LONGITUDINAL MEASUREMENT INVARIANCE ANALYSES OF THE STUDENT ENGAGEMENT INSTRUMENT – BRIEF VERSION

by

CHRISTOPHER ANTHONY PINZONE

(Under the Direction of Amy L. Reschly)

ABSTRACT

This study evaluated the psychometric properties of the Student Engagement Instrument – Brief Version (SEI-B) longitudinally across three time points with high school students in the Southeastern United States. Two subsamples of one time point were analyzed to validate the factor structure by exploratory factor analysis (40% of the sample) and confirmatory factor analysis (60% of the sample) revealing a five-factor structure in congruence with the full form of the Student Engagement Instrument (SEI). Longitudinal measurement invariance analyses were performed on each of the five imputed datasets following the suggestions and recommendations of Vandenberg and Lance (2000). The SEI-B demonstrated configural, metric, scalar, and uniqueness invariance with acceptable levels and changes of model fit across all time points and datasets suggesting it may be used as part of a comprehensive progress monitoring effort to predict students that may be at-risk to drop out of school.

INDEX WORDS: student engagement, dropout, school completion, longitudinal, confirmatory factor analysis, measurement invariance

LONGITUDINAL MEASUREMENT INVARIANCE ANALYSES OF THE STUDENT ENGAGEMENT INSTRUMENT – BRIEF VERSION

by

CHRISTOPHER ANTHONY PINZONE

B.A., Stony Brook University, 2010

A Thesis Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2016

LONGITUDINAL MEASUREMENT INVARIANCE ANALYSES OF THE STUDENT ENGAGEMENT INSTRUMENT – BRIEF VERSION

by

CHRISTOPHER ANTHONY PINZONE

Major Professor: Amy L. Reschly

Committee: Scott P. Ardoin
Stacey Neuharth-Pritchett

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

The Implications of Education and Environmental Context

Educators, researchers, parents, and all other relevant stakeholders are interested and invested in the development of successful, educated youth. A student is part of an interconnected system that, when all stars align, works reciprocally to fulfill their personal and social needs within an educational context. The importance of viewing the school as a developmental context is clear considering the amount of time an individual spends in school throughout their lives. In the United States, the majority of states require at least 990 hours of instructional time per year (Education Commission of the States [ECS], 2011) typically over the course of 13 years (i.e., K-12th grade), which does not factor in additional homework and learning support time spent outside of school.

Familial access to resources, early academic performances, and quality of social resources across the family, school, and community are only some of many significant predictors of school completion (Rumberger & Rotermund, 2012). When considering familial access to and quality of resources, one must consider how the high school graduation rate differs by ethnicity (i.e., 85% White, 67% Black, 71% Hispanic, 87% Asian, 64% American Indian; Alliance for Excellent Education, 2013). Critically, about 10% of high schools account for over 40% of high school dropouts. Native American students and students of color are roughly four times as likely to be enrolled in such schools compared to White peers (Alliance for Excellent Education, 2013). It is likely that this reflects, in part, the disproportionate representation of non-White children under the age of 18 living in poverty in 2012 (i.e., 39% Black, 36% American Indian/Alaska Native, 33% Hispanic, 25% Pacific Islander, 22% two or more races, 14% Asian, 13% Caucasian; Kena et al., 2014).

Students who fail to graduate high school report lower earnings or unemployment, are disproportionately represented in prison, are more likely to have health problems, and have increased chances of living within low socioeconomic status or on government assistance programs (Christenson et al., 2001; Rauscher, 2010; Wirt, 2004). The cost of such negative outcomes has been estimated at approximately $260,000 per dropout, totaling to over $250 billion dollars to the United States of lost earning and taxes throughout their lifetime (Rouse, 2005).

What may be even more pressing is the increased socio-economic necessity for educational attainment beyond high school graduation. The 2008-2013 economic recession had less of an impact on employment for those graduates with a bachelor's degree than those who had completed high school, with the most severe impact on those who did not complete high school at all (Kena et al., 2014). Compared to high school graduation, the outlook for four-year college graduation rates is much worse but follows similar racial-ethnic patterns (i.e., 60% White, 38% Black, 48% Hispanic, 68% Asian, 39% American Indian; Alliance for Excellent Education, 2013). It can be expected that the same skills, mindsets, and contexts that foster successful high school completion are also requisite for and related to positive post-secondary outcomes but the demands and support for students likely differ in this context.

The Importance of the Developmental Perspective for Student Engagement

This acknowledgement of differences in educational attainment and completion being attributable, in part, to factors outside of the individual is in line with other developmental metatheories, such as Bronfenbrenner's (1979) ecological model and Overton's (2013) relational developmental systems paradigm, which attempt to understand individuals through their embedded relationships within, and reciprocal interactions with, relevant environmental contexts including culture and history. Contexts such as culture and history are less frequently explicitly considered in research and practice. Relatedly, student engagement has often been viewed with developmental contexts in mind (Reschly & Christenson, 2012). Intervention and improvement of developmental contexts and relationships should enhance student learning, achievement, and identification with school. In fact, successful student engagement interventions such as Check & Connect focus on factors beyond the school environment that impact school performance and behavior in meaningful ways. Check & Connect assigns mentors to intervene not only on the student level, but with their families as well (Christenson & Reschly, 2010). In other words, student engagement needs to be viewed as a "system of systems" that cannot be separated from, and must be studied in relation to, one another (Crick, 2012). Systems-wide action and improvements may protect individuals from the negative outcomes for individuals associated with school dropout. As other researchers have noted, this means that aspects of the school environment alone are not sufficient to accomplish the goals of schooling (Reschly & Christenson, 2012). In fact, the operationalization of engagement itself may differ across contexts (i.e., engagement with school vs. engagement in learning activities) which implies different types of outcomes, determinants, and intervention based on the context of one's engagement (Janosz, 2012).

There are, of course, demographic variables associated with school completion; however, other variables may also be found within family, school, and community levels. Table 1.1 highlights many key alterable variables across contexts which correlate with high school dropout and completion (Reschly & Christenson, 2006b; Rosenthal, 1998). An important distinction to be made when considering all of the variables related to dropout and completion are whether they are alterable or amenable to intervention. Although status or other demographic variables may be useful to guide identification procedures, these variables do little to inform intervention efforts. However, alterable variables are those characteristics at different levels (i.e., individual, family, school) which directly impact behavior and prepare students for success (Reschly & Christenson, 2012).

Developmental Models of Student Engagement

Developmental models of student engagement have primarily evolved from Finn's (1989) seminal Participation-Identification Model. According to this model, participation and identification with school are on-going long-term processes rather than isolated occurrences. In this model, participation is considered to be students' behavioral (e.g., homework and classwork completion, answering questions during class, paying attention) and social (e.g., following rules, appropriately interacting with peers, attending class and school) engagement in classroom and school activities, their initiative-taking behaviors (e.g., seeking help, doing more than is required), and whether they attend academic extracurricular activities (Finn & Zimmer, 2012). When students are adequately equipped with necessary starting skills and experience success with early participatory behaviors, it begins a cycle which forms an affective bond (i.e., identification) with school that encourages continued participation, success, and autonomy (Finn, 1989). Many studies have noted the associations between participation and student success across grade levels (e.g., classroom and extracurricular participation) and beyond to postsecondary outcomes (Feldman & Matjasko, 2005; Finn & Cox, 1992; Finn, 2006). In addition, students' identification with school develops in the early grades and crystalizes over time. Identification is a strong motivator of school and classroom behavior and is a protective factor which may offset many of the deleterious effects of other contexts on school performance and contribute to the process of school completion (Finn & Rock, 1997; Voelkl, 2012).

These processes are viewed on a continuum which includes non-participation and lack of identification with school. Thus, student-level variables indicative of early school-withdrawal include poor attendance and behavior, low levels of belonging or identification with school, and general disinterest in learning (Finn, 1989). There has been considerable accuracy in the prediction of school dropout or completion which include many of these behavioral indicators from time points as early as elementary and middle school (Alexander, Entwisle, & Horsey, 1997; Barrington & Hendricks, 1989; Bowers et al., 2012; Reschly & Christenson, 2006; Schoeneberger, 2012). These early signs of disengagement from school often precede more severe learning, attendance, and behavior problems that culminate in various negative outcomes including dropping out (Reschly & Christenson, 2012).

Importantly, Finn's Participation-Identification model also noted that the skills, behaviors, and attitudes that children have acquired before they enter schools play an important role in the process of engagement, disengagement, and future identification with school. As research on student engagement has continued, it has extended beyond direct student-level processes to include contexts outside of the school as points amenable to intervention (Reschly & Christenson, 2012). Finn's model already recognized that student engagement was beyond unidimensional, as behavior and affect were both theorized to take part in the participation-identification process. Additionally, researchers have come to a consensus that the construct of student engagement is a multi-dimensional construct although they vary in definition by theoretical perspective (Fredericks, Blumenfeld, & Paris, 2004; Yazzie-Mintz & McCormick, 2012). The malfunctioning or dysfunction of any of these systems, dimensions, or relationships may impact students within the school environment and in general.

School completion and dropout are arguably the most researched outcomes within the field of student engagement. However, as many engagement researchers have noted, these ongoing developmental processes extend beyond high school completion to post-secondary outcomes making them relevant to all students (Christenson, Reschly, & Wylie, 2012a; Finn, 2006; Voelkl, 2012; Yazzie-Mintz & McCormick, 2012;). Their implications extend beyond schooling to having lifelong consequences, with higher levels of engagement being associated with social-emotional well-being, a lower likelihood of participating in risky sexual and health behaviors, future work success, and allowing the acquisition of basic proficiencies for successful social integration (Christenson et al., 2012a; Griffiths, Lillies, Furlong, & Sidhwa, 2012).

Although there are widespread differences in the theoretical conceptualizations of engagement, most current student engagement theories include students' affective (e.g., belonging, identification), behavioral (e.g., attendance, suspensions, participation), and cognitive (e.g., self-regulation, investment in learning) engagement in some form (Appleton et al., 2008; Fredricks et al., 2004; Reschly & Christenson, 2012). Affective engagement is an internal state which partly results from interactions and experiences, including those across any given student's history which contribute to students' feelings about the school, teachers, and/or peers (Appleton et al., 2006; Jimerson, Campos, & Greif, 2003; Voelkl, 2012). Behavioral engagement often consists of observable indicators that have been regularly collected by schools and may often significantly predict whether students are less likely to complete school, but are high inference indicators when used to represent students' cognitive and affective engagement (Appleton, 2006). Cognitive engagement refers to students' perceptions and beliefs toward themselves and others, such as the school, teachers, and peers (Jimerson et al., 2003). Cognitive engagement variables are associated with students' meaningful strategy use, perceived self-efficacy, achievement, and their type of goal orientation (Archambault, Janosz, Morizot, & Pagani, 2009; Greene, Miller, Crowson, Duke, & Akey, 2004).

Although behavioral and academic indicators have received the most scholarly attention, there have been a number of studies determining the unique contribution of cognitive and affective engagement to positive student outcomes (Appleton, 2006; Fredericks et al., 2004; National Research Council & Institute of Medicine, 2004). Personal connectedness and affective engagement are associated with positive outcomes despite status risk factors (Connell et al., 1994). In other words, being affectively engaged fosters resilience from negative outcomes in students who may be determined at-risk. Cognitive and affective engagement tends to be indirectly related to outcomes through their effect on behavior (Reschly, Pohl, & Appleton, 2014). Students who perceive a classroom to support the use of elaborative strategies over rote, to be delivering content that is instrumental to their future goals, and to promote their personal competence rather than only a demonstrable competence are more likely to demonstrate cognitive engagement and achievement in school through more meaningful strategy use and the adoption of a mastery goal orientation (Greene et al., 2004). Students who believe they are competent and capable, while reciprocally working in an environment which builds that competency in a way that is relevant to their future goals, appear to demonstrate behaviors which contribute to success in school.

Engagement variables, above and beyond any other risk factors for dropout, significantly differentiate the most and least successful students (Finn & Rock, 1997; Reschly & Christenson, 2006). When considering precision, sensitivity, and specificity of dropout indicators, student engagement and student achievement longitudinal growth trajectories are the most accurate malleable predictors of students who will not complete school (Bowers, Sprott, & Taff, 2012). Consequently, student engagement has emerged as a promising theoretical model for intervention efforts to aid in promoting school completion or for school dropout prevention (Appleton, Christenson, Kim, & Reschly, 2006; Fredricks, Blumenfeld, & Paris, 2004; Reschly & Christenson, 2012), as well as high school reform (National Research Council, 2001, 2011).

Student engagement is important in the development of knowledge and the skills to acquire knowledge (Janosz, 2012). Researchers have found that engagement behaviors are those seen by parents and practitioners as being essential to the learning process, are important to post-secondary and future employment success, and are amenable to intervention on individual and school reform levels (Christenson et al., 2008; Finn & Zimmer, 2012). The "quality and quantity of effort" that a person displays is associated with the pursuit of higher education, entering the workforce, and having a higher quality of life (Janosz, 2012). Given that there is a considerable benefit for society to be comprised of highly engaged individuals and that student engagement variables are associated with long-term outcomes, it is not surprising that there is an interest in understanding how student engagement functions, can be measured, and can be intervened upon across time.

Measuring Longitudinal Data Accurately

The goal of assessment, especially within an educational context, is to be able to accurately and reliably measure where an individual lies on a given construct (e.g., what is a particular student's level of engagement with school?). The accuracy of our theoretical models and understanding of underlying, latent constructs often determines the proper targeting and efficiency of our intervention efforts and programs, as well as our prediction of outcomes across groups and individuals. The precision and consistency afforded to us by accurate measurement gives us a useful way to communicate about our constructs of interest (Edwards & Wirth, 2012). If an assessment does not demonstrate invariance, our ability to make inferences about traits across groups, or across times, becomes diminished by incomparable between-group mean levels or item correlation patterns (Reise, Widaman, & Pugh 1993). As defined by Millsap (2011), "Measurement invariance is built on the notion that a measuring device should function in the same way across varied conditions, so long as those varied conditions are irrelevant to the attribute being measured" (p. 1). To give an example, we would consider a student engagement measure to be invariant if we knew that those students with identical levels of engagement in school were not reporting different scores because of other factors like race, socioeconomic status (SES), or gender. The implications of such an example are important when considering the breadth of resource investment in education on local, state, and national levels, as well as the ever-strengthening relationship of education to future outcomes (Kirsch et al., 2007).

The "varied conditions" that may affect measurement invariance have been traditionally looked at as between-group differences. However, when measuring a construct for a long enough period of time, or through any major developmental changes, there is a possibility that the construct will evolve (Edwards & Wirth, 2012). Longitudinal measurement invariance concerns this type of measurement change within individuals across time. In order to make accurate inferences about intervention effectiveness over multiple time points, such as from an intervention program for a cohort of students determined to be at-risk for dropout, one must first understand any underlying changes in stability of the instrument over time (de Jonge, van der Linden, Schaufeli, Peter, & Siegrist, 2008; Golembiewski, Billingsley, & Yeager, 1976). For example, particular items may no longer hold relevance or the dimensionality of the construct may change. Additionally, new dimensions can evolve or change out of currently existing ones (Edwards & Wirth, 2012). According to Pitts and And (1996), “researchers need to show that the same construct(s) has been measured: (a) at each measurement wave; and (b) in randomized experiments and nonequivalent control group designs in both the treatment and control groups (p. 334).” These demonstrations are thought to create the highest potential for making accurate inferences based on the data, including pre- and post-test comparisons (Pitts & And, 1996).

An example of this measurement approach may be found in a study by Bowers and colleagues (2010). The authors conducted longitudinal measurement invariance analyses on the Five C’s Model of Positive Youth Development tested across middle adolescence (Bowers et al., 2010). The Five C’s are Competence, Confidence, Connection, Character, and Caring. Their data showed that athletic competence was no longer a relevant domain from their childhood measurement model, whereas perceptions of physical appearance became much more important between the ages of 13-16 (Bowers et al., 2010). The knowledge of intra-individual changes and level of measurement invariance grants many opportunities for researchers and stakeholders. More accurate prevention and intervention efforts can be better targeted and informed based on those changes and it can provide useful information in conjunction with theoretical understandings of adolescent development (Bowers et al., 2010).

In addition to education and typical development, longitudinal measurement invariance has implications in the realms of mental health and psychopathology. The idea of beta-change in measurement invariance research is analogous to response shifts in mental health research (Fokkema, Smits, Kelderman, & Cuijpers, 2013). Fokkema et al. (2013) discuss the common clinical practice of routine outcomes monitoring, a process by which the same self-report measures are given over fixed intervals during the course of treatment as a means to understanding changes in a construct of interest. We would expect these kinds of longitudinal self-report measures to be inherently stable as matching with one's self should be the most accurate, but research exploring longitudinal measurement invariance demonstrates that this is not always the case. The problem with self-report measures lies directly within their subjective nature, a problem that cannot be avoided in the mental health field where there may be no useful, valid, or reliable objective or observable measures (Fokkema et al., 2013).

As an example, Fokkema et al. (2013) explained the often psychoeducational component of depression treatment. The therapist explains to their client what depression is and how to recognize symptoms that commonly occur within the disorder. The way that a client views a self-report measure on depression may have fundamentally changed from their psychoeducational understanding of it, which may have an effect on scores taken post-treatment. Furthermore, clients who are being treated exclusively by antidepressants may have no such change when taking the same self-report measures (Fokkema et al., 2013). They found that measurement of individuals with the Beck Depression Inventory (BDI) who were under a randomized control trial of receiving medication or medication and psychotherapy led to a violation of scalar invariance. In other words, clients' responses to items had changed depending on their condition. The manner of intervention, or whether a group received intervention at all, may bring about differences in the responding of subjective self-report items. The nature of having multiple approaches in treatment serves as a potential confound to systematic measurement.

Progress Monitoring Students

Progress monitoring, the systematic collection of student data over time to aid in evidence-based and data-driven decision making, has become a cornerstone of educational practice across the country. This is following the enactment of federal legislation such as No

Child Left Behind (NCLB, 2001), the Education Sciences Reform Act (ESRA, 2002), and the reauthorization of the Individuals With Disabilities Education Act (IDEA, 2004) which have brought about a national focus on improving academic outcomes (Lemons, Fuchs, Gilbert, & Fuchs, 2014). We monitor progress to make decisions and predictions about outcomes, often in conjunction with, or as the result of, universal screening to determine the presence of a problem (Kamphaus, 2010). These data are often utilized within tiered frameworks of service delivery, known under a variety of names across the country including, but not limited to, response-to-intervention (RTI), multi-tiered systems of support (MTSS), or instructional decision making (IDM) which rest on the presumption of prevention and early intervention (Harlacher & Siler, 2011).

Progress monitoring within tiered systems differ from traditional assessment frameworks. Typically, progress monitoring data are frequently administered, produce feedback of immediate use to educators, often increase students' goal awareness, and are used to make improved adaptive and formative decisions for instruction and intervention (Fuchs & Fuchs, 2011). Widespread tools used for the purposes of progress monitoring, such as Curriculum Based Measurement for Oral Reading (CBM-R) have faced entirely new challenges in comparison to the traditional assessments which it has largely replaced. For example, research and simulation studies have recently been conducted to better determine issues in relation to the necessary schedule, duration, and dataset quality of CBM-R measures when being used in different decision-making situations (Christ, Zopluoglu, Monaghen, & Van Norman, 2013). Despite methodological concerns and questions, CBM-R is an important exemplar for the numerous benefits of progress monitoring across many disciplines.

When students are aware of their progress, they become part of a goal striving process which may impact their engagement with academic tasks. There is evidence to suggest that higher-order processing of progress monitoring goal-relevant information, such as when students are made aware of their reading progress toward reading goals when using CBM-R, shows neuronal activations in regions associated with attention and working memory above and beyond those found when only error-monitoring a task (Benn et al., 2014). Not only is progress monitoring beneficial to educators' decision making about students, but there is a direct benefit to making students aware of their goals and progress as they may change their perception of the relevancy of a task to future outcomes, which is thought to be key to promoting engagement and eventual school completion (Reschly et al., 2014).

The benefits of progress monitoring for academics have widespread recognition and are becoming a focus for behavior and mental health within schools, as well as within therapeutic practices as a way to bridge the research-practice divide (Fitzpatrick, 2012; Merrell, 2010). With the example of school behavior, some researchers are directly seeking to bridge the gap with a "CBM analogue" when measuring students' response to social behavior interventions (Gresham, Cook, Collins, & Rasethwane, 2010). Similar to changes in academic decision making through the use of progress monitoring, behavior and mental health treatments are hoping to lessen teacher and clinician decision processes through the use of systematic screening and progress monitoring (Goodman, McKay, & DePhilippis, 2014; Kamphaus et al., 2010). These methods are an attempt to bridge the divide between assessment and intervention (Merrell, 2010).

Measuring Student Engagement

Consistent with the focus within mental health systems and in school behavior, it may be in an educational system's best interest to collect progress monitoring information on levels of student engagement. Although many behavioral indicators are used to determine levels of students' disengagement with school, affective and cognitive engagement variables have been shown to have incremental predictive utility for determining student outcomes (Lovelace, 2013). Educational research, including student engagement, must often include subjective self-response information from students as there may be important psychological variables which may not be as observable or identifiable, or less objectively measured. Those observable indicators (e.g., behavioral, academic) are those which have been the subject to the majority of research in student engagement (Appleton et al., 2006). It is important that we provide clear descriptions of theory-driven measures, especially when considering subjective, latent indicators such as cognitive and affective engagement.

After all, student engagement is a field with considerable conceptual haziness that can be viewed and defined from many different perspectives (Reschly & Christenson, 2012). The perspective from which we view student engagement affects how we organize and select the information we wish to discern from our target population and may differ from different approaches. In other words, epistemological and metaphysical concerns guide the types of questions we attempt to answer so it is important to thoroughly describe one's aims and goals (Godfrey-Smith, 2003). For example, even if information is gathered from the same raters (i.e., self-report information about the student) using the same type of assessments (i.e., rating scales, behavioral observations) to collect them, we can obtain different information on how engaged a student is, or how effective an intervention was at changing a student's level of engagement, depending upon what theoretical perspective we subscribe to as researchers and what questions we intend to answer through scientific inquiry. This is not to say that there is one right way to obtain information on student engagement, but to say that it is important that researchers and stakeholders are familiar with the intents, purposes, and background of the instruments they use and how well it aligns with their intended use.

The Student Engagement Instrument (SEI; Appleton, Christenson, Kim, & Reschly, 2006) is a 33-item self-report measure of engagement designed to tap students' cognitive and affective engagement with school and learning. There are five subtypes measured within the SEI: Teacher–Student Relationships (TSR), Control and Relevance of School Work (CRSW), Peer Support for Learning (PSL), Future Aspirations and Goals (FG), and Family Support for Learning (FSL). TSR, PSL, and FSL are affective engagement factors, CRSW and FG are cognitive engagement factors. Items which comprise these factors were created following an exhaustive literature review and piloted through diverse student focus groups.

Following its inception, the SEI has been adapted and administered to students in elementary schools (see the SEI-E; Carter et al., 2012), college settings (SEI-C; Grier-Reed et al., 2012; Waldrop, 2012), and has validated its factor structure and invariance from grades 6-12 (Betts et al., 2010; Lovelace et al., 2014). It has been implemented across the United States as well as cross-culturally (see Reschly, Betts, & Appleton, 2014) with more than 1500 requests for its use in the last year alone.

Purpose of the Study

The purpose of this study was to examine a brief version of the SEI (SEI-B) for use as a progress monitoring measure. The SEI-B was constructed by removing one item from each of six item pairs with correlated residuals from the SEI (see Betts et al., 2010), with a total of 27 remaining items. In this study the following research questions are posed:

1. What is the factor structure of the SEI-B?

a. Does the five-factor structure of the SEI replicate when items are removed for a briefer instrument?

2. Does the five-factor structure or resulting factor structure remain invariant across all three time points?

Table 1.1 – Alterable variables by context

Alterable variables by context (adapted from Reschly & Christenson, 2012)

	Protective factors	Risk factors
Student	Homework completion Class preparation High locus of control High self-concept Expectations for school completion	High rates of absence Behavior problems Poor academic performance Grade retention Employment
Family	Academic and motivational support for learning (e.g., parent support with homework, high expectations) Parental monitoring	Low educational expectations Mobility Permissive parenting styles
School	Orderly school environments Committed, caring teachers Fair discipline policies	Weak adult authority Large school size (>1,000 students) High pupil–teacher ratios Few caring relationships between staff and students Poor or uninteresting curricula Low expectations and high rates of truancy

Sources: Reschly and Christenson (2006b); Rosenthal (1998)

CHAPTER 2

METHOD

Participants

Dataset

The sample was drawn from a population of 9th grade students within a school district in the southeastern U.S over three different administrations taken at one month intervals. There were 6118 timepoint responses recorded for the SEI-B from a total of 2799 unique students. Of these students, 1037 had responses at each of the three timepoints. Resulting from data inclusion parameters followed (described below), the final dataset included responses from 915 unique students across three time points for a total of 2745 timepoint responses. All data were archival, collected throughout 2011, as part of a district-wide initiative geared toward student engagement. As no systemic student engagement interventions were being implemented, engagement data collected are considered to be baseline data (i.e., business-as-usual besides gathering survey data).

Demographic data were collected from participants. Participants were ethnically diverse, with students from the total sample (n=915) identifying as Asian 10.8% (n=99), Black 9.6% (n=88), Hispanic 12.9% (n=118), American Indian/Alaska Native 0.4% (n=4), Multiracial 3.1% (n=29), and White 63.1% (n=577). Students were predominantly fluent English speakers, with 97.5% (n=892) not receiving ELL services, 1.4% (n=13) receiving ELL services, and 1% (n=10) receiving monitoring services for reclassification out of ELL. Of these participants, 6.6% (n=61) represented students with disabilities. The majority of students were not eligible for free or reduced lunch 77.9% (n=713).

Measures

Student Engagement Instrument – Brief Version

The SEI-B is a 27-item self-report measure of student engagement which has been adapted from the validated full form Student Engagement Instrument through the removal of six items with correlated residuals (Appleton et al. 2006; Betts et al., 2010). The SEI-B survey items are designed to measure the same five factors (i.e., Teacher-Student Relationships, Control and Relevance of Schoolwork, Peer Support for Learning, Future Goals and Aspirations, and Family Support for Learning) and contain a 5-point Likert-type scale (i.e., "1" indicates "strongly disagree," "2" indicates "disagree," "3" indicates "neither agree nor disagree," "4" indicates "agree," "5" indicates "strongly agree").

As the SEI-B is a truncated version (i.e., quicker to administer) of the SEI and is theorized to contain similar levels of psychometric stability, it is a candidate for use in repeated administration for progress-monitoring student engagement levels. In order to determine whether the SEI-B could be used for such purposes it is important to identify whether the factor structure is comparable to the SEI and whether the structure of the SEI-B remains invariant over repeated administrations.

Procedures

Data Inclusion Parameters

Participants were excluded under one of several conditions to preserve the integrity of the dataset for optimal comparisons across time: individuals were excluded who a) were not present at each of the three survey administrations (though respondents were allowed to skip a small percentage of items at each administration), b) did not fall within a restricted range of dates to ensure relatively equidistant responding within and between individuals (i.e., responding to the

SEI-B in roughly 1-month intervals from March to May), and c) did not respond to at least 75% of items on each factor. Similarly, any duplicate entries (n=46, or 23 individuals with two responses) were randomly deleted where the duplicate with the lowest generated value was retained. Random removal prevents the introduction of systemic bias for the duplicate exclusion. Assignment to each time point was determined by assessing the frequency of administrations, using them as midpoints, and applying cut-off dates 15 days on either side of the midpoint. We cannot rule out systematic bias in our final sample due to these exclusions. There may be meaningful differences between those students who responded regularly and those who responded inconsistently for the purpose of our analyses. However, as the 23 cases represent such a small proportion of the sample they are unlikely to exert meaningful differences on parameter estimates.

Cross-Sectional Factor Analysis

Two subsamples of the initial time point were described by factor analysis to validate the factor structure of the SEI-B. The first sub-sample comprised 40% (sub-sample A) of the total cases, while the remaining sub-sample comprised 60% (sub-sample B). The initial time point was selected to rule out any potential bias from fatigue effects. Sub-sample A was explored using exploratory factor analysis (EFA) to determine whether the items removed from the SEI impact its latent structure. The strongest resulting model from the EFA on sub-sample A was cross-validated by confirmatory factor analysis (CFA) on sub-sample B. Consistent with Brown (2006), the acceptability of the CFA solution was determined by 1) overall goodness of fit, 2) specific points of poor fit in the model, and 3) interpretability, size, and statistical significance of model parameter estimates.

Longitudinal Multivariate Analysis

Following the cross-sectional factor analyses, remaining missing responses between time points from the full sample were multiply imputed five times within the R programming language using the *Amelia II* package (Honaker, King, & Blackwell, 2011; R development Core Team, 2009). Then, longitudinal measurement invariance (MI) CFA were estimated using the *lavaan* package (Rosseel, 2012) on each of the imputed datasets following the suggestions and recommendations of Vandenburg and Lance (2000):

1. Configural (weak) invariance: equal factor loading patterns across occasions.
2. Metric (strong) invariance: equal factor loadings across occasions.
3. Scalar invariance: equal item intercepts across occasions.
4. Uniqueness invariance: equal residual variances across occasions.

While the measurement invariance analyses are typically performed as a multi-group CFA, longitudinal MI analyses are best operationalized in a single group CFA framework. This modification allows variables of interest to correlate over time intervals as the same participants are responding to the same items over time (Fokkema et al., 2013).

Data obtained from the SEI-B are ordinal; therefore, it is recommended that mean- and variance-adjusted least squares (WLSMV) estimation be used (Reeve et al., 2007). This operation is carried out in *lavaan* by estimating the model parameters by diagonally weighted least squares (DWLS) and using the full-weight matrix for robust standard errors and a mean- and variance-adjusted test statistic (de Beurs et al., 2015). Such analyses have been shown to result in unbiased parameter and standard error estimates, and satisfactory type-I error rates when handling skewed ordinal data (de Beurs et al., 2015; Flora & Curran, 2004; Lei, 2009).

Responses on the SEI-B within each of the three measurement waves were regressed onto the five-factor structure of the SEI-B. Each factor was allowed to correlate across three measurement waves: Teacher-Student Relationships (TSR), Control and Relevance of School Work (CRSW), Peer Support for Learning (PSL), Future Aspirations and Goals (FG), and Family Support for Learning (FSL) at T0, T1, and T2, respectively. As missing data were multiply imputed, longitudinal MI analyses were performed on each of the five imputed datasets per measurement wave.

<u>Assessing Model Fit</u>

Guidelines for model goodness of fit were established by following other research performing factor structure and measurement invariance analyses (Chungkam et al., 2013; Fokkema et al., 2013; de Beurs et al., 2015). These studies underscored the importance of using many different fit indices when determining goodness of fit, as recommended by seminal research in the field (Bentler, 1990; Brown, 2006; Cheung & Rensvold, 2002; Hu & Bentler, 1999). Root mean square error of approximation (RMSEA) expresses poor model parsimony using model degrees of freedom. Browne and Cudeck's (1993) criteria, RMSEA $\leq$ 0.08 are acceptable while those greater than 0.10 are to be rejected. The comparative fit index (CFI) compares the hypothesized model to an incrementally more restricted and nested baseline model. CFI values which are $\geq$ 0.90 are acceptable (Bentler, 1990). The minimum function test statistic is dependent on sample-size, artificially producing significant results when N$\geq$400, leaving RMSEA and CFI as being sufficient for assessing model fit (Cheung & Rensvold, 2002). When assessing invariance, change in alternative fit indices (AFIs) are less sensitive to sample size than chi-square, are more sensitive to an LOI, and are generally non-redundant with other AFIs (Meade, Johnson, & Braddy, 2006).

When comparing nested models changes in model fit are typically assessed using RMSEA and CFI, as scaled chi-squared differences calculated by *lavaan* are subject to the same sample-size dependencies as the minimum function test statistic (de Beurs et al., 2015). When comparing nested models, a change in CFI which is ≥ -0.010 in conjunction with a change in RMSEA of ≥ 0.015, or a change in SRMR ≥0.030 for loading invariance in conjunction with a change in SRMR ≥ 0.010 for intercept invariance, would indicate poor model fit between models (Chen, 2007).

However, following the recommendations of another large sample-size study on measurement invariance, researchers have suggested that a general cutoff of 0.002 CFI can be used when assessing configural, metric, and scalar invariance (Chungkam et al., 2013; Meade et al., 2006). The variability in power when applying a 0.002 CFI is similar and favorable across many different conditions while RMSEA has a mixed performance, especially at larger sample sizes (Meade et al., 2006). Information gained from many different AFIs (e.g., CFI, IFI, RNI, etc.) tends to be redundant, making it unnecessary to report many different indices (Hu and Bentler, 1999; Meade et al., 2006). Therefore, the CFI cutoff alone was determined to be acceptable for assessing configural, metric, and scalar invariance.

CHAPTER 3

RESULTS

Exploratory and Confirmatory Factor Analyses

The EFA applied on 40% of the cross-sectional sample (n = 366), the first administration time point, with the proposed five-factor model from the full form of the SEI showed five correlated factors with acceptable fit indices (CFI = 0.982, RMSEA = 0.054). While the CFI and RMSEA appear to be acceptable (i.e., CFI is recommended to be >0.90, RMSEA is recommended to be <=.08), many items could potentially be improved through model revision to increase fit (Bowen, 2014; Chungkham et al., 2013; Hu & Bentler, 1999). However, the theoretical justification for maintaining the previously validated five-factor model with a similar sample of individuals (see Appleton et al., 2006) was deemed to be more important than altering the model for RMSEA or CFI values. Furthermore, while modification indices did present the opportunity for items to be re-organized, it is recommended that changes are made only if that modification is a) justifiable according to theory, b) are few in number, and c) are minor and do not impact other parameter estimates (Bowen, 2014). Although changes could be made based on modification indices, it would break these guidelines as they would be contradictory to the theory behind the model. Therefore, it was decided that the five-factor model (see Figure 3.1) was most appropriate for performing the CFA.

The CFA with the five-factor model was applied to the remaining 60% of the cross-sectional sample from first administration time point (n = 545) for cross-validation. It appears the model is impacted by the stricter measurement procedures required by the CFA as there is a decrement in fit (CFI = .953, RMSEA = .071) relative to the results of the EFA.

Longitudinal Measurement Invariance Analyses

The resulting five-factor model from the cross-sectional sample validation was used as the baseline model for the longitudinal measurement invariance tests. These tests were performed across each of the three measurement waves. Each measurement wave consisted of responses from the participants after data inclusion parameters (n=915) for a total of 2745 responses across the three measurement waves. Estimates were generated for each of the five datasets which had undergone multiple imputation as shown in Table 3.1. As results were consistent across datasets (i.e., when one dataset demonstrated fit, all datasets demonstrated fit) the results for the first multiply imputed dataset will be used when discussing results.

The baseline model is the configural invariance model (Model 1). To demonstrate configural invariance we compared our baseline to a model with a parameter requirement of equal factor loading patterns across our three time points. The SEI-B demonstrated acceptable fit (CFI = 0.91, RMSEA = 0.070 across all five imputations) according to the literature with acceptable CFI fit range between 0.90 and 0.95 and RMSEA <0.08 (Bentler, 1990; Browne & Cudeck, 1993). In other words, this means that the factor structure of the SEI-B is the same across administrations for the same set of respondents (Schmitt & Kuljanin, 2008).

Following the demonstration of configural invariance, the next restriction to be placed on our model is to require the magnitude and loading of items on each factor to be constant over time, a metric invariance model, and test this against our configural invariance model (Model

2 vs. Model 1). Our model comparison performed at the cutoff criteria recommended by Meade and colleagues (2006) simulation study for alternative fit indices (CFI = 0.002). Thus, the SEI-B has demonstrated full metric invariance.

The next parameter requirement is to fix the variance of each factor across time, a test of scalar invariance, in addition to the previous requirements of items loading equivalently on each factor across time for respondents (Model 3 vs. Model 2). Again, the SEI-B met requirements for full scalar invariance (CFI = 0.002). This demonstrates that the five factors of the SEI-B, and the item loadings onto those factors, are functioning similarly across respondents over time.

In addition to demonstrating similar instrument functioning for items loadings on factors, and for factors themselves, another requirement of measurement invariance is to demonstrate that the items themselves demonstrate invariant variance over time (e.g., does each item function the same way over time?). This is demonstrated by testing a model where a constraint is placed on item error variances, a test of uniqueness invariance (Model 4 vs. Model 3). In this model, the regression equation residuals for each item is proposed to be equivalent across groups (Schmitt & Kuljanin, 2008). The SEI-B demonstrated full uniqueness invariance (CFI = 0.002).

Table 3.1 - Measurement invariance analysis results across multiple imputations

Dataset	Configural Invariance (Model 1)	Metric Invariance (Model 2)	Scalar Invariance (Model 3)	Uniqueness Invariance (Model 4)
Imputation 1	CFI = 0.912, RMSEA = 0.070	CFI = -0.001	CFI = -0.002	CFI = -0.002
Imputation 2	CFI = 0.912, RMSEA = 0.070	CFI = -0.001	CFI = -0.002	CFI = -0.002
Imputation 3	CFI = 0.912, RMSEA = 0.070	CFI = -0.001	CFI = -0.002	CFI = -0.002
Imputation 4	CFI = 0.912, RMSEA = 0.070	CFI = -0.001	CFI = -0.002	CFI = -0.002
Imputation 5	CFI = 0.912, RMSEA = 0.070	CFI = -0.001	CFI = -0.002	CFI = -0.002

Note: Acceptable fit for model 1 = CFI > 0.90, RMSEA < 0.08 (Bentler, 1990; Browne & Cudeck, 1993). Models 2-4 would evidence misfit if CFI > 0.002 from the previous model (Meade et al., 2002).

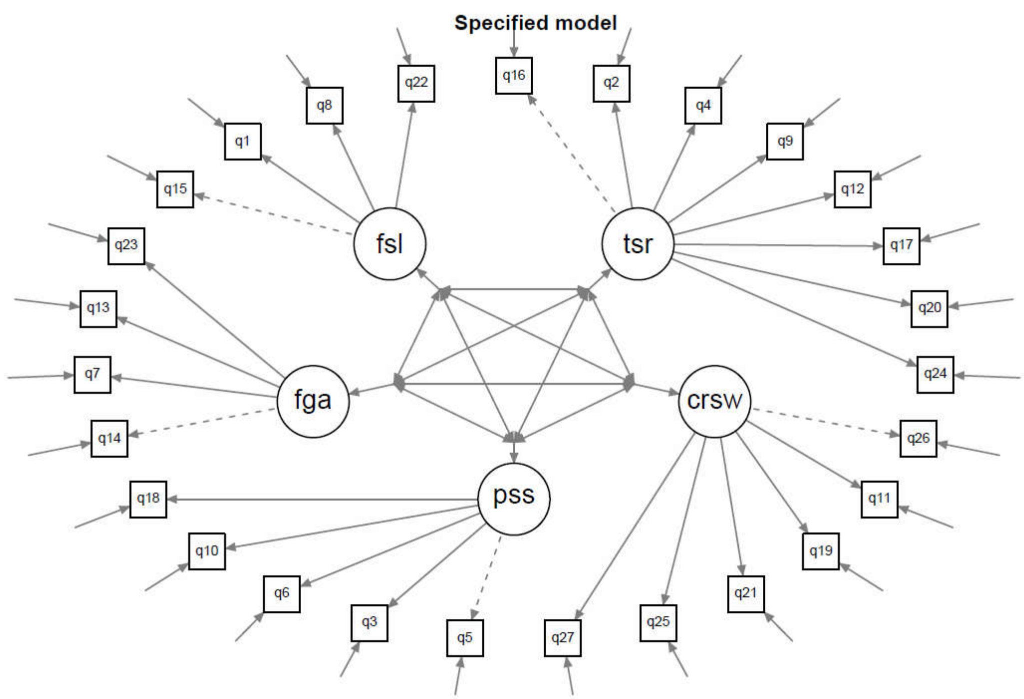

Figure 3.1 - Five Factor Model of the SEI-B

Note: TSR = Teacher-Student Relationships; CRSW = Control and Relevance for Schoolwork; PSS = Peer Support for Learning; FGA = Future Goals and Aspirations; FSL = Family Support for Learning. Items relating to question numbers may be found in Appendix A.

CHAPTER 4

DISCUSSION

Student engagement comprises psychological indicators which are actionable (i.e., amenable to intervention), matter to all students and all individuals invested in their success (i.e., those people which comprise their direct ecological networks), and extend to successes beyond the school environment. There is a compelling social and economic benefit to improving those elements which will increase the likelihood a child will complete high school, have the skills necessary for college success, and the ability to be productive in their future work and personal lives. We can modify and create systems which work to encourage, support, and develop individuals early on and continue to invest in them over time. To accomplish this, we need to develop ways to understand and monitor those important features which contribute to success.

Presently, there are few measures developed and validated to measure student engagement briefly and accurately over time. It is important that we understand and attend to student trajectories if we plan to make meaningful change through intervention. Such change cannot be inferred if we do not know if we are measuring what we are intending to measure. Measuring interventions without demonstrating measurement invariance is like trying to hit a target while blindfolded; you may have the proper techniques and the right tools, but no way of seeing and knowing what you intend to hit. The SEI-B was developed to remove the blindfold with monitoring students' engagement at the high school level by demonstrating that the instrument measures the same thing, or hits the same target, across different administrations over time.

We have demonstrated that the SEI-B retains the factor structure and validity of the full-form SEI and functions invariantly across time for students in a diverse high school in the southeastern US. These findings not only bolster the growing evidence for the developmental and contextual importance for which studies using the SEI have underscored, but helps bridge the assessment-intervention gap that is prevalent across instruments and constructs currently used for intervention in educational settings.

Limitations and Future Directions

The present study has limitations toward generalizability. Although the sample size is large and bolsters the confidence in study results, it is taken from only one school in an urban setting located in the southeastern United States. It is plausible that other factors could affect results, which are not limited to a different geographic locations (e.g., northwestern United States, or in a rural setting), school size, or different developmental periods (e.g., using brief versions of the SEI-E or SEI-C). It will be important to replicate this study across developmental periods if interventions are meant to target or span those levels of development.

This study also does not take into account many demographic factors which may be significant co-variates for the given data. With a complex longitudinal data structure, this is a difficult analysis to perform even with modern tools and is beyond the scope of the current paper. However, given that data is taken within-persons over a short period of time in a stable developmental period, it is reasonable to assume that many of these variables (e.g., sex, ethnicity) are not creating significant change within a person over that period. It may be important to include demographic covariates in future longitudinal analyses.

Fit and incremental change are not as compelling for the SEI-B when compared to its full-form predecessors which is likely due to multiple factors. The SEI-B is a shorter form which is designed to be used in repeated administrations over time. The reliability of the factor structure is likely to decrease when compared to a construct which measures additional items on each factor. Additionally, longitudinal measurement invariance analyses place even further restrictions on the model than would a CFA. Thus, the SEI-B is being analyzed to more rigorous standards when being tested for invariance over time.

REFERENCES

Alliance for Excellent Education (2013). *Fact Sheet: High School State Cards – National.* Retrieved March 25, 2015, from http://all4ed.org/wp-content/uploads/2013/09/UnitedStates_hs.pdf.

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*, 427–445. doi: 10.1016/j.jsp.2006.04.002

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*, 369-386.

Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009). Adolescent Behavioral, Affective, and Cognitive Engagement in School: Relationship to Dropout. *Journal Of School Health*, *79*(9), 408-415.

Balfanz, R., Bridgeland, J. M., Bruce, M., Fox, J. H., Civic, E., Johns Hopkins University, E. C., & ... Alliance for Excellent, E. (2013). Building a Grad Nation: Progress and Challenge in Ending the High School Dropout Epidemic. Annual Update, 2013. *Civic Enterprises*.

Benn, Y., Webb, T. L., Chang, B. I., Sun, Y., Wilkinson, I. D., & Farrow, T. D. (2014). The neural basis of monitoring goal progress. *Frontiers In Human Neuroscience*, *8*688. doi:10.3389/fnhum.2014.00688

Bentler, P.M. (1990). Comparative fit indexes in structural models. Psychological Bulletin 107, 238.

Betts, J., Appleton, J.J., Reschly, A.L., Christenson, S.L., & Huebner, E.S. (2010). A Study of the reliability and construct validity of the Student Engagement Instrument across multiple grades. *School Psychology Quarterly, 25*, 84-93.

Bowers, A. J., Sprott, R., & Taff, S. A. (2012). Do We Know Who Will Drop Out? A Review of the Predictors of Dropping out of High School: Precision, Sensitivity, and Specificity. *High School Journal*, *96*(2), 77-100.

Bowers, E. P., Li, Y., Kiely, M. K., Brittian, A., Lerner, J. V., & Lerner, R. M. (2010). The Five Cs Model of Positive Youth Development: A Longitudinal Analysis of Confirmatory Factor Structure and Measurement Invariance. Journal Of Youth And Adolescence, 39(7), 720-735.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments in nature and design*. Cambridge, MA: Harvard University Press.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY, US: Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, *154*, 136-136.

Carter, C. P., Reschly, A. L., Lovelace, M. D., Appleton, J. J., & Thompson, D. (2012). Measuring student engagement among elementary students: Pilot of the Student Engagement Instrument—Elementary Version. *School Psychology Quarterly*, *27*(2), 61.

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

Christ, T. J., Zopluoglu, C., Monaghen, B. D., & Van Norman, E. R. (2013). Curriculum-Based Measurement of Oral Reading: Multi-Study Evaluation of Schedule, Duration, and Dataset Quality on Progress Monitoring Outcomes. *Journal Of School Psychology*, *51*(1), 19-57.

Christenson, S. L., & Reschly, A. L. (2010). Check & Connect: Enhancing school completion through student engagement. In E. Doll & J. Charvat (Eds.), Handbook of prevention science. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chungkham, H. S., Ingre, M., Karasek, R., Westerlund, H., & Theorell, T. (2013). Factor Structure and Longitudinal Measurement Invariance of the Demand Control Support Model: An Evidence from the Swedish Longitudinal Occupational Survey of Health (SLOSH). *Plos ONE*, *8*(8), 1-11. doi:10.1371/journal.pone.0070541

Crick, R. D. (2012). Deep engagement as a complex system: Identity, learning power and authentic enquiry. In S.L. Christenson, A.L. Reschly, & C. Wylie (Eds). *Handbook of Research on Student Engagement*. New York: Springer.

de Beurs, D. P., Fokkema, M., de Groot, M. H., de Keijser, J., & Kerkhof, A. J. (2015). Longitudinal measurement invariance of the Beck Scale for Suicide Ideation. *Psychiatry Research*, *225*368-373. doi:10.1016/j.psychres.2014.11.075

de Jonge, J., van der Linden, S., Schaufeli, W., Peter, R., & Siegrist, J. (2008). Factorial invariance and stability of the effort-reward imbalance scales: A longitudinal analysis of two samples with different time lags. *International Journal of Behavioral Medicine*, *15*, 62–72.

Education Commission of the States (2011). http://www.ecs.org/clearinghouse/95/05/9505.pdf.

Education Sciences Reform Act of 2002, P.L. 107–279, 116 Stat. 1940 (2002).

Edwards, M. C., & Wirth, R. J. (2012). Valid measurement without factorial invariance: A longitudinal example. In J. R. Harring, G. R. Hancock (Eds.), Advances in longitudinal methods in the social and behavioral sciences (pp. 289-311). Charlotte, NC US: IAP Information Age Publishing.

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. Review of Educational Research, 75(2), 159–210.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117-142.

Finn, J. D. (2006). The adult lives of at-risk students: The roles of attainment and engagement in high school (NCES 2006–328). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Finn, J. D., & Cox, D. (1992). Participation and withdrawal among fourth-grade pupils. American Educational Research Journal, 29, 141–162.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology, 82*(2), 221-234.

Fitzpatrick, M. (2012). Blurring practice–research boundaries using progress monitoring: A personal introduction to this issue of Canadian Psychology. *Canadian Psychology/Psychologie Canadienne*, *53*(2), 75-81. doi:10.1037/a0028051

Flora, D.B., Curran, P.J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 9, 466.

Fokkema, M., Smits, N., Kelderman, H., & Cuijpers, P. (2013). Response shifts in mental health interventions: An illustration of longitudinal measurement invariance. Psychological Assessment, 25(2), 520-531. doi:10.1037/a0031669

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*, 59-109.

Fuchs, L. S., Fuchs, D., & National Center on Student Progress, M. (2011). Using CBM for Progress Monitoring in Reading. *National Center On Student Progress Monitoring*, Available from: ERIC, Ipswich, MA.

Godfrey-Smith, P. (2009). *Theory and reality: An introduction to the philosophy of science*. University of Chicago Press.

Golembiewski, R. T., Billingsley, K., & Yeager, S. (1975). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal of Applied Behavioral Science*, *12*, 133–157.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology*, *29*(4), 462-482. doi:10.1016/j.cedpsych.2004.01.006

Gresham, F. M., Cook, C. R., Collins, T., & Rasethwane, K. (2010). Developing a change-sensitive brief behavior rating scale as a progress monitoring tool for social behavior: An example using the Social Skills Rating System—Teacher Form. School Psychology Review, 39(3), 364 –379.

Grier-Reed, T., Appleton, J., Rodriguez, M., Ganuza, Z., & Reschly, A. L. (2012). Exploring the Student Engagement Instrument and career perceptions with college students. *Journal of Educational and Developmental Psychology*, *2*(2), p85.

Griffiths, A., Lilles, E., Furlong, M. J., & Sidhwa, J. (2012). The relations of adolescent student engagement with troubling and high-risk behaviors. In S. L. Christenson, A. L. Reschly, C. Wylie, S. L. Christenson, A. L. Reschly, C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 563-584). New York, NY, US: Springer Science + Business Media. doi:10.1007/978-1-4614-2018-7_27

Harlacher, J. E., & Siler, C. E. (2011). Factors Related to Successful RTI Implementation. *Communique*, *39*(6), 20-22.

Honaker, J., King, G., Blackwell, M. (2011). Amelia II: A Program for Missing Data. Journal of Statistical Software, 45(7), 1-47. URL http://www.jstatsoft.org/v45/i07/.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. doi:10.1080/10705519909540118

Individuals With Disabilities Education Improvement Act (IDEA) of 2004, P. L. No. 108–446, 118 Stats. 2647. (2004).

Janosz, M. (2012). Part IV commentary: Outcomes of engagement and engagement as an outcome: Some consensus, divergences, and unanswered questions. In S.L. Christenson, A.L. Reschly, & C. Wylie (Eds). *Handbook of Research on Student Engagement*. New York: Springer.

Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an Understanding of Definitions and Measures of School Engagement and Related Terms. *California School Psychologist*, *8*7-27. doi:10.1007/BF03340893

Kamphaus, R. W., DiStefano, C., Dowdy, E., Eklund, K., & Dunn, A. R. (2010). Determining the Presence of a Problem: Comparing Two Approaches for Detecting Youth Behavioral Risk. *School Psychology Review*, *39*(3), 395-407.

Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., & ... American Institutes for, R. (2014). The Condition of Education 2014. NCES 2014-083. *National Center For Education Statistics*.

Kirsch, I., Braun, H., Yamamoto, K., & Sum, A. (2007). America's Perfect Storm: Three Forces Changing Our Nation's Future. Princeton, NJ: Educational Testing Service.

Lei, P.W., 2009. Evaluating estimation methods for ordinal data in structural equation modeling. Quality and Quantity 43, 495–507.

Lemons, C. J., Fuchs, D., Gilbert, J. K., & Fuchs, L. S. (2014). Evidence-Based Practices in a Changing World: Reconsidering the Counterfactual in Education Research. *Educational Researcher*, *43*(5), 242-252.

Lovelace, M. D. (2013). *Longitudinal characteristics and incremental validity of the Student Engagement Instrument (SEI) [electronic resource]*, 2013.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2006, August). THE UTILITY OF ALTERNATIVE FIT INDICES IN TESTS OF MEASUREMENT INVARIANCE. In *Academy of management proceedings* (Vol. 2006, No. 1, pp. B1-B6). Academy of Management.

Merrell, K. W. (2010). Better Methods, Better Solutions: Developments in School-Based Behavioral Assessment. *School Psychology Review*, *39*(3), 422-426.

Millsap, R. E. (2011). Statistical approaches to measurement invariance. New York: Routledge.

National Research Council (2001). *Understanding dropouts: statistics, strategies, and high-stakes testing*. Washington, D.C.: National Academy Press, c2001.

National Research Council (2011). *High School Dropout, Graduation, and Completion Rates: Better Data, Better Measures, Better Decisions*. Washington, DC: The National Academies Press, 2011.

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

No Child Left Behind Act of 2001, P. L. No. 107–110, § 1–1076, 115 Stat. 1425. (2001).

Overton, W. F. (2013). Relationism and relational developmental systems: A paradigm for developmental science in the post-Cartesian era. *Advances in child development and behavior*, *44*, 21-64.

Pitts, S. C., & And, O. (1996). Longitudinal Measurement Models in Evaluation Research: Examining Stability and Change. Evaluation And Program Planning, 19(4), 333-50.

R development Core Team (2009). R Project for Statistical Computing. Vienna, Austria.

Rauscher, E. (2010). Producing adulthood: Adolescent employment, fertility, and the life course. *Social Science Research, 40(2)*, 552-571.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. Psychological Bulletin, 114(3), 552-566. doi:10.1037/0033-2909.114.3.552

Reschly, A. L., & Christenson, S. L. (2006a). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education, 27*, 276-292.

Reschly, A., & Christenson, S. L. (2006b). Promoting school completion. In G. Bear & K. Minke (Eds.), *Children's needs III: Understanding and addressing the developmental needs of children*. Bethesda, MD: National Association of School Psychologists.

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S.L. Christenson, A.L. Reschly, & C. Wylie (Eds). *Handbook of Research on Student Engagement* (pp. 3–19). New York: Springer.

Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36.

Schmitt, N., & Kulijanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, *18*, 210–222.

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. Organizational Research Methods, 3(1), 4.

Waldrop, D. M. (2012). *An examination of the psychometric properties of the Student Engagement Instrument - College Version. [electronic resource]*. 2012.

Wirt, J., Choy, S., Rooney, P., Provasnik, S., Sen, A., & Tobin, R. (2004). *The Condition of Education 2004* (National Center for Educational Statistics No. NCES 2004-077). Washington, D.C.: U.S. Government Printing Office.

Yazzie-Mintz, E., & McCormick, K. (2012). Finding the humanity in the data: Understanding, measuring, and strengthening student engagement. In S. L. Christenson, A. L. Reschly, C. Wylie, S. L. Christenson, A. L. Reschly, C. Wylie (Eds.) , *Handbook of research on student engagement* (pp. 743-761). New York, NY, US: Springer Science + Business Media. doi:10.1007/978-1-4614-2018-7_3

APPENDICES

Appendix A

Description of SEI-B items

SEI-B Item Text
1. My family/guardian(s) are there for me when I need them.
2. My teachers are there for me when I need them.
3. Other students here like me the way I am.
4. Adults at my school listen to the students.
5. Other students at school care about me.
6. Students at my school are there for me when I need them.
7. My education will create many future opportunities for me.
8. When something good happens at school, my family/guardian(s) want to know about it.
9. Most teachers at my school are interested in me as a person, not just as a student.
10. Students here respect what I have to say.
11. When I do schoolwork I check to see whether I understand what I'm doing.
12. Overall, my teachers are open and honest with me.
13. I plan to continue my education following high school.
14. School is important for achieving my future goals
15. When I have problems at school my family/guardian(s) are willing to help me
16. Overall, adults at my school treat students fairly.
17. I enjoy talking to the teachers here.
18. I have some friends at school.
19. When I do well in school it's because I work hard.
20. I feel safe at school.
21. I feel like I have a say about what happens to me at school.
22. My family/guardian(s) want me to keep trying when things are tough at school.
23. I am hopeful about my future.
24. At my school teachers care about students.
25. Learning is fun because I get better at something.
26. What I'm learning in my classes will be important in my future.
27. The grades in my classes do a good job of measuring what I'm able to do.