University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

2014

The Role of Non-Cognitive Skills in Students' Academic Performance and Life Satisfaction: A Longitudinal Study of Resilience

Rui Yang
University of Pennsylvania, rayeryoung@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Educational Assessment, Evaluation, and Research Commons, and the Quantitative Psychology Commons

Recommended Citation

Yang, Rui, "The Role of Non-Cognitive Skills in Students' Academic Performance and Life Satisfaction: A Longitudinal Study of Resilience" (2014). *Publicly Accessible Penn Dissertations*. 1507.
https://repository.upenn.edu/edissertations/1507

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/1507
For more information, please contact repository@pobox.upenn.edu.

The Role of Non-Cognitive Skills in Students' Academic Performance and Life Satisfaction: A Longitudinal Study of Resilience

Abstract

Research has shown the importance of resilience by demonstrating its significant relationship with students' academic achievement, future workplace performance, and subjective well-being. However, few studies distinguish among different definitions of resilience or distinct approaches of measuring resilience. Also there is little evidence obtained from longitudinal studies involving multiple methods in assessing resilience skills. The current study is able to overcome those limitations and make substantial progress toward the use of resilience scales and the understanding of the predictive power of resilience.

Placing resilience into a broader context of non-cognitive skills, the author identifies four groups of definitions of resilience and successfully places scales of resilience into the same four categories. Using information of nearly four thousand middle school students collected longitudinally at three time points and a resilience scale which consists of three subscales, the author explores the psychometric property of the scale, asks questions on how resilience changes over time and examines the predictive validity of resilience on various future outcomes.

In order to extract the true resilience variance from each of the scale and purify the scale from the wording effect, exploratory factor analysis and confirmatory bi-factor analysis are conducted. The author is able to obtain a single reliable factor which achieves scalar measurement invariance across time for each of the three subscales. However, the attempt to derive a general resilience factor fails because of the low correlations among the three subscale scores.

This paper also presents the results on the change of resilience over time and the relationship between each of the resilience scores and the key outcomes. By fitting different types of hierarchical linear models and growth curve models, the author finds that resilience can significantly predict students' future grade point average and life satisfaction. The relative predictive power of different resilience scores varies by outcome.

Results reveal that resilience is a promising predictor of students' academic learning and life satisfaction. Based on the results, the author provides recommendations for practitioners and researchers. Implications, limitations, and future directions of research are also discussed.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Education

First Advisor

Andrew C. Porter

Keywords

Factor Analysis, Life Satisfaction, Longitudinal Analysis, Measurement, Non-cognitive Skills, Resilience

Subject Categories
Educational Assessment, Evaluation, and Research | Quantitative Psychology

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/1507

THE ROLE OF NON-COGNITIVE SKILLS IN STUDENTS' ACADEMIC PERFORMANCE AND LIFE SATISFACTION: A LONGITUDINAL STUDY OF RESILIENCE

Rui Yang

A DISSERTATION

in

Education

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation:

____

Andrew C. Porter, George and Diane Weiss Professor of Education

Graduate Group Chairperson:

____

Stanton E.F. Wortham, Judy & Howard Berkowitz Professor of Education

Dissertation Committee:

Andrew C Porter, George and Diane Weiss Professor of Education

Robert F. Boruch, Professor of Education and Statistics

Paul A. McDermott, Professor of Education

Michael A. Rovine, Professor of Human Development, Pennsylvania State University

THE ROLE OF NON-COGNITIVE SKILLS IN STUDENTS' ACADEMIC PERFORMANCE AND LIFE SATISFACTION:

A LONGITUDINAL STUDY OF RESILIENCE

COPYRIGHT

2014

Rui Yang

ACKNOWLEGEMENT

Working on my dissertation study has been both an exciting and an agonizing journey. It is exciting because I have been constantly intellectually challenged. The enjoyment of working on an attractive topic and overcoming the difficulties and making breakthroughs are enormous. It is agonizing because I have been constantly intellectually challenged. The frustration of not being able to solve an issue, the pain of realizing the limitation of previous ideas, the anxiety due to getting stuck and not making progress are also enormous. While I am studying resilience, sometimes I also feel I am taking a resilience test—whether I am able to handle stress and how I can bounce back from setbacks are the questions on the test. I am grateful for this test. I am even more grateful for the people there to celebrate with me when the journey gets exciting and to help and encourage me when it gets agonizing.

First I want to thank my adviser Dr. Andy Porter, who has been my mentor since the beginning of my doctoral study. I was extremely nervous walking into his office as an applicant in the winter of 2009. Andy welcomed me and spoke to me with enthusiasm, shared with me his vision and wisdom, and provided me advice and guidance on my pursuit of knowledge and of happiness, which is exactly what he has been doing for me in the past four years. I am glad that I did not hesitate to come to Penn and become Andy's student. It is one of the best decisions I ever made in my life.

I want to thank Dr. Robert Boruch. As a teacher, Bob is knowledgeable and inspirational. As a friend, Bob is caring, humorous, and always makes people feel comfortable. If I have not seen Bob for a week, I will start missing the smell of tobacco on him and I will wish to come across him on the hallway so that I can see his smile and hear his jokes. I feel so lucky to have Bob as a teacher and also as a friend. Bob, your lecture on randomized controlled trials in 1988 was the best video I have seen at GSE and you did not change at all since then.

The next professor deserving a lot of thanks is Dr. Paul McDermott. Room 322 is my favorite room because of Paul. It was in that room I learned item response theory and hierarchical linear modeling. It was also in that room I heard the second best story during my years at Penn—the story of Paul waking up on the beach and squeezing water from a towel (I bet you never know it is a story about exploratory factor analysis with orthogonal rotation). Paul is also the winner of the best story award—the story on how he met his wife. Paul, thank you for the countless help you gave me and the stories you told me!

I will also never thank Dr. Michael Rovine enough for his ideas and help when I faced challenges during the dissertation study. It is an honor to serve as Mike's teaching assistant for a semester and to learn about how he approaches teaching with passion and pleasure and how he treats his students sincerely and wholeheartedly. I am grateful to be able to attend Mike's classes and to have him in my dissertation committee. Thank you, Mike!

The study would not be possible without the generosity and support from Dr. Richard Roberts at Educational Testing Service. I want to thank Rich for his time and guidance in my dissertation. Rich is an outstanding scholar, a true leader, and a great friend in life. I also want to thank my friends at ETS who has helped me with my study: Dr. Kevin Petway and Dr. Sam Rikoon. You guys are amazing.

Of course I will not forget my dearest friends at Penn GSE. Jun Hwang, thank you for your company for the many nights spending in the lab fighting against project deadlines. It was sweet memory. Jill Bowden and Meredith Richards, thank you for giving me valuable suggestions to my dissertation study as well as my future career. I am so fortunately to have you here. Sebastian Cherng, Katherine Barghaus and Jennifer McMaken, thank you for giving me valuable insights on my research and setting perfect examples for me as I move toward my goal. Christine Lee, thank you for taking care of me at GSE and you make the third floor feel like home to me.

Most importantly, I want to thank my family for their unconditional love. Mom and dad, thank you for teaching me everything in life. I am so grateful that I am able to make you happy and proud. Fang, dating you and marrying you is the best decision I made in my life. Thank you for being a friend, a role model, a cheerleader, and a sweetheart for me in the past 10 years. I cannot wait to starting the new page of my life with you together. I love you.

ABSTRACT

THE ROLE OF NON-COGNITIVE SKILLS IN STUDENTS' ACADEMIC PERFORMANCE AND LIFE SATISFACTION: A LONGITUDINAL STUDY OF RESILIENCE

Rui Yang

Andrew C. Porter

Research has shown the importance of resilience by demonstrating its significant relationship with students' academic achievement, future workplace performance, and subjective well-being. However, few studies distinguish among different definitions of resilience or distinct approaches of measuring resilience. Also there is little evidence obtained from longitudinal studies involving multiple methods in assessing resilience skills. The current study is able to overcome those limitations and make substantial progress toward the use of resilience scales and the understanding of the predictive power of resilience.

Placing resilience into a broader context of non-cognitive skills, the author identifies four groups of definitions of resilience and successfully places scales of resilience into the same four categories. Using information of nearly four thousand middle school students collected longitudinally at three time points and a resilience scale which consists of three subscales, the author explores the psychometric property of the scale, asks questions on how resilience changes over time and examines the predictive validity of resilience on various future outcomes.

In order to extract the true resilience variance from each of the scale and purify the scale from the wording effect, exploratory factor analysis and confirmatory bi-factor analysis are conducted. The author is able to obtain a single reliable factor which achieves scalar measurement invariance across time for each of the three subscales. However, the attempt to derive a general resilience factor fails because of the low correlations among the three subscale scores.

This paper also presents the results on the change of resilience over time and the relationship between each of the resilience scores and the key outcomes. By fitting different types of hierarchical linear models and growth curve models, the author finds that resilience can significantly predict students' future grade point average and life satisfaction. The relative predictive power of different resilience scores varies by outcome.

Results reveal that resilience is a promising predictor of students' academic learning and life satisfaction. Based on the results, the author provides recommendations for practitioners and researchers. Implications, limitations, and future directions of research are also discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Overview of Non-cognitive Skills

Cognitive ability refers to individuals' ability to process information, abstract, reason, remember and relate. Cognitive ability can be measured by an intelligence quotient (IQ) test. Decades of research on education has shown that students' cognitive ability is predictive of their future (e.g., Deary, Strand, Smith, & Fernandes, 2007; Brody, 1992; Whalley & Deary, 2001). It has been well established that students' cognitive ability is significantly correlated with measures of academic achievement (Deary, Strand, Smith, & Fernandes, 2007; Bartels, Rietveld, Van Baal, & Boomsma, 2002), highest degree earned (Brody, 1992), employment status (Deary et al., 2004), and actual earnings (Whalley & Deary, 2001). Because of the belief that cognitive ability is the cornerstone to academic achievement and high academic achievement is the bridge to success, it is ubiquitous that school systems all over the world spend much resource in training students' cognitive abilities—memorization, calculation, abstract and critical thinking, information synthesis, and understanding of written materials, etc. For example, in the United States, under the federal No Child Left Behind Act, students' academic achievement not only determines their advancement, it also has consequences for teachers, principals, and schools (Darling-Hammond, 2004; Meier & Wood, 2004). In East Asian countries like China, Korea, and Japan, test score serves as the dominant factor in the college application process (Brown & Park, 2002; Bracey, 1996).

However, one equally, if not more, important factor that has been largely neglected by most educational researchers are students' non-cognitive skills, which usually mean skills not directly affected by intellectual capacity. For example, character strength (Park, Peterson, & Seligman, 2004), soft skills (Duncan & Dunifon, 2012), personal skills (Bradshaw, 1985), emotional intelligence (Goleman, 2006) are all concepts belonging to the general non-cognitive skills category. Nevertheless, the dichotomy of cognitive versus non-cognitive is fundamentally flawed. As others have pointed out, nobody is able to name a human behavior that involves no cognitive processes (Bransford, Brown, & Cocking, 2000; Borghans, Duckworth, Heckman, & Weel, 2008). However, since the term non-cognitive has been widely adopted, the term will be used in this study to represent abilities or skills which are not usually captured by an IQ test.

Non-cognitive skills can be placed into two clusters: intrapersonal skills and interpersonal skills. Intrapersonal skills refer to motivation, resilience, time management, self-control, self-efficacy, optimism, and emotional stability. Interpersonal skills include teamwork, communication, negotiation, and relationship building. An informative rubric is provided by Lipnevich, MacCann, and Roberts (2013). In their study, non-cognitive skills in education were classified into four broad categories: attitudes and beliefs (i.e., motivation, self-efficacy), social and emotional qualities (i.e., teamwork, self-regulation), habits and processes (i.e., time management, learning strategies), and personality (i.e., openness to experience, agreeableness).

Growing Popularity among Researchers

Recently, non-cognitive skills have received increasing attention in the field of education and more scholars have recognized their importance (Chamorro-Premuzic & Furnham, 2003; Pianta et al., 2007; Burrus et al., 2011). In 2002, with the support of the US Department of Education and the National Education Association, the Partnership for 21st Century Skills was founded. In their framework, learning and innovation skills (i.e., creativity and cooperation), life skills (i.e., adaptability, self-direction) and cross-cultural skills receive a lot of emphasis. In 2006, the Conference Board, together with three other institutions, published a report expressing concerns that employers viewed a majority of high school and college graduates as inadequately prepared to become successful in the workforce due to a lack of essential workplace skills or soft skills (Casner-Lotto & Barrington, 2006). The National Research Council (Koenig, 2011), after becoming aware of the importance of 21st century skills, further discussed the assessment of interpersonal and intrapersonal skills, e.g., communication, teamwork, self-management, and time management. Researchers proposed to take the results of the non-cognitive skills assessments into account in the higher education admission process (Kyllonen, Walters, & Kaufman, 2005). The Organization for Economic Cooperation and Development (OECD) also realized the importance of non-cognitive skills and planned on testing collaborative problem solving skills in the next wave of the Program for International Student Assessment (PISA) (OECD, 2013).

Non-cognitive skills have also been experiencing a growing popularity among researchers in other fields: economists, psychologists, and sociologists. Economists study non-cognitive skills from the view of education production function; psychologists in early childhood care more about the processes shaping the development of children's non-cognitive skills; industrial and organizational psychologists pay more attention to personality traits predicting job performance. Sociologists focus on gender and ethnicity gaps in non-cognitive skills. Although researchers from different disciplines pursue particular research questions, their studies all suggest that non-cognitive skills are critical to students' success.

Multiple institutions have conducted systematic reviews of studies devoted to non-cognitive skills. The Research Triangle Institute, collaborating with the Spencer Foundation, conducted a review of studies that had examined relationships between various non-cognitive skills and academic outcomes (Rosen et al., 2010). In their review, Rosen et al. (2010) focused on seven non-cognitive skills: motivation, effort, self-regulated learning, self-efficacy, academic self-concept, antisocial and pro-social behavior, coping and resilience. The University of Chicago Consortium on Chicago School Research, in partnership with Lumina Foundation and Raikes Foundation, conducted another review where they categorized non-cognitive factors into five groups and provided a framework on how the groups related to academic performance as well as how the groups were connected to each other (Farrington et al., 2012). A third review was funded by the Education Endowment Foundation and the Cabinet Office and led by researchers from the Institute of Education at the University of London (Gutman & Schoon, 2013). In addition to summarizing findings related to seven non-cognitive skills (Self-perceptions, motivation, perseverance, self-control, metacognitive strategies, social competencies, and coping), Gutman & Schoon (2013) evaluated each factor on its quality of measurement, malleability, and the associated strength of evidence, which provided helpful guidance to researchers, policy makers, and practitioners.

Literature on Non-cognitive Skills

As discussed above, non-cognitive skills have attracted researchers from diverse background. There is mounting evidence that non-cognitive skills are not only key to students' academic achievement, their impacts are crucial throughout life. Instead of structuring the following short review according to the specific non-cognitive factor under study, the evidence is organized by outcomes area.

On academic achievement

Non-cognitive skills are linked to academic achievement. As mentioned above, Farrington et al. (2012) introduced a framework to categorize non-cognitive skills and manifest their connections with each other and with academic achievement. There were five factors in the framework: academic mindsets (e.g., believing in the value of study, believing that abilities can grow with effort), academic perseverance (e.g., grit, delayed gratification, self-control), academic behaviors (e.g., participating, doing homework, organizing materials), learning strategies (e.g., study skills, goal setting), and social skills (e.g., cooperation, empathy). They argued that academic mindsets were the most fundamental factor because it affected perseverance, learning strategies and social skills. Those three factors plus academic mindsets had an impact on academic behavior.

Academic behavior served as a mediating variable between the prior four factors and the academic achievement outcome.

The links between different factors and academic achievement were supported by several studies. For example, children with good social skills were more likely to maintain healthy relationships with their peers and teachers, hence having more positive feelings about school and being more motivated toward learning (Ladd, Birch, and Buhs, 1999). Moreover, teachers might have different expectations for students with different levels of non-cognitive skills, which would transfer to differentiate trajectories in students' academic growth (Espinosa and Laffey, 2003; Arnold & King, 1997). Also, non-cognitive skills like conscientiousness and emotional stability affected students' learning styles and thus influencing their GPAs (Komarraju et al., 2011). On the opposite side, children with more behavioral problems were more likely to be inefficient in the classroom setting and suffered in their learnings (Duncan et al., 2007; Claessens, Duncan, & Engle, 2009).

Research has demonstrated that general non-cognitive skills explain a salient amount of variance of students' learning outcomes (Robbins et al., 2004; Tracey & Sedlacek, 1984; Campbell, Voelkl, & Donahue, 1997; Crede & Kuncel, 2008). More recently, Duckworth (Duckworth, 2013) showed that a measure of students' grit accounted for about 4% of the variance in their GPA and students' self-regulation skills successfully predicted their grade change (Duckworth, Tsukayama, & May, 2010). Duncan et al. (2007) revealed that attention skills significantly predicted achievement scores above and beyond cognitive abilities. Bowden (2013), in her analysis of the NELS data, found that 12% of the black-white achievement gap could be accounted by the gap in their non-cognitive skills. Borghans, Meijers, & Ter Weel's (2008) study found that students' performance in cognitive tests depended significantly on non-cognitive skills, especially on their levels of motivation to do well on the test. MacCann, Fogarty, & Roberts (2012) found that time management skills were significant predictors of achievement. Blackwell, Trzesniewski, & Dweck (2007) studied educational interventions which targeted the concept of growth mindsets (the belief that talents and abilities can be developed through effort and persistence). They found that students with growth mindsets had better academic outcomes years later.

On workforce and life outcomes

Effects of non-cognitive skills extend far beyond academic achievement obtained during school years. Non-cognitive skills are related to people's employment status, job performance, and success in life. Heckman & Rubinstein (2001), by comparing high school drop-outs who passed the General Education Development (those who passed would be awarded a certificate of high school equivalency) with high school graduates, found that the graduates, despite showing no difference in cognitive abilities with drop-outs that passed the GED, were more successful in the labor force. In another study, Heckman, Stixrud, & Urzua (2006) showed that disadvantaged children who participated in the Perry Preschool Intervention Program, which aimed at raising the children's intelligence, had better life outcomes. However, further investigation found that the program's effects on students' learning outcomes vanished very soon. What the program successfully improved were children's non-cognitive abilities: personal behavior and social development. Based on the findings from the two studies, Heckman, Stixrud, & Urzua (2006) concluded that non-cognitive skills, e.g., conscientiousness, perseverance, sociability, and curiosity, mattered for children's later success in life (Heckman & Rubinstein, 2001; Heckman, Stixrud, & Urzua, 2006). Barrick and Mount (1991), by investigating the relationship between the big-five (extraversion, conscientiousness, agreeableness, emotional stability, openness to experience) personality factors and job performance, claimed that conscientiousness, agreeableness, and emotional stability were all positively related to job success. Their study stimulated a wave of studies examining the connection between the five-factor model and job performance (Baker, Victor, Chambers, & Halverson, 2004; Brunello & Schlotter, 2011; Abe, 2005; Poropat, 2009; Hough & Oswald, 2008). Blanden, Gregg, & Macmillan (2006) also established a significant relationship between non-cognitive skills and social-economic status.

On subjective well-being

Students' academic achievement, future job performances and earnings are all important outcomes when it comes to evaluating the quality of education and the effectiveness of an educational system. However, earning high scores, degrees, and money are only parts of the goals. Education is believed to improve people's well-being and to promote happiness, which has been ignored by many education researchers. In the United States, mental health and psychological well-being have been raised as issues of the education system (Ryff, 1989). Not all the students hold a positive attitude toward school and they incur all kinds of mental problems despite the development of cognitive skills. Similar situations prevail in other countries (Hu, 1994). While subjective well-being depends on a variety of factors, including income (Easterlin, 1974; Shin & Johnson, 1978; Diener & Oishi, 2000), marriage status (Andrews & Withey, 1976; Stutzer & Frey, 2006), health status (Campbell, 1976; Van Praag, Frijters, & Ferrer-i-Carbonell, 2003), and ethics (James, 2011; Frey & Stutzer, 2002), nurturing students' non-cognitive skills (i.e., social skills, stress management, emotional stability) seems a promising solution to the problem.

Steel, Schmidt, & Shulz (2008) demonstrated that personality variables accounted for 40% of the variance in subjective well-being. Singh & Jha (2012) showed that faculty members' emotional intelligence was significantly related to their well-being, as measured by occupational stress. Other non-cognitive factors that were proven to carry significant correlations with well-being include emotional reactivity (Tellegen, 1985; Rusting & Larsen, 1998), extraversion and neuroticism (Headey & Wearing, 1992; DeNeve & Cooper, 1998; Lucas & Fujita, 2000), adaptation to environment (Dienier, Oshi, & Lucas, 2003), and striving for goals (Emmons, 1986).

Experimental evidence

As discussed, non-cognitive skills are believed to be updated for human beings because they affect external criteria (e.g., academic achievement, income) and internal criteria (well-being). However, most of the studies described above were correlational or quasi-experimental studies that did not establish causal relationships. Described below are interventions tested with true experimental studies.

Those interventions aimed to improve one or more areas of non-cognitive skills. The focus is on examining whether the change of one or more non-cognitive factors resulted in changes in outcomes. Therefore, experimental studies which did not measure external criteria are not included.

The first intervention is a mindfulness-based intervention developed to improve behaviors and mental health conditions through the enhancement of children's attention (Semple, Lee, Rosa, & Miller, 2010). Three months after the intervention, participants randomized into the treatment group were found to have significantly less attention problems and less behavioral problems. Moreover, there was a significant reduction in the anxiety symptoms for the children in the treatment group who had elevated levels of anxiety before the intervention (Semple, Lee, Rosa, & Miller, 2010). A second intervention is a one-on-one mentoring program intended for 10- to 16-year-olds to improve their self-concept, attitudes, and pro-social skills (Grossman & Tierney, 1998). Grossman & Tierney (1998) found that students randomized into the treatment group were less likely to get involved in anti-social activities, had better relationship with their peers, and had modest but significant gains in GPA. Since the mentoring relationship might have had a positive impact on unmeasured areas as well, the improvement in academics could not be attributed to the enhancement of social skills. A third study was a meta-analysis involving 62 service learning programs and 11,837 students. The goal of the service learning programs was to improve students' social skills, attitudes toward school, and civic engagement. Out of the 62 programs, 21 used randomized controlled designs. It was found that for those 21 studies, students in the treatment group had significant gains in their social skills and their attitudes toward school and learning, as well as their academic performance (Celio, Durlak, & Dymnicki, 2011). The fourth study was a resilience program for children with depressive symptoms (Yu & Seligman, 2002). Because Yu & Seligman (2002) found that a pessimistic explanatory style significantly predicted depressive symptoms, the intervention focused on teaching children optimism—how to use more optimistic explanatory style when facing difficult situations. They found that children in the treatment group used the optimistic explanatory style significantly more than children in the control group and it mediated the prevention of depressive symptoms. Children in the treatment group had less depressive symptoms in the follow-ups.

As discussed, accumulating evidence suggests that non-cognitive skills can be critical to students' academic achievement, job performance and well-being. The current study focuses on one specific area of non-cognitive skills—resilience.

Resilience

Definition of resilience

Different researchers give different definitions of resilience and each definition focuses on one certain aspect of resilience. The definitions of resilience can be placed into four groups to highlight differences and connections among the definitions. The four categories are the trait, the process, the coping, and the outcome.

In the first category, resilience is defined as a set of personal characteristics. For example, Jacelon (1997) defines resilience as the ability to spring back in the face of adversity. Ahern et al. (2006) defines resilience as positive personality characteristics that enhance individual adaptation. Researchers in the second group define resilience as a process which involves the interaction between risk factors and protective factors. For example, Luthar, Cicchetti, & Becker (2000, p. 543) defines resilience as “a dynamic process encompassing positive adaptation within the context of serious adversity”. Egeland, Carlson, & Sroufe (1993) also treat resilience as a capacity that develops over time in the context of person-environment interactions. Category two definitions, which consider the effects of both internal characteristics and environmental factors, can be viewed as a generalization of the category one trait definition. Category three definitions focus on the coping aspect of resilience. Coping, according to Lazarus & Folkman (1984, p. 141), can be defined as “constantly changing efforts to manage specific external and internal demands that are appraised as taxing or exceeding the resources of a person”. For category three, resilience can be referred to as a wide set of skills and purposeful strategies to cope with stress. For example, Wagnild and Young (1993) define resilience as effectively coping with change and misfortune and Wolchik & Sandler (1997) view resilience as successfully coping with stress in everyday life. Finally, in category four resilience is treated as an outcome. Tugade & Frederickson (2004) define resilience as the ability to bounce back from negative experiences. Sapienza & Mastern’s (2011) define resilience as at risk people achieving better than expected outcomes. Rosen et al.’s (2010) definition (positively adapting under stressful situation, and Martin’s (2013) definition of overcoming challenges and difficulties that are part of everyday life both fit this category.

Definitions of resilience can also be distinguished one from another in terms of the target population. Some researchers restrict the possession of resilience to only a group of individuals who are at risk or are facing serious trauma or adversity. Other researchers argue that resilience is a capacity to overcome challenges and difficulties that are part of everyday life (Martin, 2013; Wolchik & Sandler, 1997).

In the current study, resilience is defined as how well an individual deals with stressful situations, challenges and setbacks. The author does not restrict resilience to a special group of at-risk people or require the existence of adversity as a prerequisite for people to show resilience. In today's world, competition is ubiquitous; all students are facing a higher level of pressure than ever before. Pressure comes from various sources: physical changes (Steinberg, 2008), the emergence of a sense of identity (Hankins, Roberts & Gotlib, 1997), family crisis, the burden to do well in class, and undesirable interactions with peers and teachers (Lazarus & Folkman, 1984; McMahan, 2009), etc. Success in school and later in life requires people to effectively handle stressful situations and frustrations. Not all children get a chance to fight against adversity in early stages of their lives. However, they all face difficult circumstances and setbacks. Whether students overcome challenges and utilize setbacks as step-stones for improvement separates them from those who do not. Whether children persist or give up in challenging situations can cumulatively make a material difference to their learning (Boekarts, 1993; Skinner & Pitzer, 2012). If students fail to bounce back under pressure, it might cause problems related to their learning and their psychological well-being (Masten & Coatsworth, 1998; Tinsley & Spencer, 2010).

Figure 1 below summarizes the relationships among different aspects of resilience. The hypothetical framework is not tested in this study and therefore only serves to illustrate the connections among different aspects of resilience. On one end of the causal chain are the trait aspects (personal characteristics and beliefs) and process aspects (contextual and environmental factors) of resilience. These do not directly measure resilience but to some degree may shape resilience through effects on more specific strategies and behaviors, which are the coping aspects of resilience. On the other end of the causal chain are the outcome aspects of resilience, which synthesize all other aspects of resilience and have a direct effect on outcomes. Besides the direct effect, the resilience outcome also moderates the relationship between stressors and outcomes, indicating that for students with various levels of resilience, the effects of setbacks and challenges on their learning and well-being can differ. The coping aspect of resilience fit the central position of the framework. Not only does it mediate the effects of personalities/belief and protective factors on the outcome, it contributes to the outcome itself.

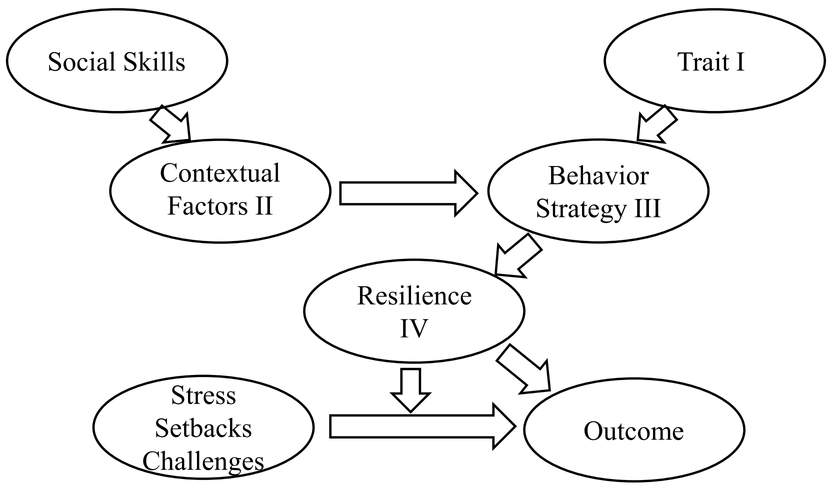

Figure 1: Relationship among different aspects of resilience

Importance of resilience

As can be seen from the previous definition section and the later measurement section, resilience is affected by a variety of factors, including individuals' personality characteristics, their beliefs and self-perception, their coping strategies, social skills and also their environmental factors (Chan, Yeh, Peng, & Yen, 2009; Rak & Patterson, 1996). Below I summarize the evidence on the importance of resilience according to the different aspects examined. Within each aspect of resilience, there is evidence on the relationship between resilience factors and two important outcome measures—academic achievement and subjective well-being.

Based on the teacher-rated conscientiousness, agreeableness, and ego-resiliency of 445 ethnically diverse children, Kwok, Hughes, & Luo (2007) discovered that a latent construct of resilience significantly predicted students' achievement one year later controlling for previous achievement and IQ. Studies (Benson, 2002; Brooks, 2006; Henderson & Milstein, 2003) also found that strengthening the resilience of students could help them reach their potential and even prevent dropout.

Unlike the above studies which directly measure people's abilities to deal with stress, research also reveals traits related to resilience. For example, Gerber et al. (2013) studied the construct of mental toughness (the quality which determines how people respond to stress and challenges) and concluded that after controlling for confounds, baseline mental toughness predicted depressive symptoms and life satisfaction over time. Other traits studied by researcher include goals and aspirations (Dickson & MacLeod, 2004), emotional intelligence (including self-control, and the ability to regulate emotions)

(Tugade & Frederickson, 2008; Garmezy, 1974), self-efficacy (Ehrenberg, Cox, & Koopman, 1991; Benard, 1991), and problem solving (Frye & Goodman, 2000), all of which help promote positive development of adolescents and prevent depression. Contextual factors also play a vital role in resilience. Studies demonstrated that a healthy and supportive relationship between students and their families and peers could help those students better deal with stress and challenges (Hamre & Pianta, 2001; Jackson & Warren, 2000).

Besides the studies demonstrating the importance of the personal characteristic and the process aspects of resilience, there were also studies which established predictive relationships between the coping aspects of resilience and key outcomes (Newman et al., 2000; Plybon et al., 2003; Crean, 2004). Newman et al. (2000), by interviewing 29 urban, low income and minority students who had success in academics, found a common characteristic was good coping strategies (e.g., be dedicated, keep up with homework, hang with the right people, etc.). Similar findings were obtained by Plybon et al. (2003). They analyzed a sample of 84 African American girls from urban families and showed that the use of supportive coping was significantly linked with better academic achievement. Successful coping skills were also related to well-being. With a sample of 304 inner city Latino students, Crean (2004) used structural equation modeling to test the mediating effects of adaptive coping between social support/social conflict and students' academic competence as well as psychological well-being. He found that adaptive coping strategies were negatively related to mental and behavioral problems.

As shown above, resilience research explores personalities, attitudes and beliefs, protective factors, and behaviors that result in positive outcomes despite the risk for maladjustment (Luthar, Sawyer, & Brown, 2006; Masten, 2004). However, there are several weaknesses in the studies mentioned above. First, researchers did not distinguish among different aspects of resilience and treated them all in the same way. Second, all the studies used only a single measure of resilience, therefore capturing only part of resilience. Third, most of the studies measured resilience at only one point in time and measured the outcomes either at the same time or at a later time point, resulting in a lack of ability to examine the change of resilience across time or the longitudinal effects of resilience on outcome measures. The current study, due to features of both the design and the measurement of resilience, is able to address all the weaknesses mentioned above.

Measurement of resilience

Before introducing the instrument used in the current study, some popular resilience scales that have been utilized by researchers and practitioners in recent decades will be reviewed. Two efforts to review and compare different resilience scales have been undertaken (Ahern et al., 2006; Windle, Bennett, & Noyes, 2011). Both focused on comparing the concurrent and predictive validity of the resilience scales while ignoring the theoretical foundations on how those scales were constructed.

As discussed previously, there are four different categories of definitions of resilience. Those four definitions each focus on one aspect of resilience: the trait aspect, the process aspect, the coping aspect, and the outcome aspect. Corresponding to the four aspects of resilience, there are four different approaches to measure resilience: the trait approach, the process approach, the coping approach, and the outcome approach.

The trait approach is the most common one. It aims to measure personal characteristics that are strongly related to resilience. Usually items under such scales contribute to different factors affecting resilience. Connor-Davidson Resilience Scale (CD-RISC) (Connor & Davidson, 2003), Adolescence Resilience Scale (ARS) (Oshio et al., 2002), and Resilience Scale (RS) (Wagnild and Young, 1993) are all constructed using this approach. The second way to build a resilience scale is to focus on the resilience process—how the protective factors help individuals deal with pressure and setbacks. It has been well documented that protective resources can interact with risk factors to influence health-enhancing behaviors (Davey, Eaker, & Walters, 2003; Hunter, 2001). Protective factors refer to environmental factors, for example, family bond, friendship, support in the community, and caring in the school. They sometimes include personal traits too as internal protective factors. Scales in this category include the Resilience Scale for Adults (RSA) (Friborg et al., 2003), the Healthy Kids Survey (HKS) (Hanson & Kim, 2007), the Resilience Scale for Children and Adolescents (RSCA) (Prince-Embury, 2005), and the Baruth Protective Factors Inventory (BPFI) (Baruth & Caroll, 2002). Thirdly, the coping approach to measure resilience focuses on respondents' specific set of skills and purposeful strategies in response to stress and challenges. As discussed before, coping is defined as constantly changing efforts to manage demands that exceed the resources of a person (Lazarus & Folkman, 1984). Scales grouped into this category include the brief resilience coping scale (BRSC) (Sinclair & Wallston, 2004), Coping Responses Inventory—Youth Form (Moos, 1995), and the Children's Coping Questionnaire (Fedorowicz, 1995). Finally, the fourth way to construct a resilience scale uses a more direct outcome approach. Items written by researchers here indicated an effect of exposure to stress. The brief resilience scale (BRS) (Smith et al., 2008) stands for scales in this category.

Trait Approach

When developing the Connor-Davidson Resilience Scale (CD-RISC), Connor and Davidson (2003) first summarized the psychological characteristics of resilient people. Their summary was drawn from three different sources. The first is Kobasa (1979)'s work with the construct of hardiness. Connor and Davidson believed that resilient people tended to view change as an opportunity and they had higher levels of commitment and control. A second group of characteristics, including self-efficacy, close and secure attachment to others, and sense of humor, came from Rutter's (1985) work. The third source was Lyon's (1991) study of people recovering from trauma from which Connor and Davidson extracted characteristics like patience and tolerance of negative effects. Besides those three sources, Connor and Davidson (2003) also added two more characteristics—optimism and hope—resulting in a list of 18 psychological characteristics, on which they developed their scale. The resulting CD-RISC was made of 25 items; the logic behind the scale was to assess resilience by measuring its characteristics (Richardson, 2002) instead of measuring the resilience process or the theory of resilience. This way of measuring resilience was based on Mrazek and Mrazek's (1987) cognitive appraisal theory of resiliency, in which they assumed resiliency emerged from personal characteristics and beliefs which enable people to use particular skills in stressful situations. Some sample items from the CD-RISC are: *Able to adapt to change; See the humorous side of things; Best efforts no matter what; Strong sense of purpose; Think of self as strong person.* Connor and Davidson (2003) also conducted a factor analysis of their resilience scale to test its construct validity. They found that a four-factor structure fit the data the best. The four factors were optimism, future orientation, belief in others, and independence, which were consistent with the characteristics they used to develop the items.

A second scale applying the same approach to measure resilience is the Adolescence Resilience Scale (ARS) developed by Oshio et al. (2002). Similar to Connor and Davidson, Oshio et al. (2002) reviewed some key earlier studies (e.g., Bandura, 1989) and created items to reflect three psychological characteristics: novelty seeking, emotional regulation, and positive future orientation. Although Oshio et al. (2002) believed their items would reflect a three-factor structure, they used a total score of all items to predict scores on several health scales. The three factor hypothesis was not tested. The scale includes a total of 21 items. Some illustrative items are: *I seek new challenges*; *I find it bothersome to start new activities*; *I can stay calm in tough circumstances*; *I lost interest quickly*; *I have difficulty in controlling my anger*; *I am sure that good things will happen in the future*; *I feel positive about my future.*

A third resilience scale using the same theoretical approach was the Resilience Scale (RS) developed by Wagnild and Young (1993). Their conceptual foundation consisted of the following five characteristics: perseverance, equanimity, meaningfulness, self-reliant, and existential aloneness. They tested the scale's internal consistency, concurrent validity and construct validity and concluded that RS performed well under those criteria (Wagnild, 2009). RS included a total of 25 items. Some of the sample items were shown below: *When I make plans I follow through them; I feel proud that I have accomplished things in my life*; *I have self-discipline*; *I keep interested in things*; *My life has meaning*; *When I'm in a difficult situation, I can usually find my way out of it.*

Process Approach

The second approach of building a resilience scale focuses on measuring the protective factors which play an important role in the resilience process. Scales in this category assess a variety of protective resources: peer relationship, teacher support, parental support and expectation, etc. Besides those environmental resources, some scales in this category attempt to measure some personal characteristics similar to scales in the first category. Researchers view those characteristics as internal resources that might help individuals fight against setbacks and challenges.

The Resilience Scale for Adults (RSA) is one scale belonging to this group. Developed by Friborg et al. (2003), RSA consists of 37 items measuring five factors that are assumed to have an impact on resilience. The five factors are personal competence, social competence, family coherence, social support, and personal structure. While personal competence, social competence, and personal structure measure the personal characteristics of individuals, the other two factors target the environmental protective factors of resilience.

The Healthy Kids Survey (HKS) (Hanson and Kim, 2007) serves as another example of scales in this category. It measures both the environmental resilience assets (i.e., caring relationships, high expectations, and meaningful participation in schools, home, and community) and the internal resilience assets for children (i.e., empathy, self-efficacy, self-awareness). Illustrative items are: *At my school, there is a teacher or some other adult who really cares about me; At home, there is a parent who always wants me to do my best; I can work out my problems; I feel bad when someone gets their feelings hurt; There is purpose to my life.*

Another scale sharing the features is the Resilience Scale for Children and Adolescents (RSCA) (Prince-Embury, 2005). The scale consists of three subscales: sense of mastery which measures optimism, self-efficacy, and adaptability; sense of relatedness which measures trust, support, comfort, and tolerance; Emotional reactivity, which measures sensitivity, recovery, and impairment. Some sample items are: *If I try hard, it makes a difference*; *If something bad happens, I can ask my parents for help*; *When I am upset, I do things that I later feel bad about*; *When I get upset, I stay upset for several days.*

The 16-item Baruth Protective Factors Inventory (BPFI) (Baruth & Caroll, 2002) is yet another scale in this category. BPFI measures the construct of resilience by assessing four primary protective factors: adaptable personality, supportive environments, fewer stressors, and compensating experiences.

Coping Approach

The third group of scales focuses on the measurement of a specific set of skills and purposeful strategies to deal with stress and challenges. The Brief Resilience Coping Scale (BRCS) (Sinclair & Wallston, 2004) is a 4-item scale designed to measure tendencies to cope with stress in a highly adaptive manner (i.e., *I actively look for ways to replace the losses I encounter in life*). The Coping Responses Inventory—Youth Form also measures resilience from the same angle (Moos, 1990). The inventory measures four factors of resilience coping: problem-solving action, positive reappraisal, emotional discharge, and cognitive avoidance. Another scale sharing similar characteristics is the Children's Coping Questionnaire (CCQ) (Fedorowicz, 1995). CCQ asks students what they would do in different stressful situations. The author developed items according to 3 coping categories (approach coping, non-constructive coping, and avoidance coping) and each category further included 3-6 sub categories. Some illustrative items are: *I try to find out more about what the problem is (approach coping); I let all my feelings out (avoidance coping); and I get mad and yell at someone (non-constructive coping).*

Outcome Approach

Unlike all the scales discussed before which assess either personal characteristics or protective resources, the Brief Resilience Scale (BRS) takes a more direct approach to measuring resilience (Smith et al., 2008). Smith et al.'s (2008) philosophy was to develop a unitary scale made up of as few items as possible instead of items that try to measure different aspects of resilience resources (Windle, Bennett, & Noyes, 2011). Some sample items are: *I tend to bounce back quickly after hard times; I have a hard time making it* *through stressful events; It is hard for me to snap back when something bad happens*; *I usually come through difficult times with little trouble.*

Current Study

Overview of the current study

Data used for the current study comes from the Mission Skills Assessment (MSA) project, a longitudinal study aiming to measure a variety of students' non-cognitive skills and to monitor their changes over time. MSA is funded by the Independent School Data Exchange (INDEX). INDEX is a not-for-profit organization through which member schools share accurate and meaningful educational and operational information in order to promote students' learning and development. MSA is administered to middle school students (6th to 8th grade) in the participating INDEX schools. The Center for Academic and Workplace Readiness and Success (CAWRS) at Educational Testing Service (ETS) supports multiple phases of the study: design, item development, pilot testing, data analysis and report. MSA was initiated in the fall of 2011 with 18 independent schools participating in the first wave of data collection. Fourteen Schools participated in the second wave of data collection in spring 2012. The number of schools increased to 22 in fall 2012 for the third wave. A fourth wave of data collection took place in winter of 2013 with more than 70 schools. The study features a multi-trait multi-method design, for which different methods (self-ratings, teacher-ratings, situational judgment tests (SJT), bio-data) are used to measure students' development in six non-cognitive areas. The six non-cognitive skills measured are: teamwork, time management, resilience, motivation, ethics, and creativity. Besides students' non-cognitive skills, the data set includes demographic information (including age, gender, and ethnicity), achievement data, life satisfaction, mathematics and science engagement, and interests in academic areas. The surveys and questionnaires are administered to students and teachers online, usually in the middle of a semester. At the end of the semester achievement data for students is collected from each participating school. The current study focuses on resilience. Demographic data, achievement data, life satisfaction information as well as all items from the resilience scale are used in the study.

Data

A total of 22 schools and 3,882 students participated in the study. For wave one, there are 2078 students from 18 schools. For wave two, there are 1678 students from 14 schools and for wave three there are 2641 students from 22 schools. Table 1 below summarizes the basic information. As can be seen from the table, about half of the participants were male and half were female. Not many students were from minority ethnicity groups. Ethnicity information was not collected from students in the second wave of data collection. For students in wave two who were also in either wave one or three, information from those waves were used to fill in the wave two missing ethnicity data. However, for those who only participated in the second wave of data collection, the information on their ethnicity was missing.

Table 1:

Descriptive statistics of participants

Descriptive Statistics	Time 1	Time 2	Time 3
# Students	2078	1678	2641
# Schools	18	14	22
Male	0.49	0.50	0.50
White	0.72	N/A	0.63
Black	0.05	N/A	0.06
Hispanic	0.02	N/A	0.02
Asian	0.06	N/A	0.07
Other	0.14	N/A	0.21

Figure 2 below shows the pattern of students' participation. Grid means participation and blank represents absence/missing. For example, the first column represents the students who participated in the study since wave 1 and stayed in the study in all three waves of data collection. Out of all 2078 students who took the survey in the first wave, 1491 remained in the study through wave 2 and 852 students participated in all three waves. A major reason for the large attrition rate was graduation. Most of the students in the 8th grade graduated in the next academic year therefore becoming ineligible for continued participation.

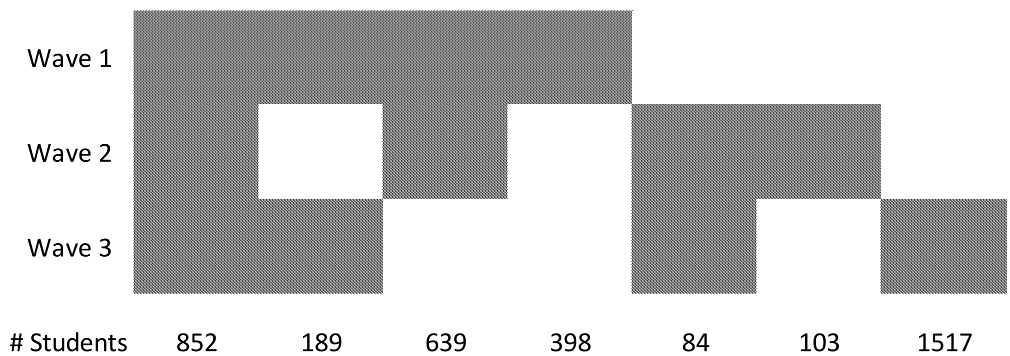

Figure 2: Patterns of Student Participation

Instrumentation

The resilience scale used in this current study is made up of three subscales. All items in the three subscales can be viewed in later sections. The first subscale is a student self-report scale of thirty-six four-point likert items. The four choices are “*Never or Rarely*”, “*Sometimes*”, “*Often*”, “*Usually or Always*”. The self-rating subscale is developed using a combination of the trait approach and the outcome approach. It includes items measuring both the personality aspect of resilience (for example, items measuring self-efficacy or emotional stability) and the direct outcome of resilience. Some sample items are: *When I try, I generally succeed*; *I remain calm under pressure*; *I get upset easily*; *I overcome challenges and setbacks*; *I give up easily when faced with an obstacle*. The second scale is a teacher-report scale of eight four-point likert items. Similar to the self-rating subscale, it combines the trait approach and the outcome approach to measure resilience. Some items under this subscale are: *(The student) remains calm under pressure*; *(The student) overcomes challenges and setbacks*. The third subscale is a student-reported situational judgment test (SJT) scale of 32 four-point likert items. In the SJT, students were presented with a hypothetical situation and were asked about their responses under the situation. An example of an SJT test is presented below:

You are feeling stressed about the amount of homework that you have been given by your teacher(s). Below are some ways that you might think, feel, or act in this situation, right at the time that you feel stressed-out. Rate how often you do each activity *when you feel stressed. How do you think, feel, or act when you are stressed from having too much homework to do?*

Some example items associated with this situation are: *I blame myself for having put off my homework*; *I go out and buy myself something*; or *I try to get organized to get on top of my homework*. All the items are on a four-point likert. The items in the SJT were designed to be aligned to one of the three general coping strategies: emotion-focused coping, avoidance-focused coping, and problem-focused coping.

Review of instrument

In what follows, the literature is used to assess the strengths and weaknesses of each of the three methods to assess resilience. Not surprisingly, the assessment of resilience and more broadly, of non-cognitive skills remains an unsettled issue among researchers. Lipnevich, MacCann, & Roberts's (2013) study provides a brief introduction of both traditional and novel ways for measuring non-cognitive skills. Conventional methods usually include self-assessment, other-ratings, bio-data, and interviews. Novel methods include situational judgment, day reconstruction method, implicit association test, forced-choice tests, and the Bayesian truth serum. Each method has its unique advantages and faces unique challenges regarding different criteria of psychometrically sound measures. Here the author focuses on the three methods applied to measure resilience in the current study.

Self-ratings

Although self-ratings have been used by people for a long time and are easy to design and administer, a considerable weakness to this method is the possibility of faking answers (Burrus et al., 2011). Viswesvaran and Ones' (1999) meta-analysis demonstrated that people could manipulate their personality scores if needed. Birkeland et al. (2006) showed that for selection purposes, between 20% and 40% of people faked their answers in personality tests. By asking job applicants to complete the same non-cognitive scale which they took a month ago during real job application process, Griffith, Chmielowski, & Yoshita (2005) found that between 30% and 50% of applicants intentionally elevated their scores when applying for a job. More detailed descriptions of the advantages and disadvantages of self-ratings can be found in Paulhus (1991) and a thorough discussion of response bias related to self-ratings can be found in Lucas & Baird (2006). Faking also poses threats to biodata measures (Schmitt & Kunce, 2002; Cullen, Sackett, & Lievens, 2006) and interviews (Ellis, West, Ryan, & DeShon, 2002; Levashina & Campion, 2006). In the data set analyzed here, no high stakes are attached to the results. Thus, the motivation to fake answers may be small.

Teacher-ratings

Kenny's (1994) work provides a comprehensive framework of interpersonal perception. His study demonstrated that by capturing the richness of social interaction, interpersonal perception enhanced the traditional measure of individual perception. Crandall (1976) used both self-ratings and other-ratings to measure respondents' quality of life. He found the convergent validity was .33 and suggested that researchers could use other-ratings as validation criteria for self-ratings. Wagerman and Funder (2007) examined the predictive power of both self-rated and peer-rated conscientiousness. They found the two measures had a correlation of .39 and both of them significantly predicted

GPA 3 years later. But the peer-rated measure had a stronger relationship than the self-rated measure. In another study conducted by Dalley, Bolocofsky, & Karlin (1994), both self-ratings and teacher-ratings were used to measure students' social competency. They found that on average, self-ratings were higher than teacher-ratings. It was not obvious which rating was consistently more valid than the other. Burrus et al. (2011) suggested people not make an arbitrary decision between the two because often the two methods each explained unique variance of the outcome variable.

Situational judgment test

Situational judgment test (SJT) is becoming increasingly popular in measuring non-cognitive constructs (Hanson & Ramos, 1996; McDaniel et al., 2001). During the administration of a SJT, individuals are presented with a specific situation and asked to select the most appropriate response from a pool of possible answers. SJTs were believed to reflect more complex judgment processes thus overcoming limitations to validity when compared with traditional assessment methods (Lipnevich, MacCann, & Roberts, 2013). Lievens & Coestsier (2002) showed that SJTs were able to better predict success in college compared with cognitive tests. Oswald et al. (2004) also demonstrated that SJTs were able to provide incremental validity when used in combination with standardized test scores. Empirical results also indicated that SJTs significantly predicted a leadership criterion related to the effectiveness of leadership skills and the initiative of seeking a leadership role. (Krokos et al., 2004) and a situational judgment test of emotion management was significantly related to psychological well-being (Burrus et al., 2012). McDaniel et al. (2001) also found that SJTs appeared to have less negative impact on minorities. Moreover, SJT were less vulnerable to faking than self-ratings (Burrus et al., 2011).

CHAPTER 2: RESEARCH QUESTIONS AND METHODS

Research Questions

As introduced previously, there are three resilience subscales in the study: a student self-rating subscale which measures resilience using the trait and the outcome approach, a teacher-rating subscale which also applies the trait and the outcome approach, and a SJT subscale which measures resilience from a coping perspective. The first research question is about the psychometric properties of the scale. How reliable are the three subscales? What is the factor structure for each of the subscales? If a subscale is unidimensional, is it possible to retain a single underlying factor? Or if a subscale is multidimensional, what do the factors represent? Is it possible to derive a composite score of resilience from all three subscales? Does the factor structure remain stable across time? After knowing the psychometric properties of the resilience measures, a second research question asks about change in resilience over time. If there is a single underlying factor for all items, how does that factor score change? If it turns out that multiple factors of resilience exist, how does each of the relevant scores change? Does resilience vary by gender and ethnicity? The third question asks about the predictive validity of the resilience score/scores. Is resilience a significant predictor of students' academic achievement? Does resilience significantly predict students' life satisfaction? If there are multiple resilience scores, which score has the strongest predicting power? The three research questions in the current study are summarized below:

1. What are the psychometric properties of the resilience scale in this study? What are the factor structures of the measures?
2. How does resilience change over time?
3. What is the relationship between respondent's resilience and their academic achievement and life satisfaction?

Methods

Research question No.1

Self-rating subscale

Item-analysis was applied to the self-rating subscale. Item-total correlation and corrected alpha were calculated for each item. Two problematic items ("I seldom get mad", "I determine what happens in my life"), indicated by having an item-total correlation outside of the .2 to .8 range or by having a corrected alpha larger than the overall alpha, were excluded from further analysis (Streiner & Norman, 2003).

As discussed before, the trait approach and the outcome approach are measuring different aspects of resilience. One advantage of measuring resilience through personal characteristics is that those constructs capture the richness of factors contributing to resilience. One weakness is that measuring some characteristics of resilience is not equal to measuring resilience itself. If a score from a scale made up of trait items is found to be significantly related to an external criterion, it is hard to tell which of the personal characteristics plays the key role. In contrast, the advantage of measuring resilience directly through outcomes is the confidence in interpreting the result as resilience. The disadvantage resides in the fact that the measure does not assess mechanisms of the resilience process.

To sum up, items fitting into the outcome approach directly measure resilience while items belonging to the trait approach assess traits which affect resilience but are not resilience. Trait items capture both resilience variance and the variance unique to the trait but not related to resilience, e.g., irrelevant variance. One option is to discard the trait items. However, the trait items are the majority of items in the subscale. This threatens the reliability of the scale. Another approach is to use bi-factor analysis. A confirmatory bi-factor analysis has the ability to divide the variance of the items into a common factor (primary factor) and unique factors (group factors) (Chen, West, and Sousa, 2006). Assuming there is a general resilience factor and several group factors measuring traits related to resilience, a bi-factor structure makes all trait items load both on their respective group factors and on the primary resilience factor. For outcome items, since they are already directly measuring resilience, they load only on the primary resilience factor. By fitting such a confirmatory bi-factor analysis model, the primary resilience factor can pick up the variance from the items that is directly related to resilience and relegate the unique variance not related to resilience to the group factors.

Before proceeding to fit a bi-factor model to the data, the factor structure was determined. Three methods were tried. The first method was an interview with the export who was involved in the development of the scale. With the expert's help, the author was able to determine the source for each item. Six items were identified as directly measuring the outcome of resilience: "I am capable of coping with most of my problems", “I overcome challenges and setbacks”, “I am resilient”, “I give up easily when faced with an obstacle”, “I am easily discouraged”, and “I give up easily”.

The second method was a Q-sort. Based on the interview and the item descriptions, the author created four groups representing four constructs: resilience, conscientiousness, self-efficacy, and emotional stability. A convenience sample of 16 individuals (mostly students in the Graduate School of Education) participated in the Q-sort. They were asked to sort items into the four categories. In case they felt none of the categories matched the item, they were allowed to put the item into an “Other” category. For each item, the percentages of respondents placing the items into each of the four categories were summarized in table 2a through table 2e below. Each table represents a group of items being categorized into a certain category by a majority of respondents (>= 50%).

Table 2a:

Summary of Q-sort results: resilience category

Item	self-efficacy	emotional stability	conscien-tiousness	resilience	other	Group based on Expert Interview
I give up easily when faced with an obstacle	6%	0%	0%	88%	6%	resilience
I am resilient	19%	0%	0%	81%	0%	resilience
I overcome challenges and set backs	25%	0%	0%	75%	0%	resilience
I give up easily	25%	0%	0%	75%	0%	resilience
I am easily discouraged	19%	13%	0%	69%	0%	resilience
I get discouraged when things go wrong	6%	19%	0%	69%	6%	resilience
There are times when things look pretty bleak and hopeless to me	13%	19%	0%	56%	13%	emotional stability

Table 2b:

Summary of Q-sort results: self-efficacy category

Item	self-efficacy	emotional stability	conscien-tiousness	resilience	other	Group based on Expert Interview
I complete tasks successfully	94%	0%	0%	6%	0%	efficacy
I am confident I get the success I deserve in life	94%	6%	0%	0%	0%	efficacy
Sometimes, I do not feel in control of my school work	81%	6%	0%	13%	0%	efficacy
When I try, I generally succeed	81%	6%	6%	6%	0%	efficacy
Overall, I am satisfied with myself	81%	19%	0%	0%	0%	efficacy
I determine what will happen in my life	75%	6%	0%	19%	0%	efficacy
I am filled with doubts about my competence	75%	13%	0%	13%	0%	efficacy
I do not feel in control of my success in school	63%	13%	6%	19%	0%	efficacy
I am capable of coping with most of my problems	56%	13%	0%	31%	0%	resilience
Sometimes when I fail I feel worthless	50%	19%	0%	31%	0%	efficacy

Table 2c:

Summary of Q-sort results: conscientiousness category

Item	self-efficacy	emotional stability	conscien-tiousness	resilience	other	Group based on Expert Interview
I forget to do things	6%	6%	81%	0%	6%	conscientiousness
I make careless mistakes	13%	6%	75%	0%	6%	conscientiousness
I avoid responsibilities	19%	0%	63%	6%	13%	conscientiousness
I am diligent	38%	0%	56%	6%	0%	efficacy
I quickly lose interest in the tasks I start	19%	6%	50%	19%	6%	grit

Table 2d:

Summary of Q-sort results: emotional stability category

Item	self-efficacy	emotional stability	conscien-tiousness	resilience	other	Group based on Expert Interview
I seldom get mad	0%	94%	0%	6%	0%	emotional stability
I get upset easily	0%	88%	0%	13%	0%	emotional stability
I am easily frustrated	0%	81%	0%	19%	0%	emotional stability
Sometimes I feel depressed	0%	81%	0%	6%	13%	emotional stability
I get annoyed by people	0%	75%	0%	13%	13%	emotional stability
I worry	13%	69%	0%	6%	13%	emotional stability
I remain calm under pressure	6%	69%	0%	25%	0%	emotional stability
I am relaxed	25%	63%	6%	0%	6%	emotional stability
I am not easily bothered by things	13%	63%	0%	19%	6%	emotional stability
I remain calm when I have a lot of homework to do	6%	63%	0%	31%	0%	emotional stability

Table 2e:

Summary of Q-sort results: items not clearly classified into any category by respondents

Item	self-efficacy	emotional stability	conscien-tiousness	resilience	other	Group based on Expert Interview
I worry about school	19%	44%	6%	19%	13%	emotional stability
My interests change quickly	6%	6%	25%	31%	31%	grit
I react slowly	6%	19%	19%	19%	38%	grit
I get stressed out easily when things don't go my way	0%	50%	0%	50%	0%	emotional stability

As can be seen, five out of the six items the expert helped to identify were also picked for the resilience category by at least 69% of the respondents. The only exception was “I am capable of coping with most of my problems”, for which 31% of the respondents identified as resilience. There were three other items categorized into resilience by a majority of respondents. The three items were “There are times when things look pretty bleak and hopeless to me”, “I get discouraged when things go wrong”, and “I get stressed out easily when things don't go my way”.

While the above two methods were based on expert judgment, a third way to determine the factor structure was not. Because all the items were on a 4-point likert scale, a polychoric correlation matrix was calculated for all 34 items (2 items were excluded due to their low item-total correlations: “I seldom get mad”; “I determine what will happen in my life”). Using the polychoric correlation matrix, exploratory factor analysis (EFA) was conducted. Squared multiple correlations were used as the initial communality estimates. Minimum average partialling (MAP) (Garrido et al., 2011; Velicer, 1976) and the scree test of eigen values were used to determine the number of potential underlying factors. Varimax, Equamax, and Promax rotation were applied in sequence to obtain the optimal factor structure. Four criteria were used to pick the final factor structure: the hyper-plane count (Yates, 1987), the total number of non-salient loadings and the number of double loader, the closeness to a simple structure (Fabrigar et al., 1999), and the meaningfulness of each factor (Gadermann, Guhn, & Zumbo, 2012). According to the MAP and scree test results, the author fit EFA models with 1 to 5 factors.

A problem was detected while fitting a 2-factor model. It was found that one factor was made up of all the positively worded items and the other factor was made up of all the negatively worded items, indicating that wording/valence had a dominant effect on the items. Therefore EFA was not able to detect the true factor structure and neither was the originally planned confirmatory bi-factor analysis for the sake of extracting the true resilience variance. Additional work was done on the CFA stage to account for the valence effects. Two more factors were created. All the positively worded items loaded on one factor and all the negatively worded items loaded on the other factor. The correlations between the two newly created factors and all other group and primary factors were restricted to zero. The author was treating those two as error factors. The hope was the two error factor would extract all the variance due to valence. Figure 3 here serves as an illustration of the final bi-factor structure adjusted by the error factors.

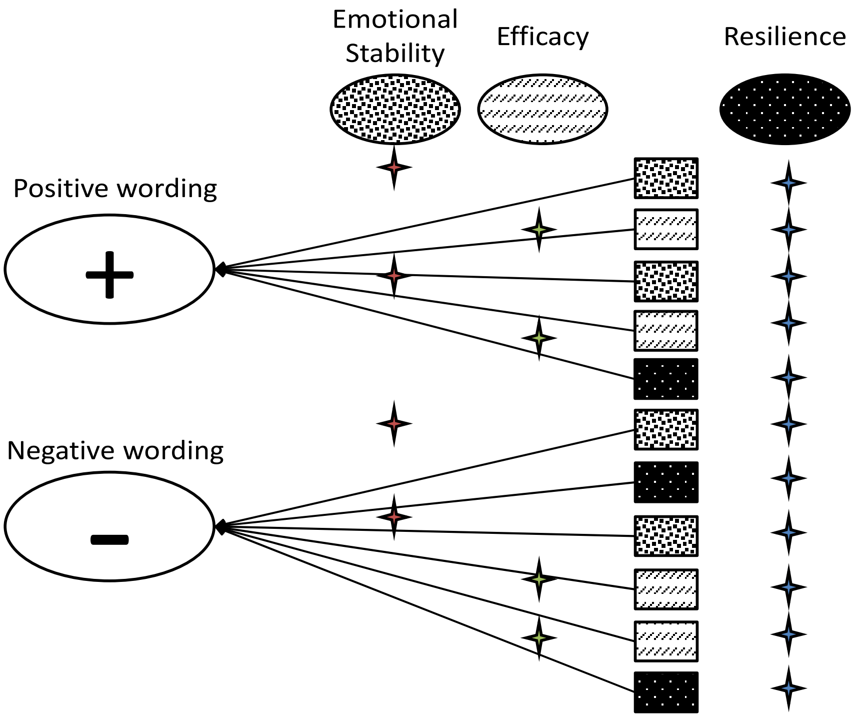

Figure 3: Illustration of the Final Bi-factor Model Structure

Each rectangle represents an item while the pattern of the fill indicates the aspect of resilience the item is intended to measure (white background with black dots or lines means trait items and black background with white dots means outcome items). A white background item loads both on its respective group factor and the primary resilience factor and a black background item loads only on the primary resilience factor. The two error factors are represented by two circles on the left side with one factor underlying all the positively worded items and the other linked with all the negatively worded ones. Those two circles pick up the error variance of valence and the two group factors on the top left pick up the unique variance unrelated to resilience, leaving the primary resilience factor on the top right absorbing all the variance that is directly related to resilience.

Due to the inability of EFA to adjust for valence effects, two separate EFAs were conducted to gather more evidence on the bi-factor structure, in addition to the evidence obtained through the interview and the Q-sort. One EFA was done for the 12 positively worded items and one was done for the 22 negatively worded items. The same procedure and the same sets of criteria, as stated previously, were followed to determine the optimal factor structure. Details about the loading and the factors can be seen from table 3 below. The panel on the top presents the EFA results for positively worded items and the bottom panel contains results for negatively worded items. A 3-factor structure was the best fitting structure for both EFAs. The three factors in the positive group each corresponded with a factor in the negative group therefore items could be collapsed together under the same 3-factor framework. No double-loaders were detected. Four items failed to load significantly on any of the factors: “I am capable of coping with most of my problems”, “I get annoyed by people”, “I react slowly”, “I do not feel in control of my success in school”. Since the first item was identified by the expert as a true resilience item and it had a close-to-salient loading on the resilience factor, the item was kept in the bi-factor analysis. The other three items were excluded from further confirmatory analyses.

Table 3:

Factor Loading for the Two Separate EFAs

		Efficacy	Emtional Stability	Resilience
+	remain_calm	0.06	0.57*	0.11
	diligent	0.24	-0.05	0.45*
	complete_task	0.49*	-0.05	0.34
	try_succeed	0.54*	0.01	0.27
	confident_success	0.67*	0.09	0.07
	satisfied	0.59*	0.19	0.04
	capable_coping	0.18	0.28	0.31
	overcome_setback	0.28	0.20	0.46*
	resilient	0.03	0.19	0.52*
	relaxed	0.18	0.58*	-0.05
	not_bothered	-0.05	0.50*	0.05
	calm_hw	0.14	0.48*	0.11
-	giveup_obstacle	0.10	0.02	0.77*
	discouraged	-0.01	0.29	0.63*
	giveup_easily	0.11	0.04	0.82*
	frustrated	0.26	0.44*	0.13
	annoyed	0.33	0.35	0.01
	avoid_respon	0.58*	-0.03	0.17
	forget	0.60*	0.09	0.02
	careless_mistake	0.50*	0.13	0.05
	lose_interest	0.62*	0.07	0.10
	react_slowly	0.34	0.13	0.17
	interest_change	0.44*	0.05	0.02
	depressed	0.21	0.57*	0.00
	worthless	0.04	0.58*	0.14
	notcontrol_school	0.42*	0.33	0.05
	doubts	0.20	0.42*	0.18
	not_in_control	0.34	0.34	0.10
	hopeless	0.30	0.54*	0.02
	discouraged_wrong	0.04	0.57*	0.23
	worry	-0.05	0.71*	0.04
	stressed_out	0.10	0.62*	0.04
	worry_school	-0.01	0.67*	0.01
	upset	0.09	0.64*	0.12

The results were compared with the results from Q-sort and results from the expert interview. Although the three did not agree perfectly, consistency was reached for a majority of the items. The final bi-factor structure was based on the separate EFA results except for two adjustments. First the item “I am capable of coping with most of my problems” was kept in the CFA despite its lack of salient loadings. Second the item “I am diligent” had neither the expert nor the Q-sort indicate it as resilience, it was excluded from further analyses.

After obtaining the factor structure, confirmatory bi-factor analysis was conducted for time 1 students. Goodness of fit statistics like Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) and Root Mean Square Error of Approximation (RMSEA) were used to determine model fit. According to Kline (2005), a CFI larger than .9 and a RMSEA less than .08 (with upper limit less than .1) signaled acceptable fit. A CFI larger than .95 and a RMSEA less than .05 suggested close fit (Hu & Bentler, 1999). The loading and model fit information for the final model are presented below.

After achieving an acceptable fit for time 1 respondents, the same model was fit to students at time 2 and time 3 to examine factor invariance across time. Configural invariance was firstly checked and based on the model fit of the configural invariance model, more restricted models were fit to examine metric invariance (weak invariance) and scalar invariance (strong invariance) across time (Brown, 2006). Configural invariance meant that the same factor structure held across time. Metric invariance indicated that not only were the factor structures the same across groups, the factor loadings were the same for every item across time. Scalar invariance represented the same factor loadings and the same item thresholds across time. A stronger case of invariance was established when there was no sizable difference considering the model fit statistics (Muthén & Asparouhov, 2002), indicated by either the non-significance of the chi-square difference test (Marsh & Grayson, 1994) or the minimal shift in CFI and RMSEA (change of CFI less than .01 and change of RMSEA less than .015) (Cheung & Rensvold, 2002).

Teacher-rating subscale

Table 4 shows the item descriptions in the teacher-rating subscale.

Table 4:

Items in the Teacher-rating Subscale

Item	Item Description
T1	Remains calm under pressure
T2	Overcomes challenges and setbacks
T3	Does not give up easily
T4	Is resilient
T5	Worries
T6	Gives up easily when faced with an obstacle.
T7	Is easily discouraged.
T8	Gets upset easily

As mentioned before, the teacher-rating subscale was made up of trait items measuring emotional stability (T1, T5, T7, T8) and outcome items directly measuring resilience (T2, T3, T4, T6). Using the polychoric correlation matrix, EFA showed clearly that valence again had a dominant effect on the items. The four positively worded items (T1-T4) loaded on one factor while the four negatively worded items (T5-T8) loaded on the other factor. Two practical concerns prevented the author from fitting a similar confirmatory bi-factor analysis model as the one fit for the self-rating subscale. First the teacher-rating subscale had relative fewer items than the self-rating subscale (8 vs 36) therefore the model would be under-identified. Secondly, due to administrative reasons, more than 40% of the students at time 1 were missing their teachers' ratings on items T5-T7. Only the four positive items were kept and a single factor structure was fit to the data for the CFA without any adjustment for unrelated resilience variance. The single factor structure was also examined for its reliability and its invariance across time.

SJT Subscale

The SJT items were assumed to measure resilience from a coping perspective. There were 32 items nested within 3 hypothetical situations. The first situation was about homework. The other two situations were about after-school activities and testing. After interviewing the expert who was involved in developing the SJT subscale, it was determined that all the items under the three hypothetical situations were developed based on three general coping styles: avoidance-focused coping, emotion-focused coping, and problem-focused coping. EFA proved that the scale manifested a three-factor structure. Table 5 shows the results of the EFA and the detailed descriptions of the three situations. The three factors were avoidance (representing avoidance-focused coping), problem (representing problem-focused coping), and emotion (representing emotion-focused coping). All items had salient loadings except for one. That item was “I call or email classmates and talk through some possible questions and answers with them” and it had a close-to-salient loading on the problem factor. The problem factor had a -.04 correlation with the avoidance factor and a -.10 correlation with the emotion factor. The correlation between the avoidance factor and the emotion factor was .38.

Table 5:

EFA Results for the SJT Subscale

Situation	Item Description	Item	Avoidance	Problem	Emotion
You are feeling stressed about the amount of homework that you have been given by your teacher(s). Below are some ways that you might think, feel, or act in this situation, right at the time that you feel stressed-out. Rate how often you do each activity when you feel stressed. How do you think, feel, or act when you are stressed from having too much homework to do?	I blame myself for having put off my homework.	hw1	-0.06	0.15	0.47*
	I treat myself to a nice meal.	hw2	0.68*	0.09	-0.11
	I try to get organized to get on top of my homework.	hw3	-0.01	0.72*	0.00
	I make the extra effort to get all of my homework done.	hw4	-0.09	0.70*	-0.05
	I go out and buy myself something.	hw5	0.78*	-0.08	0.06
	I blame my teacher(s) for having given me too much homework in the first place.	hw6	-0.03	-0.37	0.44*
	I go for a long walk.	hw7	0.65*	0.15	0.05
	I tell myself that this will never happen again.	hw8	0.13	0.27	0.52*
	I ask myself: How could I have let this happen to me?	hw9	-0.05	0.14	0.71*
	I take control and say to myself: I can do this!	hw10	0.06	0.76*	0.05
	I go out with my friends.	hw11	0.81*	-0.18	-0.06
	I convince myself: This is not really happening to me!	hw12	0.31	-0.03	0.57*
	I wish that I could change what was happening or how I feel.	hw13	-0.20	-0.04	0.74*
	I get some sleep.	hw14	0.72*	0.01	-0.09
You are feeling very stressed about the amount of after-school activities that you are participating in. You seem to be playing too many sports, have too much homework, and you still need to do your chores at home. Below are some ways that you might think, feel, or act in this situation, right at the time that you feel stressed-out. Rate how often you do each activity when you feel stressed-out. How do you think, feel, or act when you are stressed-out from doing too many after-school activities?	I blame myself for having too many activities.	ac1	-0.05	0.05	0.46*
	I treat myself to a nice meal.	ac2	0.72*	0.06	-0.08
	I try to get organized to get on top of all these activities.	ac3	0.07	0.73*	0.02
	I make the extra effort to get all of my activities completed.	ac4	0.00	0.70*	-0.01
	I go out and buy myself something.	ac5	0.82*	-0.05	0.08
	I blame others for having given me too many activities to do in the first place.	ac6	0.02	-0.21	0.52*
	I go for a long walk.	ac7	0.71*	0.12	-0.01
	I tell myself that this will never happen again.	ac8	0.15	0.16	0.58*
	I take control and say to myself: I can do this!	ac9	0.07	0.78*	0.09
You are feeling quite stressed about a class test that your teacher is giving tomorrow. You don’t think you understand all the material covered and you have had to do a lot of other things this week, so you have not studied as hard as you like or normally do. Below are some ways that you might think, feel, or act in this situation, right at the time that you feel stressed-out. Rate how often you would do each activity when you feel stressed-out. How do you think, feel, or act when you are stressed because you have a test and are not as prepared for it as you would like?	I blame my teacher(s) for having given me too many tests in the first place.	em1	0.04	-0.33	0.49*
	I go for a long walk.	em2	0.71*	0.05	0.04
	I try to get prepared for the test regardless.	em3	-0.14	0.65*	-0.03
	I call or email classmates and talk through some possible questions and answers with them.	em4	0.14	0.30	0.12
	I take control and say to myself: I can do this!	em5	0.06	0.8*	0.06
	I go out with my friends.	em6	0.78*	-0.21	0.03
	I convince myself: This is not really happening to me!	em7	0.30	-0.05	0.63*
	I wish that I could change what was happening or how I feel.	em8	-0.18	-0.06	0.76*
	I get some sleep.	em9	0.72*	0.08	-0.09

One potential problem the EFA failed to consider was local dependence: residuals for items nested under the same situation might be correlated with each other (Lee, Dunbar, & Frisbie, 2001). Therefore after moving to the CFA stage, a correlated trait correlated method confirmatory factor analysis (CTCM-CFA) model was fit to the data (Kenny and Kashy, 1992; Marsh, 1989; Kumar & Dillon, 1992). The three traits represented the three coping styles, and three methods factors represented the three hypothetical situations making up the structure. An illustration of the CTCM-CFA model can be seen in figure 4 below.

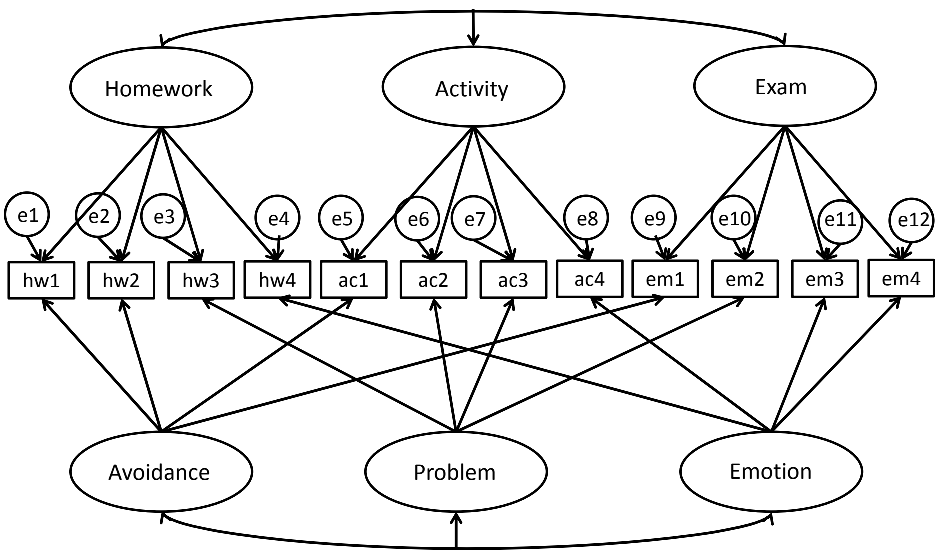

Figure 4: An Illustration of CTCM-CFA Model

Fit statistics showed that the CTCM-CFA model failed to converge to an admissible solution. One possible explanation might be that after the common variance was extracted from the items to the three "trait" factors, the remaining variance for items nested within each situation was not enough to warrant three "method" factors. Therefore a correlated trait correlated uniqueness confirmatory factor analysis (CTCU-CFA) model was fit to the data. CTCU-CFA did not assume the existence of multiple method factors (Marsh & Grayson, 1995). By allowing the residuals for items under the same situation to correlate with each other, it could account for the violation of the assumption of local independence. The model is illustrated in figure 5 below.

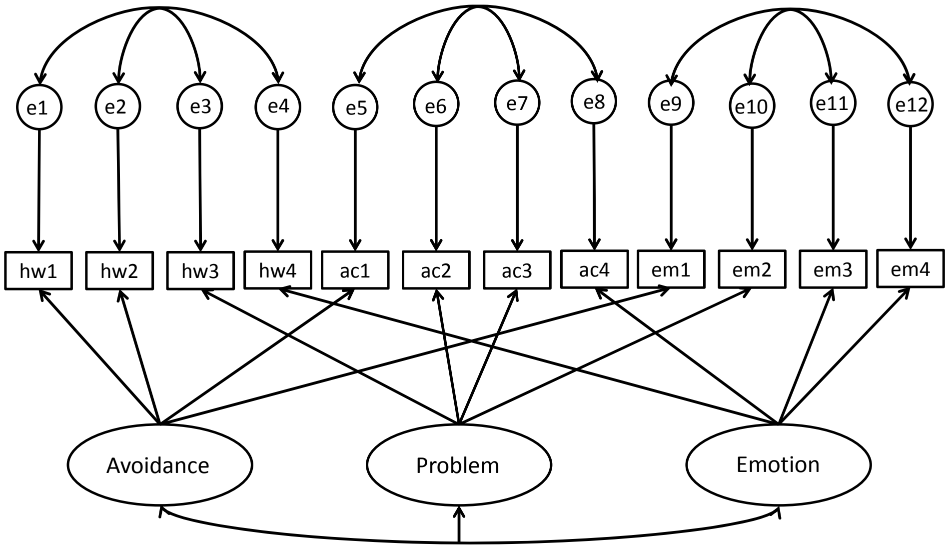

Figure 5: An Illustration of the CTCU-CFA Model

The CTCU-CFA model did not fit the data either. The items associated with the avoidance-focused coping factor and with the emotion-focused coping factor were dropped. One reason was the failure to build a model that could fit all items. Another reason was that out of the three coping factors, problem-focused coping was a better reflection of students' resilience skills. Nine items that loaded on problem-focused coping were retained and a single factor CFA was conducted. Once again the model did not perform well. Modification indices suggested that some of the residuals might be correlated. Adding the correlated residuals to the model was justified because several items under different situations were actually duplicate items. For example, “I take control and say to myself: I can do this!” was an item administered to individuals three times, once under each situation. After accounting for the correlated residuals, the CFA model had a decent fit to the data and factor invariance across time was then examined. Details on the loadings and the fit statistics are available in the results section.

Whole Resilience Scale

As described above, three factors were extracted from the three-subscales. If the three factors shared a large amount of common variance, a higher-order confirmatory factor analysis model might be an ideal model to extract a second-order factor and that factor score could serve as an overall resilience score across all methods. However, after calculating the correlations among the three resilience factors, the author did not attempt to get a composite score across methods because the correlations among the factors were not high enough to derive a reliable higher-order factor. Table 6 here presents the correlation matrix of the three factors scores. The correlations between the teacher-rating resilience score and the other two scores were low: .17 and .15 respectively.

Table 6:

Correlation Matrix of Three Resilience Scores

	self	teacher	sjt
self	1.00	0.17	0.37
teacher		1.00	0.15
sjt			1.00

Research question No.2

For each of the three subscales, one single factor was retained as the resilience factor. The score of that factor was used to represent the resilience score under each method. In order to examine the change of resilience scores over time, scalar invariance needed to be achieved (Widaman & Reise, 1997; Muthén & Asparouhov, 2002). As detailed in the results part, scalar invariance was achieved for all of the three resilience factors, one for each method. Those three factor scores were treated as the response variables in the following models. Series of 3-level hierarchical models were built (Raudenbush & Bryk, 2002), where the three levels were the within-individual temporal level, the student level, and the school level. An unconditional model was built first to obtain the variance decomposition of each resilience factor score and then a model with gender and ethnicity added as student-level predictors were fit to explain the variance of the intercept and slope of resilience on the first level. Technical details of the 3-level HLM can be seen below. The model allowed random effects at the individual and the school level for both the intercept and the slope of change.

Level 1: Temporal Level

$$Resilience\_Score_{tij}=\pi_{0ij}+\pi_{1ij}TIME_t+\epsilon_{tij}$$

Level 2: Individual Level

$$\pi_{0ij}=\beta_{00j}+\beta_{01j}*Gender_{ij}+\beta_{02j}*Ethnicity_{ij}+\xi_{0ij}$$

$$\pi_{1ij}=\beta_{10j}+\beta_{11j}*Gender_{ij}+\beta_{12j}*Ethnicity_{ij}+\xi_{0ij}$$

Level 3: School Level

$$\beta_{00j}=\Theta_{000}+\Psi_{00j}$$

$$\beta_{01j}=\Theta_{010}$$

$$\beta_{10j}=\Theta_{110}+\Psi_{10j}$$

$$\beta_{11j}=\Theta_{200}$$

$$\beta_{02j}=\Theta_{020}$$

$$\beta_{12j}=\Theta_{120}$$

Research question No.3

The third research question tackled the issue of prediction. Because of the relatively low correlations among the three different resilience scores, it was not meaningful to derive a general resilience score. Therefore it was not possible to examine the predictive validity of a single resilience measure and instead, all three different scores were used to answer the third research questions. By including all of the three scores in the model simultaneously, it allowed the author to compare the relative predictive power of the three scores. There were two groups of outcome variables. The first group of outcomes was academic achievement. It included student self-reported GPA, their mathematics and reading grades from the school, and their verbal and quantitative standardized test scores. The second outcome was life satisfaction. The life satisfaction scale was made up of 7 items. All the items were on a 6-point likert scale. Exploratory factor analysis based on the Pearson correlation matrix showed a two-factor structure was the best factor structure. The first factor was made of 5 positively worded items and the second factor was made of 2 negatively worded items. Since a 2-item factor was not reliable enough, only the first factor was retained. CFA with two factors achieved acceptable fit. The two-factor structure also achieved configural invariance across time.

Strong invariance model was also fit and was compared with the configural invariance model. The chi-square difference test was not significant, indicating the model achieved scalar invariance across time. The final scalar invariance model had an RMSEA of .052 (with an upper limit of .059), a CFI of .985, and a TLI of .983. Table 7 summarizes the final loadings for the retained factor and the factor score was used as another response variable in the models.

Table 7:

Final Loadings for the Life Satisfaction Factor

Item Description	Loading	S.E	t	p
I have a good life.	0.87	0.008	113.28	<.001
My life is just right.	0.83	0.009	96.09	<.001
My life is better than most kids.	0.56	0.014	39.20	<.001
My life is going well.	0.85	0.008	103.20	<.001
I have what I want in life.	0.71	0.012	60.63	<.001

A total of three different kinds of models were built. The three models were different regarding their complexity and they each answered the third research question from a unique perspective.

The first model was a 2-level hierarchical linear model using students' resilience scores from time 1 to predict students' outcomes at time 2 and time 3. Level 1 was student level and level 2 was school level. Pretest of the outcome was included as covariate and students' demographic variables were used as controls in the model. Three resilience scores as well as the previous outcome measure were group mean centered at level one and the average scores for each school were used as level-2 predictors. This model provided valuable information on the magnitude and the significance of the relationship between different resilience scores and the outcomes. Below are the technical details of the first model.

Level 1: Student Level

$Outcome_{3ij}=\pi_{00j}+\pi_{10j}Self\_Resilience\_Centered_{1ij}+\pi_{20j}Teacher\_Resilience\_Centered_{1ij} + \pi_{30j}SJT\_Resilience\_Centered_{1ij} +\pi_{40j}Gender_{ij} + \pi_{50j}Ethnicity_{1ij} + \pi_{60j}Outcome\_Centered_{1ij} +\epsilon_{tij}$

Level 2: School Level

$\pi_{00j}=\beta_{000} + \beta_{010}Self\_Resilience\_Mean_{0j} +\beta_{020}Teacher\_Resilience\_Mean_{0j} + \beta_{030}SJT\_Resilience\_Mean_{0j} +\beta_{040}Outcome\_Mean_{0j} +\xi_{00j}$

$\pi_{10j}=\beta_{100}$

$\pi_{20j}=\beta_{200}$

$\pi_{30j}=\beta_{300}$

$\pi_{40j}=\beta_{400}$

$\pi_{50j}=\beta_{500}$

$\pi_{60j}=\beta_{600}$

An unconditional model with no predictors was fit first to get the variance decomposition at each level and then a full model with predictors was built. Results about the variance components and about the coefficients estimates are available in the results part.

The second model was a 3-level longitudinal growth curve model. Level 1 was the temporal level (occasion level within individual). Level 2 became the student level and level 3 was the school level. Similar as model 1, only resilience scores at time 1 were used. However, unlike model 1 where only time 3 outcome was used, outcomes at all three time points were included here and the intercept and the slope at level 1 were treated as the response variables at level 2. This model not only examined the effect of resilience on the average outcome level, it also provided information on whether resilience had an impact on students' rate of change on the outcome variable (Singer, 1998). Resilience scores at level 2 were again group mean centered and the average scores for each school were grand mean centered and included as predictors at level 3. Technical details of the model are shown below.

Level 1: Temporal Level

$Outcome_{tij}=\pi_{0ij}+\pi_{1ij}TIME_t+\epsilon_{tij}$

Level 2: Individual Level

$\pi_{0ij}=\beta_{00j}+\beta_{01j}*Gender_{ij}+\beta_{02j}*Ethnicity_{ij}+ \beta_{03j}*Self\_Resilience\_Centered_{1ij} + \beta_{04j}*Teacher\_Resilience\_Centered_{1ij} + \beta_{05j}*SJT\_Resilience\_Centered_{1ij} +\xi_{0ij}$

$\pi_{1ij}=\beta_{10j}+\beta_{11j}*Gender_{ij}+\beta_{12j}*Ethnicity_{ij}+ \beta_{13j}*Self\_Resilience\_Centered_{1ij} + \beta_{14j}*Teacher\_Resilience\_Centered_{1ij} + \beta_{15j}*SJT\_Resilience\_Centered_{1ij} +\xi_{0ij}$

Level 3: School Level

$\beta_{00j}=\Theta_{000}+ \Theta_{001}Self\_Resilience\_Mean_{1j} + \Theta_{002}Teacher\_Resilience\_Mean_{1j} +\Theta_{003}\ SJT\_Resilience\_Mean_{1j} +\Psi_{00j}$

$\beta_{01j}=\Theta_{010}$

$\beta_{02j}=\Theta_{020}$

$\beta_{03j}=\Theta_{030}$

$\beta_{04j}=\Theta_{040}$

$\beta_{05j}=\Theta_{050}$

$\beta_{10j}=\Theta_{100}+\Theta_{101}Self\_Resilience\_Mean_{1j}+\Theta_{102}Teacher\_Resilience\_Mean_{1j}$

$+\Theta_{103}\ SJT\_Resilience\_Mean_{1j}+\Psi_{10j}$

$\beta_{11j}=\Theta_{110}$

$\beta_{12j}=\Theta_{120}$

$\beta_{13j}=\Theta_{130}$

$\beta_{14j}=\Theta_{140}$

$\beta_{15j}=\Theta_{150}$

The third model was also a longitudinal growth curve model. Similar to model 2, level 1 was temporal level. Level 2 was student level and level 3 was school level. The main difference between model 2 and model 3 was that not only the outcome at all three time points were used, the resilience scores at all three time points were also included. The resilience scores were treated as time-varying predictors and were included at level-1 to explain the between-individual variance (McCoach & Kaniskan, 2010). Each of the three resilience scores was group-mean centered at level-1 and the individual mean scores (also group mean centered by the school mean) were included at level-2 to explain the between-individual variance on the intercept and the slope along with gender and ethnicity. School means scores were grand mean centered and used as level-3 predictors. Technical details of model 3 are presented below.

Level 1: Temporal Level

$Outcome_{tij}=\pi_{0ij}+\pi_{1ij}TIME_t+ \pi_{2ij}Self\_Resilience\_Centered_{tij} +$

$\pi_{3ij}Teacher\_Resilience\_Centered_{tij} + \pi_{4ij}SJT\_Resilience\_Centered_{tij} +\epsilon_{tij}$

Level 2: Individual Level

$\pi_{0ij}=\beta_{00j}+\beta_{01j}*Gender_{ij}+\beta_{02j}*Ethnicity_{ij}+ \beta_{03j}*Self\_Resilience\_Mean_{1ij} +$

$\beta_{04j}*Teacher\_Resilience\_Mean_{1ij} + \beta_{05j}*SJT\_Resilience\_Mean_{1ij} +\xi_{0ij}$

$\pi_{1ij}=\beta_{10j}+\beta_{11j}*Gender_{ij}+\beta_{12j}*Ethnicity_{ij}+ \beta_{13j}*Self\_Resilience\_Mean_{1ij} +$

$\beta_{14j}*Teacher\_Resilience\_Mean_{1ij} + \beta_{15j}*SJT\_Resilience\_Mean_{1ij} +\xi_{0ij}$

$\pi_{2ij}=\beta_{20j}$

$\pi_{3ij}=\beta_{30j}$

$\pi_{4ij}=\beta_{40j}$

Level 3: School Level

$\beta_{00j}=\Theta_{000}+ \Theta_{001}Self\_Resilience\_School\_Mean_j+$

$\Theta_{001}Teacher\_Resilience\_School\_Mean_j +$

$\Theta_{001}SJT\_Resilience\_School\_Mean_j +\Psi_{00j}$

$\beta_{01j}=\Theta_{010}$

$\beta_{02j}=\Theta_{020}$

$\beta_{03j}=\Theta_{030}$

$\beta_{04j}=\Theta_{040}$

$\beta_{05j}=\Theta_{050}$

$\beta_{10j}=\Theta_{100}+ \Theta_{101}Self\_Resilience\_School\_Mean_j+$

$\Theta_{102}Teacher\_Resilience\_School\_Mean_j +$

$\Theta_{103}SJT\_Resilience\_School\_Mean_j +\Psi_{10j}$

$\beta_{11j}=\Theta_{110}$

$\beta_{12j}=\Theta_{120}$

$\beta_{13j}=\Theta_{130}$

$\beta_{14j}=\Theta_{140}$

$\beta_{15j}=\Theta_{150}$

$\beta_{20j}=\Theta_{200}$

$\beta_{30j}=\Theta_{300}$

$\beta_{40j}=\Theta_{400}$

The three groups of models examined the predictive power of resilience scores from different perspectives. Model 1 and model 2 used only resilience scores at time 1 to predict future outcomes. Model 3 used all the resilience scores at different time points. Model 1 predicted the outcome only at a certain time point in the future, while model 2 and model 3 estimated both the outcome at the baseline level and the rate of change for the outcome. Figure 6 through figure 8 below illustrate the different features for each model. To simplify the graph, only one resilience variable was shown while in case there were three resilience variables and the school level mean predictors were ignored in the graphs.

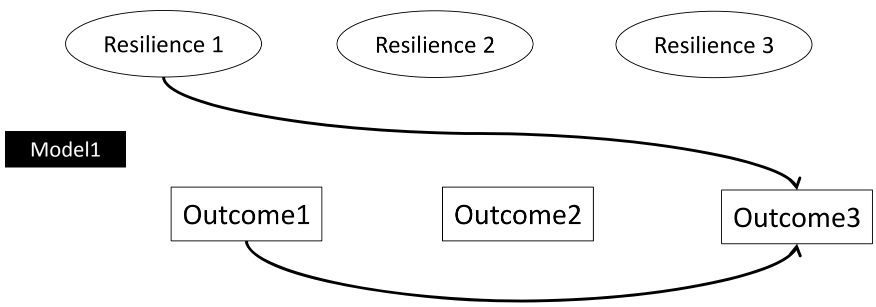

Figure 6: Illustration of Prediction Model 1

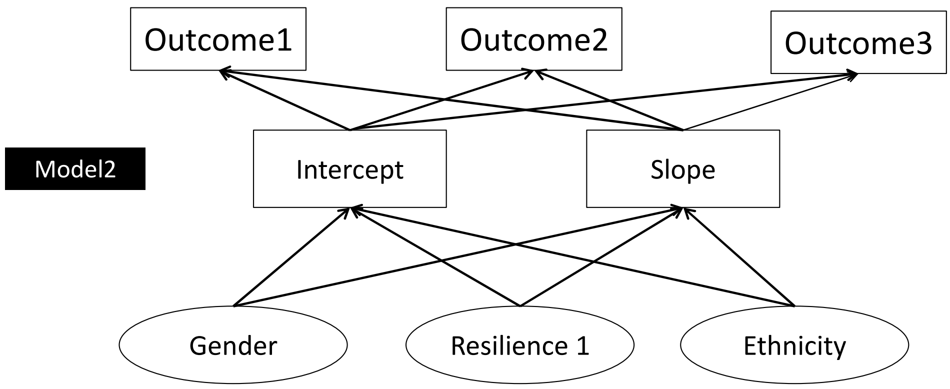

Figure 7: Illustration of Prediction Model 2

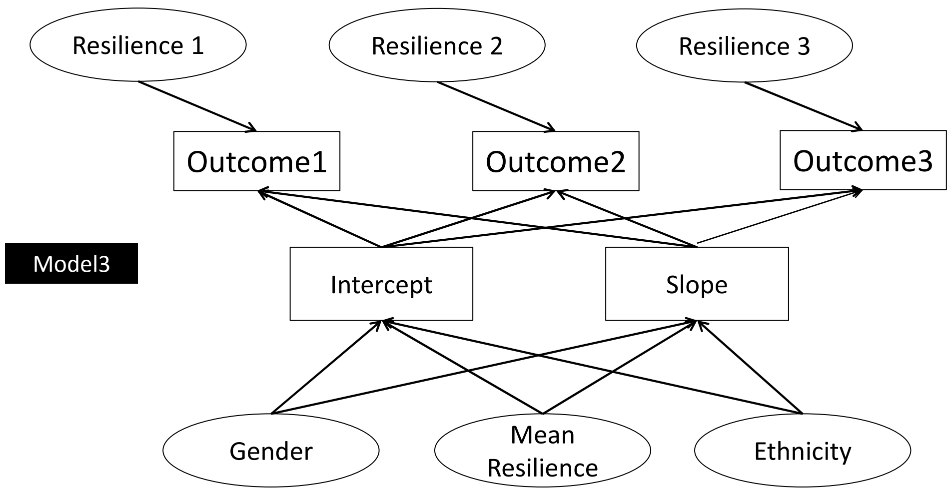

Figure 8: Illustration of Prediction Model 3

Another common feature model 2 and model 3 shared was that they both allowed random effects at the individual level and the school level. Figure 9 here shows how the random effects play a role in the model estimation. For the sake of brevity and clarity, the graph was simplified and only one individual level predictor—resilience score—was included. It was treated as a dummy variable instead of a continuous variable for the clarity of the visualization. Time was included as the sole level-1 predictor.

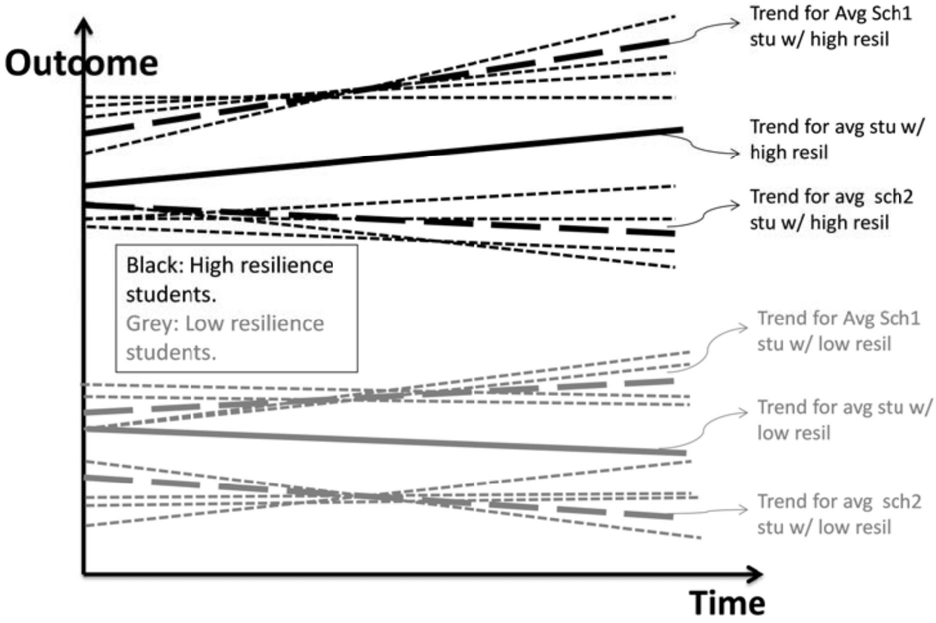

Figure 9: Illustration of the Random Effects in the Longitudinal Growth Curve Model

Figure 9 shows the level-1 relationship between time and outcome directly and illustrates the level-2 fixed effects (resilience score, treated as dummy variable here for simplification) and the random effects at both levels by varying the intercepts and the slopes of the lines. The black line represents the average linear growth curve for a student with a high level of resilience. The grey line represents the average linear trajectory for a student with low level of resilience. The dashed lines represent the average lines for students with high/low levels of resilience in different schools. They vary around the respective solid lines, representing the random effects at the school level. Since both the intercepts and the slopes of the dashed lines differ from their respective solid lines, the model allowed random effects for both intercepts and slopes at the school level. The dotted line indicates the linear trajectory for each individual student. Each cluster of dotted lines represents a group of students with either high or low level of resilience in the same school. The shifts in the slope and the intercept for the dotted lines from their respective dashed lines represent the random effects for slope and intercept at the individual level.

As discussed by Shadish, Cook, & Campell (2002), three criteria had to be satisfied for the establishment of a causal relationship: temporal precedence, correlation, and ruling out of alternative explanations. The author was handicapped to establish a causal relationship in this study because it was not an experimental study. While the first two criteria were not hard to satisfy, ruling out of all alternative explanations was impossible.

Assuming resilience significantly predicted the outcome in all three models, Model 1 and model 2 satisfied the first two criteria but was not able to rule out an alternative explanation of an unobserved between-individual third variable which affected both resilience scores and outcomes, thus weakening the causal explanation of the significant relationship between the resilience scores and the outcome. Model 3, by including the time-dependent resilience scores at the temporal level, opened the possibility of estimating the relationship between short-term changes in resilience scores and short-term changes in the outcome (Duckworth, Tsukayama, & May, 2010). If the level-1 resilience scores were significant, it mitigated the threats of the unobserved between-individual third variable explanation. Because a between-individual variable did not change, it could not cause a within-individual outcome variable to change over time. What model 3 could not rule out was a time-dependent unobserved third variable which changed in sync with the resilience scores and caused the outcome to change (Duckworth, Tsukayama, & May, 2010). Although a significant result from model 3 was unable to lead the author to the causal land, it could provide stronger evidence for a possible existence of a causal relationship.

Figure 10 below summarizes the three research questions and the corresponding methods associated with each research question.

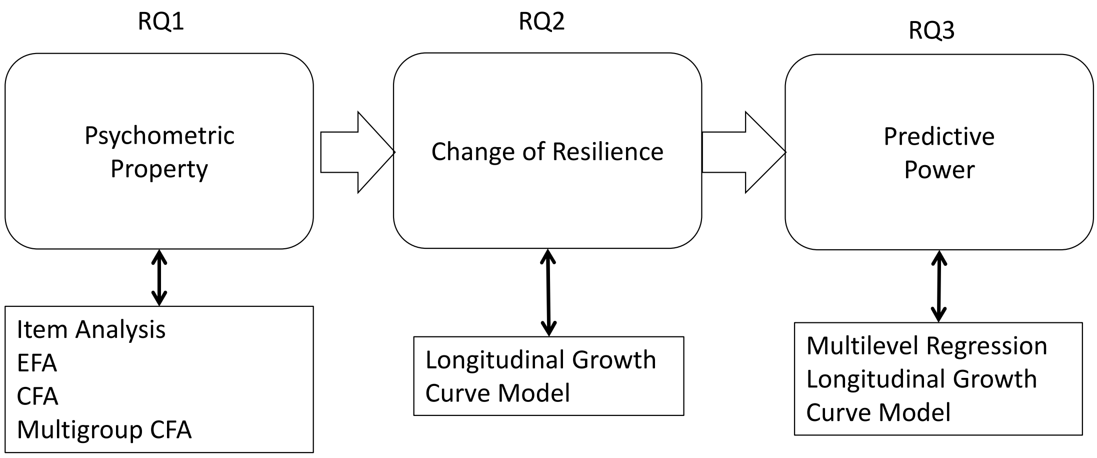

Figure 10: A Review of Research Questions and Methodology

CHAPTER 3: RESULTS

Research Question No.1: What Are the Psychometric Properties and Factor Structures of the Resilience Scale?

Self-rating subscale

The final confirmatory bi-factor analysis model with two additional valence factors fit the data well at time 1, with an RMSEA of .047 (upper limit at .051), a CFI of .958, and a TLI of .947. In order to examine factor invariance across time, a multi-group CFA was conducted and configural invariance was achieved for the primary resilience factor. After that, the strong invariance model was fit and the fit statistic was compared with those from the configural invariance model. Because the Mplus program used the weighted least squares means and variance adjusted (WLSMV) estimation for categorical variables, the chi-square statistics could not be directly compared in a chi-square difference test (Satorra, 2000). Instead, a scaling correction was required in order to correctly approximate chi-square under non-normality (Muthén & Muthén, 2012). Using the Difftest option in Mplus, the chi-square difference test was significant. However, the shift in the CFI was less than -.01 and the shift in RMSEA was less than .015. According to Cheung & Rensvold (2002), if the above two criteria were satisfied, a case of measurement invariance would still be established even if the chi-square difference test was significant. The final scalar invariance model had a CFI of .976, a TLI of .975, and an RMSEA of .032. Table 8 below shows the final item loadings for the primary resilience factor according to the scalar invariance model. The group column indicates the aspect of resilience each item measures: resilience represents the items that are directly measuring the outcome of resilience and they only load on the primary resilience factor. Emotion and efficacy represent the items that are measuring related personal characteristics and items belonging to those two groups load both on the primary resilience factor and their respective group factors.

Table 8:

Final loadings on the primary resilience factor of the self-rating subscale

Group	Item Description	Loading	S.E	t	p
resilience	I am capable of coping with most of my problems.	0.35	0.013	26.60	<.001
resilience	I overcome challenges and set backs.	0.44	0.013	35.00	<.001
resilience	I am resilient.	0.25	0.014	18.05	<.001
resilience	I give up easily when faced with an obstacle.	0.66	0.015	45.53	<.001
resilience	I am easily discouraged.	0.75	0.012	65.12	<.001
resilience	I give up easily.	0.78	0.014	53.93	<.001
efficacy	I complete tasks successfully.	0.30	0.014	21.98	<.001
efficacy	When I try, I generally succeed.	0.29	0.014	21.01	<.001
efficacy	I am confident I get the success I deserve in life.	0.39	0.013	29.94	<.001
efficacy	Overall, I am satisfied with myself.	0.53	0.013	42.26	<.001
efficacy	I avoid responsibilities.	0.53	0.013	40.05	<.001
efficacy	I forget to do things.	0.50	0.012	41.33	<.001
efficacy	I make careless mistakes.	0.48	0.012	40.34	<.001
efficacy	I quickly lose interest in the tasks I start.	0.57	0.012	47.98	<.001
efficacy	My interests change quickly.	0.39	0.013	31.03	<.001
efficacy	Sometimes, I do not feel in control of my school work.	0.61	0.012	51.53	<.001
emotion	I remain calm under pressure.	0.32	0.013	24.40	<.001
emotion	I am relaxed.	0.30	0.014	21.76	<.001
emotion	I am not easily bothered by things.	0.12	0.014	8.66	<.001
emotion	I remain calm when I have a lot of homework to do.	0.34	0.013	25.35	<.001
emotion	I am easily frustrated.	0.61	0.011	54.99	<.001
emotion	Sometimes I feel depressed.	0.69	0.013	51.65	<.001
emotion	Sometimes when I fail I feel worthless.	0.65	0.013	50.41	<.001
emotion	I am filled with doubts about my competence.	0.69	0.012	58.87	<.001
emotion	There are times when things look pretty bleak and hopeless to me.	0.72	0.013	55.73	<.001
emotion	I get discouraged when things go wrong.	0.70	0.01	66.37	<.001
emotion	I worry.	0.52	0.013	38.88	<.001
emotion	I get stressed out easily when things don't go my way.	0.59	0.012	50.16	<.001
emotion	I worry about school.	0.53	0.014	36.95	<.001
emotion	I get upset easily.	0.67	0.011	60.59	<.001

There were several notable observations on the final loadings of the primary resilience factor. First, the average loading of the outcome items was .54 and the average loading of the trait items was .47. The result that outcome items had higher average loadings was expected because those items focused on measuring resilience while the variance of the trait items was separated between their respective trait factors and the resilience factor. Second, the average loading of the negatively worded items (.61) was higher than the positively worded items (.33), suggesting that wording had a stronger effect on the positively worded items compared with the negatively worded items. That finding was consistent with the fact that the loadings on the positive error factor were more significant than the loadings on the negative error factor. A more detailed discussion of the wording effects and methods to adjust for wording is available in the discussion section. Third, the item "I am resilient" had a low loading of .25, indicating it did not perform well. It was not expected and in fact, the author assumed that item would have the highest loading. One possible explanation was that middle school students did not fully understand what resilience meant or it could be that the question was so direct that the respondents did not answerer it frankly. Last, a few items in the trait group had high loadings on the primary resilience factors. For example, item "There are times when things look pretty bleak and hopeless to me" and item "I get discouraged when things go wrong" both had loadings greater than .70. In fact, 56% and 69% of the respondents participating in the Q-sort placed those items into the resilience category.

The primary resilience factor achieved an internal consistency of .94. The standard factor score for this primary resilience factor was calculated and used as a variable in addressing research questions 2 and 3.

Teacher-rating subscale

The final one-factor-four-items structure achieved an acceptable fit at time 1. Similar to the self-rating subscale, a configural invariance model was fit across time and acceptable fit was achieved. The scalar invariance model was built next, yielding an RMSEA of .043, a CFI of .99, and a TLI of .99. Compared with the configural invariance model, the scale-corrected chi-square difference test was not significant and the shift in both RMSEA and CFI was negligible, leading to the conclusion that strong measurement invariance was established for the factor structure across time. Table 9 displays the final loadings on the teacher-rated resilience factor based on the strong invariance model. The single factor had an internal consistency of .88. Standard factor scores were calculated according to the strong measurement invariance model. As in the self-rating subscale, the factor score would be used to answer subsequent research questions.

Table 9:

Final loadings on the resilience factor of the teacher-rating subscale

Item	Item Description	Loading	S.E	t	p
T1	Remains calm under pressure	0.78	0.01	77.16	<.001
T2	Overcomes challenges and setbacks	0.95	0.003	288.67	<.002
T3	Does not give up easily	0.92	0.005	184.53	<.003
T4	Is resilient	0.93	0.005	168.75	<.004

SJT subscale

As discussed in the methods section, only the factor of problem-focused coping was retained because students with higher scores on that factor were hypothesized to be able to deal better with stressful and challenging situations. Assuming students got lower scores on the emotion-focused coping factor or the avoidance-focused coping factor, it only indicated that they seldom used those coping strategies; it did not mean they turned to effective coping strategies more often. After verifying that the one-factor structure with correlated residuals achieved acceptable fit at time 1 and configural invariance across time, the strong measurement model was tested. The RMSEA was .057. The CFI was .99 and the TLI was .99. The scale-corrected chi-square difference test was not significant and the shift in CFI and RMSEA was negligible, both evidence pointing to the existence of scalar invariance across time. Table 10 displays the final item loadings on the resilience coping factor. As before, standard factor scores were calculated in each wave to represent the resilience score from the SJT subscale and used as a variable in subsequent analyses.

Table 10:

Final loadings on the resilience factor of the SJT subscale

Item	Situation	Item Description	Loading	S.E	t	p
HW3	Homework	I try to get organized to get on top of my homework.	0.78	0.011	72.66	<.001
HW4	Homework	I make the extra effort to get all of my homework done.	0.76	0.012	62.28	<.001
HW10	Homework	I take control and say to myself: I can do this!	0.61	0.014	42.61	<.001
AC3	After-school Activity	I try to get organized to get on top of all these activities.	0.79	0.011	73.03	<.001
AC4	After-school Activity	I make the extra effort to get all of my activities completed.	0.77	0.011	66.99	<.001
AC9	After-school Activity	I take control and say to myself: I can do this!	0.62	0.014	43.59	<.001
EM3	Exam	I try to get prepared for the test regardless.	0.69	0.015	47.25	<.001
EM5	Exam	I take control and say to myself: I can do this!	0.63	0.014	45.21	<.001

Research Question No.2: How Does Resilience Change Over Time?

As discussed in the method section, the resilience scores were calculated using the maximum a posteriori method (MAP, also known as the regression method for continuous items). Because all three resilience scores achieved scalar measurement invariance, change of the three scores across time could be assessed.

Before presenting the results of the longitudinal growth curve models, descriptive statistics are shown. Table 11 summarizes the correlations across time for the three resilience scores. As can be seen in the three diagonal rectangles, the average correlation between two consecutive measures of the same resilience score was about .5, indicating that resilience changed differentials across students. The average correlation among the self-rating resilience scores and the SJT scores was higher than that of the teacher-rating scores. The result was not surprising since typically different teachers were rating the same student over time. As mentioned in the methods section, the correlations between the teacher-rated score and the other two scores were low and it would not be meaningful to try to derive a composite score using a high-order factor.

Table 11:

Correlations among the three resilience scores across time

	Self time1	Self time2	Self time3	Teacher time1	Teacher time2	Teacher time3	SJT time1	SJT time2	SJT time3
Self time1	1.00	0.53	0.47	0.18	0.10	0.16	0.37	0.29	0.27
Self time2		1.00	0.54	0.14	0.10	0.16	0.26	0.36	0.28
Self time3			1.00	0.13	0.14	0.19	0.27	0.32	0.39
Teacher time1				1.00	0.50	0.34	0.18	0.12	0.11
Teacher time2					1.00	0.34	0.05	0.09	0.16
Teacher time3						1.00	0.16	0.19	0.16
SJT time1							1.00	0.53	0.46
SJT time2								1.00	0.57
SJT time3									1.00

Table 12 shows the means, minimum, maximum, and the standard deviations of the three subscale scores across time points.

Table 12:

Descriptive statistics for the three resilience scores

	Time	N	Min	Max	Mean	Std Dev
Self-rating Resilience	1	2028	-1.16	0.84	0.000	0.31
	2	1673	-1.34	0.79	-0.081	0.33
	3	2614	-1.36	0.84	-0.013	0.32
Teacher-rating Resilience	1	1633	-2.15	0.85	0.000	0.68
	2	893	-2.09	0.8	-0.005	0.66
	3	2016	-2.16	0.84	-0.045	0.69
SJT Resilience	1	1977	-2.44	1.2	0.000	0.72
	2	1672	-2.5	1.18	-0.067	0.76
	3	2614	-2.44	1.18	-0.003	0.73

As can be seen from the table, the mean score for all three subscales at time 1 was zero; Mplus forced the factor score at time 1 to be zero in order to setup the scale for scalar invariance models. For self-rating subscale, the resilience score at time 2 was significantly lower than time 1 and time 3. The difference between the score at time 2 and time 3 was not significantly different. A similar pattern was observed for the SJT subscale. The SJT resilience score at time 2 was significantly lower than at the other time points while the score at time 1 was not significantly different from the score at time 3. The teacher-rating resilience score had a different trend. Scores at time 1 and time 2 were not different from each other. Neither were scores at time 2 and time 3. However, the score at time 3 was significantly lower than the score at time 1.

The means testing did not account for the correlation among scores from the same students or among students that attended the same schools. The longitudinal growth curve model was able to take into account the correlations among repeated measures of the same individual as well as the correlations among individuals nested in the same cluster.

Unconditional models with no predictors were fit first to get the variance decomposition. For the model with self-rating resilience score as the response variable, 49% of the variance was at the temporal level, 49% of the variance was at the individual level, and 2% of the variance was at the school level. The SJT score shared similar variance components: 48% at the temporal level, 50% at the individual level, and 1% at the school level. For the teacher-rating score, more variance was found at the temporal level and less was found at the individual level: 59% and 38%, respectively. Only about 3% of the total variance lay at the school level. Although relatively little variance was at the school level for all three dependent variables, the variance was significant in all three cases.

Table 13:

Variance decomposition for the three resilience scores

	Temporal	Individual	School
Self-Rating	49%	49%	2%
Teacher-Rating	59%	38%	3%
SJT	48%	50%	1%

Conditional models with time as the only predictor in the model were then fit. Random effects for both the intercept and the slope of time were allowed at the individual level and at the school level. After introducing time into the model, the variance of the intercept at both the individual and the school level remained significant. The estimate of time was negative and significantly different from zero for the self-rating score, indicating there was a downward linear trend for self-rating resilience. The variance of the slope was significant at both the individual and the school level, meaning that the trend varied among individuals and among schools. The coefficient estimate for time was not significantly different from zero for both the teacher-rating score and the SJT score, which signaled that no linear trend existed for those two scores. Z-test results also showed that neither the variance of slope among individuals nor among schools was significant. Table 14 summarizes the random effects estimates for all three scores.

Table 14:

Random effect estimates for both intercept and slope in all three models

Subscore	Intercept			Slope	
	temporal	individual	school	individual	school
Self-Rating	0.047**	0.049**	0.001*	0.002**	0.0004*
Teacher-Rating	0.265**	0.180**	0.014**	0.006	0.0008
SJT	0.252**	0.265**	0.007*	0.006	0.002

Note: *significant at .05 level; **significant at .01 level

Models with time as level-1 predictor and gender and ethnicity as level-2 predictors were built next. For models with teacher-rating score and SJT score as response variables, only main effects were included. The variance of slope was not significant at the individual level therefore no attempt was made to explain that variance. For the model with self-rating resilience score as response, interactions between time and gender as well as time and ethnicity were also included to see if the two predictors could explain the variance of the slope.

In the model predicting self-rated resilience, neither gender nor ethnicity was significant and neither of their interactions with time was significant. In the model predicting teacher-rating resilience, both gender and ethnicity were significant. Teachers were more likely to give higher scores to female students and Asian students compared to white students. The scores given to black students and other students tended to be lower than those given to white students. For the model with SJT resilience score as the dependent variable, gender was significant. Female students were more likely to rate themselves higher in the SJT than male students. No significant difference was detected among students from different ethnicity groups. Model estimates are summarized in table 15 below.

Table 15:

Summary of results for fixed effects in predicting the three resilience scores

	Self-rated Resilience	Teacher_rated Resilience	SJT Resilience
# Observations	6173	4485	6121
Female	-0.003	0.16**	0.255**
Black	-0.019	-0.135**	-0.028
Hispanic	-0.016	-0.074	0.008
Asian	-0.017	0.097*	0.025
Other	-0.017	-0.07*	-0.032
time	-0.011**	-0.004	-0.004

Note: *significant at .05 level; **significant at .01 level

Research Question No.3: What Is the Relationship Between Resilience and Outcomes Measures?

Before presenting the various model results, it is helpful to take a look at the correlations among the three resilience scores and the outcomes at different time points

(See in table 16). Different columns in the table represent different outcome variables and different rows represent time measured and which resilience score at time 1. For example, row 2 in the table shows the correlations between the self-rated resilience score measured at time 1 and different outcome variables measured at time 2. As can be seen, the further away from time 1 an outcome was measured, usually the lower the correlation would be. The teacher-rated resilience score at time 1 had higher correlations with the academic achievement measures at different time points than the self-rated score and the SJT score. The self-rated score and the SJT score correlated higher with the life satisfaction measure compared with the teacher-rated score.

Table 16:

Correlation between the three resilience scores and outcomes at different time points

Resilience Score at time1	Outcome measured at	GPA	Math	Reading	Quantitative	Verbal	Life Satisfaction
Self-rated resilience	time1	0.26	0.14	0.15	0.09	0.09	N/A
	time2	0.23	0.12	0.16	0.12	0.07	0.32
	time3	0.12	0.07	0.11	0.05	-0.01	0.18
Teacher-rated resilience	time1	0.30	0.32	0.35	0.26	0.26	N/A
	time2	0.31	0.33	0.31	0.20	0.24	0.06
	time3	0.25	0.24	0.26	0.20	0.21	0.06
SJT resilience	time1	0.28	0.19	0.24	0.03	0.08	N/A
	time2	0.31	0.19	0.22	0.02	0.04	0.27
	time3	0.28	0.12	0.19	0.02	-0.02	0.19

Model 1 results

As discussed in the methods part, model 1 examined the relationships between resilience and future outcomes when controlling for prior outcomes. As usual, unconditional models were built first to examine the variance decomposition at the student level and the school level. Outcomes at time 2 and outcomes at time 3 had similar variance components at each level. As can be seen in table 17, for GPA, grades, and life satisfaction, most of the variance was at the student level. For the two standardized test scores, school-level variance was about 20% while the student level variance decreased a little bit.

Table 17:

Variance decomposition of outcomes at time 2 and time 3

		GPA	Math	Reading	Quantitative	Verbal	Life Satisfaction
Outcome at Time 2	school level	4%	2%	7%	21%	19%	2%
	student level	96%	98%	93%	79%	81%	98%
Outcome at Time 3	school level	6%	4%	4%	18%	19%	1%
	student level	94%	96%	96%	82%	81%	99%

Table 18 displays the summary of R^2 estimates and coefficients estimates for the different response variables at time 2 and table 19 summarizes the results for response variables at time 3. The columns represent different models with different dependent variables. The first two rows include the number of observations used to build the model and the R^2 statistics at the individual level. Since the variance at the school level was relatively small compared to the individual level, the R^2 at the school level were not shown. The other rows represent the coefficient estimates for different time 1 predictors. Predictors at the school level are not shown because the focus is on the student level predictors. Most of the resilience scores at the school level were not significant.

Table 18:

Model 1: Fixed effects estimates for response variables at time 2

		GPA	Math	Reading	Quantitative	Verbal	Life Satisfaction
	# Observation	947	786	730	540	540	921
	Variance explained at individual level	58%	52%	60%	82%	76%	14%
Model estimates (all predictors were measured at time 1)	Pre-test of outcome	0.70***	0.68***	0.71***	0.91***	0.87***	N/A
	Self Resilience	0.07	0.07	0.08	1.38	-3.53	0.76***
	SJT Resilience	0.28***	0.06**	0.04*	0.61	0.90	0.18***
	Teacher Resilience	0.14**	0.12***	0.09***	-0.25	1.70*	0.02
	Male	-0.02	-0.04	-0.03	1.77	0.89	0.05
	Asian	0.30*	0.04	0.01	0.91	4.11	-0.37**
	Black	-0.15	-0.23**	-0.14	-1.14	-1.96	-0.14
	Hispanic	-0.15	0.03	0.07	2.46	-5.89	-0.27
	Other	0.07	0.02	0.02	1.50	0.00	0.10

Note: ***significant at p< .001, **significant at p< .01, *significant at p< .05

Table 19:

Model 1: Fixed effects estimates for response variables at time 3

		GPA	Math	Reading	Quantitative	Verbal	Life Satisfaction	Life Satisfaction
	# Observation	543	341	343	368	368	581	480
	Variance explained at individual level	42%	30%	52%	59%	49%	7%	31%
Model estimates (all predictors were measured at time 1)	Pre-test of outcome	0.54***	0.62***	0.72***	0.69***	0.70***	N/A	0.52***
	Self Resilience	-0.07	0.00	0.02	-0.89	-3.28	0.28*	-0.15
	SJT Resilience	0.24**	0.07	0.04	1.97	-0.57	0.12*	0.14**
	Teacher Resilience	0.23**	0.05	0.04	1.26	2.68*	0.09	0.09
	Male	-0.07	0.02	0.03	4.78*	0.45	0.04	0.05
	Asian	0.02	0.14	0.07	3.07	6.50	-0.50***	-0.25*
	Black	-0.36	-0.14	-0.20	0.80	-14.73*	-0.58**	-0.61**
	Hispanic	0.02	-0.23	0.01	-4.02	-11.75	0.22	0.26
	Other	-0.21	-0.11	-0.09	-0.37	1.18	-0.12	-0.16

Note: ***significant at p< .001, **significant at p< .01, *significant at p< .05

As can be seen from the tables, for academic achievement measures, teacher-rated resilience and SJT resilience achieved better predictive validity than the self-rated resilience. SJT resilience had slightly stronger predictive power than teacher-rated resilience when it came to self-reported GPA. But teacher resilience was better in predicting math and reading grades and predicting the standard test scores. On average, more variance was explained for the standardized test scores compared with the other achievement measures.

With regard to life satisfaction, self-rated resilience and SJT resilience were stronger predictors than the teacher-rated resilience. One thing worth mentioning was that life satisfaction was not measured at time 1, therefore the model did not include a pre-measure of the outcome for those two models predicting life satisfaction. It also explained why those models had low R^2 statistics compared with other models. In order to reduce the error variance, another model was fit with time 2 life satisfaction as a control variable in the model to predict time 3 life satisfaction (last column in table 19). After controlling for previous life satisfaction, self-rated resilience was no longer significant. But SJT resilience remained significant, indicating that the SJT resilience score could explain a significant amount of variance in addition to the variance explained by the pre-measure of the outcome.

A comparison of table 18 and table 19 suggests that more variance of time 2 outcomes was explained compared with time 3 outcomes. Predictors of time 2 outcomes were also more likely to be significant compared with predictors of time 3 outcomes.

Model 2 results

Similar to model 1, model 2 tested the predictive power of resilience scores at time 1, but from a different perspective. By including the temporal level, model 2 examined how resilience scores predicted both the overall level of outcome variables and the future trend of outcomes. At least three observations are required to build a longitudinal growth curve model (Raudenbush, 2001; Curran, Obeidat, & Losardo, 2010). Since life satisfaction was only measured twice, it was excluded from the response variables modeled here. Table 20 summarizes the random effects of the models. Variance decomposition was calculated based on unconditional models and the variance estimates for the slope of time was calculated in a longitudinal model with time as the only predictor.

Table 20:

Variance decomposition and random effects estimates for model 2

		GPA	Math	Reading	Verbal	Quantitative
Variance Decomposition	Temporal Level	30%	32%	26%	15%	12%
	Individual Level	64%	63%	68%	67%	67%
	School Level	5%	5%	6%	18%	21%
Random Intercept	Temporal Level	0.62*	0.13*	0.08*	116*	97*
	Individual Level	1.31*	0.26*	0.21*	517*	537*
	School Level	0.11*	0.02*	0.02*	136*	164*
Random Slope	Individual Level	0.06*	0.02*	0.003	49*	45*
	School Level	0.00	0.00	0.004	7	11

Note: *significant at p< .05 level

As can be seen from the table, most of the variance was at the individual level: from 63% to 68%. For GPA, math, and reading, the school-level variance was small: 5% to 6%. For verbal and quantitative scores, the school-level variance was around 20%. Based on unconditional models, the random effects at all three levels for all the models were significant. Based on models with time as the only predictor, the variance for the slope of time was significant at the individual level for four out of five models (not reading). None of the variances for the slope were significant at the school level. Table 21 presents the R^2 estimates at the individual level and the fixed effects estimates for predictors in all models. Compared to model 1 which used the pre-measure of the outcome as a control, this group of models explained a relatively small amount of variance at the individual level.

Time was a significant predictor in models predicting GPA, math grade, and the quantitative test score. Female students did significantly better in reading and significantly worse in quantitative test than male students. With regard to ethnicity, white was used as the reference group. In most of the models, Black and Hispanic students did significantly worse than white students. Asian students did significantly better. Turning to the three resilience scores, self-rated resilience score was a significant predictor of the overall outcome level across time in all five models. SJT resilience significantly predicted the overall GPA, math and reading scores, but wasn’t significantly related to the two standardized test scores variables. Teacher resilience was significantly positively related with all the outcome variables regarding their overall value and most of the time teacher resilience possessed the strongest relationship out of the three.

Table 21:

Summary of the fixed effects estimates for model 2

		GPA	Math	Reading	Quantitative	Verbal
	# Observation	2780	2501	2351	1995	1992
	Variance explained at individual level	27%	23%	25%	9%	6%
Model estimates (all predictors were measured at time 1)	Time	-0.06*	-0.04**	0.003	0.85*	0.44
	Female	0.002	0.04	0.14***	-9.81***	0.97
	Black	-0.5*	-0.15*	-0.49***	-16.39***	-6.46*
	Hispanic	-0.49*	-0.11	-0.27**	-15.72**	-18.01***
	Asian	0.38**	0.18**	0.11*	11.08***	5.98*
	Other	-0.07	-0.06	-0.07	-1.56	-5.27*
	Self Resilience Time1	0.74***	0.12*	0.09*	5.55*	6.14*
	SJT Resilience Time1	0.34***	0.10***	0.1***	0.57	1.28
	Teacher Resilience Time1	0.49***	0.27***	0.22***	8.37***	8.55***
	Time*Self Resilience	-0.22**	N/A	N/A	N/A	N/A
	Time*SJT Resilience	0.09**	N/A	N/A	N/A	N/A
	Time*Teacher Resilience	0.003	N/A	N/A	N/A	N/A

Note: ***significant at p< .001, **significant at p< .01, *significant at p< .05

The interaction terms between each resilience score and time indicated whether resilience had predicted the trend of the outcome. All models with interactions were tested and none of the interactions were significant in the last four models, indicating that students with different resilience scores did not have significantly different trend with regard to the outcome measure. Only Self and SJT resilience scores had significant interactions with time when predicting GPA. Students with higher self-rated resilience had a higher overall GPA across time, but also experience a lower rate of increase compared with students with lower self-rated resilience scores. The interaction between SJT score and time was positive and significant, implying that students with higher SJT scores had a sharper incline compared to student with lower SJT scores.

Model 3 results

Like model 2, model 3 was a longitudinal growth curve model with three levels. Both overall level of the outcome and the rate of change for the outcome were modeled. The difference between model 2 and model 3 was that model 3 included three more time-varying independent variables (in addition to time) at level-1. The three time-varying variables were the three resilience scores. At each time point, the resilience scores were used to predict the later outcomes for the same semester. So the level-1 variables were targeting at short-term predictions rather than long-term (compared with model 1 and model 2). The level-2 average scores in model 3 played the same role as the time 1 scores in model 2, explaining the variance of intercept and slope at the lowest level. Because the unconditional models in model 3 were the same as the ones in model 2, model 3 and model 2 had identical variance decomposition as well as estimates for random effects, which were already summarized in table 20. Table 22 summarizes the R^2 estimates for both the temporal level and the individual level. Because no school-level predictors were included in the model, the R^2 at the school level was not shown. Very little temporal-level variance was explained. The largest amount of variance was explained in the model predicting GPA and it was only 3.2%. 10% to 30% of the variance at the individual level was explained by the predictors.

Table 22:

Variance Explained at the temporal and the individual level

	GPA	Math	Reading	Quantitative	Verbal
R Square at Temporal Level	3.2%	0.4%	1.1%	0.5%	0.0%
R Sqaure at Individual Level	26.7%	25.8%	31.8%	13.2%	10.8%

Table 23 shows the fixed effect estimates for variables in different models. The first four rows represent the estimates of level-1 predictors: time and the three time-varying resilience variables. The other rows represent the estimates of level-2 predictors: gender, ethnicity, and the mean resilience scores across time. The interactions between time and mean resilience scores were tested in all models. Only the GPA model had significant interactions. All other estimates were from main-effects-only models.

Table 23:

Fixed effect estimates for model 3

		GPA	Math	Reading	Quantitative	Verbal
	# Observations	3283	2794	2644	2100	2096
Level 1 Predictors	Time	0.01	-0.02	-0.01	0.53	0.26
	Self_Resilience_Centered	0.57***	0.06	0.01	0.2	-1.4
	SJT_Resilience_Centered	0.03	0.03	0.02	0.3	0.75
	Teacher_Resilience_Centered	0.04	0.01	0.02	1.5*	0.09
Level 2 Predictors	Female	0.05	0.04	0.13***	-8.49***	2.14
	Black	-0.43***	-0.18**	-0.44***	-14.04***	-6.48*
	Hispanic	-0.49**	-0.11	-0.24**	-12.92**	-16.16***
	Asian	0.38***	0.18***	0.1*	12.06***	6.45*
	Other	-0.08	-0.04	-0.06	-1.88	-4.38*
	Self_Resilience_Mean	0.73***	0.12*	0.19***	7.75**	6.86**
	SJT_Resilience_Mean	0.36***	0.1***	0.07***	-1.25	-0.43
	Teacher_Resilience_Mean	0.59***	0.34***	0.29***	8.93***	9.71***
	Time*Self_Resilience_Mean	-0.12	N/A	N/A	N/A	N/A
	Time*SJT_Resilience_Mean	0.07*	N/A	N/A	N/A	N/A
	Time*Teacher_Resilience_Mean	0.04	N/A	N/A	N/A	N/A

Note: ***significant at p< .001, **significant at p< .01, *significant at p< .05

The estimates of level-2 predictors were similar to the model 2 results. Female students did significantly better in reading while significantly worse in quantitative test. Compared with white students, black and Hispanic students did significantly worse on measures of academic achievement while Asian students did better. The main effects associated with the three resilience scores were mostly significant across different models, suggesting that the mean level of resilience was significantly positively related to the overall level of the outcomes. The teacher-rated resilience score had a stronger relationship, especially with standardized test scores.

The interaction terms between time 1 and time 2 resilience scores represented the impact of the average resilience scores on the trend of the outcomes. The only significant interaction appeared in the model predicting GPA. The interaction between time and SJT suggested that students' with higher SJT scores were more likely to experience larger increases in their GPAs compared with students with lower SJT scores.

Focusing on the three level-1 individual-mean centered resilience scores, their estimates represented the short-term impact of resilience scores on the outcome variables—how the change of resilience scores for an individual affected the change of the outcome measures. As discussed in the methods section, if level-1 resilience predictors turned out to be significant, the author would be able to rule out the between-individual unobserved third variable alternative. Results in table 23 showed that level-1 self-rated resilience was significant in predicting GPA and the level-1 teacher-rated resilience was significant in predicting the quantitative test score. None of the level-1 resilience scores were significant in other models.

CHAPTER 4: DISCUSSION

Conclusions

All the three resilience measures were multidimensional in nature. After applying different techniques, the author was able to extract a single, reliable and stable factor to represent each of the three resilience scales. No single resilience score could be derived from the three sub-scores. A single general resilience factor derived from the three subscales would not be reliable because of the little common variance shared by the three sub-scores and therefore the attempt to derive a single score for the whole scale failed.

All three resilience scales experienced significant changes across time, although the change could not be simplified as a linear trend in two of the three models. Teachers tended to give higher ratings to female students and also to white and Asian students. Female students were also more likely to rate themselves higher in SJT resilience scores. Neither gender nor ethnicity had a significant effect on the linear rate of change of the three resilience scores.

Self-rated resilience score was not able to significantly predict any outcome in the long term (from half a year to a year) when prior outcomes and other resilience sores were controlled simultaneously. The change of self-rated resilience across time was related to the change of GPA measured one to two months later. SJT resilience score significantly predicted GPA, math grade, reading grade, and life satisfaction. Students with higher SJT scores were also more likely to have a better growth regarding their GPA. However, change of SJT resilience score across time was not able to predict the change of any outcome variables. Teacher-rated resilience was not related to life satisfaction but it had stronger relationships with grades and standardized test scores. The change of teacher-rated resilience across time was able to predict the change of quantitative test scores.

Resilience was significantly related to academic achievement and life satisfaction. Resilience scores derived from different methods tended to capture unique aspect of resilience, which was verified by their differentiated relationships with different outcomes. Self-rated resilience achieved predictive validity with GPA. SJT resilience score was strongly related to GPA and life satisfaction. Teacher-rated resilience had stronger predictive power on grades and test scores.

Significance and Implications

Previous studies on resilience did not attempt to place the various definitions of resilience into different groups according to the aspect of resilience each definition was targeting. Neither did they try to make connections between aspects of resilience and the different ways to measure resilience. The current study adds value to the field by examining resilience in detail and by distinguishing items of traits that were influential on resilience from items of outcomes that were directly measuring resilience. By applying the bi-factor analysis, a true resilience factor was successfully separated. That factor was free of the trait variance that was unique to the traits. To the author's knowledge, none of the previous studies which used resilience scales made up of items under mixed methods tried to extract resilience-only variance. The study makes methodological contributions as well by illustrating an effective way to extract information from a resilience scale to a unidimensional factor.

Placing definitions and scales of resilience into four groups provides guidance for researchers developing resilience scales. If researchers want to measure a specific aspect of resilience, it is better not to use items borrowed from a mix of scales. If the goal is to explore what personality traits can affect resilience, a resilience scale using the trait approach is not able to provide a proper outcome measure because the scale itself is made up of trait items. A significant relationship means nothing. Scales measuring particular traits and a resilience scale developed by the outcome approach will fit the objective better.

The categorization of definitions and scales of resilience are also helpful for practitioners aiming to implement interventions to improve student resilience. If the intervention's goal is to build resilience through teaching children how to use better strategies to handle stressful situations, researchers should realize that a scale solely based on the coping approach is not able to examine the real impact of the intervention. It is only good at capturing whether the program achieves its goal in changing participants' coping strategies, which can serve as a mediating variable but not the ultimate outcome. Instead, a scale developed using the outcome approach should be used to measure resilience and to provide evidence of the effectiveness of the program.

Following the same categorization, experiments can be conducted to uncover the mechanism of how resilience is built. Scales measuring the trait aspect, the coping aspect, and the process aspect can be used for the examination of their relationships with the resilience defined by the outcome approach. It will provide valuable information on what factor has the largest impact on the resilience outcome and thus give directions for building resilience interventions.

Very few resilience studies have used more than one method to measure resilience. Lai & Viering (2012) and Lipnevich, MacCann, & Roberts (2013) emphasized the importance of collecting information from multiple sources in multiple assessment modes. This study is an example of administering multiple measures (subscales) of resilience to multiple groups of raters. It provides evidence that each method captures a unique facet of resilience and the common variance shared by different methods is low. The predictive validity of each method varies with regard to the criterion used. For example, the teacher-rated method is better for predicting external academic criteria and the SJT method is better for predicting life satisfaction. The findings can help researchers better understand the strengths and weaknesses of each method.

This study also features a longitudinal design, which allows the examination of resilience over time. The fact that all three resilience scores capture differential growth across students has important implications. The low correlation over time might be a sign of poor test-retest reliability. Since most resilience studies used a scale that was similar to at least one of the three subscales used in this study, if the lack of test-retest reliability is verified, there will be a strong impetus to develop a new measure of resilience which can achieve better reliability across time.

Another possible interpretation is that resilience is subject to short-term changes of internal or external factors. If a student can be more/less resilient facing different tasks or under different situations, it will be worth examining what situations/tasks stimulate individual's resilience and what situations/tasks suppress it.

The longitudinal feature of the data also opens the possibility of fitting more complex models, which have been tried rarely in prior studies of resilience. Very few previous studies which found a significant relationship between resilience and an outcome were able to rule out the explanation of an unobserved third variable. Evidence from this study overcomes that weakness to some extent. The results contribute to the literature related to the evidence of resilience's importance and also to the general literature on the significance of non-cognitive skills.

The issue of wording discovered in this study is worth researchers' attention when assessing non-cognitive skills. Although wording has been rarely discussed when it comes to resilience scales, the effect of wording has been noticed by researchers investigating other scales. Originally, the introduction of negatively-worded items was to eliminate the acquiescence bias (Cronbach, 1946; Couch and Keniston, 1960). An unintended effect was inflated correlations among positively worded items and among negatively worded items, sometimes sufficient to yield a two-factor solution. Carmines and Zellar (1979), by conducting exploratory factor analysis, found that Rosenberg's (1965) self-esteem scale involved two factors—with one factor made up of all the positively worded items and the other of all the negatively worded items. However, they failed to realize that the two-factor structure was due to wording and claimed that the self-esteem scale had two trait factors. Marsh (1996) found the scale had only one factor after separating the effect due to wording, a finding replicated by Chen, Rendina-Gobioff, & Dedrick (2007) and DiStefano & Motl (2006).

Wording effects have been found to inflate correlations among same-worded items and result in a two-factor solution in social dominance orientation (Xin & Chi, 2010), anxiety (Motl, Conroy, & Horan, 2000), general health (Molina et al., 2014), core self-evaluation (Kennedy, 2007), and quality of life (Lin et al., 2014). Different methods under the framework of confirmatory factor analysis have been tried by researchers in order to reveal the true factor structure of the scales. Some researchers adjusted negatively worded items only and chose to include only a positive method factor due to superior model fit (Wu, 2008). Others made the adjustment by including both method factors (Lin et al., 2014; Chen et al., 2007) to test if including both led to a better solution. In the current study, a positive method factor only and a negative factor only was tried. Compared with the final bi-factor structure, the chi-square test was significant and the shift in both RMSEA and CFI exceeded the limit, indicating that the model with both method factors fit significantly better than the two alternative models. To ascertain the improved model fit associated with the two-method-factors model was not caused by the addition of random parameters (Wouters et al., 2012), the author tested five models where the items were randomly allocated to form two method factors. All five models encountered convergence problems, suggesting that the optimal fit of the final two-factor CFA model was not a coincidence.

The finding that resilience scores were able to predict academic outcomes are worth further exploring. Perhaps schools could use information related to students' resilience to select those more likely to be successful academically. Schools might also use measures of resilience to identify students at risk of poor achievement in order to intervene at an earlier stage. The finding on the predictive power of resilience on life satisfaction could be valuable in similar ways. In order to use resilience scores for selective enrollment purposes, more studies examining the longer-term predictive power of resilience are needed and a scale which is more reliable across time is highly desirable.

The finding that resilience was a significant predictor of life satisfaction but GPA was not is illuminating. Students with higher GPAs are not necessarily happier in life. But students with higher resilience are. The goal of education should not be limited to the pursuit of knowledge. Education should empower students to shape their own futures and support them in lifelong pursuit of happiness (Noddings, 2003). Strengthening children and adolescents' resilience may be key to helping them reach their full potential in life (Henderson & Milstein, 2003).

Limitations

As discussed above, some features of the current study add to its strength. But like any study, there are limitations that bear discussion. First, the measure of resilience used in the study is far from perfect. For example, there was lack of theoretical support for scale development. There were items that did not fit the concept of resilience. Although confirmatory bi-factor analysis was used to mitigate these problems, the solution might not be as good as focusing on measuring the outcome aspect of resilience.

The demand for innovative measures of resilience and other non-cognitive skills is high. The SJT subscale represents an innovative method and seems promising based on the results from this study. Nevertheless, SJT scale suffers from all the common biases shared by self-evaluation methods. For example, since the life satisfaction outcome measure comes from a self-rated scale, the significant relationship between SJT resilience and life satisfaction might reflect the social desirability bias underlying both scales (Paulhus, 1991). If researchers working on the assessment of non-cognitive skills can move beyond traditional self-surveys, and predictive validity is still obtained, the results would be more impressive.

Another limitation lies in the fact that the study has a large amount of missing data. Two types of missingness exist: non-participation of students and incomplete information on the variables. The main reason for student non-participation in wave 3 was graduation. The main reason for student non-participation in wave 1 and 2 was either the student had not matriculated into middle school or the school did not participate in the earlier waves of data collection. The main reason for the missingness on response variables was that some schools did not assign grades to students in lower grade levels or require them to participate in the standardized test. With regard to the prediction models used to answer research question No.3, neither multiple imputation nor maximum likelihood methods were tried because both require the data to be missing at random (Allison, 2001). Obviously the pattern of missingness in this current study is not at random. Regarding the longitudinal growth curve models, 40% to 50% of the total observations (person-occasion data) were from students who participated in all three waves of data collection. The rest of the observations were contributed by students who participated in the study for only one or two waves. In order to examine the robustness of the longitudinal growth curve models, analyses including only students who had participated in all three waves or who had participated in more than one wave were conducted and similar results were obtained.

One more shortcoming of the study is that grade level information was available for only a small number of students. It was possible that students in grade 6 experienced a different pattern of change in resilience compared with students in grade 7 or 8. Therefore without the ability to include grade level in the models for research question No.2, the power to account for a potential significant amount of error variance was lost.

Similar to grade level information, students' age information was not incorporated into the model. One reason lay in the fact that age information was available in years but not months. So it could not capture the differences across the three waves of data. More accurate information about age would be useful in two ways. First, age could be used as a level-2 control variable in the longitudinal growth curve models to examine if students at different ages at the time of the first data collection experience different patterns of change with regard to resilience and outcomes. Second, age could be treated as a level-1 time-varying predictor. Therefore in addition to modeling change across time, change can be modeled as students grow older. The approach would capture a more accurate picture of how student resilience as well as outcomes change during adolescence. [1]

Yet another limitation of the study was that only three waves of data were collected at the time analyses were conducted. This limited the choice of models. For example, only linear trend could be examined for longitudinal growth curve models. With

[1] The author made an attempt to transform the age information into months and age was used as level-1 time varying variable in the model. Results were very similar as presented in the previous section.

more waves of data, higher-order polynomial trends (i.e., quadratic trend, cubic trend) could be tested. Even the current linear growth curve model would benefit from more waves of data. Instead of using the time-varying resilience scores to predict the outcome collected in the same wave, the resilience scores in each wave could be linked to the outcome variable collected in the next wave. One more weakness of the current study regarding the prediction model is its inability to rule out the possibility that the significant relationship was the result of a reversed causal explanation—the outcome caused resilience to change instead of vice versa. By fitting a longitudinal growth curve model in which the outcome in each wave serves as the time-varying predictor for the resilience score from the next wave, the reversed predictive relationship alternative can be evaluated.

As mentioned in the methods section, the author is handicapped to establish a causal relationship because the study is not an experiment. No matter how much evidence can be gathered, it is impossible to disqualify all the alternative explanations of a significant predictive relationship. Different types of validity can be tested but a causal relationship cannot be verified.

Finally, there is the limitation of external validity. The sample in the study consists completely of independent school students. However, the demographic information in the participating sample is very different from a public school population regarding the percent of minority students and the percent of economically disadvantaged students. Both the measurement model and the prediction model might not fit students attending public schools. Therefore generalizing the conclusions and the implications of the current study to a broader population of public school students may not be appropriate.

Future Directions

Investigating causal relationships of resilience on future outcomes is a priority for future work. The more we know how to build and shape resilience for children and adolescents, the easier we can construct scales to measure resilience accurately and reliably and the better we can design interventions that improve students' resilience. We can further conduct experimental studies to test the effectiveness of resilience interventions and gather the strongest evidence on whether and to what extent resilience impacts student outcomes such as academic achievement.

The SJT resilience score significantly predicted life satisfaction in this study. However, both the SJT resilience score and the life satisfaction score were calculated based on a single factor composed of only positively worded items. Perhaps the positive wording effect underlying both scales inflated the common variance between the two scores. Although there were only two negatively worded items in the original life satisfaction scale, a sensitivity test was conducted using resilience to predict the average score from the two negatively worded life satisfaction items. Two models were built while the three resilience scores at time 1 were put into the model to predict the score of the negatively worded factor at time 3. One model controlled for the factor score at time 2 and the other did not. Contrary to previous findings, the SJT score was not significant in either of the two models. It could be that the factor made up of only two items was not reliable therefore attenuating the relationship between the response variable and the predictors. But the result renders the previous finding of the strong predictive power of SJT score on life satisfaction suspicious. Further research is needed to determine what makes the SJT score a significant predictor of the life satisfaction score derived from positively-worded items but not from negatively-worded items.

The finding of the low correlation between the resilience score calculated from the self-rating subscale and that derived from the teacher-rating subscale is worth further exploration. If they capture different aspects of resilience, what are those aspects? Interviews with students and their teachers whose ratings differ by a large amount might provide some insights.

A fourth wave of data collection was finished in February of 2014. If the data becomes available, life satisfaction could be modeled using model 3. Moreover, three more longitudinal growth curve models could be fit. The first model would include a quadratic term at level 1 and utilize leve-2 variables to predict the intercept, the slope for time, and the slope for time^2. The second model would examine a linear trend but the time-varying resilience scores from each wave would be linked to the outcomes in the next wave. The third model would explore the possibility of a reversed relationship where outcomes from each wave are used to predict the resilience scores from the next wave. Figure 11 through 13 are created to illustrate the three potential models.

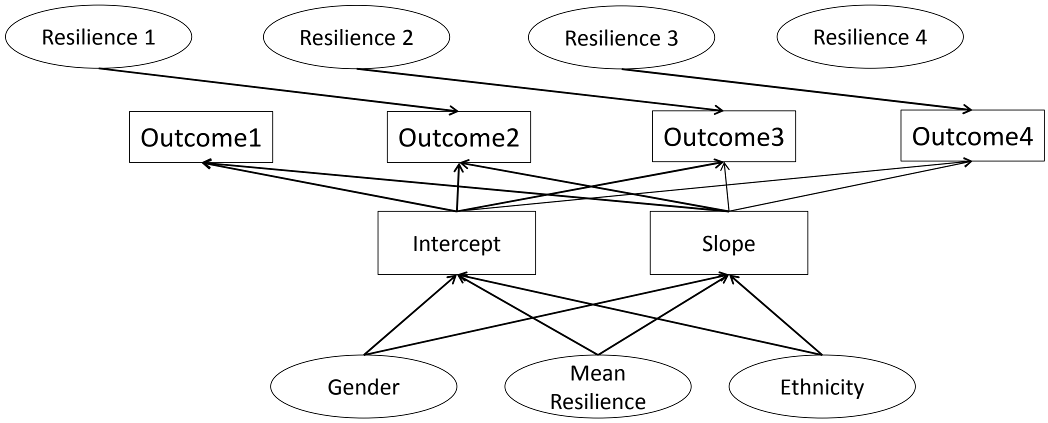

Figure 11: Illustration of the first potential model

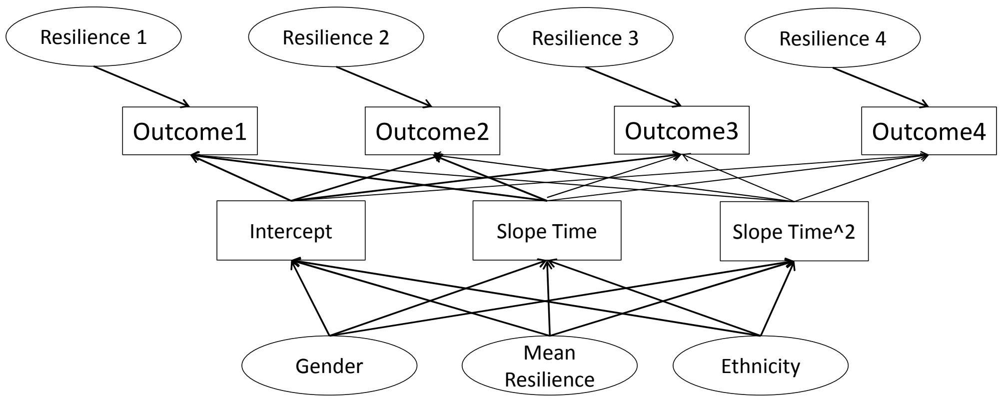

Figure 12: Illustration of the second potential model

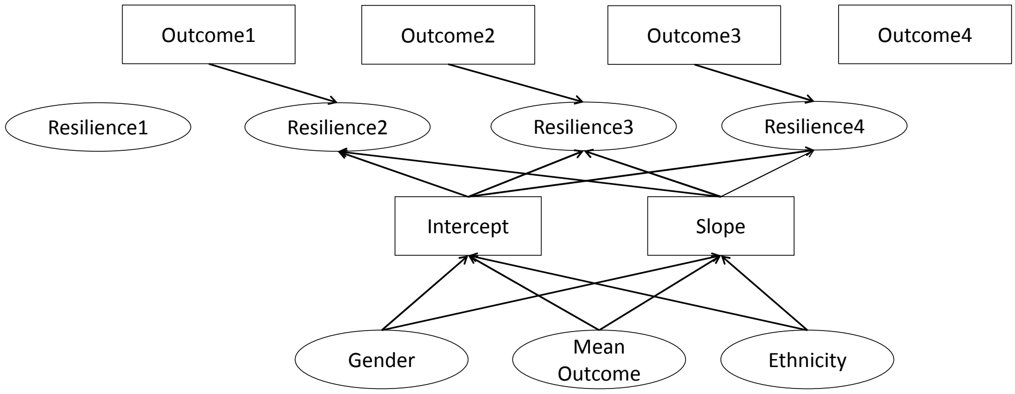

Figure 13: Illustration of the third potential model

The second and third models are both lagged prediction models, representing competing theories explaining a potential significant relationship between resilience and the outcome. Each model can also be fit under a structural equation modeling framework (Bollen, 1998) and the fit statistics can be compared to determine which theory is supported by the data. A cross-lagged prediction model could also be built. By comparing the Chi-square statistics, the RMSEA, and the CFI between the cross-lagged model and each of the nested single-direction lagged models, the author can determine if the cross-lagged model fits the data significantly better than the two less restricted models. An illustration of the three models can be seen in figure 14.

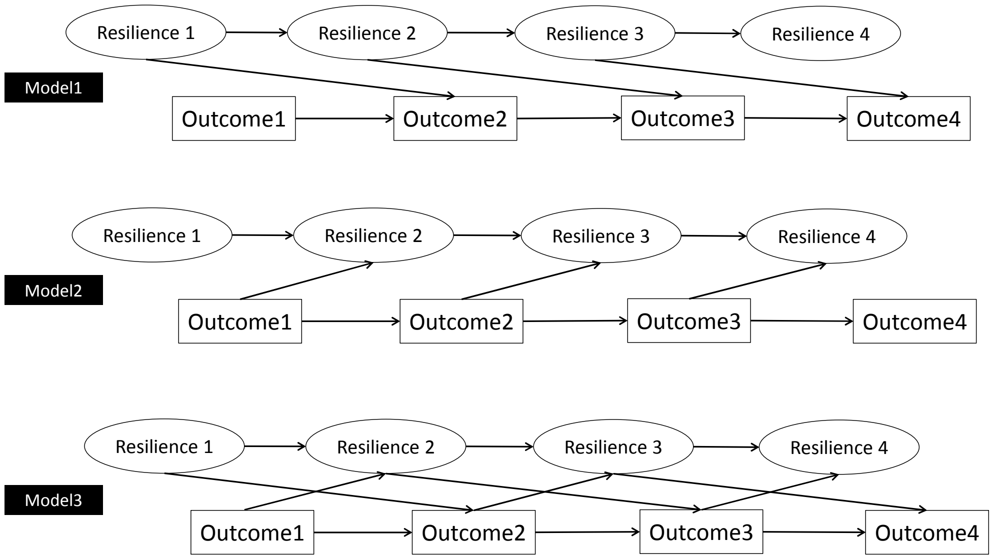

Figure 14: Illustration of SEM models

There is another advantage of using structural equation modeling. By combining the measurement part (CFA part) and the structural part (Prediction part), SEM can estimate regression coefficients after adjusting for reliability of measurement (Bollen, 1998). If a model incorporating outcomes, demographic variables, resilience factors and items could be fit, the results could become a more accurate reflection of the true relationship between resilience and the outcomes.

The current study did not examine whether the same factor structure held for students from different gender or ethnicity groups or whether resilience worked the same way for students from different groups. It might be another area worth exploring.

BIBLIOGRAPHY

Abe, J. A. A. (2005). The predictive validity of the Five-Factor Model of personality with preschool age children: A nine year follow-up study. *Journal of Research in Personality*, *39*(4), 423-442.

Ahern, N. R., Kiehl, E. M., Lou Sole, M., & Byers, J. (2006). A review of instruments measuring resilience. *Issues in Comprehensive Pediatric Nursing*, *29*(2), 103-125.

Allison, P. D. (Ed.). (2001). *Missing data* (No. 136). Sage.

Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: Americans' perceptions of life quality.* New York: Plenum.

Arnold, K. D., & King, I. C. (Eds.). (1997). *College student development and academic life: Psychological, intellectual, social, and moral issues* (Vol. 4). Taylor & Francis.

Baker, S. R., Victor, J. B., Chambers, A. L., & Halverson, C. F. (2004). Adolescent Personality A Five-Factor Model Construct Validation. *Assessment*, *11*(4), 303-315.

Bandura, A. (1989). Human agency in social cognitive theory. *American psychologist*, *44*(9), 1175.

Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2002). Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. *Twin Research*, 5, 544–553

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44*(1), 1-26.

Baruth, K. E., & Caroll, J. J. (2002). A formal assessment of resilience: The Baruth Protective Factors Inventory. *The Journal of Individual Psychology*, 58(3), 235-244.

Bracey, G. W. (1996). International comparisons and the condition of American education. *Educational Researcher, 25(1)*, 5-11.

Benard, B. (1991). *Fostering resiliency in kid: Protective factors in the family, school and community*. Portland, OR: Northwest Regional Educational Laboratory.

Benson, P. L. (2012). Adolescent development in social and community context: A program of research. *New Directions for Youth Development*, 95, 123-14

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A Meta-Analytic Investigation of Job Applicant Faking on Personality Measures. *International Journal of Selection and Assessment*, *14*(4), 317-335.

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.

Blanden, J., Gregg, P., & Macmillan, L. (2007). Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education. *The Economic Journal*, *117*(519), C43-C60.

Boekarts, M. (1993). Being concerned with well-being and with learning. *Educational Psychologist*, 28, 149–167

Bollen, K. (1989). Structural equations with latent variables. New York, NY: Wiley Interscience.

Bollen, K. A. (1998). *Structural equation models*. John Wiley & Sons, Ltd.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*,*43*(4), 972-1059.

Borghans, L., Meijers, H., & Ter Weel, B. (2008). The role of noncognitive skills in explaining cognitive test scores. *Economic Inquiry*, *46*(1), 2-12.

Bowdon, J. (2013). The origins, growth, and implications of black-white non-cognitive skill gaps during elementary school (Unpublished doctoral dissertation). University of Wisconsin, Madison, WI.

Bradshaw, D. (1985). Transferable intellectual and personal skills. *Oxford Review of Education*, *11*(2), 201-216.

Brandb, S., Feldmetha, A. K., Langa, C., Elliota, C., Holsboer-Trachslerb, E., & Pühsea, U. (2013). Adolescents with high mental toughness adapt better to perceived stress: A longitudinal study with Swiss vocational students. *Personality and Individual Differences, 54 (7),* 808-814.

Brody, N. (1992). *Intelligence (2nd edition).* New York: Academic.

Brooks, J. E. (2006). Strengthening resilience in children and youths: Maximizing opportunities through the schools. *Children & Schools*, *28*(2), 69-76.

Brunello, G., & Schlotter, M. (2011). Non-Cognitive Skills and Personality Traits: Labour Market Relevance and Their Development in Education & Training Systems. IZA Discussion Paper No. 5743. Available at SSRN: http://ssrn.com/abstract=1858066

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

Brown, A. L., & Cocking, R. R. (2000). *How people learn* (pp. 285-348). J. D. Bransford (Ed.). Washington, DC: National Academy Press.

Brown, P. H., & Park, A. (2002). Education and poverty in rural China.*Economics of Education Review*, *21*(6), 523-541.

Burrus, J., Betancourt, A., Holtzman, S., Minsky, J., MacCann, C., & Roberts, R. D. (2012). Emotional Intelligence Relates to Well-Being: Evidence from the Situational Judgment Test of Emotional Management. *Applied Psychology: Health and Well-Being*, *4*(2), 151-166.

Burrus, J., MacCann, C., Kyllonen, P. C., & Roberts, R. D. (2011). Noncognitive constructs in K-16: Assessments, interventions, educational and policy implications. *Readings on Equal Education: Diversity, Merit, and Higher Education: Toward a Comprehensive Agenda for the Twenty-First Century,* New York.

Campbell, A. (1976). Subjective measures of well-being. *American psychologist*, *31*(2), 117.

Campbell, J. R., Voelkl, K. E., & Donahue, P. L. (1997). NAEP 1996 Trends in Academic Progress. Achievement of US Students in Science, 1969 to 1996; Mathematics, 1973 to 1996; Reading, 1971 to 1996; Writing, 1984 to 1996. Report in Brief.

Carmines, E. G., & Zeller, R. A. (Eds.). (1979). *Reliability and validity assessment* (Vol. 17). Sage.

Casner-Lotto, J., & Barrington, L. (2006). *Are They Really Ready to Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century US Workforce*. Partnership for 21st Century Skills. 1 Massachusetts Avenue NW Suite 700, Washington, DC 20001.

Celio, C. I., Durlak, J., & Dymnicki, A. (2011). A Meta-analysis of the Impact of Service-Learning on Students. *Journal of Experiential Education*, *34*(2), 164-181.

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*(4), 415-427.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, *37*(4), 319-338.

Chan, Y. C., Yeh, Y. C., Peng, Y. Y., & Yeh, B. L. (2009). The development of the "inventory of adolescent resilience". *Psychological Testing,56*(4), 491－518.

Chen, Y. H., Rendina-Gobioff, G., & Dedrick, R. F. (2007). Detecting Effects of Positively and Negatively Worded Items on a Self-Concept Scale for Third and Sixth Grade Elementary Students. *Online Submission*, *14*.

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189-225.

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

Couch, A., & Keniston, K. (1960). Yeasayers and naysayers: agreeing response set as a personality variable. *The Journal of Abnormal and Social Psychology*, *60*(2), 151.

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and anxiety*, *18*(2), 76-82.

Crandall, R. (1976). Validation of self-report measures using ratings by others. *Sociological Methods & Research*, *4*(3), 380-400.

Crean, H. F. (2004). Social support, conflict, major life stressors, and adaptive coping strategies in Latino middle school students: An integrative model. *Journal of Adolescent Research, 19*(6), 657–676.

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*(6), 425-453.

Cronbach, L. J. (1946). Response sets and test validity. *Educational and psychological measurement*, *6*(4), 475-494.

Cullen, M. J., Sackett, P. R., & Lievens, F. (2006). Threats to the operational use of situational judgment tests in the college admission process. *International Journal of Selection and Assessment*, *14*(2), 142-155.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*,*11*(2), 121-136.

Dalley, M. B., Bolocofsky, D. N., & Karlin, N. J. (1994). Teacher-ratings and self-ratings of social competency in adolescents with low-and high-depressive symptoms. *Journal of Abnormal Child Psychology*, *22*(4), 477-485.

Darling-Hammond, L. (2004). Standards, accountability, and school reform. *The Teachers College Record*, *106*(6), 1047-1085.

Davey, M., Eaker, D. G., & Walters, L. H. (2003). Resilience processes in adolescents: Personality profiles, self-worth, and coping. *Journal of Adolescent Research*, *18*(4), 347-362.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: following up the Scottish mental surveys of 1932 and 1947. *Journal of personality and social psychology*, *86*(1), 130.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: a meta-analysis of 137 personality traits and subjective well-being. *Psychological bulletin*, *124*(2), 197.

Dickson, J. M., & MacLeod, A. K. (2004). Approach and avoidance goals and plans: Their relationship to anxiety and depression. *Cognitive Therapy and Research*, *28*(3), 415-432.

Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well-being across nations. *Culture and subjective well-being*, 185-218.

Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual review of psychology*, *54*(1), 403-425.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, *13*(3), 440-464.

Duckworth, A. L. (2013). True grit. *The Observer, 26(4),* 1-3.

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 9, 1087-1101.

Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. *Social Psychology and Personality Science*, *1(4)*, 311-317.

Dunbar, M., Ford, G., Hunt, K., & Der, G. (2000). Question wording effects in the assessment of global self-esteem. *European Journal of Psychological Assessment*, *16*(1), 13.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement.*Developmental psychology*, *43*(6), 1428.

Duncan, G. J., & Dunifon, R. (2012). “Soft-Skills” and long-run labor market success. *Research in labor economics*, *35*, 313-339.

Dweck, C.S., Walton, G.M., and Cohen, G.L. (2011). *Academic tenacity: Mindsets and skills that promote long-term learning*. White paper prepared for the Gates Foundation. Seattle, WA.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. *Nations and households in economic growth*, *89*.

Egeland, B., Carlson, E., & Sroufe, L. A. (1993). Resilience as process.*Development and psychopathology*, *5*(04), 517-528.

Ehrenberg, M. F., Cox, D. N., & Koopman, R. F. (1991). The relationship between self-efficacy and depression in adolescents. *Adolescence*.

Ellis, A.P.J., West, B.J., Ryan, A.M., & DeShon, R.P. (2002). The use of impression management tactics in structured interviews: A function of question type. *Journal of Applied Psychology*, 87, 1200–1208.

Emmons, R. A. (1986). Personal strivings: An approach to personality and subjective well-being. *Journal of Personality and social psychology*, *51*(5), 1058.

Espinosa, L. M., & Laffey, J. M. (2003). Urban primary teacher perceptions of children with challenging behaviors. *Journal of Children and Poverty*, *9*(2), 135-156.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Shaping School Performance--A Critical Literature Review*. Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.

Fedorowicz, A. E., & Kerig, P. K. (1995). Children's Coping Questionnaire. Unpublished measure, Simon Fraser University.

Frey, B. S., & Stutzer, A. (2002). The economics of happiness. *World Economics*, *3*(1), 25-41.

Friborg, O., Hjemdal, O., Rosenvinge, J. H., & Martinussen, M. (2003). A new rating scale for adult resilience: What are the central protective resources behind healthy adjustment?. *International journal of methods in psychiatric research*, *12*(2), 65-76.

Frye, A. A., & Goodman, S. H. (2000). Which social problem-solving components buffer depression in adolescent girls?. *Cognitive Therapy and Research*, *24*(6), 637-650.

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*,*17*(3), 1-13.

Garmezy, N. (1974, August) *The study of children at risk: New perspectives for developmental psychopathology.*

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2011). Performance of Velicer's minimum average partial factor retention method with categorical variables. *Educational and Psychological Measurement, 71*(3), 551-570.

Gerber, M., Kalak, N., Lemola, S., Clough, P. J., Perry, J. L., Pühse, U, & Brand, S. (2012). Are adolescents with high mental toughness levels more resilient against stress? *Stress and Health*.

Goleman, D. (2006). *Emotional intelligence*. Random House LLC.

Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review*,*36*(3), 341-355.

Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation review*, *22*(3), 403-426.

Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people.

Hanson, T. L., & Kim, J. O. (2007). Measuring resilience and youth development: The psychometric properties of the Healthy Kids Survey. *Issues & Answers Report, REL*, (034).

Headey, B., & Wearing, A. J. (1992). *Understanding happiness: A theory of subjective well-being*. Longman Cheshire.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child development*, *72*(2), 625-638.

Hankin, B. L., Roberts, J., & Gotlib, I. H. (1997). Elevated self-standards and emotional distress during adolescence: Emotional specificity and gender differences. *Cognitive Therapy and Research*, *21*(6), 663-679.

Hanson, M. A., & Ramos, R. A. (1996). Situational judgment tests. *Fair employment strategies in human resource management*, 119-124.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*,*91*(2), 145-149.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). *The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior* (No. w12006). National Bureau of Economic Research.

Henderson, N., & Milstein, M. M. (2003). *Resiliency in schools: Making it happen for students and educators*. Corwin Press.

Hough, L. M., & Oswald, F. L. (2008). Personality testing and industrial–organizational psychology: Reflections, progress, and prospects. *Industrial and Organizational Psychology*, *1*(3), 272-290.

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.

Hunter, A. J. (2001). A cross-cultural comparison of resilience in adolescents.*Journal of Pediatric Nursing*, *16*(3), 172-179.

Shengli, H. (1994). A study on the mental health level and its influencing factors of senior middle school students. *Acta Psychologica Sinica*, *2*.

Jacelon, C. S. (1997). The trait and process of resilience. *Journal of advanced nursing*, *25*(1), 123-129.

Jackson, Y., & Warren, J. S. (2000). Appraisal, Social Support, and Life Events: Predicting Outcome Behavior in School‐Age Children. *Child development*, *71*(5), 1441-1457.

James Jr, H. S. (2011). Is the Just Man a Happy Man? An Empirical Study of the Relationship between Ethics and Subjective Well‐Being. *Kyklos*, *64*(2), 193-212.

Jew, C. L., Green, K. E., & Kroger, J. (1999). Development and validation of a measure of resiliency. *Measurement & Evaluation in Counseling & Development, 32*(2), 75–89.

Kennedy, J. C. (2007). The effect on model fit of ignoring item wording effects in the core self-evaluations scale. Paper presented at 2007 annual meeting of the Academy of Management, Philadelphia, PA.

Kenny, D. A. (1994). *Interpersonal perception*. New York: Guilford Press.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, *112*(1), 165.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kobasa SC. 1979. Stressful life events, personality, and health: an inquiry into hardiness. *Journal of Personality and Social Psychology*, 37, 1–11.

Koenig, J. A. (2011). *Assessing 21st century skills: summary of a workshop*. National Academies Press.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, *51*(4), 472-477.

Krokos, K. J., Meade, A. W., Cantwell, A. R., Pond, S. B., & Wilson, M. A. (2004). Empirical keying of situational judgment tests: Rationale and some examples. In *annual conference of the Society for Industrial and Organizational Psychology, Chicago, IL*.

Kwok, O., Hughes, J. N., & Luo, W. (2007). Role of resilient personality on lower achieving first grade students' current and future achievement. *Journal of School Psychology, 45*(1), 61–82.

Kyllonen, P., Walters, A. M., & Kaufman, J. C. (2005). Noncognitive constructs and their assessment in graduate education: A review. *Educational Assessment*, *10*(3), 153-184.

Kumar, A., & Dillon, W. R. (1992). An integrative look at the use of additive and multiplicative covariance structure models in the analysis of MTMM data. *Journal of Marketing Research (JMR)*, *29*(1).

Ladd, G. W., Birch, S. H., & Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child development*, *70*(6), 1373-1400.

Lai, E.R. and Viering, M. (2012). Assessing 21st Century Skills: Integrating Research Findings. Retrieved from: http://www.pearsonassessments.com/hai/images/tmrs/Assessing_21st_Century_Skill_NCME.pdf

Lazarus, R. S., Folkman, S., & Stress, A. (1984). Coping, 1984. *Lazarus, RS& Folkman, S,"“Coping and adaptation”", Gentry, WD (Eds.), Handbook of Behavioral Medicine*, 11-21.

Lee, G., Dunbar, S. B., & Frisbie, D. A. (2001). The relative appropriateness of eight measurement models for analyzing scores from tests composed of testlets. *Educational and Psychological Measurement, 61*(6), 958-975.

Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment,14*(4), 299-316.

Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, *97*(2), 460.

Lin, C., Luh, W., Cheng, C., Yang, A., & Ma, H. (2014).Evaluating the wording effect and psychometric properties of the Kid-KINDL: Using the multitrait-multimethod approach. *European Journal of Psychological Assessment, 30*(2), 100-109.

Lipnevich, A., MacCann, C., & Roberts, R. (2013). Assessing Non-Cognitive Constructs in Education: A Review of Traditional and Innovative Approaches. *Oxford Handbooks Online.*

Lucas, R. E., & Baird, B. M. (2006). Global Self-Assessment.

Lucas, R. E., & Fujita, F. (2000). Factors influencing the relation between extraversion and pleasant affect. *Journal of personality and social psychology*,*79*(6), 1039.

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child development*, *71*(3), 543-562.

Luthar, S. S., Sawyer, J. A., & Brown, P. J. (2006). Conceptual issues in studies of resilience. *Annals of the New York Academy of Sciences*, *1094*(1), 105-115.

Lyons, J. A. (1991). Strategies for assessing the potential for positive adjustment following trauma. *Journal of Traumatic Stress*, *4*(1), 93-111.

MacCann, C., Fogarty, G., & Roberts, R. D. (2012). Strategies for Success in Vocational Education: Time Management is More Important for Part-Time than Full-Time Students. *Learning and Individual Differences*, *22*, 618–623.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors?. *Journal of personality and social psychology*, *70*(4), 810.

Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement, 13*(4), 335-361.

Marsh, H. W., & Grayson, D. (1994). Longitudinal confirmatory factor analysis: common, time-specific, item-specific, and residual-error components of variance. *Structural Equation Modeling, 1*(2), 116-145.

Martin, A. J. (2013). Academic buoyancy and academic resilience: Exploring ‘everyday’ and ‘classic’ resilience in the face of academic adversity. *School Psychology International*, *34*(5), 488-500.

Masten, A. S. (2004). Regulatory processes, risk, and resilience in adolescent development. *Annals of the New York Academy of Sciences*, *1021*(1), 310-319.

Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American psychologist*, *53*(2), 205.

McCoach, D. B., & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in psychology*, *1*.

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, *86*(4), 730.

Meier, D., & Wood, G. H. (Eds.). (2004). *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools*. Beacon Press.

Molina, J. G., Rodrigo, M. F., Losilla, J., & Vives, J. (2014). Wording effects and the factor structure of the 12-item general health questionnaire (ghq-12). *Psychological Assessment, 26*(1).

Motl, R. W., Conroy, D. E., & Horan, P. M. (2000). The Social Physique Anxiety Scale: an example of the potential consequence of negatively worded items in factorial validity studies. *Journal of applied measurement*.

Moos, R. H. (1995). Development and applications of new measures of life stressors, social resources, and coping responses. *European Journal of Psychological Assessment*, *11*(1), 1.

Mrazek, P. J., & Mrazek, D. A. (1987). Resilience in child maltreatment victims: A conceptual exploration. *Child Abuse & Neglect*, *11*(3), 357-366.

Muthén, B., & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. Retrieved from http://www.statmodel.com/download/webnotes/CatMGLong.pdf

Muthén, L. K., & Muthén, B. O. (2012). Mplus. *The comprehensive modelling program for applied researchers: user's guide*, *5*.

Newman, B. M., Lohman, B. J., Newman, P. R., Myers, M. C., & Smith, V. L. (2000). Experiences of urban youth navigating the transition to ninth grade. *Youth & Society, 31*(4), 387–416.

Noddings, N. (2003). *Happiness and education*. Cambridge University Press.

Noor, N. M., & Alwi, A. (2013). Stressors and well-being in low socio-economic status Malaysian adolescents: The role of resilience resources. *Asian Journal of Social Psychology*, *16*(4), 292-306.

OECD (2013). *PISA 2015 Collaborative Problem Solving Framework*. OECD Publishing.

Oshio, A., KANEKO, H., NAGAMINE, S., & NAKAYA, M. (2003). Construct validity of the adolescent resilience scale. *Psychological Reports*, *93*(3f), 1217-1222.

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology*, *89*(2), 187.

Park, N., Peterson, C., & Seligman, M. E. (2004). Strengths of character and well-being. *Journal of Social and Clinical Psychology*, *23*(5), 603-619.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson & L. S. Wrightsman (Ed.). *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Pianta, R. C., Cox, M. J., & Snow, K. L. (2007). *School readiness and the transition to kindergarten in the era of accountability*. Paul H Brookes Publishing.

Plybon, L. E., Edwards, L., Butler, D., Belgrave, F. Z., & Allison, K. W. (2003). Examining the link between neighborhood cohesion and school outcomes: The role of support coping among African American adolescent girls. *Journal of Black Psychology, 29*(4), 393–407.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, *135*(2), 322.

Prince-Embury, S. (2005). Resiliency scales for adolescents: A profile of personal strengths manual.

Prince-Embury, S. (2008). The resiliency scales for children and adolescents, psychological symptoms, and clinical status in adolescents. *Canadian Journal of School Psychology*, *23*(1), 41-56.

Rak, C. F., & Patterson, L. E. (1996). Promoting resilience in at-risk children.*Journal of Counseling & Development*, *74*(4), 368-373.

Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual review of psychology*, *52*(1), 501-525.

Raudenbush, S.W., & Bryk, A.S. (2002) *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage.

Richardson, G. E. (2002). The metatheory of resilience and resiliency. *Journal of clinical psychology*, *58*(3), 307-321.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological bulletin*, *130*(2), 261.

Rosen, J. A., Glennie, E. J., Dalton, B. W., Lennon, J. M., & Bozick, R. N. (2010). *Noncognitive Skills in the Classroom: New Perspectives on Educational Research*. RTI International. PO Box 12194, Research Triangle Park, NC 27709-2194.

Rosenberg, M. (1965). Society and the adolescent self-image.

Rubin, D. B. (2009). *Multiple imputation for nonresponse in surveys* (Vol. 307). Wiley. com.

Rusting, C. L., & Larsen, R. J. (1998). Diurnal patterns of unpleasant mood: associations with neuroticism, depression, and anxiety. *Journal of Personality*, *66*(1), 85-103.

Rutter, M. (1985). Resilience in the face of adversity. *British journal of psychiatry*, *147*(1), 598-611.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of personality and social psychology*, *57*(6), 1069.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In Heijmans, R.D.H., Pollock, D.S.G. & Satorra, A. (eds.), *Innovations in multivariate statistical analysis: A Festschrift for Heinz Neudecker* (pp. 233-247). London: Kluwer Academic Publishers.

Sapienza, J. K., & Masten, A. S. (2011). Understanding and promoting resilience in children and youth. *Current Opinion in Psychiatry*, *24*(4), 267-273.

Schmitt, N., & Kunce, C. (2002). The effects of required elaboration of answers to biodata questions. *Personnel Psychology*, *55*(3), 569-587.

Semple, R. J., Lee, J., Rosa, D., & Miller, L. F. (2010). A randomized trial of mindfulness-based cognitive therapy for children: Promoting mindful attention to enhance social-emotional resiliency in children. *Journal of Child and Family Studies*, *19*(2), 218-229.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002*). Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shengli, H. (1994). A study on the mental health level and its influencing factors of senior middle school students. *Acta Psychologica Sinica*, *2*.

Shin, D. C., & Johnson, D. M. (1978). Avowed happiness as an overall assessment of the quality of life. *Social indicators research*, *5*(1-4), 475-492.

Sinclair, V. G., & Wallston, K. A. (2004). The development and psychometric evaluation of the Brief Resilient Coping Scale. *Assessment*, *11*(1), 94-101.

Singh, I., & Jha, A. (2012). Emotional Intelligence and Occupational Stress among the Faculty Members of Private Medical and Engineering Colleges of Uttar Pradesh: A Comparative Study. *Advances in Management*, *5*(7).

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In *Handbook of research on student engagement* (pp. 21-44). Springer US.

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back.*International journal of behavioral medicine*, *15*(3), 194-200.

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological bulletin*, *134*(1), 138.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental review*, *28*(1), 78-106.

Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use* (3rd ed.). New York: Oxford University Press.

Stutzer, A., & Frey, B. S. (2006). Does marriage make people happy, or do happy people get married?. *The Journal of Socio-Economics*, *35*(2), 326-347.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum

Tinsley, B., & Spencer, M. B. (2010). High Hope and Low Regard: The Resiliency of Adolescents' Educational Expectations While Developing in Challenging Political Contexts. *Research in Human Development*, *7*(3), 183-201.

Tracey, T. J., & Sedlacek, W. E. (1987). A comparison of white and black student academic success using noncognitive variables: A LISREL analysis.*Research in Higher Education*, *27*(4), 333-348.

Tugade, M. M., & Fredrickson, B. L. (2004). Resilient individuals use positive emotions to bounce back from negative emotional experiences. *Journal of personality and social psychology*, *86*(2), 320.

Van Praag, B. M., Frijters, P., & Ferrer-i-Carbonell, A. (2003). The anatomy of subjective well-being. *Journal of Economic Behavior & Organization*, *51*(1), 29-49.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika,41*(3), 321-327.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, *59*(2), 197-210.

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality*, *41*(1), 221-229.

Wagnild, G. (2009). A review of the Resilience Scale. *Journal of Nursing Measurement*, *17*(2), 105-113.

Wagnild, G. M., & Young, H. M. (1993). Development and psychometric evaluation of the Resilience Scale. *Journal of nursing measurement*.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, *54*(6), 1063.

Whalley, L. J., & Deary, I. J. (2001). Longitudinal cohort study of childhood IQ and survival up to age 76. *Bmj*, *322*(7290), 819.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M.

Windle, G., Bennett, K. M., & Noyes, J. (2011). A methodological review of resilience measurement scales. *Health and quality of life outcomes*, *9*(8), 1-18.

Windle, S. G. West, K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281-324). Washington, DC: American Psychological Association.

Wolchik, S. A., & Sandler, I. N. (1997). *Handbook of children's coping: Linking theory and intervention*. Plenum Press.

Wouters E, Booysen FlR, Ponnet K, Baron Van Loon F (2012) Wording Effects and the Factor Structure of the Hospital Anxiety & Depression Scale in HIV/AIDS Patients on Antiretroviral Treatment in South Africa. PLoS ONE 7(4): e34881. doi:10.1371/journal.pone.0034881

Wu, C. H. (2008). An examination of the wording effect in the Rosenberg Self-Esteem Scale among culturally Chinese people. *The Journal of social psychology*, *148*(5), 535-552.

Xin, Z., & Chi, L. (2010). Wording effect leads to a controversy over the construct of the social dominance orientation scale. *The Journal of psychology*,*144*(5), 473-488.

Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany, NY: State University of New York Press.

Yu, D. L., & Seligman, M. E. (2002). Preventing depressive symptoms in Chinese children. *Prevention & Treatment*, *5*(1), 9a.