	12.5	1.7
Item28. I am able to think of new ideas or other ways to solve a problem	1.5	4.3	22.9	47.2	22.7	1.4
Item29. I come up with creative ideas that other people find useful	2.0	6.7	29.3	37.8	22.4	1.8
Item30. I think about my options before I make a decision	2.4	4.5	24.3	42.3	24.8	1.7
Item31. I make good decisions in my personal life	2.4	5.0	24.0	40.5	26.8	1.3
Item32. I have ways to help me remember or help me learn	1.8	5.2	26.3	43.3	22.1	1.4
Item33. I think about how I am growing or progressing	2.3	4.8	23.1	40.0	28.7	1.2
Item34. To overcome challenges, sometimes I need to change what I do or try a different approach	1.8	3.3	21.7	42.9	29.5	.8

Note. Total *N* = 5,983. SD = “Strongly Disagree”, D = “Disagree”, N = “Neutral”, A = “Agree”, SA = “Strongly Agree”, NR = “No Response”.

Three items had corrected item-total correlations below .30 (19, 21, and 22), and four additional items had correlations below .40 (7, 14, 15, and 16).

Reliability

Overall internal consistency of the Student Skills and Beliefs Survey as measured by Cronbach's α is .91. Internal consistency of each subscale was also calculated: academic behaviors (α = .78), future orientation (α = .72), interpersonal skills (α is .69), sense of belonging (α = .51), mindsets (α = .73), and thinking and learning (α is .84). In addition, Cronbach's α if item deleted was calculated for each subscale (see Table 5). Results show that none of the items would raise the alpha level of any subscale by .05 if removed; in fact, in most cases alpha levels would suffer if items were to be deleted.

Exploratory Factor Analysis (EFA)

Factorability. Most items had moderate correlations (.30-.60) with several other items, suggesting reasonable factorability (Tabachnic & Fidell, 2007). Keiser-Meyer-Olkin (KMO) measure of sampling adequacy was .96, while the Bartlett's Test of Sphericity was significant (χ^2 (561) = 29412.33, $p < .001$), both indicating that data is suitable for factor analysis (Tabachnic & Fidell, 2007; Williams, Brown, & Onsman, 2010).

Extraction. Principal axis factoring (PAF) method was used to extract the factors. The initial results show that 6 factors have eigenvalues over 1.0, the first factor explaining 34% of the variance, second factor for 5.7% of the variance, third factor for 4.6%, fourth for 3.5% of the variance, and fifth and sixth factors had eigenvalues very close to 1.0, each explaining 3% of the variance. Figure 1 displays the scree plot of the initial EFA analysis. The scree plot suggests that three to four factors should be retained.

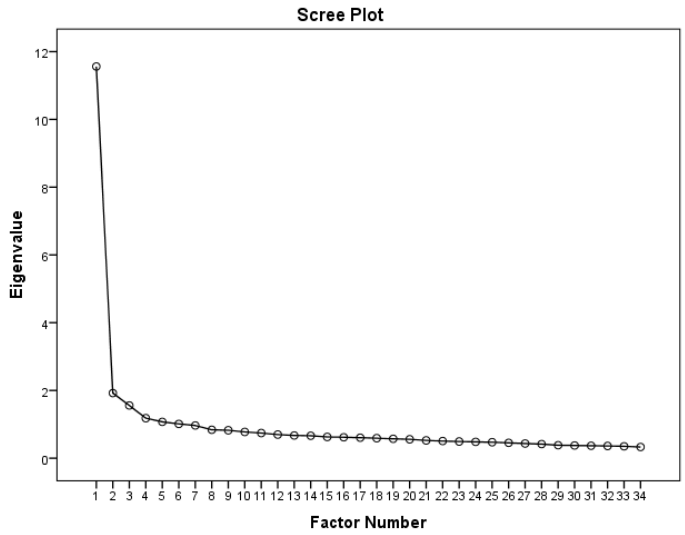

Figure 1. *Scree Test* (*SPSS Output*).

Parallel analysis was conducted next. SPSS syntax for parallel analysis has been adopted from O'Connor (2000), and can be found in Appendix A. Table 7 displays the results of the parallel analysis for the first 10 factors. Results of parallel analysis indicate that four factors should be extracted. Based on the three criteria described above, it was determined that further EFAs should be conducted to test three, four, and five-factor models.

Table 7

Parallel analysis results

Factor Number	Actual Data Eigenvalue	Random Data Eigenvalue	Retain (Y/N)
1	11.561	1.236	Y
2	1.922	1.212	Y
3	1.556	1.188	Y
4	1.181	1.169	Y
5	1.072	1.153	N
6	1.012	1.138	N
7	.965	1.123	N
8	.838	1.109	N
9	.824	1.096	N
10	.775	1.083	N

Note. Y = Yes, N = No. Factors 11-34 have been omitted.

Table 8

Factor Loadings and Communalities of the Final Four-Factor Model after Promax Rotation

Survey Item	Factors and loadings				Communalities
	1	2	3	4	
1. I concentrate on my schoolwork	.790				.573
2. I am a responsible student	.827				.568
3. I complete my schoolwork regularly	.768				.504
4. I take responsibility for working on my goals	.608				.489
5. I am hopeful about my future	.436				.340
6. I know I will graduate from high school	.387				.353
8. I am good at staying focused on my goals	.568				.496
9. I listen to other ideas even if those ideas are different from mine		.401			.321
10. I work well in a group or team		.614			.280
11. I am good at sharing ideas in front of other people		.811			.408
12. I am good at writing thoughts and ideas		.596			.369
13. I am able to handle problems with people without fighting		.384			.322
15. I am comfortable asking my teacher(s) for help		.395			.313
17. I can do even the hardest school work if I try			.742		.459
18. I'm certain I can master the skills taught in class this year			.579		.477
19. I really can't do much to change how smart I am*				.845	.689
20. I can do almost all the work in class if I don't give up			.480		.440
21. I can learn new things but I can't really change how smart I am*				.720	.521
23. How smart I am is something that I can change				.373	.360
24. Even if the work is hard, I can learn it			.568		.545
25. I'm certain I can figure out how to do the most difficult class work			.808		.611
26. I make good decisions in my school life	.809				.586
28. I am able to think of new ideas or other ways to solve a problem		.470			.390
29. I come up with creative ideas that other people find useful		.644			.443
30. I think about my options before I make a decision		.339			.337
32. I have ways to help me remember or help me learn		.371			.466
33. I think about how I am growing or progressing		.413			.375
34. To overcome challenges, sometimes I need to change what I do or try a different approach		.407			.350

Note. *Items have been reverse coded. Items 7, 14, 16, 22, 27 and 31 have been removed.

Rotation. Promax rotation method was used in this study. The communalities and rotated factor loadings of all items were examined for each model. In an iterative process, items with low communalities (< .20), items with low factor loadings (< .32), and crossloading items (items that have loadings > .30 on two or more factors) were removed to improve each model.

After conducting separate EFAs, the four-factor model emerged as the most interpretable. It had the most number of loadings above .32, fewest crossloading items, and provided the best theoretical fit of all the models tested. Table 8 on the previous page provides the final results of the four-factor EFA, including the items, factors and factor loadings from the factor pattern matrix, and communalities after extraction. Table 9 displays the correlations among the four factors. First factor consists of items 1-6, 8, and 26; second factor of items 9-13, 15, 28-30, and 32-34; third factor of items 17-20, 24, and 25; and fourth factor includes items 19, 21, and 23. Items 7, 14, and 16 have been removed due to low communalities, items 22 and 27 were removed because of low item loadings, and item 31 was a crossloading item.

Table 9

Correlations among four latent factors

Factor	1.	2.	3.	4.
1. Factor 1	---			
2. Factor 2	.71	---		
3. Factor 3	.73	.75	---	
4. Factor 4	.05	.11	.23	---

Note. $N = 2991$.

Confirmatory Factor Analysis (CFA)

Initially, the following models were tested using all 34 items: a unidimensional model in which all items load on to one factor, an a priori defined six-factor model in which items load on

to the six subscales specified by the creators of the survey, and a four-factor model suggested by the EFA prior to the removal of the items.

Model Fit. Table 10 displays model fit indices for the models tested. The results indicate that the one-factor model ($\chi^2(527) = 6982.35$, $p < .001$, CFI = .80, RMSEA = .06, SRMR = .05) and the six-factor model ($\chi^2(512) = 5423.33$, $p < .001$, CFI = .85, RMSEA = .06, SRMR = .05) are not good fits for the data, with only one of the pre-established criteria met (SRMR). The four-factor model with all the items included is the best fit out of the three models initially tested ($\chi^2(521) = 4435.80$, $p < .001$, CFI = .88, RMSEA = .05, SRMR = .06); however, it only met two out of four fit indexes criteria (RMSEA and SRMR). At this point, theoretically and statistically justifiable modifications to the four-factor model were considered.

Table 10

Model fit statistics for various models tested

Model	χ^2	df	CFI	TLI	RMSEA (95% CI)	SRMR
One-factor	6982.35***	527	.80	.79	.064 (.063-.065)	.051
Six-factor	5423.33***	512	.85	.84	.057 (.055-.058)	.047
Four-factor	4435.80***	521	.88	.87	.05 (.049-.051)	.066
Three-factor (modified)	2309.60***	296	.92	.92	.048 (.046-.05)	.037

Note. $N = 2992$. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.
*** $p < .001$.

Modification. Initial modification of the four-factor model involved removal of the six items (7, 14, 16, 22, 27, and 31) that were flagged during the EFA analysis. After that, modification considerations of the model included the examination of modification indices,

residuals, and parameter estimates. Modification indices and residuals indicated that item 23 ("How smart I am is something that I can change") would improve the model fit if it was loaded on the third factor, rather than on the fourth factor. Items 19 and 21 in the fourth factor are negatively worded items with low corrected item-total correlations. In addition, item 21 ("I can learn new things but I can't really change how smart I am") is a "double-barreled" question that consists of two separate statements, which can cause confusion among participants (DeVellis, 2012). After consideration, a decision was made to remove the two negatively worded items (19 and 21) from the analysis, and move item 23 to the third factor. This modification yielded a three-factor model with favorable improvement in the model fit, with four out of the five fit criteria met (see Table 10). Table 11 presents the correlations among the three factors.

Table 11

Correlations among three latent factors

Factor	1.	2.	3.
1. Factor 1	---		
2. Factor 2	.79***	---	
3. Factor 3	.77***	.84***	---

Note. $N = 2992$. *** $p < .001$.

Table 12 presents the factors and survey items, factor loadings, standard errors, and Cronbach's alpha coefficients for the new factors. All the standardized loadings in the modified three-factor model fall within fair to excellent range (.46 - .73), and the internal consistency measured by Cronbach's α is good for all three factors (.79 - .85). Figure 2 shows a diagram of the final measurement model.

Table 12

CFA modified six-factor model unstandardized (standardized) loadings, standard errors and p-values

Factors (Cronbach's α)/ Items	Unstandardized (Standardized) Loadings	*SE*
Factor 1 (α = .85)		
1. I concentrate on my schoolwork	1.00 (.73)***	.00
2. I am a responsible student	1.00 (.72)***	.03
3. I complete my schoolwork regularly	1.13 (.68)***	.03
4. I take responsibility for working on my goals	.90 (.67)***	.03
5. I am hopeful about my future	.73 (.53)***	.03
6. I know I will graduate from high school	.73 (.50)***	.03
8. I am good at staying focused on my goals	.99 (.68)***	.03
26. I make good decisions in my school life	1.06 (.72)***	.03
Factor 2 (α = .85)		
9. I listen to other ideas even if those ideas are different from mine	1.00 (.57)***	.00
10. I work well in a group or team	.98 (.48)***	.05
11. I am good at sharing ideas in front of other people	1.14 (.50)***	.05
12. I am good at writing thoughts and ideas	1.13 (.55)***	.05
13. I am able to handle problems with people without fighting	1.09 (.54)***	.05
15. I am comfortable asking my teacher(s) for help	1.08 (.52)***	.05
28. I am able to think of new ideas or other ways to solve a problem	1.05 (.61)***	.04
29. I come up with creative ideas that other people find useful	1.09 (.58)***	.04
30. I think about my options before I make a decision	1.11 (.60)***	.04
32. I have ways to help me remember or help me learn	1.15 (.65)***	.04
33. I think about how I am growing or progressing	1.15 (.61)***	.04
34. To overcome challenges, sometimes I need to change what I do or try a different approach	1.01 (.56)***	.04
Factor 3 (α = .79)		
17. I can do even the hardest school work if I try	1.00 (.61)***	.00
18. I'm certain I can master the skills taught in class this year	1.03 (.65)***	.04
20. I can do almost all the work in class if I don't give up I can do almost all the work in class if I don't give up	.89 (.59)***	.03
23. How smart I am is something that I can change	.83 (.46)***	.04
24. Even if the work is hard, I can learn it	1.14 (.72)***	.04
25. I'm certain I can figure out how to do the most difficult class work	1.18 (.72)***	.04

Note. N = 2992. CFA = confirmatory factor analysis. Items 7, 14, 16, 19, 21, 22, 27 and 31 have been removed.
*** $p < .001$.

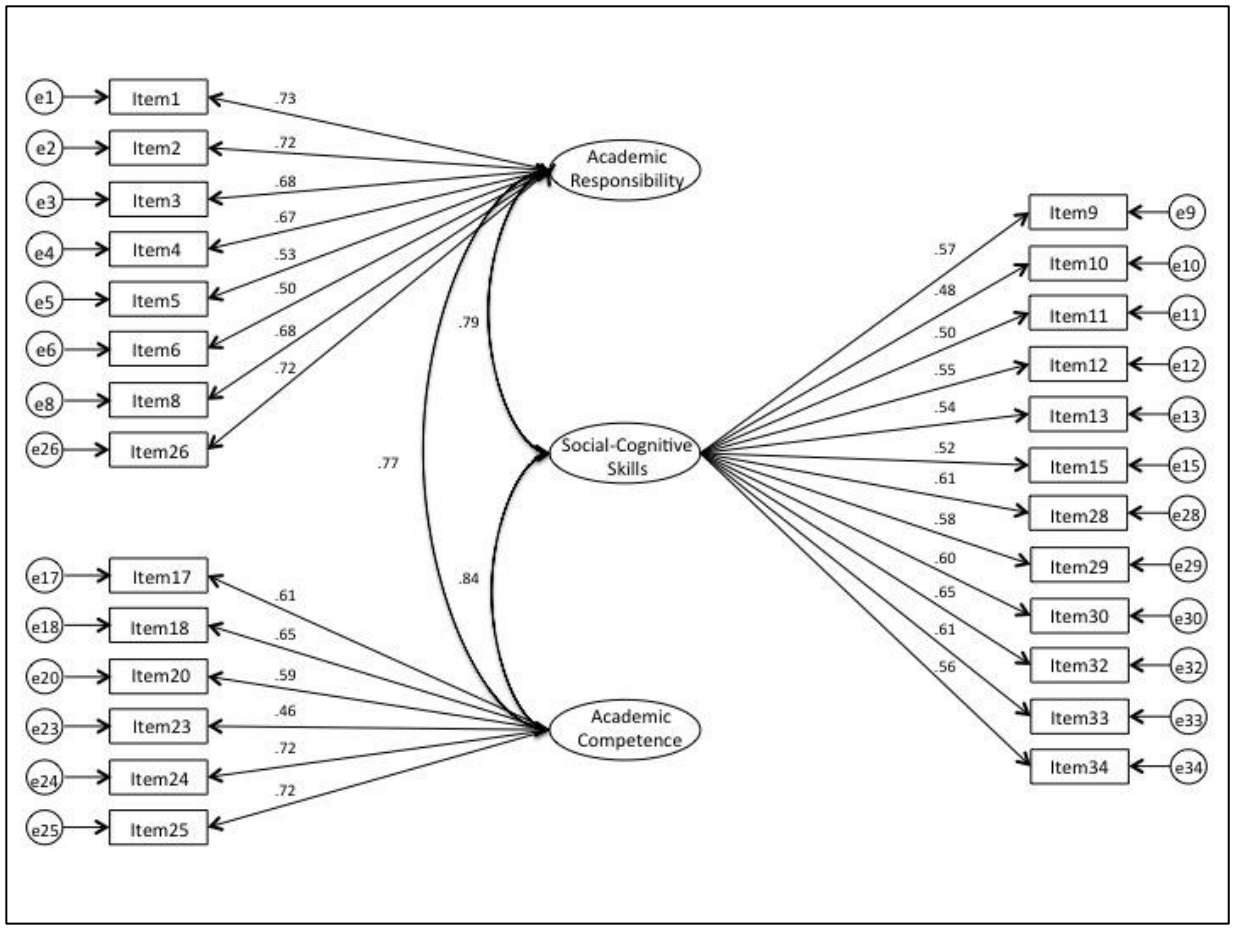

Figure 2. *Modified three-factor model of 26 items from "Student Skills and Beliefs" survey with standardized loadings.*

Measurement Invariance

Measurement invariance was tested for gender, ELL-status, and grade band by conducting multiple-group CFAs in Mplus version 7. Table 13 presents the results of this analysis. Configural invariance models were tested by specifying groups (e.g. males = 1, females = 2), while leaving all the parameters unrestricted. To test metric invariance, factor loading parameters were restricted, in other words, the factor loadings for different groups were forced to be equal among the groups.

Table 13

Fit indices for invariance tests of the three-factor modified model

Model	χ^2	df	$\Delta\chi^2$	Δdf	CFI	ΔCFI	RMSEA	SRMR
Gender: Males (n = 1440), Females (n = 1447)								
Configural invariance	2850.61***	618			.91		.05	.04
Metric invariance	3684.77***	621	834.16***	3	.88	-.03	.06	.30
ELL status: ELL (n = 728), nonELL (n = 2213)								
Configural invariance	2809.42***	618			.91		.05	.04
Metric invariance	3810.95***	621	1001.53***	3	.88	-.03	.06	.28
Grade: Elementary (n = 838), Middle (n = 613), High School (n = 1499)								
Configural invariance	3868.22***	940			.89		.06	.06
Metric invariance	-	-	-	-	-	-	-	-

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.
*** $p < .001$.

Results of the multiple-group CFA show that factor structure is stable across gender (CFI = .91, RMSEA = .05, SRMR = .04) and ELL status (CFI = .91, RMSEA = .05, SRMR = .04), although the overall chi-squared tests were significant, $\chi^2(618) = 2850.61$, $p < .001$ and $\chi^2(618) = 2809.42$, $p < .001$, respectively. Configural invariance for different grade bands did not hold, $\chi^2(940) = 3868.22$, $p < .001$, CFI = .89, RMSEA = .06, SRMR = .06.

Gender results for metric invariance show that $\Delta\chi^2(3) = 834.16$, $p < .001$, ΔCFI = -.03, RMSEA = .06, SRMR = .30, indicating poor fit. Similar pattern can be seen in multi-group comparison for ELL and nonELL groups, $\Delta\chi^2(3) = 1001.53$, $p < .001$, ΔCFI = -.03, RMSEA = .06, SRMR = .28. Metric invariance for grade level comparisons was not tested due to the fact that configural invariance was not established.

Partial invariance. Factor loadings and modification indices in unrestricted models were examined to determine which items are potentially non-invariant (Van de Schoot et al, 2012). Next, backwards elimination was used (Millsap 2007), whereby the factor loading restrictions

were imposed on all but one of the items and the model fit was examined after each one to determine which items were non-invariant. Items 1, 9, and 17 were identified as non-invariant. When factor loading constraints on these items were relaxed, metric invariance model fit greatly improved for both gender ($\Delta\chi^2(23) = 50.58$, $p < .001$, ΔCFI = -.001, RMSEA = .05, SRMR = .06) and ELL comparisons ($\Delta\chi^2(23) = 50.38$, $p < .001$, ΔCFI = -.001, RMSEA = .05, SRMR = .05). Thus, partial invariance has been established (see Table 14).

Table 14

Fit indices for metric invariance tests with factor loading constraints relaxed on non-invariant items (1, 9, & 17)

Model	χ^2	df	$\Delta\chi^2$	Δdf	CFI	ΔCFI	RMSEA	SRMR
Gender: Male (n = 1140), Female (n = 1447)	2901.19***	641	50.58***	23	.91	-.001	.05	.06
ELL status: ELL (n = 728), nonELL (n = 2213)	2859.80***	641	50.38***	23	.91	-.001	.05	.05

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.
*** $p < .001$.

Discussion

The purpose of this study is to help advance further development of Student Skills and Beliefs Survey instrument. In order to achieve this goal, several analyses were performed: item and reliability analyses were performed to obtain initial psychometric information about individual items and the six subscales; exploratory and confirmatory factor analyses were conducted to check the dimensionality of the survey instrument, and measurement invariance analyses by gender, ELL status, and grade band was performed to check whether measurement model is stable across groups. This section describes the results and recommendations for further development of the survey instrument, as well as possible future analysis and limitations.

Item Analysis and Reliability

Analyses of item means, standard deviations, skewness and kurtosis, and of item response rates did not raise any concerns. All items were slightly negatively skewed as all means were above 3.0, however this is quite common in the literature, and the values of skew and kurtosis were within the acceptable range of +/- 2 (Bentler & Wu, 2003), suggesting that it is acceptable to conduct further psychometric analysis on this data. Corrected item-total correlations revealed that items 19, 21, and 22 had item-total correlations below .30, and items 7, 14, 15, and 16 had item-total correlations below .40. These low values indicate that these items may not be measuring the same construct as the rest of the survey.

Overall internal consistency of the Student Skills and Beliefs Survey is excellent, Cronbach's α = .91. Internal consistency of most subscales is within the acceptable level (α ranging from .69 to .84), with one exception of the sense of belonging subscale (α = 0.51). Low internal consistency of this subscale could be due to low number of items (n = 3), low corrected item-total correlations of items (14, 15, and 16), or because the items belong to different subscales. As was later revealed by factor analysis, two of these items had to be removed from analysis due to low communalities and factor loadings (14 and 16), and item 15 loaded on to the new social-cognitive factor together with interpersonal skills items.

Factor Analysis

While the makers of the survey had preconceived notions about the structure of the instrument, exploratory and confirmatory factor analyses revealed that in this current version, Student Skills and Beliefs survey does not measure six different factors. EFA analysis revealed that a four-factor model was the best fit for the data after considering item loadings, communalities, the number of crossloading items, and theoretical implications. However, the

fourth factor consisted of only three very similarly worded items (19, 21, and 23), indicating that it is not a very strong factor. In addition, items 19 and 21 were flagged during the item analysis because of low corrected item-total correlations.

CFA results confirmed our suspicions about the fourth factor. Results indicated that model fit would be improved if item 23 was to load on to the third factor. Therefore, a decision was made to disassemble the fourth factor by deleting items 19 and 21 and moving item 23 to the third factor. These modifications resulted in significantly improved model fit of the final three-factor model.

The first factor includes items from academic behavior and future orientation subscales, as well as one item from the thinking and learning subscale ("I make good decisions in my school life"). It makes sense that students who view school and learning as important for their future would make good decisions about academic activities, which in turn results in exhibiting good academic behaviors such as completing school work regularly. Contrary to belief that the items in academic behavior and future orientation measure two separate things, the results of the factor analysis show that they load on to one factor, renamed academic responsibility.

The second factor includes items from interpersonal skills and thinking and learning subscales. Items in these subscales share common themes, such as coming up with and sharing new ideas, finding new and creative solutions to problems, and working with others. This factor has been renamed social-cognitive skills.

The third factor is comprised of the five items that were borrowed from PALS academic efficacy subscale (Midgley et al., 1998), and one item (Item 23 – "How smart I am is something that I can change"), which was taken from the Theory of Intelligence Scale (Blackwell et el., 2007). This factor has been renamed academic competence.

Although the new factors do not map out neatly to the framework of school engagement described by Fredricks et al. (2004), multidimensional nature of the survey instrument was supported by factor analysis. The results of this study reflect the difficulties in measuring school engagement described in the literature. Student behavior, emotion, and cognition are all interconnected within each individual, creating difficulties when trying to separate them into different factors.

Several items were identified as problematic during factor analysis. Items 7, 14, and 16 were removed due to low item communalities. Items 22 and 27 were removed due to low factor loadings. Item 31 was a crossloading item, with factor loadings above .32 on first and second factors. Items 19 and 21 were removed due to low item-total correlations and the fact that they loaded onto an unstable fourth factor.

Measurement Invariance

Multiple-group CFA analysis revealed that configural invariance is not stable across different grade levels. This means that the instrument is not measuring the same constructs across elementary, middle, and high-school grade bands. This is problematic, as it precludes us from making any meaningful comparisons between different age groups. This finding is consistent with existing literature on school engagement, and may explain why most studies choose to focus on either high school, elementary, or middle school students. It is possible that students of different ages perceive these items in different ways.

Configural invariance of the instrument is stable across gender and ELL-status, however, metric invariance did not hold. That means that while the same constructs are being measured for gender and ELL groups, the scale of the measurement (units) are not the same across different groups. This creates a problem and precludes us from meaningfully comparing these groups on

their levels of school engagement. Fortunately, partial invariance can still be established when metric invariance does not hold. By constricting factor loadings of all but one item in an iterative process and comparing model fit indices, non-invariant items were identified. The following items were identified as non-invariant across both gender and ELL status comparisons:

- Item1 – I concentrate on my schoolwork
- Item9 – I listen to other ideas even if those ideas are different from mine
- Item17 – I can do even the hardest school work if I try

Results show that by relaxing factor loadings on these items, partial invariance can be reached for both gender and ELL comparisons. It seems that students interpret these items in different ways depending on which group they belong to (e.g. males and females can have different ideas on what it means to concentrate on one's schoolwork).

Summary

Contrary to the belief of survey developers, the Student Skills and Beliefs Survey in its' current version does not measure six different constructs. Factor analysis revealed that a three-factor structure is the best fit for the data, with the following factors: academic responsibility (8 items, α = .85), social-cognitive skills (12 items, α = .85), and academic competence (6 items, α = .79).

Measurement invariance testing revealed that the factor structure was found to be stable across gender and ELL groups, however it is unfortunately not stable across different grade bands. This precludes any meaningful comparisons on the results of the survey among elementary, middle, and high school age groups. Further testing revealed that metric invariance does not hold for gender and ELL group comparisons, but fortunately partial metric invariance could be achieved by removing three non-invariant items from the analysis (1, 9, and 17).

Recommendations

Items 1, 9 and 17 were found to be non-invariant and therefore it is recommended that they be dropped from future versions of the survey. In addition, items 19 and 21 should be removed due to the fact that they have very low item-total correlations and were worded similar to item 23, only negatively ("How smart I am is something that I can change"). Negative wording, as well as the double-barreled nature of item 21 may have created confusion. Item 31 is also recommended for removal due to being a crossloading item.

In addition I recommend the developers either drop items 7, 14, 16, 22, and 27, or to add more items that measure those constructs. These items were removed during analysis due to low factor loadings and/or communalities, which indicate that they do not measure the same constructs as the rest of the items. Nevertheless, if those constructs are important, it may be necessary to keep these items and add more items that measure the same things.

It is also recommended that different forms of Student Skills and Beliefs survey will need to be developed for elementary, middle, and high school students, due to the absence of configural invariance. It would be necessary to conduct cognitive interviews of these forms prior to administration, to ensure that students interpret the items in an intended way.

Limitations and Future Research

Given that the data is from only one school district in South King County, WA, the results of this study may not be generalizable to other populations. Fortunately, the sample is quite diverse, with a large percentage of ELL students, as well as good representation of various ethnicities. Nevertheless, it will be important to administer the survey to several districts to ensure the findings can be generalized to a broader population.

Further analysis needs to be conducted to determine the source of configural invariance for different grade bands. Once metric invariance is established across all groups of interest, further invariance testing can be conducted. This includes testing for intercept invariance together with metric invariance (also known as scalar invariance), and a test for strict measurement invariance, in which factor loadings, intercepts, and errors are constrained to be equal among all groups (Wu et al., 2007).

In addition, future analysis needs to be conducted to examine the relationship between student academic achievement and the scores on this survey, as well as the relation between other variables important for school success, such as attendance and graduation rates.

Conclusion

According to Fredricks et al. (2004), conceptual clarity can sometimes be sacrificed for practical reasons when it comes to measuring school engagement. In this current version, the results of Student Skills and Beliefs survey would preclude us from comparing differences in emotional, behavioral, and cognitive levels of engagement. However, if results of this simple instrument help identify students at risk for school disengagement, or predict academic achievement, its' use may still be justified.

References

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, *44*(5), 427-445.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359-373.

Bandura, M., & Dweck, C. S. (1985). The relationship of conceptions of intelligence and achievement goals to achievement-related cognition, affect and behavior. *Unpublished manuscript, Harvard University*.

Bentler, P. M., & Wu, E. J. C. (2003). *EQS for Windows User's Guide*. Encino, CA: Multivariate Software, Inc.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, *78*(1), 246-263.

Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis.* Lawrence Erlbaum Associates. Hillsdale, NJ.

Connell, J. P. (1990). Context, self, and action: A motivational analysis of self-system processes across the life span. *The self in transition: Infancy to childhood*, 61-97.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes.

Costello, A. B., & Osborne, J. W. (2005). *Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis.* Retrieved April 23, 2014 from http:/pareonline.net/pdf/v10n7.pdf.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihalyi and I. Csikszentmihalyi (Eds.), *Optimal experience* (pp. 15-35). Cambridge, UK: Cambridge University Press.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method.* Hoboken, NJ: John Wiley & Sons, Inc.

Dornbusch, S. M., & Steinberg, L. (1990). *Measures of school engagement*. Unpublished manuscript, Temple University, Philadelphia.

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

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on motivation in education, 3,* 139-186.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: the impact of stage-environment fit on young adolescents' experiences in schools and in families. *American psychologist*, *48*(2), 90.

Eftekhari-Sanjani, H. (2008). Study of Item Nonresponse for a Sample of Questionnaires from the 2007 Commodity Flow Survey (CFS). *Section on Survey Research Methods.* Washington, D.C..

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. Oxford University Press.

Finn, J. D. (1989). Withdrawing from school. *Review of educational research*, *59*(2), 117-142.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of applied psychology*, *82*(2), 221.

Field, A., (2005). *Discovering Statistics Using SPSS*. 2nd ed. London: Sage.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109.

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In *Handbook of research on student engagement* (pp. 763-782). Springer US.

Geiser, C. (2012). *Data analysis with Mplus*. New York, NY: Guilford Publications.

George, D., & Mallery, P. (2006). *SPSS for windows step by step: A simple guide and reference.* Boston, MA: Allyn Bacon.

Guthrie, J.T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. Kamil & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 3, pp. 403-422). Mahwah, NJ: Lawrence Erlbaum.

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

Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist*, *8*(1), 7-27.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of school health*, *74*(7), 262-273.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Lamborn, S., Newmann, F., & Wehlage, G. (1992). The significance and sources of student engagement. *Student engagement and achievement in American secondary schools*, 11-39.

Leong, F. T., & Austin, J. T. (Eds.). (2006). *The psychology research handbook: A guide for graduate students and research assistants*. Sage.

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American educational research journal*, *37*(1), 153-184.

McCoach, D. B. (2002). A Validation Study of the School Attitude Assessment Survey. *Measurement and Evaluation in Counseling and Development*, *35*(2), 66-77.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525-543.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., ... & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. Ann Arbor, MI.

Millsap, R. E. (2007). Invariance in measurement and prediction revisited. *Psychometrika*, *72*(4), 461-473.

National Research Council and Institute of Medicine. (2004). *Engaging Schools: Fostering high school students' motivation to learn*. Washington, DC: The National Academies Press.

Newmann, F. M. (1992). *Student Engagement and Achievement in American Secondary Schools*. New York, NY: Teachers College Press.

O'connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments, & computers*, *32*(3), 396-402.

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of educational research*, *70*(3), 323-367.

Perry, J. C., Liu, X., & Pabian, Y. (2010). School engagement as a mediator of academic performance among urban youth: The role of career preparation, parental career support, and teacher support. *The Counseling Psychologist*, *38*(2), 269-295.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, *82*(1), 33.

Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and psychological measurement*, *53*(3), 801-813.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common core standards the new US intended curriculum. *Educational Researcher*, *40*(3), 103-116.

Psychological Corporation. (1970). *Stanford Achievement Test*. San Diego, CA: Harcourt, Brace, Jovanovich.

Raymond, M., DeBacker, T., & Green, B. (1999). Perceived instrumentality and academics: The link to task valuing. *Journal of Instructional Psychology*, *26*(4).

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: examining reasons for acting in two domains. *Journal of personality and social psychology*, *57*(5), 749.

Sedaghat, M., Abedin, A., Hejazi, E., & Hassanabadi, H. (2011). Motivation, cognitive engagement, and academic achievement. *Procedia-Social and Behavioral Sciences*, *15*, 2406-2410.

Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of educational psychology*, *82*(1), 22.

Tabachnick, B. G., & Fidell, L. S. (2007). Multivariate analysis of variance and covariance. *Using multivariate statistics.* Boston, MA: Allyn & Bacon.

Traub, R. E. (1994). *Reliability for the social sciences: Theory and applications* (Vol. 3). Sage.

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486-492.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate behavioral research*, *25*(1), 1-28.

Voelkl, K. E. (1996). Measuring students' identification with school. *Educational and Psychological Measurement, 56,* 760-770.

Voelkl, K. E. (1997). Identification with school. *American Journal of Education, 105*, 294-318.

Wang, M. T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child development*, *83*(3), 877-895.

Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, *47*(3), 633-662.

Wang, M. T., Willett, J. B., & Eccles, J. S. (2011). The assessment of school engagement: Examining dimensionality and measurement invariance by gender and race/ethnicity. *Journal of School Psychology*, *49*(4), 465-480.

Wellborn, J. G., & Connell, J. P. (1987). Manual for the Rochester assessment package for schools. *Unpublished manuscript, University of Rochester, New York.*

Wentzel, K. R. (1994). Relations of social goal pursuit to social acceptance, classroom behavior, and perceived social support. *Journal of Educational Psychology*, *86*(2), 173.

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care*, *8*(3).

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 and Evaluation*, *12*(3), 1-26.

Yazzie-Mintz, E. (2007). Voices of students on engagement: A report on the 2006 high school survey of student engagement. *Center for Evaluation and Education Policy, Indiana University*, *12*.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist*, *25*(1), 3-17.

Appendix A

SPSS Syntax for Parallel Analysis

```
set mxloops=9000 length=none printback=none width=80 seed = 1953125.
matrix.
* enter your specifications here.
compute Ncases = 3000.
compute Nvars = 34.
compute Ndatsets = 100.
compute percent = 95.
* computing random data correlation matrices & eigenvalues.
compute evals = make(nvars,ndatsets,-9999).
compute nm1 = 1 / (ncases-1).
loop #nds = 1 to ndatsets.
compute x = sqrt(2 * (ln(uniform(ncases,nvars)) * -1) ) &* cos(6.283185 *
uniform(ncases,nvars) ).
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute evals(:,#nds) = eval(d * vcv * d).
end loop.
* identifying the eigenvalues corresponding to the desired percentile.
compute num = rnd((percent*ndatsets)/100).
compute results = { t(1:nvars), t(1:nvars), t(1:nvars) }.
loop #root = 1 to nvars.
compute ranks = rnkorder(evals(#root,:)).
loop #col = 1 to ndatsets.
do if (ranks(1,#col) = num).
compute results(#root,3) = evals(#root,#col).
break.
end if.
end loop.
end loop.
compute results(:,2) = rsum(evals) / ndatsets.
compute specifs = {ncases; nvars; ndatsets; percent}.
print specifs /title="Specifications for this Run:" /rlabels="Ncases" "Nvars" "Ndatsets"
"Percent".
print results /title="Random Data Eigenvalues" /clabels="Root" "Means" "Prcntyle".
end matrix
```

Appendix B

Student Skills & Beliefs Survey

V1.0 ENGLISH VERSION

Marking Instructions: You may use pen or pencil.
Please completely fill-in the bubble for your response.

This survey is confidential and anonymous.

Read each of the following statements and decide how much you agree or disagree with them.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I concentrate on my schoolwork	○	○	○	○	○
I am a responsible student	○	○	○	○	○
I can do even the hardest school work if I try	○	○	○	○	○
I take responsibility for working on my goals	○	○	○	○	○
I am hopeful about my future	○	○	○	○	○
I solve problems by first breaking them into smaller steps	○	○	○	○	○
I am able to control my emotions	○	○	○	○	○
I am good at staying focused on my goals	○	○	○	○	○
I am certain I can master the skills taught in class this year	○	○	○	○	○
I really can't do much to change how smart I am	○	○	○	○	○
I can do almost all the work in class if I do not give up	○	○	○	○	○
I make good decisions in my school life	○	○	○	○	○
I listen to ideas from other people even if those ideas are different from mine	○	○	○	○	○
I am able to think of new ideas or other ways to solve a problem	○	○	○	○	○
I complete my schoolwork regularly	○	○	○	○	○
I can learn new things but I cannot really change how smart I am	○	○	○	○	○
I understand my family culture and values	○	○	○	○	○
I feel that I belong to a group of friends	○	○	○	○	○
I am comfortable asking my teacher(s) for help	○	○	○	○	○
I am involved in activities to support my school or community	○	○	○	○	○
How smart I am is something that I can change	○	○	○	○	○
I work well in a group or team	○	○	○	○	○
I am good at sharing ideas in front of other people	○	○	○	○	○
I am good at writing thoughts and ideas	○	○	○	○	○
I am able to handle problems with people without fighting	○	○	○	○	○
Even if the work is hard, I can learn it	○	○	○	○	○
I come up with creative ideas that other people find useful	○	○	○	○	○
I know I will graduate from high school	○	○	○	○	○
I think about my options before I make a decision	○	○	○	○	○
I am certain I can figure out how to do the most difficult class work	○	○	○	○	○
I make good decisions in my personal life	○	○	○	○	○
I have ways to help me remember or help me learn	○	○	○	○	○
I think about how I am growing or progressing	○	○	○	○	○
To overcome challenges, sometimes I need to change what I do or try a different approach	○	○	○	○	○

Please give use a little information about yourself. Remember that all information is confidential and your responses will not be shared with anyone at your school or district.

What is your gender?

- ○ female
- ○ male
- ○ other

Within the last month, how often were you absent from school?

- ○ Never
- ○ 1 day
- ○ 2 - 4 days
- ○ 5 or more days

What is your race / ethnic origin?

- ○ American Indian/Alaskan Native
- ○ Asian
- ○ Pacific Islander/Native Hawaiian
- ○ Black/African American
- ○ Hispanic/Latino of any race
- ○ White
- ○ Two or more Races

In which activities do you participate?

- ○ sports
- ○ music
- ○ cultural programs
- ○ theater/drama
- ○ student council/student government
- ○ other afterschool programs
- ○ other
- ○ none of the above

Name of your school?

What grade are you in?

- ○ 4th
- ○ 5th
- ○ 6th
- ○ 7th
- ○ 8th
- ○ 9th
- ○ 10th
- ○ 11th
- ○ 12th

What were your grades like last year?

- ○ mostly 1's,
- ○ mostly 2's,
- ○ mostly 3's,
- ○ mostly 4's
- ○ mostly A's,
- ○ mostly B's,
- ○ mostly C's,
- ○ mostly D's
- ○ mostly E/F's

How often is English spoken in your home?

- ○ Always
- ○ Most of the time
- ○ Sometimes
- ○ Rarely or never

What are you most likely to do after you finish high school?

- ○ attend a four-year college
- ○ attend a community college
- ○ attend a vocational school
- ○ work full-time
- ○ join the military
- ○ other
- ○ don't know

I work at a paid job

- ○ 0 hours/week
- ○ 1-10 hours/week
- ○ 10-20 hours/week
- ○ more than 20 hours/week

Thank you for your time.

The ROAD MAP PROJECT
Supported by CCER

YOUTH DEVELOPMENT EXECUTIVES OF KING COUNTY

cee
The Center for Educational Effectiveness