University of South Carolina
Scholar Commons

2016

The Relationship between Peer-Self Concept and Academic Self-Concept in Elementary Students: A Person-Oriented Perspective

Catherine Jorgensen Snyder
University of South Carolina

Follow this and additional works at: https://scholarcommons.sc.edu/etd

Part of the Educational Psychology Commons

المنارة للاستشارات

www.manaraa.com

The Relationship between Peer-Self Concept and Academic Self-Concept in Elementary Students: A Person-Oriented Perspective

by

Catherine Jorgensen Snyder

Bachelor of Science
Michigan State University, 1982

Master of Arts
The University of Akron, 1988

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Educational Psychology and Research

College of Education

University of South Carolina

2016

Accepted by:

Christine DiStefano, Major Professor

Matthew Irvin, Committee Member

Robert Johnson, Committee Member

Kate Niehaus, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ii

DEDICATION

To my husband Tim, who not only knew what a formidable task I had undertaken, but also knew how to gently push me forward. Believe me, I noticed that you never once asked me when I would finally be finished! I am grateful for your support both emotionally and financially, as I could not have done it without you.

Also, to my parents, Ken and Sharon, who instilled in me the resolve and stamina to pursue my goals, even in the face of daunting obstacles.

ACKNOWLEDGEMENTS

I would like to acknowledge all the hard work and dedication my committee members contributed in helping this dissertation come to fruition. I appreciate all the insightful feedback and the hours spent reviewing this manuscript and the accompanying data. I feel lucky to have had the opportunity to learn from all of you.

Special thanks go out to my committee chair, Dr. Christine DiStefano. Her sense of humor in class was always something to look forward to. Yet, it was her belief in me that kept me going through the design dilemmas, data runs, and countless edits. I am grateful for her calming spirit and endless support as I have trudged through the challenges that I will fondly look back on as my graduate career.

ABSTRACT

Self-concept is one of the most researched constructs in educational psychology (Marsh, Xu, & Martin, 2012). It has been established that there are different domains of self-concept which fall into two main categories, academic and non-academic (Shavelson, Hubner, & Stanton, 1976). Unfortunately, little research has been conducted on the non-academic areas of self-concept and little is known about how social self-concept interrelates with academic self-concept. The focus of this study was to contribute to the research base by investigating the relationship between a facet of nonacademic self-concept, namely peer self-concept, and achievement self-concept among elementary students.

The current research utilized person-oriented methodology to study peer self-concept. A nationally representative sample of students in elementary school (from the NCES ECLS-K database) was followed to examine changes in perceived academic and peer self-concept over the course of two years (from grades three through five). Latent class and latent transition analyses (person-oriented research approaches) were conducted to determine intra-individual changes in academic and peer self-concept over time and how these changes predicted academic performance in grade five.

Results of the latent class analysis revealed that students with positive peer self-concept tended to have positive self-concept in reading, math, and other school subjects. Latent transition analysis showed that most students move to the next higher latent class over time, reflecting improvements in self-concept. The domain of academic self-concept that appeared to vary the most over time was that of math. Implications for school and classroom interventions and areas for further research are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I
Introduction

Understanding the factors that motivate individuals academically has been a quest for educators since William James and John Dewey identified the learner as an active force in the educational process (see Dewey, 1929; James, 1926). Since this pioneering research, many established motivational theories have been applied to education in an attempt to explain differences in achievement. These theories have, for example, focused on students' expectancies for academic success (e.g., self-efficacy theory; Bandura, 1986 and expectancy-value theory; Eccles, 1983; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), perceived competence (e.g., Kaplan & Midgley, 1997; Cocks & Watt, 2004), achievement goals (e.g., Ames, 1992; Nichols, 1984; Ford & Nichols, 1991), self-determination (e.g., Deci & Ryan, 1985, 2000; Ryan & Deci, 2000, 2009) and academic interest (e.g., Hidi, 2000; Schiefele, 2004; Silvia, 2005). Numerous studies using these theories have found that student motivation could be influenced by factors internal to the student (e.g., traits, drives, expectancies, values), as well as those which are external (e.g., rewards, social approval, expectations, environmental constraints).

Self-concept is one of the oldest and most studied constructs in educational psychology, addressing both cognitive and social domains. However, the focus of self-concept research has been largely one-sided, primarily addressing academic self-concept and largely ignoring the nonacademic domains.

1

While social self-concept was a major subdivision of the original model proposed by Shavelson, Hubner, and Stanton (1976) and refined by Marsh and Shavelson (1985), it has not received much consideration since then. In fact, social self-concept was not fully described or investigated for construct validity until Byrnes and Shavelson's 1996 research, 20 years later. Furthermore, there has been very little research on the peer self-concept subdomain. In fact, through a comprehensive literature search, only one study was found that specifically addressed peer self-concept. Accordingly, the overarching aim of the current study was to examine the relationship between peer and academic self-concept and how peer self-concept interfaces with academic self-concept to influence academic achievement. Toward this end, a longitudinal, secondary data analysis of the large-scale ECLS-K data was conducted. In particular, this research examined measures of academic and peer self-concept in concert with math and reading achievement data for students in grades three through five.

Importance of Self-Concept

Self-concept is a multi-dimensional construct that recognizes the importance of both academic and nonacademic (i.e., social, emotional, physical) factors for influencing motivation (Byrne, 2002; Byrne & Shavelson, 1986; Marsh, Parker, & Barnes, 1985; Marsh & Shavelson, 1985; Marsh, Smith, Barnes, & Butler, 1983; Marsh et al., 2012; Shavelson et al., 1976). Shavelson et al. (1976) envisioned self-concept as a multi-dimensional cognitive construct that included "self-perceptions that are formed thorough experience with and interpretations of one's environment." They noted that these perceptions are strongly influenced by evaluations made of one's behaviors by significant others (see Marsh et al., 2012, pg. 429). Marsh found that self-concept influences our

2

actions and then our actions influence our self-concept in a continuous circular cycle. Self-concept has been extensively researched and validated, as a way to understand student motivation to achieve academically (Marsh et al., 2012; Wylie, 1961, 1974, 1979).

Marsh and Craven clarified the distinction between self-concept and self-esteem in their 2006 publication. They noted that self-concept can be broken down into specific components which directly relate to outcomes like achievement. In contrast, self-esteem is a global construct much the same as general self-concept (the top level of the hierarchy). There have also been studies conducted to differentiate self-concept from Albert Bandura's concept of self-efficacy (1986). The main focus of self-efficacy rests on the individual's perception of their ability to accomplish a task and how this information is used to set goals, monitor performance, and adjust behavior. Research findings have shown that academic self-concept actually influences a student's specific feelings of self-efficacy. Bong and others concluded that self-concept is heavily based on social comparisons while self-efficacy is more competency-based and context-specific. Research also found that self-concept is a better predictor of affect, while self-efficacy is better for predicting actual achievement (Bong, 1998; Bong & Clark, 1999; Bong & Skaalvik, 2003; Ferla, Valcke, & Cai, 2009; Marsh, Dowson, Pietsch, & Walker, 2004; Marsh, Walker, & Debus, 1991).

Self-concept has been found to be an important construct for describing and understanding individual's beliefs and behaviors. Not only is it an important factor in decision-making and goal-directed behavior, but it is also a desirable outcome on its own. For example, individuals have been found to invest more effort when they have a positive

self-concept and have confidence in their ability (Marsh & Craven, 2006; Marsh et al., 2012; Shavelson et al., 1976). In fact research has shown that self-concept predicts academic achievement even beyond SES and prior achievement (Marsh, 1990a, 1993; Marsh, Byrne, & Yeung, 1999; Marsh & Craven, 1997, 2006). Most research on academic self-concept has involved some examination of academic achievement outcomes. Marsh and Martin (2011) summarized the cyclical relationship between achievement and academic self-concept. They found from reviewing numerous research studies that higher academic self-concept led to improvements in achievement and that the resulting gains in achievement further fueled higher academic self-concept. Additional research has supported the conclusions drawn by Marsh and Martin and have found even stronger relationships between domain-specific self-concept and corresponding achievement measures (Marsh et al., 2012; Marsh et al., 1999; Marsh & Craven 2006; Marsh, Hau, & Kong, 2002; Marsh & O'Mara, 2008; Pinxten, De Fraine, Van Damme, & D'Haenens, 2010).

Results have also been consistent in finding that academic achievement measures are un-related to nonacademic domains of self-concept (Arens, Yeung, Craven, & Hasselhorn, 2011; Marsh & Craven, 2006; Marsh & O'Mara, 2008; Marsh et al., 1985; Marsh, Parker et al., 1983; Marsh & Shavelson, 1985; Marsh, Smith, & Barnes, 1984; Marsh et al, 2012; Pinxten et al., 2010; Shavelson & Bolus, 1982). However, all of this research was conducted using variable-oriented methodology. It is the aim of this study to revisit the relationships between achievement and peer self-concept by using a person-oriented approach. Given the consistent results found on the influence of social competence on achievement performance it is expected that supportive relationships with

4

peers will have a positive impact on achievement motivation. For example, Kathyrn Wentzel found in her research that socially competent students were also higher achieving. Her research consistently found that social responsibility (being dependable, following rules, getting assignments done on time) distinguished low achieving from high achieving students. Specifically, Wentzel found that socially responsible students have positive interactions with other students and teachers that enhance learning (Wentzel, 1989, 1991a, 1993, 2009; Wentzel & Watkins, 2002).

Contributions of the Present Study to Self-Concept Research

As previously mentioned, self-concept has been one of the most studied constructs of self-perception. However, little research has been conducted on the nonacademic domains of self-concept. There has been even less research in the peer self-concept subdomain. The literature search undertaken as part of this dissertation, found only one study that specifically addressed peer self-concept.

This study goes beyond simply addressing the neglected area of peer self-concept, by also focusing on students in the upper grades of elementary school, grades three through five. In the limited research that has considered both perceptions of academic and social competence, most work has been done in middle school, grades six through nine (Ferla et al., 2009; Goodenow, 1993; Hamm & Faircloth, 2005; Kindermann, 1993, 2007; Kindermann & Skinner, 2009; Ryan, 2001; Wentzel, 1991a, 1991b, 1996, 1997).

Research on the mid- to late-elementary school age group is informative for two important reasons. First, it provides information about students' perceptions of their competence in peer relationships, at this stage of development. Many researchers have noted that peers become the most important role model for students when they reach

5

adolescence, but little research has been conducted to investigate the relationships between peers in the middle years of elementary school (Altermatt, 2011; Dowson & McInerney, 2001; Hamm & Faircloth, 2005; Hamm & Zhang, 2010; Higgins, 2007; Kindermann, 1993, 2007; Molloy, Gest, & Rulison, 2011; Oberle &Schonert-Reichl, 2013; Schunk & Meece, 2006; Steinberg, Brown, & Dornbusch, 1996; Véronneau & Dishion, 2011; Wentzel, 2003, 2005, 2009). Also, researchers studying the hierarchical nature of academic self-concept have found that it begins to differentiate into separate subject area domains beginning at fifth grade (Marsh Parker et al., 1983; Marsh, Smith et al., 1984; Marsh, Tracey, & Craven, 2006). A couple of studies were done with students from kindergarten to second grade (Marsh, Craven, & Debus, 1991, 1998; Marsh, Ellis, & Craven, 2002) but there is no research on domain specificity between second and fifth grades. Therefore, the present study's emphasis on grades three through five affords a more complete developmental perspective related to changes in the structure of academic self-concept as children mature.

Existing research on self-concept has almost exclusively relied upon methodology that compares individuals using summary scores (e.g., means). This variable-oriented approach, focuses on the relationships between measures (i.e., group differences, inter-correlations). While there may be advantages to this approach, a deficit is that the variables are assumed to have the same impact across all unique individuals. So, instead of looking for changes in cognitions and abilities in individuals over time, variable-oriented research focuses on making comparisons between groups (inter-individual analyses). Variable-oriented approaches involve traditional methodology such as ANOVA, regression, and correlational research. Even when studies have utilized newer

6

analytical approaches, they have been limited to variable-oriented approaches like confirmatory factor analysis, multi-level modeling, or structural equation modeling (Arens et al., 2011; Ireson & Hallam, 2009; Lindner-Muller, John, & Arnold, 2012; Marsh, Ellis et al., 2002; Marsh et al., 2006; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). To better understand individual differences, research investigations conducted with a person-oriented approach are warranted.

A review of the literature found three studies that used latent variable, person-oriented methodology when studying self-concept (De Fraine, Van Damme, & Onghena, 2007; Marsh, Ludkte, Trautwein, & Morin, 2009; Van de Schoot & Wong, 2012). These studies each used a slightly different latent variable methodology (i.e., Bayesian latent class analysis, latent profile analysis, and multilevel latent growth analysis, respectively).

The Marsh et al. (2009) study was the only one to investigate the possibility that latent profiles could better describe the complexity of self-concept. This study aimed to determine whether differences in self-concept profiles could be explained as qualitative differences (students have different shaped profiles of self-concept over a number of domains) or quantitative differences (students have more or less self-concept with similar profile shapes). Through latent profile analysis, Marsh and colleagues found that both qualitative and quantitative differences occurred between different profiles. Their results highlighted the importance of both the level of self-concept in addition to the differential importance of domains for individual students.

Unfortunately, there were some weaknesses with the Marsh study. It did not include any measure of social or peer-self-concept. Also, the self-concept instrument that was used was a German adaptation of the *SDQ-III* with five newly developed scales.

7

There was no validity or reliability evidence reported on these new scales. In addition, this investigation was not longitudinal, so it did not provide any information about changes in classes or profiles over time.

Both the Van de Schoot & Wong (2012) and the De Fraine et al., (2007) studies were conducted with samples of Dutch students using adapted self-concept instruments. The De Fraine et al., study only looked at academic self-concept as it related to Dutch language achievement. Van de Schoot and Wong (2012) examined global self-concept and 12 domain-specific areas of self-concept among young Dutch adults. Three of these specific areas of self-concept appeared to have some relevance to peer self-concept: social acceptance, close friendships, and romantic relations. However, the results of this study were inconclusive in that researchers found that men and women with highly delinquent behaviors formed two groups, one with high self-concept and another with low self-concept. The same results were obtained for men and women with no evidence of delinquent behaviors, some had high self-concept and others had low self-concept. While both of these Dutch studies were informative, it is not known how their results might generalize to other countries.

The fact that the research on self-concept is dominated by variable-oriented methodology makes it difficult to gain a clear understanding of the complexities that play out between student perceptions of their competence with respect to academic and peer relationships during individual development (Bergman & Trost, 2006; Sterba & Bauer, 2010). Research in developmental science has emphasized that investigations should look at the sum total of the complex interactions within the biological, behavioral, cognitive, and social-emotional subsystems of an individual within their environment. Furthermore,

8

examining the sum total of complex interactions within subsystems can be undertaken within each of these areas and/or across them in order to clarify how the constituent parts work together during the course of development (Bergman & Magnusson, 1997; Bergman, Magnusson, & El-Khouri, 2003; Bergman & Wangby, 2014; Molenaar, 2004; von Eye & Bogat, 2006). However, such complexities cannot be adequately modeled within the constraints of variable–oriented research methodologies. Instead the interactionist focus of person-oriented research methodologies is more appropriate, as these techniques allow researchers to model unique relationships within individuals in terms of cognitive and social perceptions as well as belief and motivational subsystems.

Research in developmental science also emphasizes that research methodology must be logically tied to the study's problem of interest. When the aim is to further an understanding of the complexities of human behavior then the methodology must be holistic and interactionist (Bergmann and Trost, 2006; Raufelder, Jagenow, Hoferichter, & Drury, 2013; Sterba & Bauer, 2010). For the current study, the problems of interest are the development of peer self-concept and its relationship with academic self-concept over time. The research questions address changes in the students' perceptions of peers and academic competency as children age and progress from one grade level to another. These study objectives require a person-oriented approach in order to understand how the complex interactions differ for individuals.

An additional weakness of extant research related to perceptions of cognitive and social competence and their relative impact on achievement motivation is the scarcity of true longitudinal research designs. Studies have often covered just the time period of a single school year, from fall to spring (see Hamm & Faircloth, 2005; Kindermann, 1993,

9

2007; Kindermann & Skinner, 2009; Ryan, 2001). It has only been common for researchers to include multiple years of data in their designs when studying reciprocal effects of self-concept (Guay, Marsh, & Boivin, 2003; Marsh, Hau et al., 2002; Marsh & O'Mara, 2008; Pinxten et al., 2010). Using cross sectional data makes it difficult to model developmental changes in academic and peer self-concept (Arens et al., 2011; Marsh & Ayotte, 2003). Without multiyear longitudinal designs, the ability to demonstrate causation is compromised (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007). Results from short-term longitudinal and cross-sectional analyses may be determined by the nuances present in the particular sample selected for study. For research to be able to demonstrate causation, it is recommended that designs include multiple measures to represent the latent constructs and that the analyses extend over multiple time periods (Kirk, 2013; Kline, 2011; Marsh & Martin, 2011). Even person-oriented methodologies benefit from incorporating a longitudinal component so that the analyses are not simply identifying clusters of individuals underlying a set of data at a single point in time.

For educators, it is equally important to understand changes occurring within a child over the course of time as it is to understand the differences between students. There are a number of ways in which this study will help advance knowledge about students' perceptions of their interest and competence with peer relationships and how this relates to their academic achievement. First, by using person-oriented analyses it will be possible to model the complex interactions students are faced with from internal cognitions and perceptions of their interpersonal competencies. In addition, knowledge of how cognitive and social perceptions interact will help educators tailor educational

experiences to maximize benefit to individual students. For some students this may mean establishing time during the school day to cultivate peer relationships. Peer tutoring and peer-led learning groups may be another beneficial intervention. Finally, using Marsh and Shavelson's (1985) multi-dimensional self-concept as a guiding framework will have a simple benefit; it will make it easy to communicate research results and their implications throughout the educational community.

11

<center>**Chapter II**</center>

<center>**Review of the Literature**</center>

This chapter will provide an overview of self-concept as a motivational construct. As part of this summary, areas where there are gaps in the existing literature base will be highlighted. A brief overview of person-oriented methods and corresponding latent class techniques will also be provided to set the stage for the methodology used in the current study.

Historical Origins of Self-Concept

The development of the Marsh/Shavelson model (1985) of multidimensional self-concept was an outgrowth of a disorganized smattering of research on the construct, much of which is detailed by Barbara Byrne, Shavelson and colleagues, and Ruth Wylie (Byrne, 1996; Shavelson et al., 1976; Wylie, 1961, 1974, 1979). These authors' critiques of the then current research were, that there was no consistent operational definitions of self-concept and that the methodological approaches to studying the construct were flawed (i.e., lack of appropriate measurement instruments, and failure to test for alternative interpretation of results). These critiques spurred Richard Shavelson and colleagues (1976) to establish a standardized definition of the construct of self-concept. Organized research followed and led to the development of the full multidimensional, hierarchical model (Marsh et al., 2012).

<center>12</center>

The hierarchical structure of the Marsh/Shavelson model starts with an all-encompassing perception of the self, known as the global self-concept. This overarching construct is then broken down into two distinctly separate branches: academic and non-academic self-concepts. Both of these divisions are then further differentiated. Academic self-concept is broken down into specific content-related domains such as math and science. Non-academic self-concept is further separated into three major areas: social, emotional, and physical self-concept. Each of these three areas is further differentiated as shown in the full multidimensional, hierarchical model in Figure 2.1.
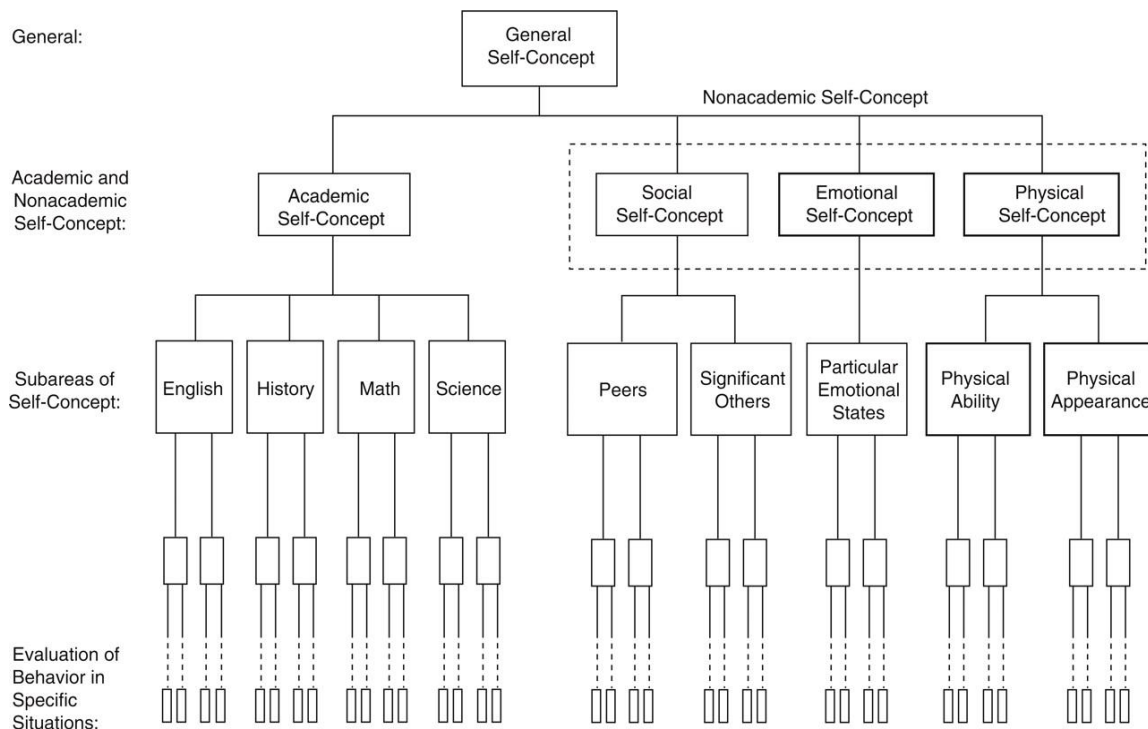
Figure 2.1. Structure of self-concept. (From "Self-Concept: Validation of Construct Interpretations" by R. J. Shavelson, J. J. Hubner, & J. C. Stanton, *Review of Educational Research,* 1976,*46,* 407-441. Copyright 1976 by the *Review of Educational Research,* Reprinted by permission.)
[Note: Shavelson and colleagues have boxed off the nonacademic factors in their model to indicate that they are not seen to be organized in a hierarchy as the academic factors are.]

13

The Structure of Self-Concept

Research aimed at testing the construct validity of the Shavelson model with its differentiated, hierarchical structure led to a couple of key refinements to the original 1976 model. The first refinement grew out of research into the existence of a single higher-order factor. Results from a number of studies showed that there was little correspondence between math and verbal self-concept, pointing to the possibility of two or more distinct areas of academic self-concept. Newly developed covariance structure modeling analyzed the structures of student self-report responses and tested alternatives to the single, general self-concept factor (Byrne & Shavelson, 1986; Marsh & Hocevar, 1985; Marsh & Shavelson, 1985; Marsh, Byrne, & Shavelson, 1988; Marsh 1990b; Shavelson & Marsh, 1986). Results from these studies concluded that for academic self-concept the hierarchical model structure could be more accurately represented with two higher order factors. The refined Marsh/Shavelson model (Marsh & Shavelson, 1985), shown in Figure 2.2, is believed to provide a more accurate representation of academic self-concept. Here specific domains are paired with the corresponding broader academic factors. The dotted lines indicate crossover subject areas that relate to both math and verbal skills.

The second refinement to the original Shavelson model involved the nonacademic portion of self-concept, specifically social self-concept. This portion of the original Shavelson model was admittedly rudimentary. Byrne and Shavelson (1996) undertook research with student responses to refine and further specify social self-concept.
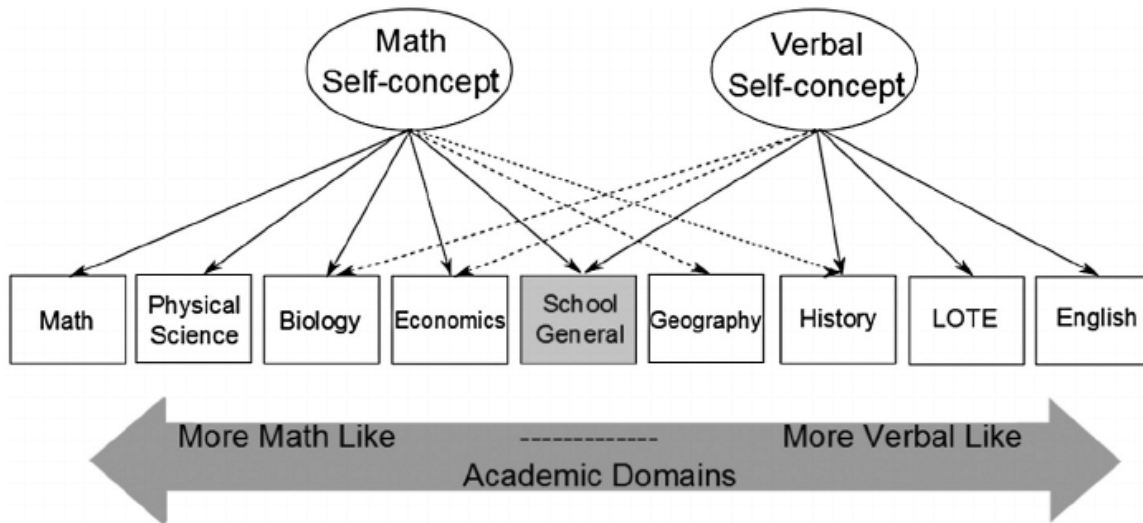
Figure 2.2. An elaboration of Marsh and Shavelson's (1985) revision to academic self-concept (that includes a wider variety of specific academic facets). From "A Multifaceted Academic Self-Concept: Its Hierarchical Structure and Its Relation to Academic Achievement" by H. W. Marsh, B. M. Byrne, *&* R. J. Shavelson, 1988, *Journal of Educational Psychology, 80,* p. 366-380. Copyright 1988 by the American Psychological Association. Reprinted by permission.)

The resulting expanded model, see Figure 2.3, breaks general social self-concept into both school and family facets. These two facets are further specified to differentiate between classmates and teachers under school social self-concept and siblings and parents under family social self-concept.

Below the dotted line in the model are the actual behaviors corresponding to that specific social context. Byrne and Shavelson believed that behaviors partially determined an individual's self-concept in an area, in line with Shavelson's definition of self-concept as self-perceptions related to interactions within a specific social context (Marsh et al., 2012).

15
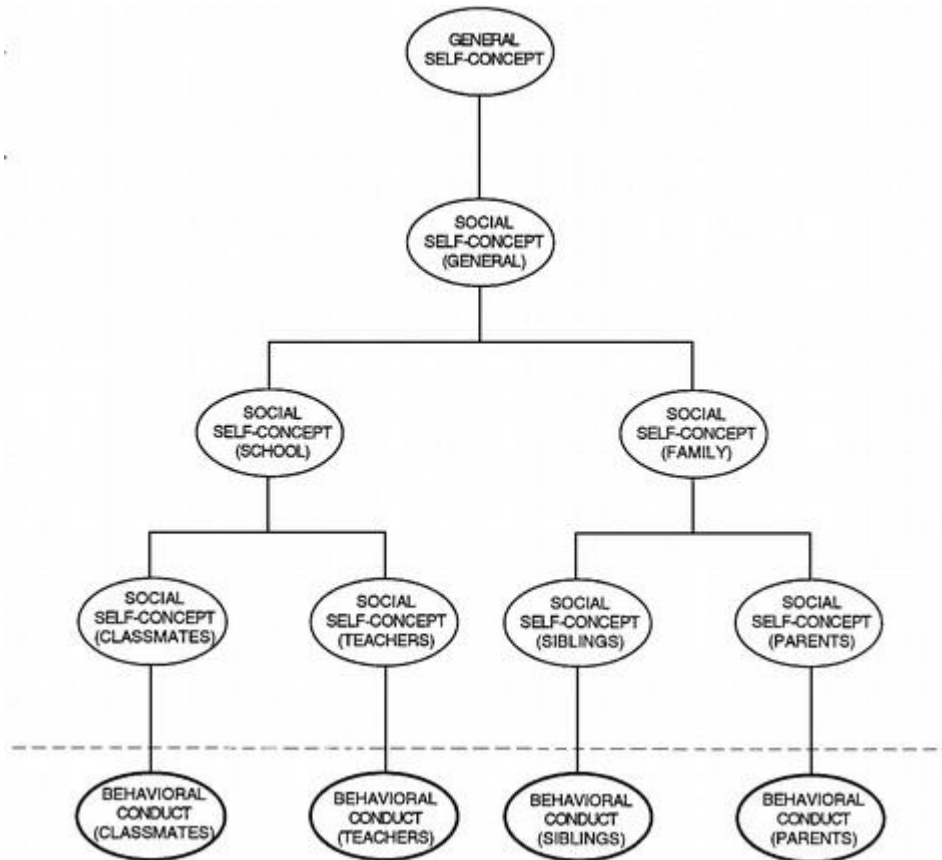
Figure 2.3. Hypothesized model of social self-concept. From "On the Structure of Social Self-Concept for Pre-, Early, and Late Adolescents: A Test of the Shavelson, Hubner, and Stanton (1976) Model" by B. M. Byrne, & R. J. Shavelson, 1996, *Journal of Personality and Social Psychology, 70,* p. 599-613. Copyright 1996 by the American Psychological Association. Reprinted by permission

Existing Research Related to Peer and Social Self-Concept

In order to identify the research related to peer self-concept that has already been conducted, a comprehensive literature search was conducted by the author (i.e., Google Scholar, Eric, PsychInfo, and Web of Science). Unfortunately, only a handful of studies were returned and just one actually addressed peer self-concept. As a follow-up, additional searches were conducted to include social self-concept, as well as, academic or

16

general self-concept research related to peer relationships. The results of these additional searches turned up only a few more publications, showing that additional research in this area is warranted.

Connolly and Konarski (1994) authored the only peer-reviewed study related to peer self-concept as a construct. Through the use of Harter's *Self-Concept Profile for Adolescents* (SCPA, 1988) the authors conducted a CFA and identified a 3-factor model of peer self-concept. They concluded that the resulting factor structure provided evidence that peer self-concept is multi-dimensional and consists of facets related to three distinct groups: (1) peers, (2) close friends, and (3) romantic relationships. The authors then used hierarchical multiple regression and correlational analysis to relate the three different factors of peer self-concept to self-report responses about students' experiences with peers, friends and romantic others. They concluded that since correlations were highest between corresponding groups (e.g., experiences with a large peer group will contribute to peer self-concept but not close friend or romantic relationship self-concept) that this served as evidence that peer self-concept is multi-dimensional and that differentiation occurs as students have different types of experiences with others. However, the study was cross-sectional, thus directionality could not be established, nor could it be proven that differentiation occurred in peer self-concept over time. While Connolly and Konarski's research worked to define the construct of peer self-concept, it did not attempt to show how this domain interfaces with academic self-concept domains. There was also no attempt to relate peer self-concept to any educational outcomes.

A similar study to that conducted by Connolly and Konarski was undertaken on the construct of social self-concept. Authors Lindner-Muller, John, and Arnold examined

the structure of social self-concept through the use of cross-lagged panel analysis in their 2012 research. The authors found that two facets of social self-concept could be clearly differentiated in their model, contact and empathy. Lindner-Muller and colleagues found these construct distinctions existed in children as young as second grade.

How peer factors relate to general self-concept. A number of the studies recovered through the literature search looked at general self-concept as an outcome of peer-related interventions such as: reciprocal peer tutoring (Fantuzzo, Davis, & Ginsburg, 1995), and peer-assisted learning (Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006). Other studies looked at different characteristics of peers and how these related to general self-concept. Peer academic reputation (Gest, Rulison, Davidson, & Welsh, 2008), same-sex and opposite-sex peer relationships (Hay & Ashman, 2003), isolated peer status (Lawler-Prince & Grymes, 1990), peer acceptance (Obiakor, Stile, & Muller, 1987), peer and self-ideals (Grymes & Lawler-Prince, 1993), and parent-adolescent relationship quality's impact on peer relations (Dekovic & Meeus, 1997) were all examined for their impact on general self-concept.

Social self-concept research. While outside of the focus of the current study, there were some pertinent results from the literature search found within the broader domain of social self-concept. One of these studies found that social self-concept (operationalized by the *SDQ-III* scales of Same-Sex and Opposite-Sex Relations) increased as a result of a summer program for gifted students (Rinn, 2006). Others found differences in social self-concept between hearing and hearing-impaired students and between students diagnosed with learning disabilities (Schmidt & Cagran, 2008; Zahra,

18

Arif & Yousuf, 2010). These studies examined social self-concept specifically, instead of looking at general self-concept as a result of social groupings or interventions.

The studies detailed above, investigated social self-concept or the impact of peer-related interventions on general self-concept and found that self-concept could be increased based on interpersonal relationships. For example, general self-concept increased in situations where students had positive relationships with parents and when they were accepted by others in the school.

Peers Influence on Achievement Motivation

There is a large body of research on the impact that peers have on each other with respect to achievement and motivation. While the details are beyond the scope of this review, a brief summary of what is known about peer influences in the classroom is beneficial.

Peer relationships have the potential to affect the academic achievement of students when peers serve as comparative referents. Students not only learn effective ways to complete tasks from peers but simply observing peers' performances can change students' ability beliefs and expectancies (Bandura,1986, 1989, 1997, 2012; Bouchey & Harter, 2005; Fan, 2011; Freiberger, Steinmayr, & Spinath, 2012; Schunk, 1987, 1995; Schunk & Meece, 2006; Schunk & Zimmerman, 1997; Wentzel, 1993, 2009). Peers also impact school performance by clarifying instructions, providing assistance, and sharing effective learning strategies (Lim & Kim, 2011; Molloy et al., 2011; Wentzel, 1991a, 1993, 2003, 2009). Peers encourage others to try new skills and follow classroom rules (Bandura, 1986, 1997; Hamm & Faircloth, 2005; Han, Hu, Liu, Jia, & Adey, 2013; Lim & Kim, 2011; Oberle & Schonert-Reichl, 2013; Schunk, 1995; Véronneau & Dishion,

19

2011; Wentzel, 2003, 2009; Wentzel & Watkins, 2002**).** Peers self-select into groupings based on the ability level of the group and then adopt academic values from the group which become part of the their identity (Darensbourg & Blake, 2014; Dowson & McInerney, 2001; Eccles, 2005; Hamm, Schmid, Farmer, & Locke, 2011; Kindermann, 1993, 2007; Kindermann & Skinner, 2009; Masland & Lease, 2013; Osterman, 2001; Ryan, 2001; Steinberg et al., 1996; Van Ryzin, Gravely, & Roseth, 2009; Véronneau & Dishion, 2011; Wentzel, 1993, 2005, 2009; Wentzel & Asher, 1995). Peers impact student motivation through caring peer relationships where students strive to perform academically when their close friends value achievement (Altermatt, 2011; Baker, 2006; Deci & Ryan, 2008; Dowson & McInerney, 2001; Fan, 2011; Finn, 1989; Goodenow, 1993; Goodenow & Grady, 1993; Connell & Wellborn, 1991; Diseth & Samdal, 2014; Han et al., 2013; Martin & Dowson, 2009; Molloy et al., 2011; Oberle & Schonert-Reichl, 2013; Osterman, 2001; Ryan & Deci 2000; Skinner, Wellborn, & Connell, 1990; Van Ryzin et al., 2009; Wentzel, 1996, 2003, 2005, 2009; Wentzel et al., 2012; Wentzel & Caldwell, 1997; Wentzel & Watkins, 2002; Wentzel & Wigfield, 1998). Peer expectations, values, and goals are adopted from others through a desire to please and fit in, especially as children near adolescence (Cocks & Watt, 2004; Darensbourg & Blake, 2014; Fan, 2011; Liem, Lau, & Nie, 2008; Hamm et al., 2011; Masland & Lease, 2013; Ryan, 2001; Wigfield, Tonks, & Klauda, 2009). Peers' levels of social competence impact their engagement and achievement (Wentzel, 1989, 1991a, 1991b, 1993, 1996, 1997, 1998, 2009; Wentzel et al., 2004, 2010; Wentzel & Watkins, 2002). As a whole, this body of research has established that peer-related factors can influence student beliefs, values, expectations, motivation and perceptions about their academic ability. It

20

is surprising, given all the research conducted on peer influences, that research involving the peer self-concept domain remains lacking.

Self-Concept as a Theoretical Framework.

The present study will use multi-dimensional self-concept as a theoretical framework for examining the association between perceptions of interest and competence in academic areas and peer relationships. As part of the study it will be possible to examine whether profiles of peer and academic self-concept change over time and if these changes are accompanied by changes in academic achievement. This information will add to current research by modeling complex interactions between student perceptions of their interest and competency in academic and social areas. The use of multi-dimensional self-concept as a theoretical framework is appropriate in that it recognizes not only students' academic perceptions but also their perceptions of relationships with peers in the school.

There are other benefits to using the Marsh/Shavelson model (1985) as a theoretical framework for the current study. This model was instrumental in focusing research on the development of measurement instruments tailored to assessing the self-concept domains. Herbert Marsh and Susan Harter both developed a number of self-report instruments for use in assessing self-concept in individuals across the developmental spectrum, (*Self-Description Questionnaire-SDQ*; Marsh 1990d, 1990e, 1990f; *SPPC-Self-Perception Profile for Children*; Harter, 1985; *The Perceived Competence Scale for Children*; Harter, 1982). Prior to the development of these scales there were just a couple of self-concept instruments available and these had not been

21

subjected to reliability and validity studies (Self-Concept of Ability, Brookover, Thomas, & Paterson, 1964; Affective Perception Inventory, Soares & Soares, 1979).

The family of *Self-Description Questionnaires (SDQ's)* have firmly established reliability and validity and have been the measurement tools used in many studies on the construct of self-concept (Berndt & Bergy, 1996; Marsh, 1984; Marsh, Barnes, Carins, & Tidman, 1984; Marsh & O'Neill, 1984; Marsh et al., 1985; Marsh, Relich, & Smith, 1983; Marsh, Smith et al., 1983) including the research utilizing NCES's ECLS-K data files (Kim, Schwartz, & Cappella, 2014; Froiland & Oros, 2014; Niehaus & Adelson, 2013, 2014). The *Self-Description Questionnaire-I (SDQ-I)* was used as the primary measurement instrument for this study (Marsh 1988, 1990d, 1990e, 1990f). It should be noted that specific details about the technical qualities of the *SDQ-I* scale can be found in Chapter III.

Academic and Social Self-Concept Research Relevant to the Current Study

Kathyrn Wentzel and Allan Wigfield found that different types of goals, both academic and social, interacted and jointly influenced achievement motivation (see Wentzel & Wigfield, 1998). They concluded that a single goal framework was inadequate for describing student motivation in schools. Others have supported their findings and have determined that socially-derived goals (e.g., receiving acceptance from others, being a valuable part of a peer group), work in concert with academic goals to impact both students' feelings of social competence and academic performance (Dowson & McInerney, 2001; Hijzen, Boekaerts, & Vedder, 2006; Wentzel 1989, 1991a, 1991b, 1996, 1998; Wentzel, Baker, & Russell, 2012; Wentzel, Barry, & Caldwell, 2004; Wentzel, Battle, Russell, & Looney, 2010; Wentzel & Wigfield, 2009; Wentzel, et al.,

22

2012). However, few studies of academic self-concept also consider the influence of social or peer self-concept. The aim of the current study is to investigate these relationships. To this end, some of the research that has been conducted with academic self-concept is relevant and is detailed in the following sections.

Development of self-concept in children. When self-report responses on the *SDQ* from students at different ages became available, tests of construct validity and invariance at different ages began to emerge (Marsh, Parker, et al., 1983; Marsh, Relich, et al., 1983; Marsh, Smith, et al., 1983; Marsh & Shavelson, 1985; Shavelson & Bolus, 1982). Conflicting results were found when looking at differentiation of academic self-concept by developmental age. Shavelson and Bolus (1982) found that over a six-month time lag self-concept measures remained stable at both global and domain-specific levels of the hierarchy. These finding resulted from a study of a study of 130 7th and 8th grade students. However, their results were not in agreement with research from Marsh, Smith et al. (1983) and Marsh, Parker et al., (1983), who found that specific academic self-concept dimensions did change, even in preadolescents. Marsh and colleagues concluded that only global self-concept was stable. Both of these studies were conducted on students in the fifth grade. Only the Marsh, Smith et al., study used repeated testing with a similar six-month time lag as used with the Shavelson and Bolus study. This dissertation will attempt to add to this research base in hopes of finding more consistency in results through the use of a two-year longitudinal assessment of academic and peer self-concept on the same students.

Across many studies, research results have agreed that self-concept is more global at early ages (kindergarten through second grade) and becomes more differentiated as

23

children mature (Marsh, 1990a; Marsh & Hocevar, 1985; Marsh & Shavelson, 1985). The hierarchical structure of academic self-concept appears to become weaker with age, as the influence from global self-concept weakens and there is more influence from domain specific areas (Marsh, Parker, et al., 1983; Marsh, Relich, et al., 1983; Marsh, Smith, et al., 1983). Research has indicated that the differentiation of self-concept into more subject-specific domains starts to occur at grade five. However, there has not been research conducted on students between grades two and grade five. The current study will contribute to knowledge about how and when development of more differentiated domains of self-concept occurs.

Byrne and Shavelson (1996) applied correlational multi-trait, multi-method (MTMM) analyses to study the structure of social self-concept. Through these analyses they found that social self-concept was also multidimensional and hierarchical, becoming more differentiated as children matured. The Byrne and Shavelson study was the only study investigating the hierarchical structure of social self-concept over time, beyond research conducted for social self-concept scale development (see Berndt & Bergy, 1996). Then in 2012, the Lindner-Muller et al., study examined changes in the factor structure of social self-concept finding that this broad domain could be broken down into subdomains of empathy and contact. This study involved the use of a German self-report instrument with German students. It is not clear whether these results would transfer to another culture.

A similar line of research involved looking at bias in student self-concept with maturation. Marsh and Craven (1997) found that as children aged they evidenced a more realistic self-concept. At first, children display globally favorable self-concepts. As they

24

learn to self-evaluate more realistically, their self-concepts become more differentiated and the correlations between different domains dramatically decreases. Byrne and Shavelson (1996) also found that younger children overinflate their self-concept, in the same way that they do their academic competency. That is, young children feel good about themselves and their performances in every aspect of their lives (Berndt & Bergy, 1996; Harter, 1983, 1988; Lindner-Muller et al., 2012; Marsh et al., 1998; Wigfield & Eccles, 2002). Researchers have also found that students' achievement expectations become more accurate as they mature (Schunk, 1995; Schunk & Meece, 2006; Schunk & Pajares, 2009; Schunk & Zimmerman, 1997).

The current investigation will fill a void in the research base by studying the latent class structure of both academic and peer self-concept across a two-year period (from grades three to grade five) with the use of a more appropriate person-oriented methodology. This design will afford examination of the relationship between student's peer and academic self-concepts and how these change as students mature.

Reciprocal effects model (REM) and causal ordering. Another area of research relevant to the current study is one addressing the reciprocal relationship between self-concept perceptions and achievement performance. Issues of causal ordering between self-concept beliefs and academic achievement behaviors have been researched by many, even before the Shavelson et al., model (1976). The classic discussion from Calsyn and Kenny (1977) noted two different theoretical postulates: the self-enhancement model and the skill development model. The self-enhancement model stated that evaluations of significant others shape our self-concept of academic ability. It is this self-concept that then determines our achievement performance. Implications from this model imply that

25

to change achievement, one's self-concept would need to be bolstered. In contrast, the skill development model argues that self-concept is a direct consequence of one's academic ability and achievement. This model implies that to improve achievement performance, academic skills need to be improved. These improved abilities would heighten confidence and self-concept, which would further improve academic performance. The differences between these two theories has fueled flurries of research, continuing to the current day (Marsh, 1990b; Marsh, Byrne, & Yeung, 1999; Marsh & Craven, 2006; Marsh, Hau et al., 2002; Marsh & Martin, 2011; Marsh & O'Mara, 2008; Marsh et al., 2005; Marsh & Yeung, 1997, 1998; William & Williams, 2010).

Marsh conducted an early study, using panel analysis, aimed at finding the relationship between academic self-concept and GPA (1990b). The results of this study indicated that neither academic self-concept nor achievement was the causative factor. Instead he found that they were mutually reinforcing and reciprocally related. As a result he termed this cyclical effect the reciprocal effects model (REM). Guay, Marsh, and Boivin (2003) contributed to the refinement of the theory when they tested the reciprocal effects model with three different age cohorts and found that reciprocal effects were invariant. Their results debunked the original claim that the prevalence of skill development or self-enhancement models vary with the age of the student. Guay and colleagues found that the key to the relationship between academic self-concept and achievement is that there must be a match in domain specificity between the achievement measure and the self-concept indicator.

One of the largest efforts to test the generalizability of the reciprocal effects model was actually conducted within the construct of self-efficacy. Bandura noted the

26

same cyclical effects between self-evaluations and performance and termed this "reciprocal determinism" in his social cognitive theory (1986). To determine how widespread the reciprocal effects were, Williams and Williams (2010) tested the relationship between self-efficacy and performance in math with an SEM feedback loop model. Their results showed that reciprocal determinism was confirmed in 30 out of 33 different countries, making it a consistent phenomenon across cultures.

Person-Oriented Perspective and Research Methodology

One of the limitations inherent in much of the research that has been conducted on motivation is that studies have largely been variable-and not person-oriented. Instead of looking for changes in individual cognitive abilities, behaviors, and other characteristics over time (intraindividual), variable-oriented research focuses on either making comparisons between groups (e.g., ANOVA) or finding inter-individual relationships between predictors and dependent variables (e.g., regression, correlation).

Variable-oriented approaches focus on the relationships between variables (i.e., group differences, prediction of outcomes, inter-correlations between measures) that are largely assumed to be the same across all unique individuals. To obtain a clear understanding of the unique social and environmental factors impacting the development of academic achievement, it is necessary to follow an interactionist research focus. Instead of assuming that relationships between variables impact all individuals in the same way, an ecological fallacy, research should focus on the complex, more individualistic relationships within individual motivational systems. To achieve this level of in-depth investigation, person-oriented research is necessary. Person-oriented research looks at sorting and classifying individuals instead of sorting and classifying variables to

27

explain constructs. People are grouped into classes based on characteristics that they have in common with others. These characteristics distinguish those within the group from people outside the group. In essence, person-oriented methodologies minimize intragroup (within-group) differences and maximize intergroup (between-group) differences.

For example, when looking at student learning, there is a group of students who excel when math problems are presented with hands-on manipulative tools. Then there are others who excel by memorizing formulas. Both groups could have the same test scores, but approach the learning task in very different ways. Looking at grouping students by test scores assumes that they all learn in the same manner, simply because they have the same final score. In fact, there are very large differences in the learning strategies and processes that people with the same test score display. Uncovering these underlying differences is instructive in understanding the complexities of learning and motivation.

Besides failing to acknowledge the importance of complex interactions within and between individuals, the variable-oriented approach fails to recognize the significant relationship between the parts and the larger whole. Higher order interactions of behaviors and processes are simplified or wholly disregarded. Instead the variable-oriented approach primarily focuses on linear relationships across people. Even with model-based techniques (i.e., structural equation modeling) researchers focus on reproducing lower level relationships (mean structures and variance-covariance matrices) instead of attending to higher-order complex interactions (Bergman & Trost, 2006; Sterba & Bauer, 2010).

In contrast, person- oriented research looks at holistic and interactionist development specific to the individual (Bergman & Magnusson, 1997; Bergman & Trost, 2006). These methods examine the complex relationships within and/or between the person's behavior, goals, values, and the environmental context as well as how these relationships change over time (e.g., latent class growth analysis, cluster analyses, latent transition, latent Markov models, and pattern analysis approaches). Not only do these complex interactions involve the environment, they also involve other people with their own complexities. When studying human development, a holistic focus with supporting methodology should guide both research activities and interpretation of results. Multifaceted person-oriented research not only aids in understanding complex patterns of behavior but also provides a common terminology for communicating and furthering research (Bergman, 1998; Bergman & Magnusson, 1997).

Assumptions of the person-oriented approach. The centerpiece of the person-oriented approach is the state of the system and how it changes over time. This approach adheres to five basic assumptions: (1) the state is unique to the individual, (2) the state is a complex interaction of many factors, (3) states are characterized by having a meaningful structure that shows growth and differences between individuals, (4) processes occur as patterns in an orderly fashion, and (5) there are a small number of actual observed patterns. The fact that there are only a relatively small number of behavior patterns that are actually utilized, leads to stability in the system (Bergman & Magnusson, 1997). These assumptions apply both between different individuals and within the same individual over time.

Person-oriented approaches to research look to find differences between people in a population and find meaningful groups within this population that contain individuals who are similar. These similarities are discovered from their responses to survey or test items, or other measures of observed behavior. Critics of the person-oriented approach focus on the fact that the data being used still consists of variables. However, in person-oriented approaches variables serve only as building blocks for creating profiles and the variables have no significant meaning by themselves without considering how they interface with other variables. Thus, several focal variables are used all together at the same time and the resulting profile is what is used in ensuing analyses (Bergman & Magnusson, 1997; Bergman & Trost, 2006).

For this research study, the following definitions for person-and variable-oriented approaches will apply. Variable-oriented approaches will include methods which are based on relations between variables and the use of linear statistical models (e.g., ANOVA, MANOVA, regression techniques, SEM). The use of more sophisticated longitudinal models like panel, time-series, cross-lag, and analysis of difference scores are also variable-oriented since they view individuals as interchangeable units, ignoring their unique multivariate makeup. Person-oriented approaches differ in that they examine all variables or measures as interrelated components forming a profile, which is the unit of analysis. Person-oriented methodologies include cluster, latent class, latent transition, and latent profile analyses (Bergman & Magnusson, 1997; Bergman & Trost, 2006). These methods consider characteristics of the individual when creating profiles, and result in a distinguishable "location" for each individual within the profile.

30

Different Person-Oriented Research Approaches

There are a number of different person-oriented methods for determining unique individual profiles. Two different approaches, cluster analysis and latent class cluster analysis, have been commonly used to examine intraindividual changes and variations within dynamic environmental and social systems. Both of these approaches concentrate on categorizing or classifying individuals based on a profile of multivariate characteristics in an attempt to better understand commonalities and differentiation in personal characteristics.

Cluster analysis (CA). This data reduction technique is similar to factor analysis and focuses on creating smaller homogeneous groups from a larger data set of individual responses (Aldenderfer & Blashfield, 1984; DiStefano, 2012; DiStefano & Mîndrilă, 2013). Cases within a cluster are very similar with respect to a specific characteristic and are different from the cases that are outside of the cluster. Multivariate response profiles are used to identify like cases, which are then used to create the clusters. Proximity or distance measures (Euclidean and Mahalanobis' distances) are used to evaluate how dissimilar two cases are from each other across a set of variables.

A number of algorithms have been developed to use dissimilarity information to form clusters, with final results dependent on the procedures used (see DiStefano, 2012 for a discussion of these techniques). With cluster analysis the challenge lies in finding the appropriate number of homogenous groups for the data and then interpreting the underlying construct. Both demographic characteristics of grouped cases and cluster centroid information aid in interpretation of clusters. Internal and external validation procedures are undertaken to make sure that the resulting clusters are not an artifact of

31

the specific data set. However, even with these procedures in place, identification and validation of clusters is largely a subjective judgment since there are no statistical significance tests for the adequacy of cluster solutions.

Since variables from any type of measurement scale can be used for forming clusters, it is critical that the researcher selects variables that are relevant and well represent the research questions. In addition, the variables used must be independent and uncorrelated. One of the major drawbacks of cluster analysis is that since the variables are observed, they are tied to their scale and need to be on same metric level. Variables can be standardized to allow comparisons but this can mask the very uniqueness that is most useful and relevant in creating groupings. In fact, researchers can obtain wholly different results when using standardized instead of raw, scale-dependent values for variables (DiStefano, 2012; DiStefano & Mîndrilă, 2013). Another drawback is that the clustering distance measures are influenced by differences in variability across the relevant variables, as well as the differences in height between profiles. Variables with larger variance can exert undue influence on the final distance calculations.

Latent class cluster analyses (LCCA). Another statistical method that identifies subsets of individuals with similar responses to multivariate data elements is latent class cluster analysis. For examining developmental, intraindividual change, both LCCA and CA can be used either alone or in sequence as person-oriented research approaches (Bergman et al., 2003).

As with cluster analysis the number of distinct groups underlying the data is unknown. However with LCCA, groupings are based on unobserved, latent categories, not the observed measurement variables. LCCA assumes that the latent class variable is

32

free of measurement error and is the only element that causes the observed variables to be related to each other. Therefore, when latent class membership is removed, randomness is all that remains (Collins & Lanza, 2010). This is the essence of the local independence assumption (see "Assumptions" below).

LCCA is similar conceptually to factor analysis in that it undertakes data reduction through focusing on larger unobserved constructs. Factor analysis examines relationships between continuous, normally-distributed variables through correlations. In contrast, LCCA looks at structural relationships by examining qualitative differences between cases (taxonomies) in terms of categorical variables.

Assumptions of LCCA. The assumptions for CA and LCCA are similar, the variables must be independent and uncorrelated. Beyond this common assumption, LCCA assumes each case has a specific probability distribution for membership in the latent class. Students are classified in classes they are most likely to be affiliated with, yet the classification is based on probability, and thus, is not fully certain (Collins & Lanza, 2010; DiStefano, 2012; DiStefano & Mîndrilă, 2013). Finally, LCCA assumes local independence (i.e., that the indicators within a class are independent of each other and the latent class explains all the observed relationships between the indicators) and that the variables within a latent class are all uncorrelated.

Latent Class Cluster Analysis includes both Latent Class (LCA) and Latent Profile (LPA) analyses. The distinction is that latent class analysis involves only categorical observed measures while LPA deals with continuous data. These models are flexible and accommodate a mixture of variables from different measurement scales, without requiring standardization.

33

Benefits of LCA. Latent class analyses categorize cases based on a statistical model which describes the population. This model groups individuals based on their similarity with respect to a latent variable. With LCA the measurement error inherent in observed variables is partitioned into residual variance.

LCA models use a top down approach (DiStefano & Mîndrilă, 2013) where the expected distribution of the data is first described and then probabilities that cases are members of the different latent classes are determined. LCA, does not rely on the use of distance measures between cases. Cluster analysis, in contrast, uses a bottom up approach that looks for similarities in cases (through distance measures) and then tries to find corresponding explanations for the data.

Additional benefits from using LCA are that these model-based clustering methods support the use of statistical significance testing (Chi-square referenced) and fit indices for finding the optimal solution. Finally, LCA also allows the inclusion of covariates when predicting individual's latent class membership, supports modeling change over time in the latent class structure (latent transition analysis), and can be used to predict related outcomes or additional similarities among cases within the groupings.

Model Restrictions. Latent class methods allow flexibility in deciding which parameters to estimate. This makes it possible to test a large number of different solutions and to allow solutions that have unequal variances between indicators or groups. Being able to impose restrictions on the number of parameters to be estimated also facilitates model identification by ensuring there is sufficient available information. Since LCA models have the flexibility to free restrictions on parameter equalities, they can achieve better fit with more parsimonious models. Because of this LCA avoids a common

problem found with CA, that of overestimating the number of clusters underlying the data (DiStefano, 2012; DiStefano & Kamphaus, 2006).

Model Estimation. To determine the explanatory value of the model for the given data, a number of model parameters must be estimated (e.g., centroids, variances, covariances). Estimations are conducted through either maximum likelihood (most common) or maximum posterior methods. During estimation cases are classified into groups based on the likelihood of belonging to the group given score values for the case across a set of variables. The process used for both of these methods is iterative where an initial value for the parameters is compared to a model estimate. Re-estimation of values for the parameters continues until the best estimate for the given data is obtained. The criteria for stopping the iterative estimation process is set by either total number of iterations or by reaching a set convergence criteria (Collins & Lanza, 2010). This final optimal set of parameters has the largest likelihood of reproducing the actual observed data (Collins & Lanza, 2010; DiStefano, 2012; DiStefano & Kamphaus, 2006). As an additional check of the superiority of these values, estimation procedures should be repeated with alternative estimation start values. This will determine the stability and internal validity of the final parameter values for the latent class (Collins & Lanza, 2010).

Model Selection. The best way to find the number of distinct classes that underlie the observed data is to test a number of alternative models and find which best fits the data. This process starts with the extraction of only one class, the independence model. This baseline model assumes that there is no relationship between variables and simply reports the set of observed means (Nylund, 2007; Nylund, Bellmore, Nishina, & Graham, 2007). After estimating this model, additional class are added one at a time and the

35

parameters are re-estimated and fit is re-evaluated. This process continues until the model fails to converge. The optimal model is one that is the most parsimonious, with the most favorable fit statistics, and one that can be interpreted in line with existing research.

There are more methods for evaluating the adequacy of LCA solutions than there are with CA and solutions can be assessed even when the model specifications or groups are different. The relative fit indices for LCA include both the Akaike Information Criteria (AIC; Akaike, 1987) and the Bayesian Information Criteria (BIC; Schwartz, 1978). The BIC measure has also been modified to consider relative sample size in the Sample-Size Adjusted BIC (SABIC; Sclore, 1987). All of these indices evaluate fit based on the parsimony of the model. The more parsimonious model is one that has acceptable fit with a smaller number of latent classes. Models with lower AIC, BIC, and SABIC values are better fitting for the data. These relative fit indices support significance testing of the changes in fit resulting from adding additional classes to the model.

Another class of fit indicators for LCA looks at the certainty with which cases are allocated to different latent classes. After cases are assigned to classes, posterior probabilities are calculated based on the model and actual observed scores. A classification table is developed to show how well the model performed at accurately categorizing cases within classes. The diagonal of the table contains the probabilities for making a correct classification. The off-diagonals show misclassifications. A good fitting model will have high values on the diagonal and small values on the off-diagonals.

A measure similar to posterior probabilities is that of entropy. Here the model is again evaluated for how well it accurately classifies cases. However, in this case, the measure tells how randomly classifications have occurred. Entropy values range from .00

36

to 1.00 and are a single value reflection of the results in the classification table of posterior probabilities. Entropy values closer to 1.00 indicate more accurate and certain classifications.

Assessing the overall fit of different solutions should include multiple criteria to ensure that the decision is not based on the nuances inherent in the particular sample of data being used. To this end, additional factors beyond the relative fit and certainty measures should be considered. These additional factors include: (1) descriptive information about the cases falling within groups, (2) class centroid values, (3) matches of latent classes and membership groups to existing theory, and (4) the usefulness of latent classifications for practical applications (DiStefano & Kamphaus, 2006; Nylund, Bellmore et al., 2007). Ultimately, the superior model will produce parameters that should generalize across time and situations and also, will be simple and easy to understand (Bergman & Trost, 2006).

The process for consulting fit indices and descriptive information detailed above should be conducted on individual models after each extraction and estimation. After settling on the best final model additional validation should be conducted to make sure that different classes represent distinctions in individual characteristics or behavioral constructs. One simple method is conducting ANOVA's to determine whether there are group differences between the classes on observed measures that were not used to classify cases (DiStefano, 2012; DiStefano & Mîndrilă, 2013).

37

Longitudinal Cluster and Latent Class Cluster Analyses

One of the goals of this research is to determine how achievement motivation characteristics change as students mature. Given that person-oriented analyses are more appropriate for studying developmental changes than variable-oriented approaches (see Bergman & Magnusson, 1997; Bergman et al., 2003), this study will apply person-oriented research approaches to the study of achievement motivation in the middle and later years of elementary school. In addition to using person–oriented analyses, this investigation will follow a longitudinal, multi-year design instead of being cross-sectional in nature. This methodology will make it possible to determine whether motivational characteristics are stable or whether they change over time. While both cluster analytic and latent class techniques support longitudinal research methodology, the focus will be on finding the best CA or LCA technique to address this study's research questions.

Longitudinal CA. There are a number of different approaches to longitudinal cluster analysis. One of the simplest methods is to conduct a descriptive analysis of the characteristics of individuals in each cluster, at every time period. Examining group centroids and differences in variability within groups provides useful information about the stability of emergent groupings. Data can also be examined to determine whether the number of resulting groups differs between time periods, across ages, or between different cultures (DiStefano, 2012; DiStefano, Kamphaus, & Mîndrilă, 2010).

Cluster analysis results can also be physically linked across different time periods. The independent cross-sectional cluster analyses at each time period are first examined for extreme profiles, or outliers cases. These cases are removed and then hierarchical algorithms like Ward's method are applied to each separate grouping (Bergman et al.,

2003; DiStefano, 2012). Differences are allowed to occur in the number of emerging clusters at each time period. This method is known as linking of clusters after removal of residual (LICUR). The resulting data is linked across adjacent time periods through cross-tabulation of resulting clusters and their defining centroids. Cross-tab results show the prevalence of membership in different groups across the time periods. A benefit of this method is that frequencies of group membership at each time can be tested for being significantly different from chance.

Another longitudinal analysis that can be applied to clusters is referred to as "I-states as objects analysis" (ISOA). Bergman and El-Khouri (1999) developed this approach and describe it as follows. They define a person's profile of characteristics at a specific point in time as their "I-state." This approach allows the researcher to examine how an individual's profile changes over time, as well as how similar the profiles between different individuals are, either at the same time or at different measurement periods. Instead of using separate cluster analyses and linking them through cross-tabulations, ISOA conducts the comparisons within one analysis (see Bergman et. al, 2003; DiStefano, 2012). A dissimilarity index is developed from data points across all variables and all individuals. This index is then used to generate clusters which are examined for stability and for change over time.

Longitudinal LCA. When examining changes in latent class membership over two distinct time periods, a useful procedure is latent transition analysis (LTA). LTA focuses on examining whether individuals change class membership from Time A to Time B, or remain stable (Collins & Lanza, 2010; DiStefano, 2012; Lanza & Collins, 2008; Nylund-Gibson, Grimm, Quirk, & Furlong, 2014). The model is also used to

39

predict what class an individual will belong to at a later time, based on current class membership. Latent transition procedures start by using LCA to determine the number of underlying classes represented in the data. Different latent models are evaluated through the use of a combination of criteria. Models are examined for statistical fit, classification accuracy, interpretability, and match to existing theory. The best model is selected not only based on traditional fit measures, but also based on predictive accuracy. Models that have higher posterior probability and entropy values indicate more certainty and more accuracy in predicting class membership at a later time. The clarity of the distinctions between groups and interpretability of centroid values (i.e., simple structure of the model) also impact on the selection of the best model. Finally, correspondence with existing theory and the utility of using the latent class designations for interventions and program placements are also considered. LTA analyses are very flexible and can compare different models, even when they are not nested. Researchers can relax restrictions on error and covariance structures and can incorporate covariates or outcome measures into the design.

When there are more than two time points involved in the longitudinal analysis of latent classes, latent class growth modeling (LCGM) is often used. This technique is used to identify groups of people with a similar pattern of change over time. These individuals are said to be following the same trajectory (Andruff, Carraro, Thompson, Gaudreau, & Louvet, 2009; Hancock, Harring, & Lawrence, 2013; Nagin, 1999, 2005). Intercepts and slopes are estimated for each discrete trajectory in the model. LCGM fixes the intercepts and slopes to be equal for all individuals within a distinct trajectory. This allows flexibility and degrees of freedom for estimating non-linear relationships (quadratic,

cubic) across three or more time periods. Trajectories can then be modeled as linear (stable or steadily increasing or decreasing), quadratic (stable, increasing, or decreasing until a point when there is a change in direction or rate of change), or cubic (trajectories with two changes in direction or rate of change). LCGM analyses require the researcher to specify the number of different trajectories represented in the data. Different models are compared based on fit statistics, posterior probabilities and existing theory. The most parsimonious model with the best fit to the data is preferred.

Appropriate Methodology for Addressing Research Questions

As noted previously, when examining developmental, intraindividual change, either LCA/LPA or CA can be used as person-oriented research approaches. The dataset used for this study has student self-report survey responses from two different time points, supporting a longitudinal analysis. The observed survey measures are identical at both grade levels. Since the items are on the same metric level, this solves the problem of CA being scale-dependent. However, using CA would present a problem in that the clustering distance measures will be influenced by differences in variability across variables, and across the two different grade levels. If there are variables with large variances they could exert undue influence on the final distance calculations.

The benefits from using LCA/LPA are the ability to use statistical significance testing and fit indices for finding the optimal solution. Latent class methods also allow flexibility in deciding which parameters to estimate, making it possible to test many different solutions. Imposing restrictions on the number of parameters to be estimated also facilitates model identification by ensuring there is sufficient available information. Since LCA/LPA models have the flexibility to free restrictions on parameter equalities,

41

they can achieve better fit with more parsimonious models. LTA analyses are also very flexible and can compare different models, even when they are not nested. Equally important is the fact that covariates and outcome measures can be incorporated into the design. This was an important benefit since one of the goals of this study was to examine the relationship between latent classes and academic achievement measures. For these reasons LCA/LPA and LTA were deemed to be the best methodologies for addressing the goals of the study and the corresponding research questions, which are detailed below.

Study Objectives and Research Questions

To address the shortcomings of existing research methodology being used to study self-concept, the current research study was conducted. The main objectives of the current study were to:

- Identify the latent class structure underlying student responses from four of the scales from the *SDQ-I* self-report instrument, which was used as part of the national ECLS-K longitudinal study.

- Compare the resulting latent class structure to multi-dimensional self-concept theory domains to see how peer and academic self-concepts relate to each other within class structures (Byrne, & Shavelson, 1996; Shavelson et al., 1976; Marsh et al., 1988).

- Conduct a longitudinal analysis to examine the change in resulting latent class structure as students mature from grade 3 to grade 5,

- Examine the relationship between students' latent class status and academic performance in both grades three and five (proximal outcomes), and

42

- Examine the predictive relationship between latent class status at grade three and academic achievement at grade five (distal outcomes).

By utilizing person-oriented research with the theoretical framework of multi-dimensional self-concept, this study will address the following research questions.

Research Questions

1. Do student responses from the self-report scales used with the ECLS-K (*SDQ-I*, grade 3 and grade 5) form latent clusters that relate to the multi-dimensional structure of the Marsh/Shavelson conceptual model (1985) of self-concept? In other words, are there resulting class structures where some students have higher reading self-concept, while others have higher math self-concept? Is there a resulting class where some students have higher self-concept related to peer relations and lower self-concept in the academic areas?

2. Is there a significant association between the outcomes of math and reading achievement and the resulting latent classes at each grade level?

3. Do the latent class profiles identified in grade 3 change or remain stable over time (i.e., from grade 3 to grade 5)?

4. Do latent class statuses at grade 3 predict achievement performance in math and reading at Grade 5?

43

Chapter III

Method

This chapter gives a detailed description of the data and methodology utilized for investigating the research questions and addressing the study goals. The dataset, participant characteristics, and sampling frame are defined, followed by an explanation of the instruments and methodology used to conduct the person-oriented, latent transition analysis.

Dataset

The present study utilized data which has been publicly released by the National Center for Education Statistics (NCES) for use in conducting research about child development, school readiness, and early school experiences. The specific data used (the Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS-K) data file) was collected as part of NCES's Early Childhood Longitudinal Studies Program.

The ECLS-K was initiated in the 1998-99 school year and followed a nationally representative sample of children from kindergarten through the eighth grade. This was the first large-scale, nationally representative study focusing on a cohort of children over a lengthy period of development. The design of the ECLS-K provided wide-ranging, reliable data describing children's experiences in elementary and middle school, and related these early experiences to later development and learning in school.

The ECLS-K data were collected directly from multiple sources: students, their parents, teachers, and school administrators. Information was gathered on children's

44

home (e.g., educational activities), school, classroom environments (e.g., curriculum materials), as well as teacher qualifications, to fully assess children's cognitive, social, emotional, and physical development. Direct participant responses were gathered from objective cognitive assessments, social and behavioral self-report surveys, interviews and school records. The breadth and depth of information gathered through the ECLS-K data analysis design has allowed educators and researchers to examine how home, school, and environmental factors and programs affect child development and students' performance in school.

Sampling procedures

The design of this nationally representative study not only allowed detailed cross-sectional analyses but included rigorous longitudinal linkages by following a random subsample of students who transferred from their base year schools in subsequent rounds of data collection. The total sample consisted of 21,260 kindergartners. When these kindergarteners were in the spring of first grade, students who were not in kindergarten in the United States during the previous year were added to the sample. Supplementing the sample with additional students was done to make it more representative of the demographics of kindergarteners in the nation. As a result of adding these students, researcher were able to obtain estimates relative to first grade students. There were no additional students added to the sample in either the third or the fifth grade data collection rounds (Tourangeau, Lê, Nord, Sorongon, & Chapman, 2009).

The ECLS-K data were collected through a complex, multi-stage probability sampling designed to represent 1994 population estimates. The first stage (Primary Sampling Unit) was based on counties within geographic regions. Within these regions

45

the second stage units were public and private schools. The final stage consisted of students within schools, stratified by gender and ethnicity. Sampling strategies were developed to compensate for non-response and differential probabilities of being selected for the different waves in the study design. To accomplish this, students from private schools, Asians, and Pacific Islanders were oversampled. The resulting design followed 21,357 children from approximately 1,000 kindergarten programs across the nation, representing different types of schools, ethnicity, gender, and SES status. Comparisons of the weighted population of students from the study with the weighted population of eighth graders from the 2006 Population Survey of the Bureau of the Census shows that the ECLS-K represented about 80% of all of the eighth graders in the 2006-07 school year (Tourangeau et al., 2009).

Sampling weights are used to make data drawn from samples representative of the larger target population. Weights adjust sample data for both nonresponse and differential probabilities of being selected into the sample. Since the present study is looking at student data from across two different waves of the study, waves 5 and 6, a longitudinal sampling weight (C56CW0) was used in all the analyses conducted for this study. Using this weight provides estimates that are better approximations of those that would have been obtained from the actual population of third and fifth grade students across the country (roughly 3,936,880 students). However, all sample sizes reported in these analyses are actual unweighted frequencies, to provide transparency in terms of the exact number of respondents. Of importance in the interpretation of the data in this study is the fact that,

the sample of children in the final data collection, the eighth-grade round, is only representative of the cohort of children who were in kindergarten in 1998-99 or first grade in 1999-2000.

Participant Characteristics and Sample Size

The ECLS-K data utilized for this study were obtained from the pool of students who were kindergarteners in the fall of 1998 and spring of 1999, had a valid response in any round of the longitudinal study, and had non-missing gender and ethnicity information. The kindergarteners in the pool attended public and private schools, and full- and half-day programs. Longitudinal analyses followed these same children through five additional waves of data collections at grades 1, 3, 5, and 8 and ended when most participants were in the eighth grade. Of these students 51.6% were male and 48.4% were female. White students made up 57.3% of the sample, 16.3% were black, 19.1% were Hispanic, 2.8% were Asian, and 4.5% were some other ethnicity or described themselves as multi-racial. A total of 15,285 students were in the sample at the time of the third grade round of data collection and 11,803 of these same students remained in the sample at time of the fifth grade data collection.

Data from the total sample were split to form both an exploratory and a confirmatory group. Analyses and model building was completed on the exploratory group and then applied to the confirmatory sample. By replicating the analyses from the model-building process on an independent confirmatory sample, confidence can be drawn in the generalizability of the findings to the larger population (Browne & Cudeck, 1992; Collins, Graham, Rousculp, & Hansen, 1997; Cook & Campbell, 1979; Kirk, 1995; Kline, 2011; Pastor, Barron, Miller, & Davis, 2007).

47

The exploratory group consisted of 10,657 students and the confirmatory group consisted of 10,606 students (after listwise deletions required by the analysis design). Demographic breakdowns across the two samples are shown in Table 3.1.

Table 3.1. *Percent of Students in Exploratory and Confirmatory Samples*

	Exploratory Sample(10,657)	Confirmatory Sample(10,606)
Male	.514(5474)	.530(5408)
Female	.486(5183)	.470(5198)
White	.577(5929)	.584(5850)
Black	.167(1562)	.147(1549)
Hispanic	.189(1911)	.198(1959)
Asian	.027(678)	.027(682)
Other	.041(577)	.045(566)

Note: N-sizes in parentheses

Measures

The measures available from the ECLS-K data set include a large number of cognitive and affective variables--in addition to a battery of academic performance measures. All children were assessed with these measures regardless of whether they were found to be performing on grade level.

Self-Concept. To measure socioemotional development students completed a modified version of the *Self-Description Questionnaire (SDQ-I,* Marsh, 1990a). This self-report instrument was originally developed by Herbert Marsh to measure different facets of the Marsh/Shavelson model of self-concept (Marsh & Shavelson, 1985; Shavelson et al., 1976). This version of the *SDQ* was designed for use with children ages 5-12. The original *SDQ-I* instrument asked children about their academic, physical, and social self-perceptions across seven different scales: Physical Abilities, Physical Appearance, Relations with Peers, Relations with Parents, Reading, Mathematics, School Subjects. For the resulting scales, School Subjects, Reading, and Mathematics scales had 10 items

48

in each scale, the other four scales consisted of a total of 8 items each. Several studies

have used factor analysis to verify the multidimensional nature of self-concept through

use of the *SDQ* (Marsh, Relich et al., 1983; Marsh et al., 1983). Research has also

supported the reliability and validity of the instrument (Marsh, Barnes, et al., 1984;

Marsh & Holmes, 1990; Marsh et al., 1985).

NCES shortened the original *SDQ-I* instrument for use with the third and fifth

graders in the ECLS-K sample. Only four of the seven originally published scales were

administered. The scales administered were renamed (i.e., Perceived Interest and

Competence in Reading, Math, All Subjects, and Peer Relations) and the total number of

items per scale was modified (Tourangeau et al., 2009). Two new scales (i.e.,

Internalizing Problems and Externalizing Problems) were also developed to assess

student self-report of adjustment problems. All of the items for these two new scales were

field tested and examined for reliability and validity through factor analytic studies (e.g.,

Atkins-Burnett & Meisels, 2001; Pollack, Najarian, Rock, & Atkins-Burnett, 2005). The

resulting *SDQ-I* instrument consists of a total of 42 items across 6 scales for both the

third and the fifth grades. These six scales represent factors from the multidimensional

model of self-concept (Marsh & Shavelson, 1988; Shavelson et al., 1976) with academic

self-concept being measured by the first three scales and socioemotional self-concept by

the final three scales. The actual wording of the items in each scale are contained in

Appendix A.

To make the SDQ scale acceptable for use with younger children, the original

scale items were also modified to contain wording that was easier for young children to

understand (Pollack et al., 2005). Readability indices using the Flesh-Kinkaid index were

at a 1.1 grade reading level for the revised scales. In addition, for students in the third grade, the survey administrator read the items so that reading ability would not be a factor in the responses. The number of rating scale points was also reduced from five to four, to make it easier for young students to make distinctions. Students responded by rating each item as "very true," "mostly true," "a little bit true," or "not at all true."

To address the research questions for this study, only four of the six *SDQ-I* scales were used, Perceived Interest and Competence in Mathematics, Reading, All Subjects, and Peer Relations. The other two scales, Externalizing Problems and Internalizing Problems weren't included because the items found in these two scales were not relevant to the research questions of the current study. The scales were used as developed by NCES, as composites of items asking for self-perceptions related to both perceived interest and competence. Use of the scales as they were modified by NCES has been the approach used by other researchers (Adelson, McCoach & Gavin, 2012; Froiland & Oros, 2014; Niehaus & Adelson, 2013) and maximized the appropriateness of the corresponding reliability and validity evidence.

NCES examined measures of internal consistency using Cronbach's coefficient alpha for the revised scales and found values that were consistent with those of the original instrument. These values are reported by grade level in Table 3.2 for the scales being used in this study (see Ellis, Marsh, & Richards 2002).

50

Table 3.2. *ECLS-K Modified Self-Description Questionnaire Scale Reliabilities (SDQ I) --Grades 3, and 5.*

Modified SDQ Composite Measure	Alpha Grade 3: Original *SDQ-I**	#Items	Alpha-Grade 3	Alpha-Grade 5
Perceived Interest and Competence-Math	.94	8	.90	.92
Perceived Interest and Competence-Reading	.93	8	.87	.90
Perceived Interest and Competence-All Subjects	.89	6	.79	.83
Perceived Interest and Competence-Peer Relations	.86	6	.79	.82

* reliability values for all public school students

Internal consistency measures for these composites across the two grades were high indicating acceptable reliability across the respective scales. These high reliability values indicate students were consistent when responding with their self-perceptions. Scale range, weighted mean, and standard deviations from the ECLS-K national sample were also reported for the third and fifth grade *SDQ-I* assessments. Table 3.3 displays these descriptive statistics by grade.

Weighted means indicate that students' interest and perceived competence across content areas and peer relations appear to have declined from grade 3 to grade 5. Student perceptions of competence and interest in grade 5 were also more variable than they were in grade 3. While these changes were not tested for statistical significance, they are supported by prior research indicating that for many students, feelings of efficacy and competence in school declines as they mature (Schunk & Pajares, 2005; Wigfield and Eccles, 2000, 2002).

51

Table 3.3. *ECLS-K Modified Self-Description Questionnaire (SDQ I) Scale Range, Weighted Means and Standard Deviations--Grades 3 and 5.*

Grade Level/Composite Measure	Range of Values	Weighted Mean*	Weighted Standard Deviation
Grade 3 *SDQ* Perceived Interest and Competence-Math	1-4	3.16	0.79
Grade 3 *SDQ* Perceived Interest and Competence-Reading	1-4	3.26	0.66
Grade 3 *SDQ* Perceived Interest and Competence- All Subjects	1-4	2.92	0.66
Grade 3 *SDQ* Perceived Interest and Competence- Peer Relations	1-4	3.03	0.65
Grade 5 *SDQ* Perceived Interest and Competence-Math	1-4	2.92	0.79
Grade 5 *SDQ* Perceived Interest and Competence-Reading	1-4	3.00	0.74
Grade 5 *SDQ* Perceived Interest and Competence- All Subjects	1-4	2.71	0.65
Grade 5 *SDQ* Perceived Interest and Competence- Peer Relations	1-4	2.98	0.63

*Child weights C5CW0 and C6CW0 were used to obtain these statistics.

Direct cognitive assessments. Students in rounds five and six (grades three and five) of the data collection were administered direct cognitive assessments in reading, math, and science. The ECLS-K cognitive assessment batteries were designed to assess children's academic achievement growth since the baseline assessment in kindergarten. A two-stage approach was used to make the assessment more accurate and efficient.

52

Students were first given a short routing test which indicated what difficulty level of the actual comprehensive assessment to administer. Child development and education experts consulted on the design and development of the assessment instruments. They detailed the important cognitive knowledge and skills that should be assessed by the end of grades 3 and 5 to address the standards from elementary schools' curricula from across the nation.

Pools of test items in each of the content domains were developed by a team of elementary education specialists and were designed to extend the longitudinal scales initiated in kindergarten with grade-appropriate changes in content and format. Test items were reviewed by elementary school curriculum specialists for appropriateness of content, difficulty, and for sensitivity issues. The cognitive assessments were designed to have overlapping items, ones that also appeared on the assessment for the adjacent grade level. This created a longitudinal scale which allowed the instruments to assess growth across time.

Vertical equating of scale scores for the kindergarten/first-grade to the third grade data was achieved through a small sample of grade 2 students who were administered the same items as those in grades 1 and 3. The longitudinal scores necessary for measuring gain over time were estimated by pooling the data from kindergarten/first-grade with the data from the third grade. As further support for the equating results, a small sample of student in the second grade were administered items from both the first and third grade assessments. It was not necessary to conduct a bridging of items between grades three and five since there was sufficient overlap between high performing 3rd graders and low performing 5th graders.

Results for the cognitive assessments were reported in proficiency, cluster, criterion-referenced, and standardized (T) scores. T-scores reported performance relative to that of peers while criterion-referenced (IRT), proficiency, and item cluster scores evaluated performance in reference to specific skill standards. For this study IRT scores were used, since they make it possible to create a common scale allowing longitudinal comparisons of achievement scores (Crocker & Algina, 2008; Hambleton, 1983; Lord, 1980). Therefore, IRT scores can be used to calculate gains, even when the assessments from different grades are not identical. By using the pattern of correct responses with item difficulty and discrimination information, each student can be placed on a continuous ability scale. It is then possible to estimate the score the child would have achieved if all of the items in all of the assessment forms had been administered.

Table 3.4 shows the alpha coefficients and descriptive statistics for the math and reading IRT scale scores for the grades used in this study.

Table 3.4. *ECLS-K Math and Reading IRT Scale Range, Means, and Standard Deviations--Grades 3, and 5.*

Grade Level	Range of Values	Weighted Mean	Standard Deviation	Alpha Coefficient
Grade 3 Math IRT	0-174	98.77	24.96	.95
Grade 5 Math IRT	0-174	122.94	25.18	.95
Grade 3 Reading IRT	0-212	125.70	28.57	.94
Grade 5 Reading IRT	0-212	148.67	26.85	.93

NCES provided validation documentation for their achievement assessments through a number of different methods (Tourangeau et al., 2009). First, they established content validity through including content standards from a number of state and national assessments. Secondly, they compared their assessments to assessments from state accountability systems and national test publishers. They also established content validity

54

through the judgments of curriculum experts. Convergent validity evidence was provided through comparing their field test results with those of normed and validated published instruments. They also had experts who were familiar with the National Assessment of Educational Progress (NAEP) verify that the assessments were similar, based on a review of the content blueprint and field test results.

Student demographic variables. Table 3.5 gives descriptive information for all of the variables being used in the current study.

Table 3.5. *Description of Variables*

Variable	Operational Definition
Criterion Variable	
Reading and Math IRT scores	Continuous scale score ranging from 0 to 212
Direct Child Assessments	
SDQ-I Measures of Perceived Interest/Competence in Reading, Math, All Subjects, and Peer Relations	Likert scale survey items with response categories ranging from 1 to 4 points
Covariates SES status	Continuous composite variable, standardized, ranging from -3 to 3
Gender	0= Male; 1=Female
Ethnicity	White, Non-Hispanic; Black or African American, Non-Hispanic; Hispanic; Asian; and Other. Dummy coded variables with White as the reference group.

Socio-economic status (SES) is a composite from the ECLS-K data which consists of five different variables: father's education level, mother's education level, father's occupation, mother's occupation, and household income. Hot deck imputation was used to construct missing values. Hot deck imputation uses the value reported by a similar respondent for each individual missing item (Enders, 2001; Little & Rubin, 2002).

55

The SES composite is a continuous variable that has been standardized with scores ranging from -3.0 to 3.0. Higher values on this variable indicate higher levels of SES (Tourangeau et al., 2009).

These student-level demographic variables and measures of cognitive and socioemotional development were used as the data elements for the current study and served to provide operational definitions of the study research questions.

Research Questions Operationalized

Latent profile and latent transition analyses were conducted to answer each of the study's research questions. Mplus (version 7.4) was used for all analyses.

Research Question 1. *Do student responses from the self-report scales used with the ECLS-K (SDQ-I, grade 3 and grade 5) form latent clusters that relate to the multi-dimensional structure of the Marsh/Shavelson conceptual model (1985) of self-concept? In other words, are there resulting class structures where some students have higher reading self-concept, while others have higher math self-concept. Is there a resulting class where some students have higher self-concept related to peer relations and lower self-concept in the academic areas?*

Subscales scores from the four *SDQ-I* scales for grade 3 and grade 5 were used to conduct two independent latent profile analyses. The following procedures were followed for each analysis. A series of alternative models designating the number of latent classes that describe student responses on the *SDQ-I* were examined. First a model was fit that represented the presence of just one latent class. Additional classes were added and model fit was assessed with each addition. The process of adding classes was continued until there was no improvement in model fit.

56

Model fit was assessed with both absolute (the log-likelihood value, ℓ) and relative model fit statistics (e.g., Lo-Mendell-Rubin likelihood ratio test, AIC, BIC and SABIC information criteria). In addition, measures of certainty, posterior probabilities and entropy, were considered during model selection. Assessing the overall fit of different solutions included additional factors beyond these relative fit and certainty measures: (1) descriptive information about the cases falling within groups (demographics such as gender and ethnicity), (2) class centroid values, (3) matches of latent classes and membership groups to multi-dimensional self-concept theory, and (4) the usefulness of latent classifications for practical applications (DiStefano & Kamphaus, 2006).

Research Question 2. *Is there a significant association between the outcomes of math and reading achievement and the resulting latent classes at each grade level?*

Resulting latent classes will be modeled with IRT achievement scores at each grade to determine how individual class status relates to achievement performance.

Research Question 3. *Do the latent class profiles identified in grade 3 change or remain stable over time (i.e., from grade 3 to grade 5)?*

Latent transition analysis will be conducted using the best models resulting from the latent profile analyses in grade 3 and grade 5. Latent profiles for grades 3 and 5 will be compared to determine whether the same number of classes and same class definitions are present in the responses at each grade level. Response patterns will be examined to determine whether perceptions of positive peer relationships have become more prevalent for a group of students (i.e., from grade 3 to grade 5).

Research Question 4. *Do latent class statuses at grade 3 predict achievement*

performance in math and reading at Grade 5?

Grade 5 IRT scores will be regressed on latent classes from grade three to

determine whether latent class status predicts distal outcomes of academic achievement.

Data Analysis

For all variables, missing values were recoded as (-9). Also, the demographic

variables of gender (Males =0, Females=1) and ethnicity (Whites =0 and serve as the

reference group for the other ethnic categories) were dummy coded. *SDQ-I* survey items

were reverse coded when appropriate so positive responses all had the highest scale point

value.[1].

Missing data. Missing values for analytic variables are a common problem with

longitudinal analyses, primarily due to attrition. With the ECLS-K study, a subsample of

students who moved after the kindergarten base year (Wave One) were followed across

subsequent data collection waves. But many students who changed schools were lost and

not included in the full data design. For this study, problems with data analysis due to

missing data were minimal. By grade 3 (wave 5), there were 15,285 students

participating. Of these students, 14,386 (94.11%) responded to the *SDQ-I* self-report

survey. Only 11 students failed to respond to one or more survey items. In grade 5 (wave

6), there were 11,803 students still participating. Of these, 10,144 (85.94%) completed

the *SDQ-I.* Only 4 students omitted responses to one or more item. As the ECLS-K study

progressed over time, researchers faced greater difficulties in getting students to complete

the full battery of cognitive and socioemotional direct child assessments.

---

[1] Full MPlus coding and data analysis output are available upon request.

Given the longitudinal design and concomitant attrition of this study, the assumption is made that any missing data is essentially missing at random (MAR). Given this assumption, missing data was handled either through full-information maximum likelihood (FIML) estimation (latent profile analysis) or listwise deletion (generating factor scores from CFA measurement models and covariate analyses).

When assessing the stability of latent class profiles over time (latent transition analysis) only students who had latent class status from grade 3 were used in the analysis (i.e., listwise deletion). Finally, to determine the relationship of class status to distal outcomes, only those students with valid IRT data and latent class status at both grades 3 and 5 were included (i.e., listwise deletion). The reduction in sample size from the final model of the LPA to the LTA for grade three was a total of 395 (7.75%) students and a total of 491 (9.45%) students for grade five. Since the sample size was still quite large and more than adequate for maintaining power, multiple imputation was not conducted for *SDQ-I* or IRT responses.

Data analysis specifications. Given that the ECLS-K data is derived from a complex, stratified sampling scheme (i.e., students within schools, schools within counties), special attention was given to matching the analysis type and estimation methods to appropriately address these constraints. Since the present study did not examine effects related to students nested within schools, the analysis Type = TWOLEVEL was not selected. While the ECLS-K sampling design does include students from the same school, there were limitations to this design. Specifically, the original number of kindergarten students per school yielded an average size of 23 students. In addition, students were randomly selected for participation in the study from all the

59

kindergarteners across the school. As the study progressed from wave to wave, fewer students were found in the same school. NCES reports that half the students had changed schools at least once by the time they were in the third grade (NCES, 2009). Given these considerations, the GENERAL analysis type in Mplus was used for the CFA analysis. Further analyses were conducted to determine the effect of stratification in the complex sampling design of the ECLS-K data. The final LPA model in grade five was re-analyzed using the TYPE=COMPLEX analysis with the fifth grade school ID as the stratification variable. The class means obtained from this analysis were very similar in value to those obtained using only the student sample weights. When using the TYPE=COMPLEX analysis, the stratification variable was used to correct for non-response and differential probability of selection on another level above that of the student responses. Since the current study is using a person-oriented analysis at the student level, the impact of school and geographical representations are not essential for interpreting the results.

With respect to the estimation technique, Robust Maximum Likelihood (MLR) was selected in order to better accommodate the characteristics of the sample design and the measurement indicators. MLR uses the same iterative procedures as ML but provides an adjustment for both non-normality and non-independence through the use of a sandwich estimator (Muthén, & Muthén, 1998-2015). MLR was preferred for testing the adequacy of the measurement model in this study since ML could not be used with the sample weights being used to approximate population characteristics.

Since the *SDQ-I* scale has a four-point ordered measurement scale, an exploratory analysis was conducted to compare the model fit statistics generated by MLR and the Weighted Least Squares Means and Variance Adjusted (WLSMV) estimator. The

60

WLSMV estimator is generally recommended for use with categorical data, specifically when the score points are unordered or the number of response categories is less than five (Finney & DiStefano, 2013). The comparison between MLR and WLSMV estimators was conducted on grade five *SDQ-I* responses because this group involves more mature students who are better able to make distinctions related to the abstraction of self-concept. The grade five data also displayed greater differences in model fit with re-specification of error variances. The CFA results comparing the two estimation methods both converged and yielded factor loadings that were all significant. There were 90 free parameters to be estimated in the MLR model and 118 in the WLSMV model (3 thresholds per indicator for WLSMV instead of 28 errors and intercepts in the MLR model). The model fit statistic chi-square for the two models were X^2_{WLSMV} (344) = 3071.087 and X^2_{MLR} (344) = 2690.723. The CFI and TLI approximate fit indices were more favorable for the WLSMV model (CFI=.953, TLI=.949) than the MLR model (CFI=.864, TLI=.851). Finally, the RMSEA values were slightly lower for the MLR model (RMSEA = .037) than the WLSMV model (RMSEA = .038), indicating better correspondence between the predicted and observed covariance matrices. The comparison between the two models favored MLR in terms of the chi-square model test and the RMSEA, but favored WLSMV in terms of approximate fit indices. Since the fit statistics were largely comparable, the MLR estimator was selected for use based on the fact that it is more efficient and that it provides a correction for both non-normality and non-independence (Muthén, 2005; Muthén, & Muthén, 1998-2015).

61

Establishing the Measurement Model

The general process for establishing a measurement model is to test the model implied by theory by specifying alternate models. The process starts with specification of only one latent variable. This model is followed by others with increments in the number of latent variables and parameters, in a method that is theoretically logical. Estimation and re-specification occur until the implied model accounts for a significant portion of the total variance and is consistent with a theoretical framework.

As stated earlier, the construct validity of the *SDQ-I* has been established through the original work of Marsh and colleagues (Marsh, Relich et al., 1983; Marsh, Smith et al., 1983; Marsh & Holmes, 1990; Tashakkori & Kennedy, 1993), and for the modified version being used as part of the ECLS-K study (Atkins-Burnett & Meisels, 2001; Ellis et al., 2002; Pollack, et al., 2005; Tourangeau, et al., 2009). In addition, most researchers that use the *SDQ-I* with the ECLS-K data elements, have used the standard six factor configuration (Adelson et al., 2012; Froiland & Oros, 2014; Niehaus & Adelson, 2013). Given the ample evidence supporting the construct validity of the *SDQ-I* instrument and the corresponding factorial model of self-concept, the current study also utilized the traditional 6 factor model. However, since the research questions addressed in this study do not address internalizing or externalizing problems, only a four factor model was used. The standard four factors model is shown in Figure 3.1. A complete listing of the factors (latent constructs) and their corresponding items (manifest scale indicators) are included in Appendix A.

62

Latent Profile and Latent Transition Methodology

Both latent profile analysis (LPA) and latent transition analysis (LTA) were used to support a person-oriented approach to examining individual differences and change over time in traits related to class membership. Both of these models are part of a larger group of models known as mixture models. Mixture models are all based on the premise that there are many subgroups in a population, and that these subgroups differ in characteristics from each other. These characteristic differences are not directly observable and are believed to be the result of one or more latent variables. Latent class and latent profile analysis use patterns of responses to observed variables to define the latent variable construct. Latent profile analysis explains the relationship between the observed variables in much the same way that factor analysis explains common variance between items. By examining the different trait profiles within the sample, a greater understanding of the complexities of self-concept was gained. Person-oriented analyses accommodated examination of these complexities, while traditional variable-oriented analyses would not. In fact, variable-oriented methodology assumes that the relationships between characteristics are the same for all members (Bergman & Magnusson, 1997; Bergman & Trost, 2006; Nylund-Gibson et al., 2014; Sterba & Bauer, 2010; von Eye & Bogat, 2006).

63
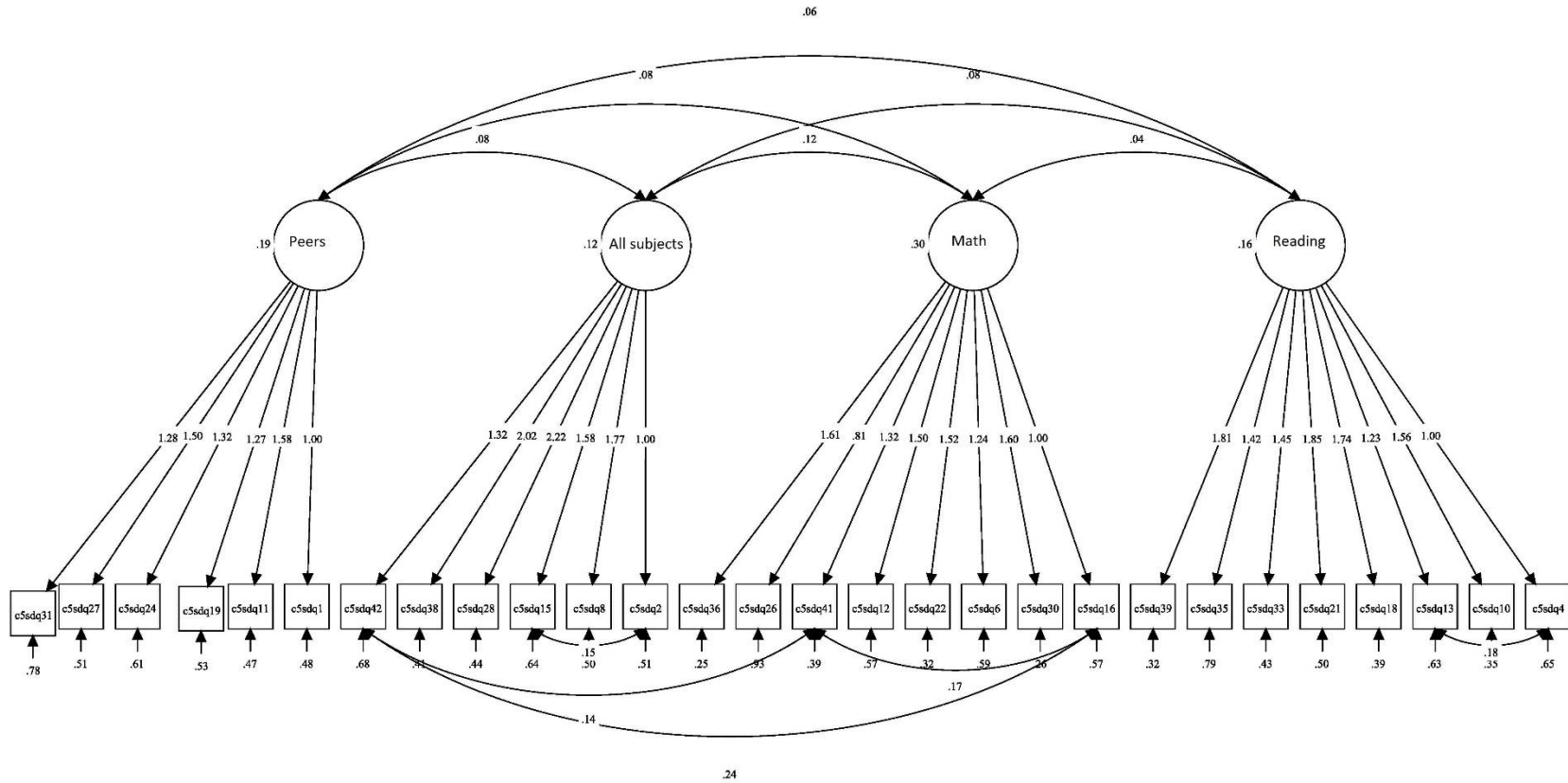
Figure 3.1. Measurement model for grade 3 and grade 5 *SDQ-I.*

The assumption with LPA is that the shared characteristics that define groups are influenced by an underlying categorical latent variable. Additionally, each individual, by definition, belongs to only one of the final set of mutually exclusive and exhaustive latent classes (Lanza & Collins, 2008) The process of LPA is exploratory in nature since the researcher rarely has an a priori hypothesis about the number of classes underlying the data (Collins & Lanza, 2010; Lanza & Collins, 2008; Nylund, 2007; Nylund-Gibson et al., 2014). LPA was used in this study to empirically define groups based on profiles or clusters of self-report responses from the *SDQ-I*.

Latent profile analysis began by looking at a series of alternative models designating the number of latent classes that describe the range of student responses in the *SDQ-I* data. MLR with Satorra-Bentler (S-B) scaling corrections, was used to generate model fit, parameter estimates, and standard errors. The final model was considered to be the best empirical representation of the latent structure underlying the observed student responses to the *SDQ-I*. Ultimately, the superior model should produce parameters that generalize across time and situations and are simple and easy to understand (Bergman & Trost, 2006; Collins & Lanza, 2010; DiStefano & Kamphaus, 2006). The relationships between classes and the constructs of multi-dimensional self-concept theory were examined to assess the explanatory value of this theoretical framework for understanding individual differences in class membership.

The second stage of data analysis involved conducting the Latent Transition Analysis (LTA), assessing the change in latent class status over time. Class membership for each grade level was estimated with LPA and final models were selected. Then the LTA was conducted. LTA is a longitudinal extension of LCA that calculates the change

65

in latent status by time in a matrix of transition probabilities between the two consecutive time points. The LTA model provides three sets of parameters: (1) latent status membership probabilities at grade 3 (Time 1), (2) transition probabilities between latent statuses over time, and (3) item-response probabilities conditional on latent status membership and time. (Cosden, Larsen, Donahue, & Nylund-Gibson, 2015; Lanza & Collins, 2008; Nylund, 2007; Nylund-Gibson et al., 2014). LTA uses an auto-regressive technique to calculate the regression of the grade 5 *SDQ-I* latent status memberships onto the latent statuses from grade 3. The resulting LTA regression coefficients indicate both the strength and the direction of the relationship between the two adjacent class profiles. Results from the two-stages of analyses are fully described in Chapter Four.

Chapter IV

Results

Descriptive Statistics

　　To set a context for the analysis, univariate and bivariate descriptive statistics for the study variables are detailed in the following tables.

　　Univariate statistics. Socioeconomic measures are reported in Table 4.1 for both the exploratory and the confirmatory subsample. Students' SES status was used with gender and ethnicity characteristic as a covariate when modeling latent profiles.

Table 4.1.*Measures of Socioeconomic (SES) Status*

Subsample	Grade Level	N	Weighted Mean	Weighted SD	Skew	Kurtosis
Exploratory	**Grade 3**	6767	-.109	.824	.323	-.115
	Grade 5	5497	-.115	.825	.314	-.065
Confirmatory	**Grade 3**	6720	-.080	.792	.358	-.043
	Grade 5	5490	-.095	.780	.348	.021

Note SES is measured on a continuous scale ranging from -3.00 to +3.00

　　SES measures for individual students ranged from -2.49 to 2.58 in both the exploratory and confirmatory subsample, showing substantial diversity across the standardized scale. Skew and kurtosis values do not depart from the absolute value cutoffs recommended by Kline of +/- 3.00 for skew and +/10.00 for kurtosis (Kline, 2011). While, statistical tests for the skew and kurtosis values were significant this is

likely to be due to the very large sample sizes. As with chi-square statistics, when

samples are large even small, trivial differences can be found to be significant.

Distributions of the SES variables do not indicate a departure from normality. Given the

large sample size in both subsamples, it is reasonable to conclude that the data does not

violate the assumptions required for the use of normal theory estimators.

Table 4.2 provides descriptive statistics for both reading and mathematics IRT

scores for the subsamples of students[2].

Table 4.2. *Descriptive Statistics: Reading and Mathematics IRT Scores*

Subsample/ Grade Level	IRT Subject	N	Mean	Std Dev	Skew	Kurtosis
Exploratory						
Grade 3	Reading	7080	123.865	28.671	-.133	-.577
	Math	7124	95.982	25.262	.072	-.721
Grade 5	Reading	5598	146.439	27.439	-.402	-.302
	Math	5599	119.745	26.172	-.446	-.515
Confirmatory						
Grade 3	Reading	7100	124.487	28.287	-.143	-.552
	Math	7144	97.386	24.920	.024	-.671
Grade 5	Reading	5594	146.973	27.033	-.440	-.273
	Math	5599	120.938	25.932	-.498	-.426

Means for the reading and math scores are very close to those from the full

ECLS-K sample, as detailed in Table 3.4. Mean IRT assessment scores increase as

students progressed from grade three to grade five. In addition, the variability of scores

also decreased from grade three to grade five for reading scores. Skew and kurtosis

values are still below recommended absolute value cutoffs, not surprising given the large

sample sizes.

---

[2] Range of Possible Score Point Values for IRT scores, Gr.3: Reading=0-174, Math=0-212; Gr.5: Reading=0-174, Math=0-212.

Table 4.3. *Descriptive Statistics: SDQ-I Scales of Perceived Interest and Competence Measures*

Subsample/ Grade Level Exploratory	SDQ-I Scales	N	Mean	Std Dev	Skew	Kurtosis
Grade 3	Reading	7133	3.257	.652	-.957	.466
	Math	7135	3.092	.772	-.768	-.293
	All Subjects	7133	2.895	.651	-.418	-.285
	Peer Relations	7133	3.000	.658	-.531	-.283
Grade 5	Reading	5602	2.986	.726	-.427	-.626
	Math	5602	2.887	.796	-.386	-.76
	All Subjects	5603	2.679	.661	-.243	-.480
	Peer Relations	5603	2.947	.644	-.464	-.303
Confirmatory						
Grade 3	Reading	7147	3.261	.658	-.966	.451
	Math	7154	3.102	.770	-.740	-.354
	All Subjects	7147	2.915	.656	-.399	-.410
	Peer Relations	7146	3.025	.647	-.531	-.247
Grade 5	Reading	5602	2.994	.734	-.478	-.567
	Math	5602	2.936	.767	-.407	-.717
	All Subjects	5603	2.718	.643	-.264	-.437
	Peer Relations	5603	2.979	.646	-.546	-.094

Table 4.3 describes the distribution of SDQ-I responses by subsample and by grade. Skew and kurtosis for the *SDQ-I* values are still below recommended cutoffs for non-normality. Interestingly, student responses to the various measures of perceived competence and interest declined from grade 3 to grade 5 for all scales, but only slightly for the peer relations scale. This occurred as mean IRT scores increased from grade 3 to

grade 5. These results suggest that while student academic performance has improved, for many students their confidence and interest in the subject has declined.

Bivariate statistics. Further information about the relationships between variables is shown in the first-order correlations shown in Table 4.4. Correlations are reported for the covariate variables (i.e., gender, ethnicity, SES), outcome variables (i.e., Math and Reading IRT scores) as well as the scale scores for each of the four *SDQ-I* scales being used in the present study. Correlations are reported separately by grade level.

As expected, correlations between composite IRT scores from one grade to another were quite large and positive, (.903 between grade 3 and grade 5 IRTs). Correlations are moderately high between IRT scores and SES measures (.492 for grade 3 and .498 for grade 5). Correlations were at least .50 for *SDQ-I* responses within the All Subjects scale and the independent scales of both Reading and Math. Also, as expected, correlations between *SDQ-I* measures at grade 3 and grade 5 were moderately high (ranging from .393 in All Subjects to .455 in Reading). It should be noted that moderate correlations also exist between *SDQ-I* ratings for Peer Relations and Perceived Interest and Competence perceptions in All Subjects (.424 at grade 3, .404 at grade 5). These correlations between perceptions of peer relations and general perceptions of school support previous research indicating that school settings meet students' needs for feelings of belonging and impact their social self-concept (Goodenow & Grady, 1993; Dowson & McInerny, 2001; Wentzel et al., 2012).

Inferential Statistics

To establish the adequacy of the *SDQ-I's* four factor measurement model a solution was specified and estimated using MLR with Mplus, version 7.4. The model was evaluated

70

for fit using criteria selected from different families: those assessing absolute fit, comparative/incremental fit, parsimony-adjusted fit and predictive fit. In addition, the chi-square model test statistic, parameter SMC's, and residuals were examined. Using a balance across these different measures is in line with recommendations from Kline (2011), Tanaka (1993), and the combination approach of Hu and Bentler (1999). Since the factor scores that are used for later analyses are determined by the parameter estimates of the model, considerable attention was given to ensuring the best possible fit for the final models at each grade.

Model fit chi-square statistic. The likelihood ratio chi-square statistic is both the most popular and most essential index, since it tests the exact fit hypothesis that there are no discrepancies between the model-predicted population covariance matrix and the observed matrix from the sample. However, there are many known problems with this statistic. First the chi-square statistic is a dichotomous decision rule that does not quantify the degree of fit along a continuum. Secondly, the chi-square statistic is influenced by sample size, model complexity, and violation of multivariate normality assumptions. The sample size directly affects the power of the test statistic to detect actual correspondence between the theoretical model and observed data. In essence, with a small sample a poor model could be found to be nonsignificant simply due to power. On the other hand, with large samples the power is so sensitive that even slight differences between the proposed and observed model are found to be significant. To combat the sample size problem some researchers have advocated using a ratio statistic that adjusts the chi-square value by the degrees of freedom for the model. This normed chi-square value (NC) is then considered

71

Table 4.4. *First-Order Pearson Correlation Coefficients between Measures*

	gender	ethnicity	SES	IRT gr 3	IRT gr 5	g3 SDQ R	g3 SDQ M	g3 SDQ All	g3 SDQ Peer	g5 SDQ R	g5 SDQ M	g5 SDQ All	g5 SDQ Peer
gender	1												
ethnicity	0.012	1											
SES	-0.006	-0.235	1										
IRT gr 3	-0.003	-0.220	0.492	1									
IRT gr 5	-0.017	-0.200	0.498	0.903	1								
g3 SDQ R	0.131	0.008	0.041	0.128	0.110	1							
g3 SDQ M	-0.128	0.036	-0.023	0.041	0.031	0.18	1						
g3 SDQ All	0.050	0.036	-0.020	0.0258	0.007	0.522	0.549	1					
g3 SDQ Peer	0.051	-0.060	0.021	-0.047	-0.068	0.324	0.309	0.424	1				
g5 SDQ R	0.134	-0.022	0.116	0.215	0.216	0.455	0.028	0.236	0.119	1			
g5 SDQ M	-0.098	0.019	0.041	0.143	0.151	0.076	0.434	0.280	0.147	0.158	1		
g5 SDQ All	0.086	0.005	0.082	0.165	0.160	0.288	0.260	0.393	0.214	0.541	0.566	1	
g5 SDQ Peer	0.075	-0.084	0.101	0.071	0.048	0.159	0.147	0.215	0.411	0.277	0.258	0.404	1

Note: SDQR=Self-Description Questionnaire Reading
SDQM=Self-Description Questionnaire Math
SDQAll=Self-Description Questionnaire All Subjects
SDQPeer=Self-Description Questionnaire Peer Relations

to be evidence of good model fit if the resulting ratio is 3.0 or greater. However, Kline warns against the use of the NC statistic since degrees of freedom for the model are not influenced by sample size and there is a lack of clear guidelines for interpreting the resulting NC ratio (see Kline, 2011, pg. 204).

Approximate fit indices. These are alternative measures of fit that should also be consulted when determining the adequacy of a model. These indices are able to quantify the degree of model fit in much the same way that R^2 quantifies the amount of variance accounted for in multiple regression. These are continuous measures as opposed to the dichotomous reject/retain decisions indicated by the chi-square model test statistic.

Absolute fit indices. These approximate fit statistics compare the fit of the obtained covariance matrix to that of the covariance matrix implied by the model. The present study reports the results of the Standardized Root Mean Residual (SRMR) as an absolute fit index. This index is the average of the standardized residuals obtained when comparing the covariance matrix specified by the model with the covariance matrix obtained from the sample. Ideally, these residual values should be near 0.0, which indicates perfect fit. The higher the SRMR value the worse the model fit. Hu and Bentler (1999) suggest a cut-off value of <.08 for SRMR as an indication of acceptable model fit, with good fit requiring values <.05.

Incremental/comparative fit indices. The Tucker-Lewis Index is a comparative fit measure computed by comparing the chi-square values between the null or worst fitting model and the model being proposed. Value range from 0 to1.0, with larger values indicating better fit. . A value of .95 or higher is needed to indicate good fit (see Schreiber, Stage, King, Nora, & Barlow, 2006). This index includes an adjustment for

73

the complexity of the model, by penalizing for additional parameters. The Comparative

Fit Index (CFI) is also a comparative, or incremental fit index. As with the TLI, the CFI

measures the improvement in the fit of the model over that of a baseline or null model.

The CFI places a penalty of one for each parameter estimated. Value range from 0 to1.0,

with larger values indicating better fit. A value of .95 or higher is needed to indicate good

fit (Schreiber et al., 2006). For this study the CFI is being used as the TLI is not available

in Mplus, version 7.4.

Parsimony adjusted fit indices. The Root Mean Square Error of Approximation

(RMSEA) is an absolute measure of the fit between the covariance matrix predicted by

the model and the sample covariance matrix. Calculations favor models with fewer

parameters through weighting by model degrees of freedom. One of the advantages of

this measure is that it is less sensitive to large sample sizes. The RMSEA also produces a

corresponding 90% confidence interval. The lower bound is used to test the close-fit

hypothesis that the discrepancy between the predicted and observed covariance matrices

is ≤ .05 (Browne & Cudeck, 1992; Schumacher & Lomax, 2010; Steiger 1990). The

upper bound is used to conduct a less common test of poor fit, which states that a poorly

fitting model is indicated by a discrepancy of .10 or greater between predicted and

observed models.

Predictive fit indices. The present study examined two different indices that

predict the fit that would likely be obtained if a replication was conducted with the same

size sample from the population. The Akaike Information Criterion (AIC) and the

Bayesian Information Criterion (BIC) are both used to select between different models

when there is no nesting. For both of these indices the model with the lowest AIC or BIC

index has the better fit. When testing different models the researcher usually includes both the null model (one with no associations between variables) and the saturated model (one that perfectly reproduces all of the variances and covariances) as reference points. The AIC index also imposes a penalty for the number of parameters being estimated.

Table 4.5 and 4.6 display the resulting tests of the alternative measurement models for both grades using model test statistics and fit indices from across different families. In addition a summary of chi-square difference tests between models are provided. A complete discussion of these model comparisons follows.

The initial results for the CFA measurement models in both grades show acceptable fit across some indicators. For example, in both grades the RMSEA is well below the Hu and Bentler suggested cutoff of <.06. Similarly, the SRMR values across both grades are lower than Hu and Bentler's suggested cutoff value of <.08. The close fit hypothesis (i.e., H_0: $\varepsilon_0 < .05$) cannot be rejected for any of the models tested, across both grades. Failure to reject this hypothesis provides more support for the conclusion that the proposed model closely fits the observed data matrix. A similar test is often conducted where the hypothesis is of poor fit. Here the hypothesis being tested is H_0: $\varepsilon_0 > .10$, where, ε_0 is defined as the RMESA or error of approximation in the population. The aim is to reject the hypothesis and be able to conclude that the model being proposed is not as bad as a model with poor fit. For all the models tested, the poor fit hypothesis was

75

Table 4.5. *Confirmatory Factor Analysis Fit Information for the SDQ-I Scale (MLR Estimator), grade 3*

Grade 3 Measurement Model (N=5,699)								
	X_M^2	df_M	X_M^2	df_M	RMSEA (90% CI)	CFI	AIC	SRMR
Initial 4 factor model	1950.68**	344	----	----	.029 (.027-.030)	.873	378953.73	.054
Model with one error covariance	1812.84**	343	29.97*	1	.027 (.026-.029)	.884	378224.06	.053
Model with two error covariances	1738.48**	342	26.87*	1	.027 (.026-.028)	.890	377814.95	.053
Model with three error covariances	1657.69**	341	24.71*	1	.026 (.025-.027)	.896	377392.81	.053
Model with four error covariances	1585.95**	340	22.68*	1	.025 (.024-.027)	.902	377007.58	.052
Model with five error covariances	1521.09**	339	20.93*	1	.025 (.023-.026)	.907	376664.15	.053

*p<.05; **p<.001

Table 4.6. *Confirmatory Factor Analysis Fit Information for the SDQ-I Scale (MLR Estimator), grade 5*

Grade 5 Measurement Model (N=5,592)								
	X_M^2	df_M	X_M^2	df_M	RMSEA (90% CI)	CFI	AIC	SRMR
Initial 4 factor model	2960.72**	344	----	----	.037 (.036-.038)	.864	340855.23	.060
Model with one error covariance	2816.82**	343	70.68*	1	.036 (.035-.037)	.872	340166.04	.060
Model with two error covariances	2667.33**	342	64.05*	1	.035 (.034-.036)	.879	339480.17	.060
Model with three error covariances	2528.00**	341	57.82*	1	.034 (.033-.035)	.887	338853.46	.061
Model with four error covariances	2390.91**	340	52.13*	1	.033 (.032-.034)	.894	338225.33	.062
Model with five error covariances	2302.98**	339	47.77*	1	.032 (.031-.033)	.898	337803.99	.060

*p<.05; **p<.001

rejected, indicating closer fit than a poor fitting model. Both the close fit and poor fit tests are based on the RMSEA confidence intervals with cutoff values taken from work on model fit indicators by Browne and Cudeck (1992, see also Kline, 2011).

The chi-square test of global fit was significant, indicating that the implied covariance matrix based on theory is significantly different from the observed matrix. However, as Kline (2011) notes, the chi-square measure is overly influenced by sample size. With large samples, the test is often significant, even when the absolute difference between the two matrices is very small. The samples in the present study are much larger than typical CFA and SEM samples of 100 or 200 cases.

The items from the *SDQ-I* are very general and have quite a bit of overlap in meaning, both within and across scales. For example, items in both math and reading ask students to self-report on the following: "I like reading" or "I like math," and "I enjoy doing work in reading" or "I enjoy doing work in reading math." Similarly, items in the Perceived Interest and Competence in All School Subjects simply ask students their feelings about all subjects, not excluding reading or math. These vague questions are likely to generate very similar responses. Modification indices were examined for items at both grade levels. The modification indices for the items detailed below had the highest change in chi-square values as a result of allowing correlation of their error terms (decrease in chi-square values ranged from 70.96 to 141.40).

Before considering letting these error terms co-vary, identification rules were examined. Since there were still more than three indicators with uncorrelated error terms within each factor, the alternative models were all identified (see Kline, 2011). Given this, additional models were run at each grade level. These alternative model allowed the

78

error variances for different pairs of items to co-vary. The items corresponding to these

error covariances were:

Grade 3

Item #42 —"I get good grades in all school subjects," and Item #16 —"I get good
grades in math."
Item #4 —"I get good grades in reading," and Item #13 —"Work in reading is
easy for me."
Item #41—"I am good at math," and Item #16 —"I get good grades in math."
Item #41 — "I am good at math," and Item #42 — "I get good grades in all
school subjects."
Item #2 — "I am good at all school subjects;" and Item #15 —"Work in all school
subjects is easy for me."

Grade 5

Item #42 — "I get good grades in all school subjects," and Item #16 —"I get
good grades in math."
Item #6 — "Work in math is easy for me," and Item #41 —"I am good at
math."
Item #4 —"I get good grades in reading," and Item #33 — "I am good at
reading."
Item #2 —"I am good at all school subjects," and Item #15 —"Work in all
school subjects is easy for me."
Item #10 —"I like reading," and Item #18 —"I am interested in reading."

Each set of item pairs were added to the models, one at a time. After each addition

the models were rerun and fit was re-examined. As additional errors were allowed to co-

vary, RMSEA decreased and CFI values improved and were closer to the recommended

level of .95 from Hu and Bentler (1999). The chi-square model fit statistic decreased, as

did the AIC, indicating better model fit.

To test the significance of the decrease in chi-square values for models with

added error covariances, the Satorra-Bentler scaled chi-square difference test was

calculated. As shown in Table 4.5, the results were significant between each of the

different sets of nested models at both grade levels. These significant results indicate that

the fit of the model with error covariances for related items were statistically superior to the models where errors were constrained to be independent of each other.

The information from these nested tests, along with changes in approximate fit indices clearly indicate that a model that allows errors to co-vary is better at approximating the observed variances in the sample data. This conclusion is further supported from examining the content of the actual items and their high degree of overlap in meaning and wording. Making distinctions between these items can be difficult at any level, and even more so when students are young and less adept at making abstract decisions.

Factor variances and covariances. Estimates for each grades' final model, including error covariances, are displayed in Table 4.7. Variance estimates for the four latent variables show that there is slightly more variability in the responses to the math indicators. While not displayed due to space limitations, factor loadings across all models for both grade levels were moderate to high within factors (ranging from .414 to .895). Correlations of this magnitude indicate both convergence within and discrimination between factors.

Correlations between the four different factors in the model are also informative. The correlations between reading and math were low (.111 in grade three and .182 in grade five). However, the relationships between math and all school subjects (correlation of .626 in grade three and .648 in grade five) and reading and all school subjects (correlation of .593 in grade three and .606 in grade five) were relatively strong. This emphasizes the overlap and ambiguity students perceive about what is being addressed in the all school subjects items.

Table 4.7. *Factor Variance/Covariance and Error Covariances from Final Measurement Model (MLR Estimator)*

	Unstandardized (Gr3/Gr5)	SE (Gr3/Gr5)	Standardized (Gr3/Gr5)
Factor Variances and Covariances			
Reading	.213/.158	.023/.021	1.00/1.00
Math	.356/.297	.027/.028	1.00/1.00
All Subjects	.151/.121	.016/.016	1.00/1.00
Peer Relations	.246/.193	.024/.025	1.00/1.00
Reading with Math	.031/.039	.009/.009	.111/.182
Reading with All Subjects	.106/.084	.010/.009	.593/.606
Reading with Peers	.061/.061	.010/.009	.268/.351
Math with All Subjects	.145/.123	.013/.012	.626/.648
Math with Peers	.065/.082	.011/.011	.219/.342
All Subjects with Peers	.089/.081	.011/.010	.461/.530
Error Covariances			
Grade 3			
Item 42 with Item 16	.241	.024	.384
Item 4 with Item 13	.177	.023	.277
Item 41 with Item 16	.169	.022	.360
Item 41 with Item 42	.145	.024	.283
Item 2 with Item 15	.147	.021	.258
Grade 5			
Item 42 with Item 16	.187	.016	.379
Item 6 with Item 41	.152	.018	.369
Item 4 with Item 33	.155	.017	.330
Item 2 with Item 15	.150	.016	.345
Item 10 with Item 18	.114	.015	.340

Relationships between perceptions of peer relations and academic subjects were also informative. Correlations were low between perceptions of peer relations and interest and competence in reading and math (.268-gr.3 and .351-gr.5 for reading and .219-gr.3 and .342-gr.5 for math). However, correlations were approximately .50 when looking at perceptions of interest and competence in all school subjects and peer relations (.461-gr.3

81

and .530-gr.5). This lends support to the idea that students perceive school as a social place and that their perceptions of interest and competence at school are influenced by the peer relationships they develop there.

Residual variance estimates. Kline (2011) recommends that the pattern and magnitude of standardized residual values be examined as a more informative diagnosis of misspecification. The rule of thumb he proposes is looking at the absolute number and the pattern of standardized residuals with values greater than .10. The percentage of the 378 estimated residuals that were over .10 in grade three was 6.1% and 5.8% in grade five. The pattern of residuals showed that many were associated with items containing ambiguous or overlapping meaning, the same items identified during the examination of modification indices during model re-specification analyses. The following items had higher residuals (>.10) in both grades:

- Item #2: "I am good at all school subjects;"
- Item #4: "I get good grades in reading,"
- Item #6: "Work in math is easy for me,"
- Item #13: "Work in reading is easy for me," and
- Item # 15: "I get good grades in math."

These items with high standardized residuals, are seen as poor measures for explaining the sample covariances in the observed data. That is, there is significant variability left unexplained after regressing the indicator on the underlying latent factor in the model. One conclusion is worth noting, none of the indicators with high residuals or large modification indices related to student perceptions of interest and competence with peer relations. Perhaps students see the items in this scale as more clear-cut and distinguishable.

82

Testing for Measurement Invariance

As previously discussed, the items across the different scales for the *SDQ-I* are vague and ask students to self-report on very similar feelings. These issues often lead researchers to conduct invariance testing to examine differential interpretation of item content across groups. The present study did not conduct invariance testing for two reasons. First, the purpose of running the CFA on the *SDQ-I* scales is to verify the adequacy of the measurement model for creating latent classes and profiles, not to validate or create tests or scales. Secondly, research has already been conducted related to invariance testing for the *SDQ-I* scales. Niehaus and Adelson, in their 2013 study, looked at differential interpretation of *SDQ-I* items across third grade English proficient and English language learners. The results of their multi-group CFA showed that self-concept was measured by the *SDQ-I* similarly across three different language groups.

Generating Factor Scores

The measurement model has established that the *SDQ-I* indicators are associated with four different latent factors. Scores are generated for individuals, in order to identify their ranking or position on these factors (DiStefano, Zhu, & Mîndrilă, 2009). Factor scores were generated by Mplus version 7.4, based on the best model from the CFA results at each grade level (both models had four factors and five error covariances). Mplus generates factor scores through a regression-based method. The resulting scores are standardized and considered to be *z*-values. While the raw data from the *SDQ-I* is represented on a 4 point Likert scale, the factor scores generated through confirmatory factor analysis are treated as continuous. The factor scores are estimated as the maximum of the posterior distribution of the factor, and use a model to predict the optimal score.

83

Specifically, this method weights the observed scores within the factors by regression coefficients that take into account correlations between indicators and factors, among factors, as well as correlations between all observed variables. This regression-based method results in standardized factor score estimates for each individual case (Muthén, 2005; Muthén & Muthén, 2015). The resulting scores indicate each student's approximate position with respect to the latent variable. These factor scores were used for all further analyses within the present study.

After factor scores were generated the resulting distributions for each grade were analyzed for normality, factor determinacy, univocality, correlational accuracy, and validity. Table 4.8 and the accompanying histograms (Figures 4.1 through 4.8) show the factor score distributions for both grades.

Table 4.8. *Skew and Kurtosis Values for Regression-Based Factor Scores*

Factor	Grade 3		Grade 5	
	Skew	**Kurtosis**	**Skew**	**Kurtosis**
Reading	-1.046	.552	-.435	-.628
Math	-.909	-.183	-.482	-.788
All Subjects	-.579	-.243	-.240	-.542
Peer Relations	-.529	-.246	-.456	-.191

Skew and kurtosis values were not out of the range of acceptable values, thus indicating a normal distribution of factor scores in each grade. All skew and kurtosis values were negative, which is confirmed by the histograms. Each distribution was negatively skewed to some degree, indicating more favorable responses by students to the *SDQ-I* questionnaire. This was most pronounced for both reading and math responses in grade three, and to a lesser degree in grade five. For grade 3, levels of negative skew were double that found in grade 5, across all factors. This suggests that younger children

84

may have more positive perceptions of their performance across different subjects and in their social relationships with peers. While these differences were not tested for significance, this inference is supported by the research studying maturation of self-concept and self-perception of competence in children (Harter, 1988; Marsh et al., 1998; Wigfield & Eccles, 2002).

Negative kurtosis values indicate platykurtic distributions; in other words, scores that cluster closer around the mean, with fewer outliers and values in the tails. These distributional results indicate a greater tendency to give responses from the middle of the scale, avoiding extreme responses.

The distributions that conform to the normal curve most are from the All School Subjects and Peer Relations factors in both grades. The math factor scores distribution in grade five is multimodal at three points, .5 *SD* below the mean, at the mean, and just beyond the .5 *SD* mark above the mean. The largest percent of responses occurs at the mode above the mean. This indicates that students describe themselves most often as liking math and doing well in the course work. However, there is another large group of students reporting that they do not like math and do not do well in math class.

While the distributions for the factor scores appear to deviate from a normal distribution, characteristics of subgroups which may underlie this broader population is what we are attempting to explain with mixture models (Muthén, 2005). In addition, the large sample sizes in the present study allow the central limit theorem to affect the distributions. Therefore, we can consider the factor scores as appropriate for use with the MLR estimator.
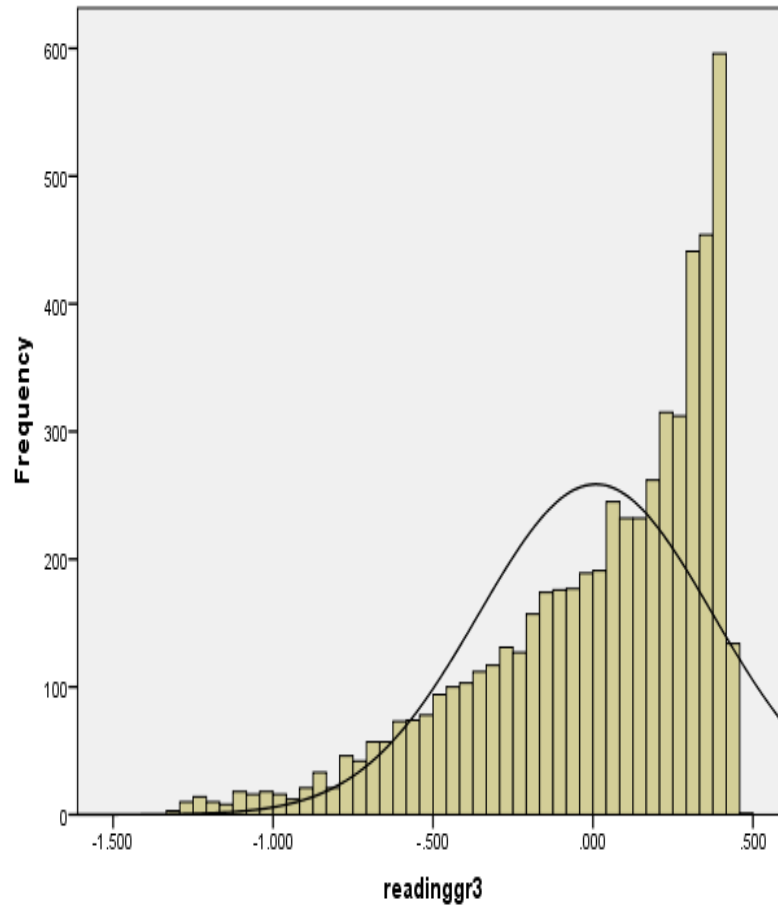
Figure 4.1 Grade 3 Reading Factor Scores
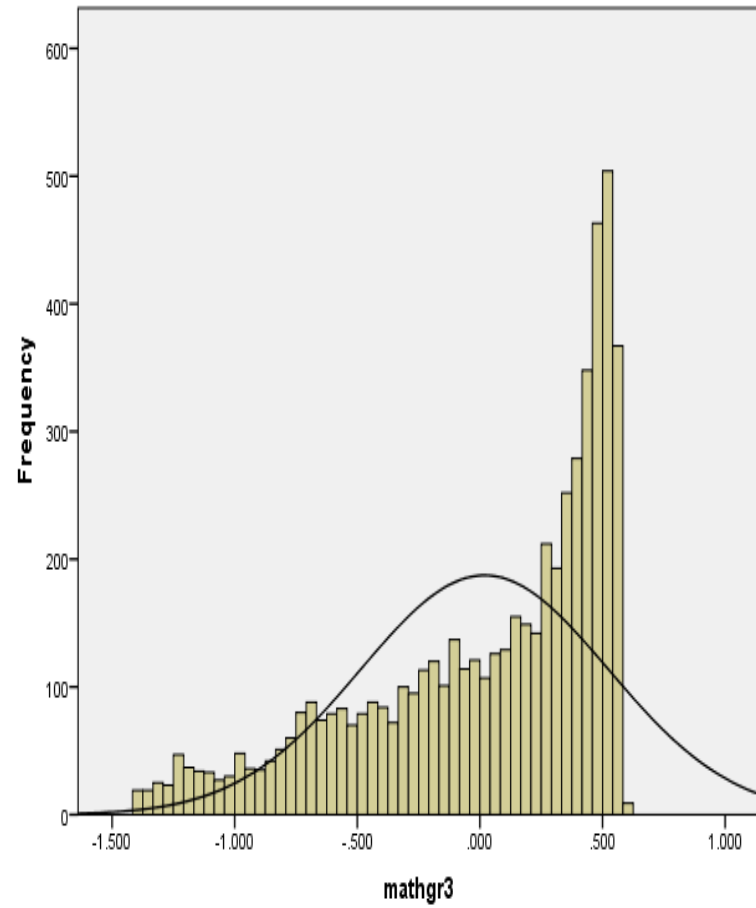
Figure 4.2 Grade 3 Math Factor Scores
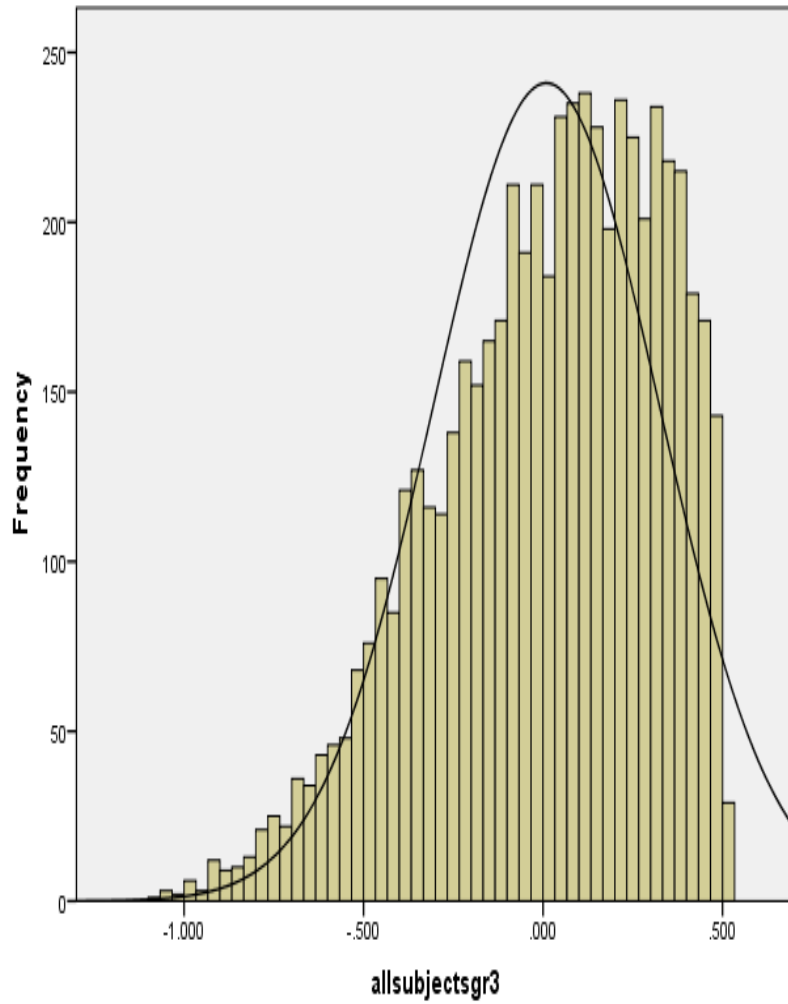
Figure 4.3 Grade 3 All Subjects Factor Scores
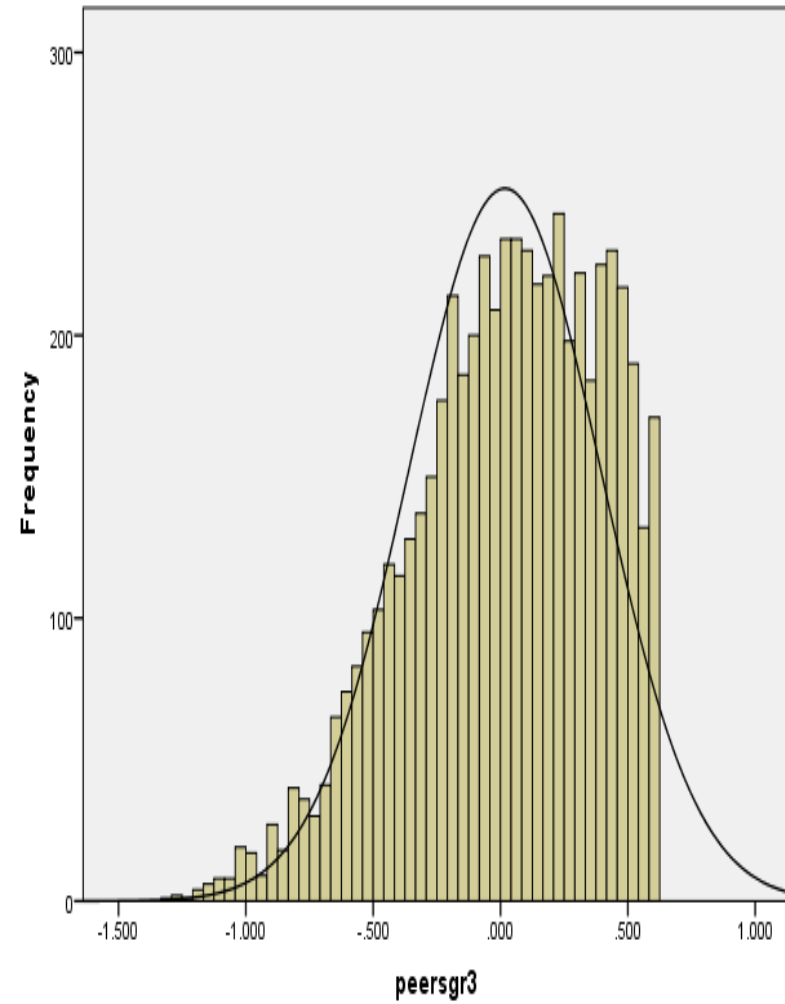
Figure 4.4 Grade 3 Peer Relations Factor Scores
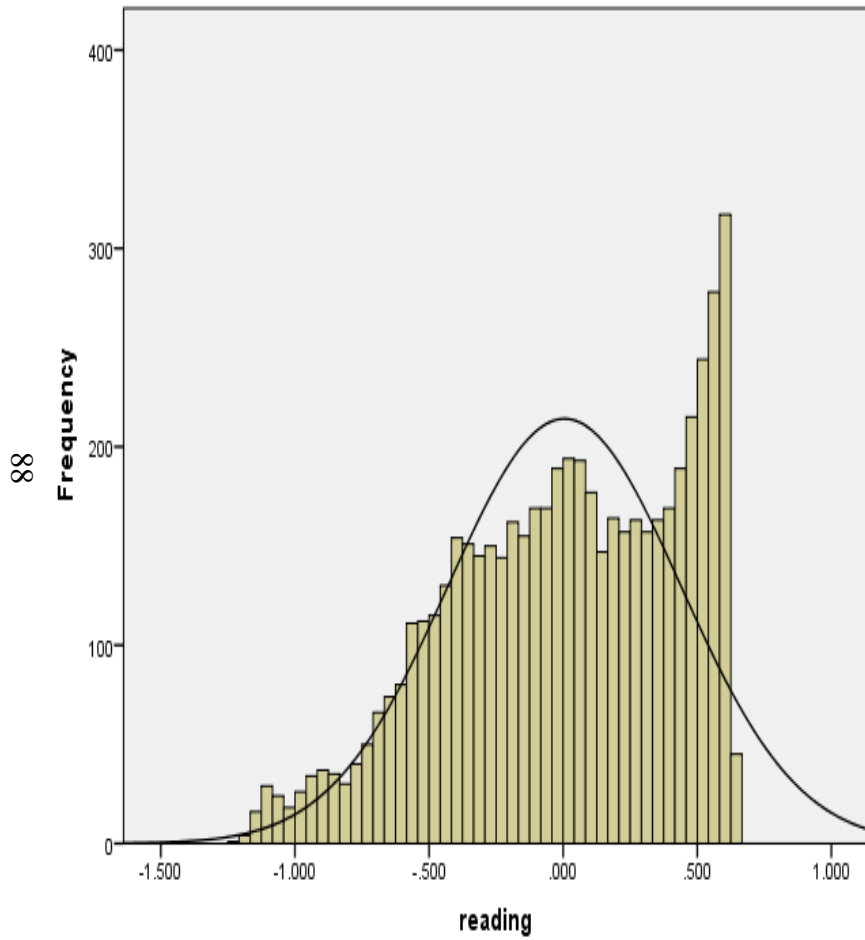
Figure 4.5 Grade 5 Reading Factor Scores
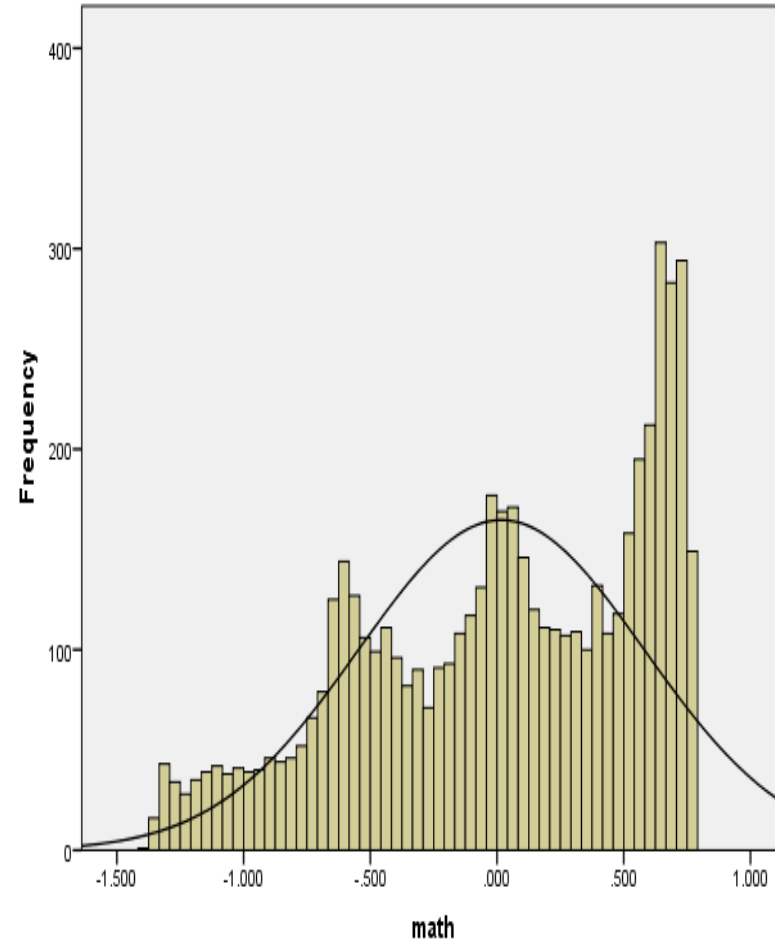
Figure 4.6 Grade 5 Math Factor Scores

88

Frequency

300

200

100

0

-1.500   -1.000   -.500   .000   .500   1.000

allsubjects

Frequency

250

200

150

100

50

0

-1.500   -1.000   -.500   .000   .500   1.000

peers
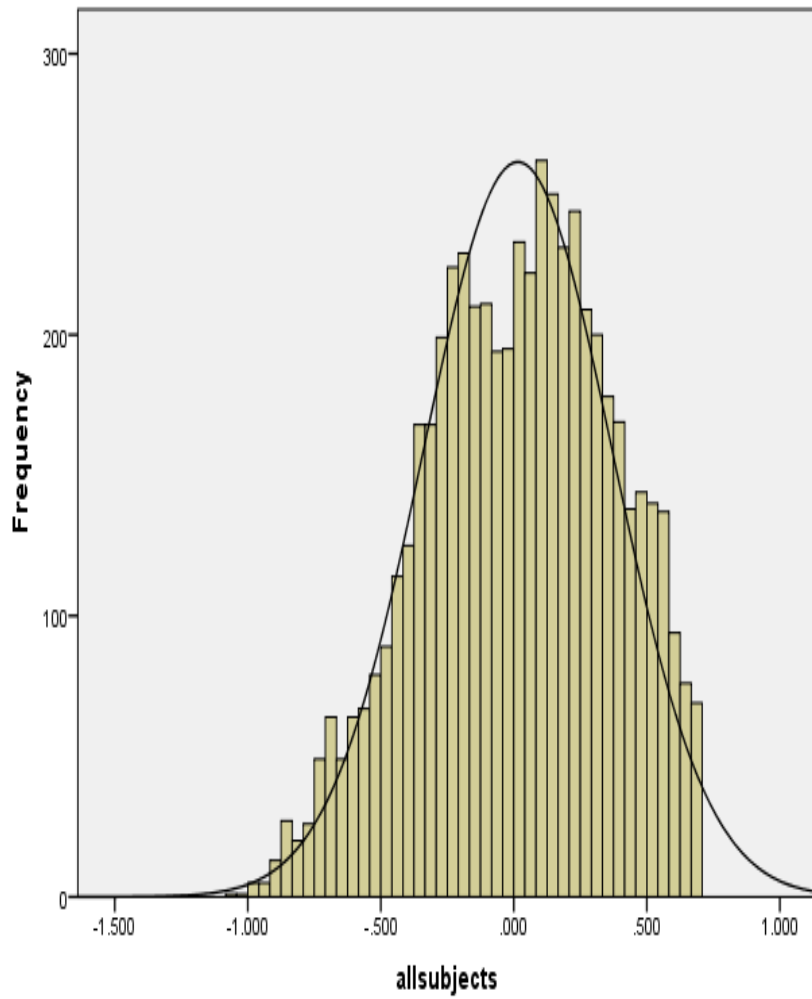
Figure 4.7 Grade 5 All Subjects Factor Scores
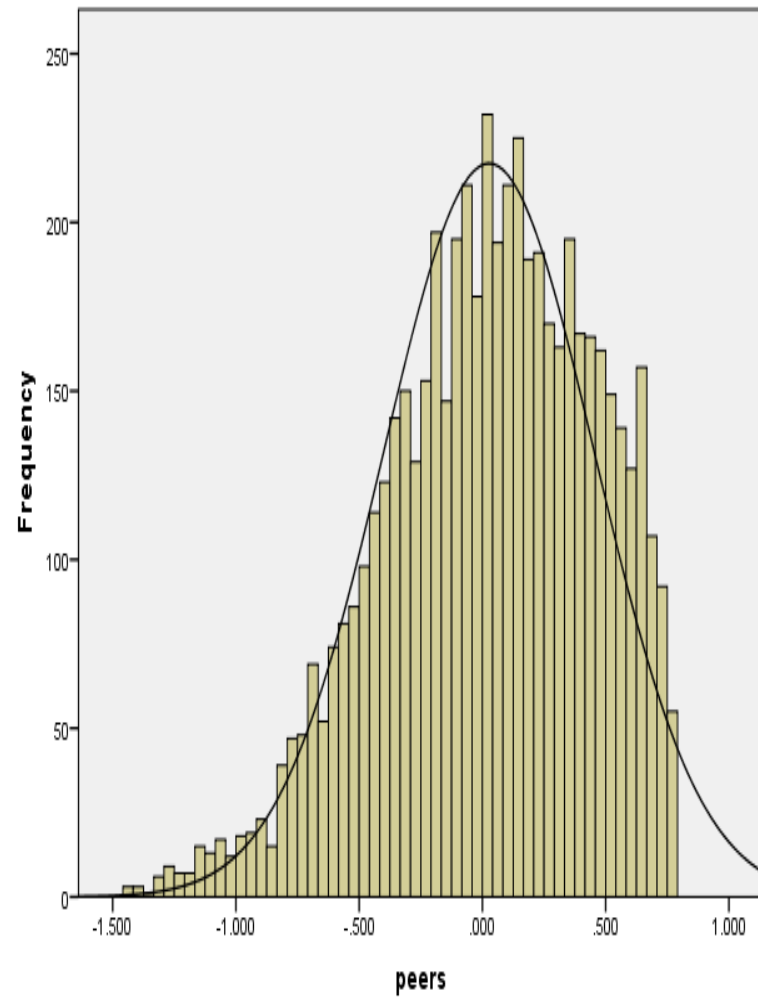
Figure 4.8 Grade 5 Peer Relations Factor Scores

Factor scores also need to be examined in terms of their correlations. Specifically, factor loadings for individuals should have the strongest correlations with their assigned factors (validity), and low correlations with other factors, when the different factors are seen to be independent and uncorrelated (univocality; Distefano et al., 2009). Within this analysis, factors score coefficients were examined to determine if there were indicators that had high coefficients (regression weights) for factors they were not identified with. Indicators aligned with the All Subjects factor had moderate coefficient values across the other academic factors in the model. This provides more evidence of ambiguity in the focus and meaning of the indicators in the All School Subjects factor. This is a limitation that needs to be considered as results from the latent profile and latent transition procedures are analyzed.

Finally, factor determinacies were examined to discover the extent to which the estimated factor scores approached true factor scores. Determinacies range from 0.0 to 1.0 and result from multiple correlations between the items and the factor. The higher the determinacies value, the more validity in the factor scores. The determinacies (multiple correlations) are also higher when the items within a factor have better reliability. In the present study, determinacy values ranged from .898 to .962 for grade three and .919 to .971 for grade five. The lowest values were for the Peer Relations factor (.898 in grade three and .919 in grade five). These values, taken with the other distributional and correlational information, support the conclusion that the validity of the factor scores are adequate to proceed with the latent variable analysis procedures.

Latent Profile Analysis: Cross-Sectional Results

To test the hypothesis that perceptions of peer self-concept (in addition to those for academic self-concept) have utility value for describing achievement motivation, latent profile analysis was undertaken. By using a person-oriented, latent variable approach, students were classified into categories based on similar profiles of self-concept characteristics, as evidenced through their *SDQ-I* response patterns.

The expectation is that the *SDQ-I* responses are influenced by an underlying latent construct broadly defined as self-concept. Based on the Marsh 2009 research, it was anticipated that latent classes resulting from LPA would show differences in the level of overall self-concept (i.e., classes would form around high, medium, and low levels of self-concept across all four domains). Difference in profiles shapes were also expected. Specifically, it was expected that some students would have higher reading self-concepts, while others would have higher self-concepts in math. Finally, it was expected that some students would have higher self-concepts related to peer relations and lower levels of self-concept in the academic areas. This class structure was tested by looking at student response patterns and determining the classes or distinct groupings across the self-concept latent construct that may explain the underlying variance and covariance patterns in the data.

Latent profile analyses were conducted using the factor scores from the final CFA measurement model at two time periods--when students were in grade three, and later when they were in grade five. To begin the analysis of the general model, a one factor, independence model was estimated. This model served as the baseline for comparing other class solutions. If the independence model would have been found to have good fit,

91

www.manaraa.com

this would have suggested that there was no relationship between the variables and that there was not an overarching latent construct (Nylund et al., 2006; Nylund et al, 2007).

After the independence model was estimated additional classes were added consecutively, with re-estimation of parameters and fit occurring after each addition. Specifically, factor means, class probabilities, log-likelihood values, and improvement of fit indices were examined.

Some general comments about the data:

- Missing data was not considered to be problematic since there was only one missing data pattern identified and relatively few missing data points. The covariance coverage matrix showed the full proportion of data was present with a value of 1.000 for all variances and covariances.

- Default estimation for latent profile analysis with Mplus version 7.4 allows the means across classes to be freely estimated. Variances were held equal across classes with covariances set to 0.0 (i.e., local independence).

- To prevent problems with convergence, different start values and iterations were selected. Models were re-estimated with varying values to make sure the same solution was obtained. It should be noted that during the development of the LPA measurement models for both grades, convergence and replication of the best log-likelihood values were achieved, indicating stable solutions.

92

The model-building process. During the analyses conducted to determine the optimal class structure for the data, three different categories of information were routinely consulted: (1) mean factor values across classes (class centroids), (2) absolute and relative fit statistics, and (3) classification accuracy (conditional probabilities and entropy values). For ease of reading, the model-building process is described in terms of each class addition.

Analysis of class centroids (mean structure). Essential to the task of interpreting the meaning of resulting classes is the examination of the resulting mean structure between and across classes.

Two class models. In comparison to the baseline model, the two factor model resulted in a better fit to the data in both grade levels. Patterns of mean scores indicated that there were two groups of students, one with high scores (those that have positive perceptions of their interest and competence with school subjects and peer relationships) and the other with low scores (those having negative perceptions of their interest and competence with school subjects and peer relationships).

Three class models. When three classes were specified, all fit indices and classification accuracy indicators improved. However, the distinction between classes as revealed by their means became less clear cut. In both grades, one class still had high mean scores across all factors (47% of the students at grade 3 and 44% at grade 5). There was a second smaller group of students (14% of the students in grade 3 and 14% in grade 5) that showed negative perceptions of their interest and competence towards math and reading, and an even stronger negative response to all school subjects. Interestingly, their perceptions of interest and competence with peer relations were more positive for this

group. Finally, the third class at both grade levels was found near the mean across all factors. This is a large class of students that appear to be ambivalent towards school (39% of the students at grade 3 and 42% at grade 5).

Four class models. The four class solutions in both grades had complex structures. In both grades, the largest class still involved students with positive perceptions across school subjects (51% of the students at grade 3 and 41% at grade 5). This group also had lower mean scores with respect to peer relations at both grade levels. This class may be described as "achievement-oriented," including students who have positive perceptions of their interest and competence in academic areas as opposed to peer relationships. A second class (20% of the students at grade 3 and 26% at grade 5) appears to be ambivalent with respect to school content and peer relations. The third class that emerged in both grades, displayed favorable perceptions in math (10% of the students at grade 3 and 18% at grade 5). In grade three this class has very negative perceptions of all school subjects. This class could be described as one of "math lovers." The last class in both grade levels was fairly difficult to interpret.

Five class models. These models were essentially the same as the four class models, only more complicated and thus difficult to interpret. Parsimony concerns make these models difficult to defend.

Model fit--absolute and relative fit statistics. An examination of model fit and classification certainty for the different solutions included the following measures.

Log likelihood values are reported as part of the class analyses with mixture modeling. These values are based on the log likelihood function which is maximized during the estimation algorithm used to determine class membership. Log likelihoods are

also the basis of the information criteria measures (AIC, BIC, SABIC) used to compare models with different numbers of classes. For the log likelihood values and the various information criteria, a smaller value indicates better fit of the model to the observed variance/covariance matrix. These values are consulted when making comparisons between models, even when different parameter structures are being estimated. Nylund, Asparouhov, & Muthén (2007) conducted a simulation study to determine which of these statistics was most efficient and found that the SABIC is the most reliable and thus, the preferred measure.

Normally when determining the best model, log likelihood differences between nested models are examined. The differences are referenced to the chi-square distribution and if the chi-square value is significant, then the model with additional parameters being freed is preferred. However, it has been discovered that with latent class analyses the differences between log likelihood values with nested models is not distributed as a chi-square distribution. Therefore, the resulting *p*-values are not accurate. Given this problem, researchers have advised against using the chi-square difference test and have proposed two alternative likelihood ratio tests (Lo, Mendell, & Rubin, 2001; Nylund et al., 2007).

The Lo-Mendell-Rubin Likelihood Ratio Test is used to compare the model currently being estimated with a model having one fewer classes. The *p*-value is used to indicate the probability that the observed data came from the models with the fewer number of classes. A small *p*-value indicates that the model with more classes is preferred. Larger *p*-values indicate the more parsimonious model should be retained. A second likelihood ratio involves the use of bootstrapping with likelihood ratio differences.

The Bootstrap Likelihood Ratio Test (BLRT) also returns a *p*-value that indicates whether the currently estimated model fits better than one with one fewer classes. While the BLRT statistic has been found to be reliable (Nylund, Asparouhov et al., 2007), it is not available for use with sampling weights.

Classification accuracy. All models were evaluated for the accuracy of their classifications. There are two main measures used to evaluate classification accuracy. Posterior classification probabilities give the full set of probabilities for each classification decision in terms of averages. The matrix of probabilities is examined to make sure that the probabilities for the assigned class (accurate classifications) are high and that the classifications for other classes (inaccurate classifications) are very low. The information from this matrix of probabilities is also collapsed into a single value called "entropy." Entropy values range from .00 to 1.00 and tell how randomly classifications have occurred. Values closer to 1.00 indicate more accurate and certain classifications.

Results of the combined statistical indices, along with the patterns of factor means, and classification accuracy were used to determine the optimal number of classes in the latent variable of self-concept. While the information criteria and fit indices decreased in value with the addition of classes, the Lo-Mendell-Rubin (LMR) test were significant when testing the 4 class against the 3 class models, as well as the 5 class against the 4 class models. Significant *p*-values indicate that the simplest models (two classes) are preferred for both grades. Entropy values increased until the five class models, when the entropy values decreased. This indicates that the classifications of individuals into classes became more random and less accurate with a five class model. In addition, interpretation of the classes based on means was more difficult with the five

class models in each grade. Given these results, it was decided that additional analyses should be conducted to relax some of the constraints imposed on the models through the Mplus program defaults. At the end of this first stage of model development the best constrained models for each grade were named based on the conditional probabilities, classification certainty, and fit indices.

Stage one--best constrained models:

- Grade Three: 3-class model
- Grade Five: 4-class model

Releasing model constraints. Given the lack of simple structure from the first phase of the modeling building process (i.e., with more than three classes in grade three and four classes in grade five), additional analyses were conducted with constrained variances being freely estimated.

Freeing variances. As part of this process, the best models for each grade from the first stage of analysis (i.e., with constrained variances) were reanalyzed. These models were tested to determine how they compared with corresponding models with freed variances. In both grades the models with freed variances had better fit statistics, conditional probabilities, and entropy values. Mean structures also had more separation between classes making them easier to interpret. Adjacent models with one additional class and one less class were also examined within the freed variance parameterization. Final results indicated that the three class model was the optimal model in terms of interpretive meaning, fit, and classification certainty for grade three. The four class model was found to be optimal for grade five.

97

Stage two--best freely estimated variance models:

- Grade Three: 3-class model

- Grade Five: 4-class model

Estimating covariances. After testing the models with freed variances at each grade, one more analysis was undertaken. This analysis freely estimated the covariances between selected factors. The covariances that were freed were between the factor scores for all school subjects and those for both math and reading. Given the high correlations between these three factors and the ambiguity in the content of the corresponding survey items, it appeared that the relationships between these factors should be modeled. When freely estimating variances and covariances, the number of estimated parameters grew considerably. The complexity of these models resulted in convergence problems in both grades. Lack of convergence indicates model misfit and misspecification (Muthén & Muthén, 1998-2015).

As noted earlier, it is not possible to achieve accurate results through chi-square difference testing between these nested models. Additionally, the LMR likelihood difference test cannot be used with models differing in parameterization (Lo et al., 2001; Pastor et al., 2007). Therefore, determining the appropriate number of classes relied on examination of relative values of fit indices, classification accuracy, as well as substantive interpretability of the classes, usefulness in practice, and ties to existing theory (Bergman & Trost, 2006; DiStefano & Kamphaus, 2006; Nylund, Bellmore et al., 2007). Therefore, the final models selected retained the simpler structures found when only the variances were allowed to be free across the different classes.

98

Stage three--best models with free variances and covariances constrained to 0.0

- Grade Three: 3-class model

- Grade Five: 4-class model

Covariates as predictors of class membership. To further describe the latent classes, covariate analyses were conducted to determine whether class membership could be predicted by characteristics of individuals. Three different types of demographic variables were used: gender, SES and ethnicity. Covariates were added to the final models at each grade, one variable at a time. After each addition, the models were examined to make sure the fit statistics and classification probabilities continued to improve. Satorra-Bentler log likelihood difference tests were also used to make sure that each covariate was a significant predictor of class composition in the model. The results of these significance tests for covariates are shown in Table 4.9.

Table 4.9. *Satorra-Bentler Log Likelihood Difference Tests for Covariates*

Grade Level	Covariate	Parameters Estimated	Satorra-Bentler Difference
Grade 3	Gender	28	23.274
	SES	34	-867.973
	ethnicity	28	2428.483
Grade 5	Gender	38	2.179*
	SES	47	-1855.675
0	ethnicity	38	-1762.679

*$p>.05$

The only demographic variable that was not found to be nonsignificant as a covariate was gender in grade five. The final models including covariates were determined to be:

Stage four--best models with covariates, free variances, & covariances

constrained to 0.0

- Grade Three: 3-class model with covariates of gender, ethnicity, and SES

- Grade Five: 4-class model with covariates of ethnicity and SES

The final models with covariates were used to generate conditional probabilities of membership in each latent class and the final class assigned for each individual. These were used to report cross-sectional differences in class membership between grades three and five. Those results are reported in the section on latent transition analysis.

The final models. The tables and graphs which follow show the final results for the four phases of analysis conducted to determine the class structure of the latent self-concept construct. Table 4.10 and 4.11 display the absolute and relative fit statistic, and entropy values for the final models at each stage. To establish the generalizability of the results for the final models, the analyses were cross-validated on an independent hold-out sample. Table 4.12 and 4.13 display absolute and relative fit statistic, and entropy values for the confirmatory models.

Log likelihood values decreased considerably as variances were freely estimated and covariates were added to the models. For example, in grade three the 3 class model had a log likelihood value of -6859.804 when variances were set to be equal across the

Table 4.10. *Model Fit and Class Probabilities-Exploratory Sample, Grade 3*

Classes and Model Paramaterization by Grade Level	# of Para-meters	Entropy	Log-likelihood	AIC	BIC	SABIC	Lo- Mendell-Rubin test
Grade Three (N=5699)		26	38	23	35	50	
C#1	8	NA	-11255.618	22527.236	22580.452	22554.999	NA
C#2	13	.822	-7924.355	15874.711	15961.135	15919.825	6511.924 p=0.0000
C#3	18	.827	-6859.804	13755.608	13875.272	13818.074	2080.977 p=0.2277
C#4	23	.861	-5884.691	11875.381	11968.287	11895.199	1906.143 p=0.2847
C#5	28	.823	-5348.394	10752.787	10938.932	10849.957	1048.350 p=0.3887
3 Classes Free Variances	26	.845	-4432.25	8916.499	9089.348	9006.728	NA*
3 Classes Free Variances w/covariates (N=5098)	38	.853	-4058.121	8192.242	8440.633	8319.882	NA*

*Lo-Mendell-Rubin results cannot be compared between these models due to different parameterization.

Table 4.11. M*odel Fit and Class Probabilities-Exploratory Sample, Grade 5*

Classes and Model Paramaterization by Grade Level	# of Para-meters	Entropy	Log-likelihood	AIC	BIC	SABIC	Lo- Mendell- Rubin test
Grade Five (N=5592)							
C#1	8	NA	-14064.751	28145.502	28198.534	28173.113	NA
C#2	13	.790	-11437.746	22901.492	22987.671	22946.361	5134.993 *p*=0.0000
C#3	18	.792	-10763.428	21562.856	21682.179	21624.981	1318.087 *p*=0.113
C#4	23	.805	-9990.592	20027.185	20179.654	20106.567	1510.658 *p*=0.0459
C#5	28	.791	-9636.324	19328.648	19514.263	19425.287	692.487 *p*=0.3972
4 Classes Free Variances	35	.827	-9359.703	18789.407	19021.425	18910.206	NA*
4 Classes Free Variances w/covariates (N=5194)	50	.829	-8641.477	17382.954	17710.717	17551.834	NA*

*Lo-Mendell-Rubin results cannot be compared between these models due to different parameterization.

Table 4.12. *Model Fit and Class Probabilities-Confirmatory Sample, grade 3*

Classes and Model Paramaterization by Grade Level	# of Para-meters	Entropy	Log-likelihood	AIC	BIC	SABIC	Lo- Mendell-Rubin test
Grade Three (N=5479)							
C#1	8	NA	-10592.396	21200.792	21253.662	21228.240	NA
C#2	13	.837	-7423.089	14872.177	14958.090	14916.780	6194.698 *p*=0.0000
C#3	18	.835	-6442.921	12921.841	13040.798	12983.599	1915.827 *p*=0.6267
C#4	23	.857	-5593.969	11233.937	11385.937	11312.850	1659.353 *p*=0.2461
C#5	28	.876	-5111.211	10278.422	10463.465	10374.490	943.594 *p*=0.2392
3 Classes Free Variances	26	.848	-4231.872	8515.744	8687.570	8604.950	NA*
3 Classes Free Variances w/covariates (N=5098)	38	.858	-3633.296	7342.592	7589.400	7468.650	NA*

*Lo-Mendell-Rubin results cannot be compared between these models due to different parameterization.

Table 4.13. *Model Fit and Class Probabilities-Confirmatory Sample, grade 5*

Classes and Model Paramaterization by Grade Level	# of Para-meters	Entropy	Log-likelihood	AIC	BIC	SABIC	Lo- Mendell-Rubin test
Grade Five (N=5592)							
C#1	8	NA	-13432.902	26881.804	26934.692	26909.270	NA
C#2	13	.779	-11006.256	22038.512	22124.456	22083.146	4743.128 *p*=0.0000
C#3	18	.765	-10371.719	20779.439	20898.437	20841.239	1240.267 *p*=0.1789
C#4	23	.798	-9634.376	19314.752	19466.806	19393.719	1441.213 *p*=0.0269
C#5	28	.772	-9374.467	18804.935	18990.044	18901.069	508.018 *p*=0.6596
4 Classes Free Variances	35	.822	-8907.614	17885.228	18116.614	18005.395	NA*
4 Classes Free Variances w/covariates (N=5194)	50	.817	-8318.285	16736.570	17063.557	16904.674	NA*

*Lo-Mendell-Rubin results cannot be compared between these models due to different parameterization.

three classes. When variances were freely estimated, the log likelihood value decreased to -4432.25. After adding covariates to the model likelihood value decreased further to -4058.121. Similar decreases were obtained with all the information criteria (i.e., AIC, BIC, SABIC). Entropy values also increased when variances were allowed to be estimated and covariates were added, indicating that parameterization was more accurate and less random.

Class probabilities for the different models and the parameterizations in each step of the model building process are shown in Table 4.14. The smallest classes were found in the five-class models, which is to be expected. The final results are further detailed in the distribution of means (class centroids) in Table 4.15 and figures 4.9 and 4.10.

It was expected that there would be class differences in profiles shapes, where some students would have higher reading self-concepts and others would have higher self-concepts in math. However, that was not how classes were formed. For grade three, there was a class of students with high self-concept across all factors. The proportion of students assigned membership to this high self-concept class ("High SC") is only 17%. The second class is near the mean across all factors, an ambiguous group labelled as "Average SC." Both the average and high achiever groups have slightly higher math self-concept factor scores. The final group is low performing, but particularly in math. This

Table 4.14. *N-size & Latent Class Probabilities for Most Likely Class Membership*

Classes in Model	C#1	C#2	C#3	C#4	C#5
Grade 3 (N=5699)					
Two Class Model	2055 (0.361)	3644 (0.639)			
Three Class Model	793 (0.139)	2651 (0.465)	2255 (0.396)		
Four Class Model	549 (0.096)	1104 (0.194)	1138 (0.200)	2909 (0.510)	
Five Class Model	515 (0.090)	486 (0.085)	1021 (0.179)	1021 (0.179)	2003 (0.352)
3 Classes Free Variances	2324 (0.408)	2288 (0.402)	1086 (0.191)		
3 Classes Free Variances w/covariates (N=5098)	2087 (0.410)	2148 (0.421)	863 (0.169)		
Grade 5 (N=5592)					
Two Class Model	2380 (0.426)	3212 (0.574)			
Three Class Model	2378 (0.425)	762 (0.136)	2452 (0.439)		
Four Class Model	1425 (0.255)	855 (0.153)	2300 (0.411)	1011 (0.181)	
Five Class Model	652 (0.117)	776 (0.139)	1068 (0.191)	1185 (0.212)	1911 (0.342)
4 Classes Free Variances	585 (0.105)	762 (0.136)	2127 (0.380)	2118 (0.379)	
4 Classes Free Variances w/covariates (N=5194)	2036 (0.392)	1624 (0.313)	842 (0.162)	836 (0.133)	

Table 4.15. *Means for Final Class Solutions with Covariates and Variances Freely Estimated*

	Low SC, Low Math (C#1)	Average SC (C#2)	High SC (C#3)	
Grade 3 (N=5699)				
Reading	-.252	.089	.358	
Math	-.393	.197	.462	
All Subj.	-.313	.131	.389	
Peer Rel.	-.264	.090	.389	

	Low SC (C#1)	Average SC, Low Math SC(C#2)	Average SC, High Math (C#3)	High SC (C#4)
Grade 5 (N=5592)				
Reading	-.297	.191	-.026	.503
Math	-.400	-.040	.644	.538
All Subj.	-.359	.127	.224	.518
Peer Rel.	-.262	.155	.080	.403
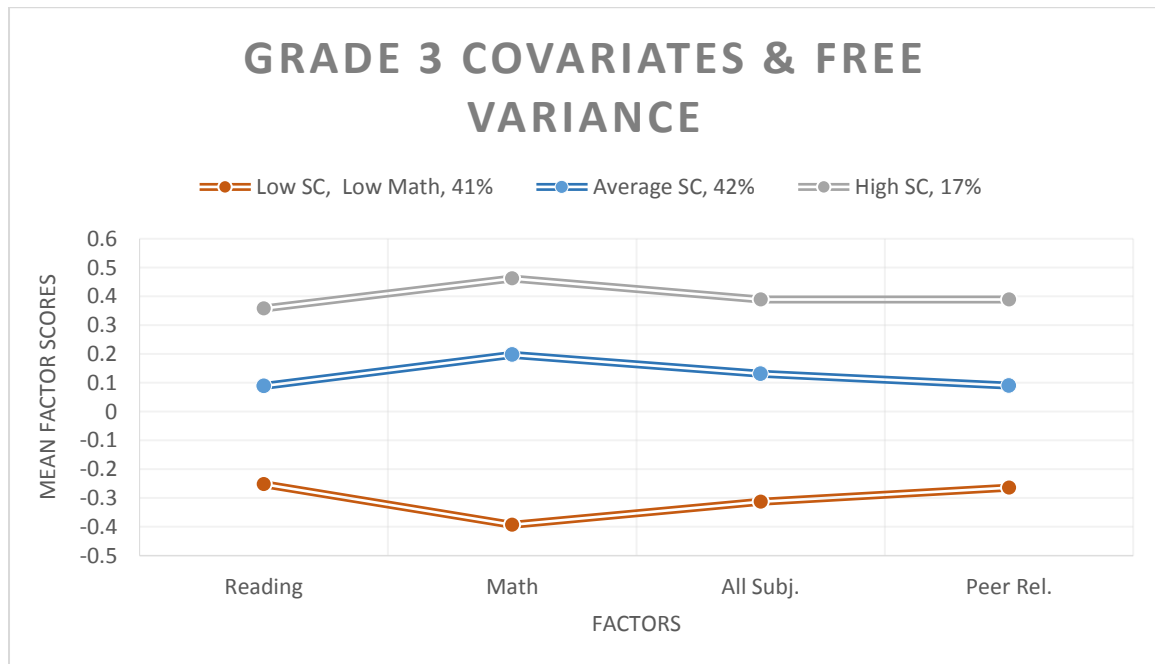
Figure 4.9 Latent profile analysis results for grade three.

group was named "Low SC, Low Math." Overall, the classes have simple structure and clearly discriminate between groups, without cross over.

The profiles for fifth graders are marked by a low class ("Low SC") and a high class ("High SC"). The low performer class consists of close to half of the students at this grade (39%). The third class is an average group with less favorable perceptions of
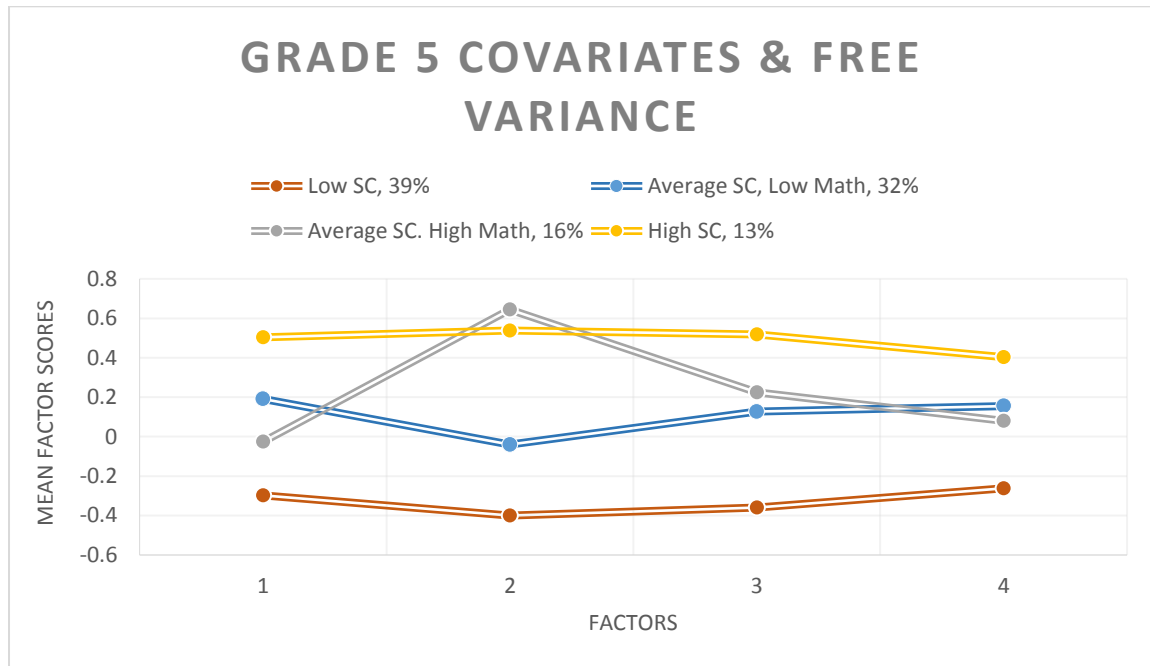
Figure 4.10. Latent profile analysis results for grade five.

interest and competence in math (named the "Average SC, Low Math" group). The fourth class is the most interesting with strong positive perceptions in math, a "math lovers" group. This "Average SC, High Math" class actually has more positive math self-concept than the "High SC" group. Finally, evidence of the certainty and accuracy of the classifications that determined the final class structure are shown in Table 4.16.

Table 4.16. *Classification Accuracy-Conditional Probabilities*

Grade 3	Low SC, Low Math	Average SC	High SC	
Low SC, Low Math	**0.948**	0.052	0.000	
Average SC	0.066	**0.915**	0.019	
High SC	0.000	0.054	**0.946**	

Grade 5	Low SC	Average SC, Low Math	Average SC, High Math	High SC
Low SC	**0.943**	0.047	0.010	0.000
Average SC, Low Math	0.075	**0.893**	0.015	0.017
Average SC, High Math	0.038	0.079	**0.860**	0.023
High SC	0.000	0.049	0.067	**0.884**

Classification of students into classes is probabilistic in latent class/profile

methodology. The certainty with which students have been classified based on their

SDQ-I responses is very accurate, especially for third grade and the large group of low

performers in grade five.

Covariate Analyses

The results of the covariate analyses from both grades are shown in Table 4.17.

Table 4.17. *Logistic Regression Coefficients and Log Odds for Covariates, Grade Three*

Grade 3	Covariate	Coefficient	SE	*Z*	*p*-value	Log Odds
Low SC, Low Math class	Gender*	-.511	.160	-3.201	.0001	.600
	Black**	-.996	.249	-4.000	.0000	.369
	Hispanic**	-.617	.193	-3.195	.0001	.540
Average SC class	Gender*	-.433	.158	-2.743	.0006	.648
	Black**	-.848	.247	-3.432	.001	.428
	Hispanic**	-.422	.192	-2.198	.028	.656

*Gender (0= male, 1=female)
**For each ethnic category 0=white as the reference group

109

Results of the covariate analysis for grade three showed that males were more likely to be in the "Low SC, Low Math" and the "Average SC" class (log odds of .600 and .648 respectively) when compared to the referent of "High SC." This means that females were more likely to be in the "High SC" group. Blacks and Hispanics were more likely than whites to be in the first and second classes ("Low SC, Low Math" and the "Average SC" groups), with log odds of .369 and .428 for blacks and .540 and .656 for Hispanics. Average SES was roughly equivalent across the three classes (coefficients were not significant).

Results of the covariate analysis for grade five are displayed in Table 4.18. In grade five, low SES, blacks, Hispanics and Asians were all more likely than whites to be in class 1 ("Low SC") instead of the reference class ("High SC"). Blacks were also more likely than whites to be in class 2 ("Average SC, Low Math") instead of the reference class ("High SC").

Table 4.18. *Logistic Regression Coefficients and Log Odds for Covariates, Grade Five*

Grade 5	Covariate	Coefficient	SE	Z	*p*-value	Log Odds
Low SC	SES	-.414	.129	-3.212	.001	.661
	Black*	-.686	.302	-2.268	.023	.504
	Hispanic*	-.441	.223	-1.971	.049	.644
	Asian*	-1.060	.299	-3.543	.000	.346
Average SC,	Black*	-.827	.342	-2.417	.016	.438
Low Math	Asian*	-.676	.322	-2.097	.036	.509
Average SC,	SES	-.329	.163	-2.023	.043	.719
High Math	Asian*	-.661	.329	-2.009	.044	.516

*For each ethnic category 0=white as the reference group

110

The covariate analysis serves as validity evidence for the final latent profile structure in both grades. Results show that minorities are more likely to be in the lower classes, as are males. These results are consistent with the results of research studying minority and gender differences in interest and feelings of competence in academic areas (Biddle, 2014; Kao & Thompson, 2003; Voyer & Voyer, 2014).

Latent Transition Analysis

Latent transition analysis (LTA) is an extension of latent class and latent profile analyses, and as such is part of the family of mixture models. However, LTA is unique in that it conducts a longitudinal examination of the changes in latent class membership between different measurement periods. Latent transitions involve qualitative changes that are reflective of moving from one stage to another. In essence, latent transition tracks individuals who have been placed in subgroups based on a common profile. These individuals are followed over time to see if they remain in that same subgroup or if they transition to another subgroup with different characteristics.

Latent transition models use the results of time-specific, cross-sectional latent profile analyses as the measurement models. The steps in the transition analyses involve regressing outcomes from a later time point on the variables from an earlier time point. LTA relationships can be analyzed without specifying the exact time period, and therefore are discontinuous. Growth models, by contrast, focus on quantitative changes that are occurring continuously over time to assess change at the latent level. Growth models also focus on observable variables instead of latent constructs.

There were a number of distinct steps undertaken to conduct the LTA for the present study. First, the demographic characteristics of the sample and the distribution of

111

the observed indicators were described. Then the adequacy of the measurement model was established for both time periods, grades three and five (see descriptive and CFA results earlier in this chapter). After establishing the adequacy of the measurement model, LPA was conducted and the number of classes underlying the data was determined. The resulting cross-sectional classes were named and then related to covariates of gender, SES, and ethnicity. Finally, the reliability and validity of the two latent profile solutions were verified through cross-validation. Analysis of a hold-out sample that was fit to the same data verified the generalizability of the latent profile solutions for each grade. The use of cross-validation is recommended whenever feasible to ensure that the final model is not capitalizing on unique nuances from the sample used to estimate models (Browne, 2000; Collins et al., 1997; Kline, 2011).

If the same number of classes would have been identified at each time point, testing for measurement invariance would have been advised. Invariance testing could determine whether the measured indicators were functioning in the same manner at each time point, and thus could be interpreted in the same way across time. If there are a different number of classes at the different time points then it is not possible to test for measurement invariance (Nylund, 2007; Nylund et.al, 2006).

Given that the number of classes at grade three (3 classes) and grade five (four classes) differ, no invariance testing was conducted. In fact, for the present study it was hypothesized that as children mature there will be more specialization and domain-specificity found in perceptions of academic self-concept. This would make it likely that an increase in the number of classes observed would occur in grade five. This increase in specialization has been verified with prior research (Marsh, 1990a; Marsh & Hocevar,

1985; Marsh & Shavelson, 1985). However, person-oriented research methodology has not yet been used to determine how classes related to self-concept may split and differentiate. Given the need for research in this area, this model was freely estimated at both time points (i.e., 3 classes at grade three and four classes at grade five). Modeling complete measurement non-invariance allowed class structures to change naturally as children matured. Unfortunately, this level of free estimation came at a cost. The complexity of the time-specific models led to a failure to replicate the best log-likelihood value, even after increasing the number of random start values (i.e., 1000 random starts), optimizations, and iterations.

Alternative grade five measurement models. Given the research decision not to constrain any parameters when conducting the LTA, an investigation was undertaken to test some alternate LPA measurement models in grade five. This investigation was conducted within the latent transition modeling framework. Models with three, five, and six classes were tested to determine whether any of these would be a better fit for the data. The results of this investigation are detailed in Table 4.19.

Log Likelihood, AIC, BIC, and SABIC fit statistics all decreased with each additional class. In addition, results from traditional likelihood ratio difference tests (LRT's) were significant due to the complexity of the models. However, after examining the resulting class structures, the decision was made not to select any of the models as superior to the four class model. The resulting classes in each of these models (3, 5, and 6 class models) were not well differentiated and did not represent distinctions between the *SDQ-I* profiles. In addition, the complexity of the five and six factor models made them very difficult to estimate.

113

Table 4.19. *Alternate LTA Measurement Models for Grade Five, Fit Statistics (N=4703)*

Classes	3 Classes	4 Classes	5 Classes	6 Classes
Free Parameters	104	135	163	197
Loglikelihood	-11013.558	-10130.768	-9616.954	-9092.312
AIC	22235.116	20531.536	19565.908	18578.625
BIC	22906.535	21403.090	20637.597	19850.448
SABIC	22576.062	20974.110	20110.110	19224.455
LRT Diff test	NA	-216.381 *df*=31 *p*=0.000	3265.615 *df*=28 *p*=0.000	1510.658 *df*=34 *p*=0.000
Entropy	.851	0.850	0.859	0.862

After finalizing the measurement models for each time point, the LTA procedures were initiated. Results from the analysis are detailed in the series of tables which follow. First, in Table 4.20, is a cross-sectional display of the prevalence of students in each class at each time period.

Table 4.20. *Cross-Sectional Class Representation from LPA Results*

	Grade 3 (N=5098)		Grade 5 (N=5194)
Low SC, Low Math	41%	Low SC	39%
Average SC	42%	Average SC, Low Math	32%
		Average SC, High Math	16%
High SC	17%	High SC	13%

While these data represent only a cross-sectional summary from the LPA results, they do display the overall movement patterns in class membership across the two time periods. The data show that there is a large group of students with self-perceptions of low

interest or competence at both time points (roughly 40%). Most of the third graders that

were in the "Average SC" class, likely transitioned to one of the average categories in

grade five. Approximately 68% of the students with average perceptions in grade five had

lower perceptions of their math self-concept and ended up in the "Average SC, Low

Math" class. Twenty-four percent of the students had moved out of the "High SC"

category by the fifth grade.

Another way to look at the transition data, again based purely on the cross-

sectional data from the LPA's, is shown in Table 4.21.

Table 4.21. *Cross-Sectional Class Representation from LPA Results (N=4703)*

Grade 3 Classes	Grade 5 Classes			
	Low SC	Average SC, Low Math	Average SC, High Math	High SC
Low SC, Low Math	56.0%	28.8%	10.3%	4.8%
Average SC	29.0%	36.0%	21.1%	13.8%
High SC	15.9%	32.8%	20.5%	30.8%

These results illustrate that most of the students from the "Low SC, Low Math"

self-concept class in grade three were still in the lowest class when they got to fifth grade.

A moderate percentage transitioned to the "Average SC, Low Math" class. This transition

could indicate that math self-concept was the perception pulling down feelings of overall

achievement competence. About a third of the students who rated themselves as average

in grade three fell to the "Low SC" class in grade five. Similarly, about a third of the

students who rated themselves as high in terms of self-concept in grade three fell to the

"Average SC, Low Math" group in grade five. Again, it appears that feelings of self-

concept in math have more weight in determining overall perceptions than any of the

115

other areas of self-concept. This movement also supports the research that shows that with maturation students become more realistic when self-evaluating and are better able to understand their academic abilities (Marsh & Craven, 1997; Marsh et al., 1998; Harter, 1988; Wigfield & Eccles, 2002).

With LTA, we move from looking at the *SDQ-I* scores from grade 3 and grade 5 as two different data sets to actually treating them as repeated measures from the same individuals. Longitudinal transitional class proportions are shown in Table 4.22.

Table 4.22. *Longitudinal Class Proportions from LPA* (N=4703)*

	Grade 3 (N=5098)		Grade 5 (N=5194)
Low SC, Low Math	.355	Low SC	.212
Average SC	.454	Average SC, Low Math	.332
		Average SC, High Math	.274
High SC	.191	High SC	.183

*Based on Most Likely Latent Class Patterns

These results show that most students perceived themselves as average in terms of self-concept measures when they were in third grade. Fewer students had low self-perceptions by the time they were in fifth grade. By the time the students were in fifth grade, the large "Average SC" category split into two distinct groupings: "Average SC, Low Math" and "Average SC, High Math." Again, these results support the research that shows that academic self-concept becomes more differentiated as children mature (Marsh, 1990a; Marsh & Hocevar, 1985; Marsh & Shavelson, 1985). Instead of general perceptions of low ability, students are able to determine that they lack skill and

116

confidence in math in particular. Table 4.23 gives cross-tabulations for the matched data, showing what latent class the model predicted students would transition into by grade five. These data display the proportion of students in each latent class pattern.

Table 4.23. *Longitudinal Proportions for Latent Class Patterns from LTA Results* (N=4703)*

Grade 3 Classes	Grade 5 Classes			
	Low SC	Average SC, Low Math	Average SC, High Math	High SC
Low SC, Low Math	.067	.157	.075	.055
Average SC	.104	.121	.144	.084
High SC	.041	.053	.054	.044

*Based on Most Likely Latent Class Patterns

The LTA model predicted that the largest proportion of students would be found in the average classes, across the two time periods. Twelve percent of the sample transitioned from the "Average SC" class in grade three to the "Average SC, Low Math" class in grade five. Another 14% transitioned from "Average SC" to "Average SC, High Math." Finally, the largest proportion of the sample started in the "low SC" group and transitioned to "Average SC, Low Math" group by the fifth grade (.157).

Finally, Table 4.24 gives the probability of transitioning to a specific class at time two, given class membership at time one.

Table 4.24 *Latent Transition Probabilities Based on the Estimated LTA Model (N=4703)*

Grade 3 Classes	Grade 5 Classes			
	Low SC	Average SC, Low Math	Average SC, High Math	High SC
Low SC, Low Math	.198	.431	.214	.157
Average SC	.225	.275	.321	.178
High SC	.217	.265	.301	.216

117

These data show that the greatest conditional probability is for transitioning from the "Low SC, Low Math" class in grade three to the "Average SC, Low Math" class in grade five (.431). The least likely transitions were found from "Low SC, Low Math" and "Average SC" in grade three to "High SC" in grade five (.157 and .178, respectively). The highest probabilities occurred when transitioning up to the next highest category at time two (e.g., from "Average SC" in grade 3 to "Average SC, Low Math" and "Average SC, High Math" in grade five).

Two group gender analysis. Additional analyses were conducted to examine the proportions of males and females that transitioned from one class to another. Findings revealed that greater proportions of males moved from the "Low SC, Low Math" class in grade three to the "High SC" class in grade five (.063 for males and .035 for females). Similar results were found with more males moving from the other two classes in grade three ("Average SC" and "High SC") to "High SC" in grade five. Females, on the other hand, moved in greater numbers than males from the "Low SC, Low Math" class to only the two average classes in grade five. Fewer females than males dropped from the "High SC" class in grade three to the "Average SC, Low Math" class in grade five (.046 for males and .026 for females).

Proximal and distal outcomes. Distal outcomes can be analyzed to provide information that shows how the outcomes at time two relate to class membership at time one. In the present study, reading and math test scores were added to the latent transition model to determine if there was any relationship between self-concept perceptions and achievement. The analysis in Table 4.25 factored IRT scores into the LTA as proximal outcomes at each grade level.

Table 4.25. *Relationship between Classes and IRT Scores over Time*

Reading		Math	
Grade 3	**Grade 5**	**Grade 3**	**Grade 5**
Low SC, Low Math	Low SC	Low SC, Low Math	Low SC
125.303 (27.891)	123.433 (20.455)	97.043 (23.905)	97.062 (19.863)
Average SC	Average SC, Low Math	Average SC	Average SC, Low Math
128.148 (28.342)	159.200 (16.765)	101.629 (24.681)	131.999 (15.561)
	Average SC, High Math		Average SC, High Math
	128.803 (23.293)		103.738 (22.244)
High SC	High SC	High SC	High SC
120.792 (28.777)	174.163 (13.145)	92.770 (25.300)	146.277 (10.865)

For students who remained in the low self-concept latent class over the two year period (those showing stability), their IRT scores were nearly identical. This was true for both math and reading. On the other hand, when students started in the "High SC" class in third grade and remained stable in that class, their scores in both reading and math increased by more than 50 points. Interestingly, when students who were in the "Average" class at grade three and transitioned to the average class in grade five, their scores in both subjects only changed if their math self-concept was low.

For distal outcomes, the analysis was accomplished by estimating the mean of the grade five IRT scores for each of the grade five class statuses (Nylund, 2007, Nylund et al., 2006). The resulting means were then tested for significant differences with the Wald Test of Parameter Constraints, see Table 4.26.

119

Table 4.26. *Wald Difference Test Results, LTA with Distal Outcomes at Grade Five*

Subject	Class	IRT Score (SD)	Wald Test-Significant Results**
Math	Low SC (M1)	126.204 (25.15*)	Wald X^2_{M1-M4} (1,N=4703) = 15.576, p=.0001
	Average SC, Low Math (M2)	124.108 (25.15*)	Wald X^2_{M2-M4} (1,N=4703) = 12.106, p=.0005
	Average SC, High Math (M3)	125.136 (25.15*)	Wald X^2_{M3-M4} (1,N=4703) = 13.312, p=.0003
	High SC (M4)	115.067 (25.15*)	
Reading	Low SC (R1)	138.539 (26.25*)	Wald X^2_{R1-R4} (1,N=4703) = 5.918, p=.015
	Average SC, Low Math (R2)	151.210 (26.25*)	Wald X^2_{R1-R2} (1,N=4703) = 13.651, p=.0002
	Average SC, High Math (R3)	155.049 (26.25*)	Wald X^2_{R1-R3} (1,N=4703) = 39.531, p=.0000
	High SC (R4)	1146.607 (26.25*)	Wald X^2_{R3-R4} (1,N=4703) = 6.054, p=.0139

*Model estimation constrained the variances between classes to be equivalent.
**Only significant mean differences are reported.

The results of the distal outcomes analyses show that there are significant differences in math IRT scores between students in the "Low SC" and the "High SC" groups. Differences were also found between the "Average SC, Low Math" and the "High SC," and the "Average SC, High Math" and the "High SC" classes. For reading IRT scores, significant differences were found between the "Low SC" and "High SC" classes. There were also differences between the "Low SC" and both average classes. Significant differences were even found between the "Average SC, High Math" and the "High SC" class.

Conclusions

To address the goals of this dissertation and to answer the related research questions a series of data analyses were undertaken. Each stage of the analysis built on the previous and concluded with a full longitudinal model that represented the relationships between peer self-concept, other academic forms of self-concept, and numerous covariates and outcome measures. Taken as a whole the person-oriented techniques utilized allow examination of the complex, interrelated influences on students in classroom and how these influences shape perceptions and behavior.

The final LTA model resulting from the LPA's at each grade level with covariates and proximal and distal outcomes is shown in Figure 4.11
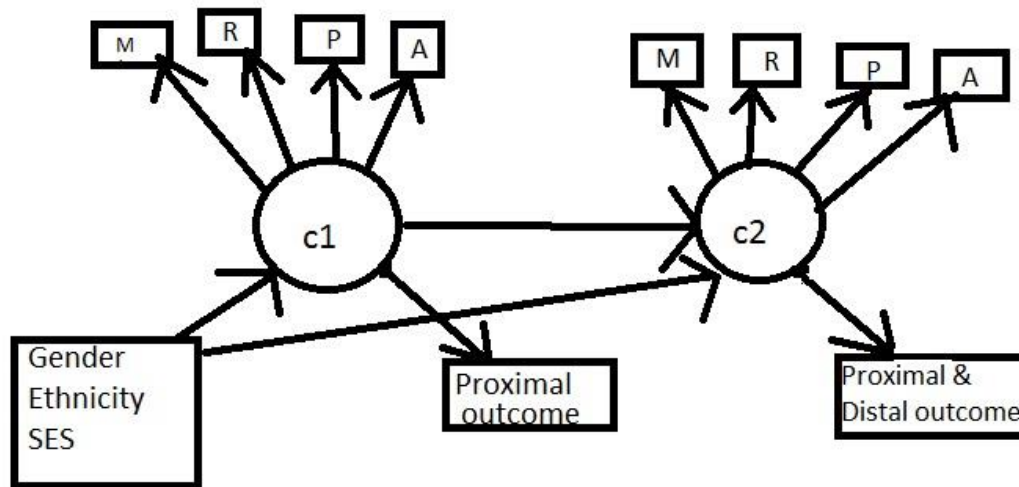

Figure 4.11 Latent Transition Model with Covariates and Proximal/Distal Outcomes

Covariates: Gender, ethnicity, SES
C1=LPA at grade 3; C2=LPA at grade 5
M=*SDQ-I* Math, R= *SDQ-I* Reading, P= *SDQ-I* Peer Relations, A= *SDQ-I* All Subjects

Chapter V

Discussion

The present study was undertaken to fill a gap in the self-concept literature base, that of peer self-concept. While, self-concept is one of the most studied constructs of self-perception, the existing research has primarily addressed academic self-concept. Studies have largely ignored the importance of peer self-concept and its relationship to academic performance.

Social self-concept was a major subdivision of the original model proposed by Shavelson et al. (1976) and refined by Marsh and Shavelson (1985), yet it has not received much consideration since then. Social self-concept was not described or investigated for construct validity until Byrnes and Shavelson's 1996 research, 20 years later. There has been even less research into the structure or functioning of the peer self-concept subdomain. In fact, the results of a comprehensive literature search only found one study that addressed peer self-concept. It is clear that little is known about how the peer self-concept construct functions and how it influences behavior. The current study was conducted to attempt to clarify how peer self-concept interacts with academic self-concept and how latent profiles across these constructs change as children in elementary school mature from 3rd to 5th grade.

Differences between Peer and Academic Self-Concept Constructs

There were a number of findings from the current study that help to describe how peer and academic self-concept relate to each other. From examining simple first-order

122

correlations between the *SDQ-I* latent factors, the current study found moderate relationships between the Peer and the All School Subjects scales at both grade levels (.424 in grade three and .404 in grade five). When looking at the Reading and Math scales, the correlations were lower and less stable. For example, correlations between the Peer and Math self-concept scales were .309 in grade three and .258 in grade five. Similarly, for the Peer and Reading scales the correlations were .324 for grade three and .277 for grade five. These differences, while cross-sectional and only descriptive in nature, point to possible differences in the academic and social self-concept constructs that warrant further study. Specifically, since correlation values are lower between the Peer scale (social self-concept) and the two academic scales, this suggests that childrens' self-perceptions in academic and nonacademic areas might be distinct. Also, since the correlations decline over time, it appears that there may be more differentiation between social and academic self-concept as students mature. These results are in line with those found by Marsh and colleagues when researching the developmental differentiation of academic self-concept (Marsh, Parker, et al., 1983; Marsh, Smith, et al., 1983, 1984).

Findings obtained when conducting analyses to establish the measurement models, indicated that academic *SDQ-I* items had high residuals and modification indices, while peer *SDQ-I* items did not. These measures show that there was more error in model estimation for the academic *SDQ-I* items. Again, this is merely descriptive information that may indicate that peer self-concept is a distinctly different construct from self-concept in the academic areas.

123

Developmental Changes in *SDQ-I* Measures

Research on academic self-concept has established that, with age, self-perceptions become more specialized. Students evidence fewer overall, global self-perceptions and display more differentiated, content-specific, self-concepts as they mature (Marsh, 1990a; Marsh & Hocevar, 1985; Marsh & Shavelson, 1985). Byrnes and Shavelson (1996) established that social self-concept was also organized in a hierarchy showing increasing differentiation with age. Unfortunately, there has been very little published since the Byrne and Shavelson study. In fact, only one study investigating the differentiation of social self-concept over time was found through a comprehensive search of the literature. The current study's findings work to fill the gap in the literature base in this area by showing support for developmental change in both peer and academic *SDQ-I* scores, both from a descriptive analysis and from person-oriented results.

Factor scores were used to represent each students' position on the latent class variable derived from the LPA's conducted at each grade. The factor scores distributions at grades three and five showed some marked differences. Negative skew was much more pronounced in grade three for all areas, except peer self-concept. Skew indicates that the factor scores were predominantly positive, with relatively few values below the mean. These results support the developmental research finding that younger children like everything and believe they are good at everything (Byrne & Shavelson, 1996; Harter, 1988; Marsh & Craven, 1997; Marsh et al., 1998; Wigfield & Eccles, 2002). By grade five, students appear to have become more realistic about their interest and competence

across the different self-concept domains. This is an area that would benefit from focused research aimed at determining the statistical significance of these differences between age groups.

In addition, through examining descriptive statistics, a decline in the level of positive self-perceptions across all *SDQ-I* scales occurred from grade three to grade five. The decline in terms of weighted mean scores (on a 4-point scale) for math was .19 and for reading was .27, from grade three to grade five. However, the decline for peer self-concept was only .08 for the same time period. Peer self-concept appeared to be more stable across the two-year period. Again, these results are only descriptive, but they suggest areas for future research into the changes in peer and academic self-concept as children mature.

Contributions from Person-Oriented Methodology

Existing research studies related to academic self-concept established construct validity through the use of confirmatory factor analysis, path analysis, structural equation and covariance modeling techniques, all variable-oriented approaches (Arens et al., 2011; Lindner-Muller et al., 2012; Marsh, 1987; Marsh, Craven et al., 1997, 1998; Marsh et al., 2002; Marsh, Parker et al., 1983; Marsh, Smith et al., 1984). Even within the area of academic self-concept, there have only been a handful of studies that have used person-oriented research approaches (De Fraine et al., 2007; Marsh et al., 2009; Van de Schoot & Wong, 2012).

The Marsh et al. (2009) study was the only one to investigate the possibility that latent profiles could better describe the complexity of self-concept across numerous domains. Marsh and colleagues aimed to determine whether differences in self-concept

125

profiles could be explained as qualitative differences (students have different shaped profiles of self-concept where one academic domain has higher levels of self-concept than others) or quantitative differences (students have high, medium, or low self-concepts consistently across all domains). Through LPA, Marsh and colleagues found that both qualitative and quantitative differences occurred between different profiles. Unfortunately, there were weaknesses with the Marsh study in that it did not include any measure of social or peer-self-concept. The self-report instrument that was used was a German adaptation of the *SDQ-III* with new scales that did not have any supporting validity or reliability evidence. In addition, this investigation did not provide any information about changes in classes or profiles over time.

Person-oriented research methodology has not been used to investigate how peer self-concept perceptions form distinct classes and how these classes differentiate over time. The present study addresses this void by being the first to use model-based methodology (LPA and LTA analyses) to study the interrelatedness of peer and academic self-concept. Given the gaps in the research base, the present study makes a significant contribution to furthering knowledge in the field.

Findings from Latent Profile Analysis

The present study provides additional support for self-concept differentiation. Specifically, findings indicate that latent class structures are more global in grade three and more domain-specific by grade five. This was found to be particularly true for math self-concept. The findings are detailed below.

Based on the Marsh 2009 research, it was anticipated that latent classes resulting from LPA would show differences in the level of overall self-concept (i.e., classes would

form around high, medium, and low levels of self-concept across all four domains).

Difference in profiles shapes were also expected. Specifically, it was expected that some students would have higher reading self-concept, while others would have higher math self-concept. Finally, it was expected that some students would have higher self-concept related to peer relations and lower self-concept in the academic areas.

The LPA results did not uncover any classes with profiles indicating high levels of peer self-concept at either of the two grade levels. Instead, the results indicated that peer self-concept varied at roughly the same level as academic self-concept. It appears that student with high profiles feel as if they excel at everything, including relationships with their friends. Similarly, those with low profiles have negative perceptions across all areas, including their relationships with others. These results parallel the findings from much of Wentzel's extensive research on social competence by showing that social competence tends to be closely related to academic competence (Wentzel, 1989, 1991a, 1991b, 1993, 1998, 2009; Wentzel & Watkins, 2002).

In addition, LPA results showed that classes formed around a specific level of interest or perceived competence across the four self-concept domains. Results supported differences in quantity (elevation) instead of quality (profile shape) of the resulting classes (Marsh et al., 2009). In grade three, results showed high, medium and low groupings with very little variation across the different self-concept domains.

Pronounced differences occurred in the resulting class profiles at grade five. Here math self-concept emerged as a specific profile. A four-class solution was the optimal model with the "average SC" class from grade three being split into two separate groups. These two groups at grade five differed primarily on math self-concept responses. The

127

two profiles display polarization, where one group is interested in and/or feels competent at math above all else (the "math lovers") and the other group is not interested in and/or does not feel competent at math (the "math haters"). The "math lover" group displayed an extreme profile where the reading self-concept mean score was below that of even the "Low SC, Low Math" class. Conversely, the math self-concept mean score was higher than even that of the "High SC" class.

The LPA results for fifth grade are also confirmed by the distribution of grade five factor scores. The math factor scores formed a multi-modal distribution indicating 3 groups: one at the mean, one that is interested in and/or feels competent at math (.5 SD above the mean) and another that is not interested in and/or does not feel competent in math (.5 SD below the mean). This type of distribution does not occur with factors scores from reading or in any other areas of grade three or grade five. The factor score distribution and LPA evidence together support the conclusion that differentiation and specialization in content area self-concept occurred as children moved from middle to late elementary school. The polarization of math self-concept scores at grade five is deserving of further research. An initial literature search did not produce any studies that specifically addressed why students feelings of confidence, competence, or interest in math change at this age. This may be because the present study is examining students' self-perceptions of competence and interest, not actual academic achievement. This is an area that is deserving of future research from a person-oriented point of view.

Latent Class Membership by Demographic Category

The results of a covariate analysis in grade three showed that males were more likely to be in the "Low SC, Low Math" and the "Average SC" classes (log odds of .600

128

and .648 respectively). This means that females were more likely to be in the "High SC" group reflecting higher levels of interest and perceived competence across self-concept domains. Blacks and Hispanics were more likely than whites to be in the first and second classes ("Low SC, Low Math" and "Average SC" groups), reflecting generally lower levels of perceived competence and interest. Average SES was roughly equivalent across the classes in grade three. In grade five, low SES, blacks, Hispanics and Asians were all more likely than whites to be in the "Low SC" class.

In terms of gender, the fact that more girls are found in the "High SC" group could be explained by gender expectations. Girls often try to perform well and follow the social rules of the classroom to please others. This could explain why girls have greater feelings of interest and competence across the different self-concept domains (Fan, 2011; Goodenow & Grady, 1993; Voyer & Voyer, 2014). The parallel between the covariate results obtained in the present study and established research on gender and ethnic differences in achievement provides support for the validity of the final LPA models and class structures.

Longitudinal Changes in Self-Concept

The latent transition analysis conducted in the present study described how class membership changed as students moved from third to fifth grade. Cross-sectional results showed that almost a third of the students (29.0%) who rated themselves as average in terms of interest and competence across the four self-concept domains ("Average SC") in grade three, fell to the "Low SC" class in grade five. Similarly, about a third of the students who rated themselves as "High SC," in grade three fell to the "Average SC, Low Math" group in grade five (32.8%). These changes suggest that self-concept might

129

decline over time or that with maturation students become more realistic and are better able to evaluate their academic abilities. Even though these are speculations, other research has supported both of these conclusions (Marsh & Craven, 1997; Marsh et al., 1998; Harter, 1988; Wigfield & Eccles, 2002).

When examining the longitudinal, model-based probability for moving from one class to another, the results are quite different. The highest transitional probabilities occurred for moving up to the next highest category in grade five from grade three (e.g., from "Low SC, Low Math" in grade three to "Average SC, Low Math" in grade five; or from "Average SC" in grade three to "Average SC, High Math" in grade five). The least likely transitions were found from "Low SC, Low Math" and "Average SC" in grade three to "High SC" in grade five (.157 and .178, respectively).

Results from multi-group gender analyses showed that greater proportions of males moved from the "Low SC, Low Math" class in grade three to the "High SC" class in grade five (.063 for males and .035 for females). Females, on the other hand, moved in greater numbers from the "Low SC, Low Math" class to the two average SC classes in grade five. Fewer females dropped from the "High SC" class in grade three to the "Average SC, Low Math" class in grade five (.046 for males and .026 for females). These results suggest that males increase their interest or perceptions of competence across areas more than females do when they move from third to fifth grade.

For students who remained in the "Low SC" latent class in grade five, their IRT scores were nearly identical at each grade. This was true for both math and reading. On the other hand, when students started in the "High SC" class in third grade and remained in that class at fifth grade, their scores in both reading and math increased by roughly 50

130

points. These results may reflect the reciprocal effects (REM) model where increases in self-concept to high levels fuels increases in academic performance (Marsh & Martin, 2008; Marsh et al., 2009, 2012).

Another interesting finding was that when students who were in the "Average SC" class at grade three and stayed in an average SC class in grade five, their test scores in both reading and math only changed if their math self-concept was low in grade five ("Average SC, Low Math" class). For this group, there was an increase of roughly 30 points in both subject areas. This may be interpreted as such, students who have more confidence or interest in math (from the "Average SC, High Math" group) are complacent or do not try to improve, thus their academic performance does not change between third and fifth grades. A more in-depth, qualitative study (e.g., focused interviewing of students and teachers, detailed student observations) could help uncover the processes that are occurring during the period between third and fifth grade to impact achievement test scores.

How the Study Results Answer the Research Questions

Research Question 1. *Do student responses from the self-report scales used with the ECLS-K (SDQ-I, grade 3 and grade 5) form latent clusters that relate to the multi-dimensional structure of the Marsh/Shavelson conceptual model (1985) of self-concept? In other words, are there resulting class structures where some students have higher reading self-concept, while others have higher math self-concept. Is there a resulting class where some students have higher self-concept related to peer relations and lower self-concept in the academic areas?*

LPA results using mean factor scores from the *SDQ-I* scales showed that the resulting classes were based on overall perceptions of competency and interest across each of the four domains. In grade five, one of the domain specific self-concept measures did differentiate class membership. Math self-concept created a differentiation between the two average competency classes. One of the resulting classes had high perceptions of interest and competency in math and the other had low perceptions.

Research Question 2. *Is there a significant association between the outcomes of math and reading achievement and the resulting latent classes at each grade level?*

In grade three, the lowest IRT scores were found with the students in the "high SC" class. This occurred in both reading and math. The "Average SC" class had the highest test scores in both areas. In grade five, the progression from lowest to highest scores in both subjects was from "Low SC" to "Average SC, High Math" to "Average SC, Low Math" to "High SC." Again, the "Average SC, Low Math" has higher achievement scores with lower self-perceptions of interest and competence in math.

Research Question 3. *Do the latent class profiles identified in grade 3 change or remain stable over time (i.e., from grade 3 to grade 5)?*

The highest transitional probability was for students to move to the next higher class in grade five. The least likely transitions were found from "Low SC, Low Math" and "Average SC" in grade three to "High SC" in grade five. Greater proportions of males moved from the "Low SC, Low Math" class in grade three to the "High SC" class in grade five. Females were less likely than males to drop from the "High SC" class in grade three to the "Average SC, Low Math" class in grade five.

Research Question 4. *Do latent class statuses at grade 3 predict achievement performance in math and reading at Grade 5?*

To answer this research question, grade five IRT scores were regressed on grade three latent class membership and were used to predict distal outcomes of achievement test scores. For all classes there was a consistent relationship between membership at grade three and test scores at grade five.

- Students who were in the "Low SC, Low Math" latent class in grade three, had the greatest probability of transitioning to the "Average SC, Low Math" latent class in grade five (.431, latent transition probability). This would translate to an increase in IRT scores from grade three to five in both reading (gain of 33.9 points) and math (gain of 35.0 points).

- Students who were in the "Average SC" latent class in grade three, had the greatest probability of transitioning to the "Average SC, High Math" latent class in grade five (.321, latent transition probability). This would translate to an increase in IRT scores from grade three to five of only .65 of a point in reading and 2.11 points in math.

- Students who were in the "High SC" latent class in grade three, had the greatest probability of transitioning to the "Average SC, High Math" latent class in grade five (.301 latent transition probability). This would translate to an increase in IRT scores from grade three to five in both reading (gain of 8.0 points) and math (gain of 10.9 points).

It appears from examining the achievement score patterns in these result that students with lower levels of self-concept may react to better academic performance with

133

increased self-concept. This conceptualization would support the existing research base from Marsh's reciprocal effects model (Marsh, 1990b; Marsh et al., 1999; Marsh & Craven, 2006; Marsh et al., 2002; Marsh & Martin, 2011; Marsh & O'Mara, 2008; Marsh et al., 2005; Marsh & Yeung, 1998; William & Williams, 2010). Additional research should be conducted in an effort to understand the processes underlying these results. This research should include qualitative methodology (e.g., interviews, case studies), so that students can be questioned about their interest and competency beliefs at various times as they progress from grade three to grade five.

Contributions from the Current Study

The present study was conducted to better understand the relationship between student perceptions of their interest and competence within math, reading and all school subjects, and how these relate to their perceptions of interest and competence with respect to peer relations. In addition, this study addressed gaps in the existing research on self-concept, by addressing elementary students with the use of a multiyear longitudinal design.

Focused on the middle to late elementary school age group. This study adds to current research by focusing on students in the upper grades of elementary school, grades three through five. Previous studies related to social competency and peer relationships have primarily been conducted in middle school, grades six through nine (Ferla et al., 2009; Goodenow, 1993; Hamm & Faircloth, 2005; Kindermann, 1993, 2007; Kindermann & Skinner, 2009; Ryan, 2001; Wentzel, 1991a, 1991b, 1996, 1997). The one study found that directly related to peer self-concept focused on students in grades 9-12 (Connelly & Konarski, 1994).

134

It is clear from the results of this study that there are changes in students' self-perceptions as they move from grade three to five. Again, this could be an indication that students become more aware of their competencies and are more realistic in their self-assessments. Older students are also more differentiated in terms of their self-concept. Instead of focusing on general perceptions of ability, they are more able to evaluate their competencies and interests within specific content domains. By focusing on the later years of elementary school a more complete developmental perspective is gained with respect to the changes in peer and academic self-concept as children mature. Research on this age group also provides needed information about the relative importance of students' perceptions of their interest and competence in academic areas as compared to peer relationships at this stage of development.

Longitudinal analysis. Many of the studies investigating the impact of self-perceptions of cognitive abilities and relationships with peers on achievement motivation have used cross-sectional (Dowson & McInerney, 2001; Masland & Lease, 2013; Oberle & Schonert-Recihl, 2013; Wentzel 1989, 1991a), or short- term longitudinal analyses that continued only from the fall to the spring of the same school year (Hamm & Faircloth, 2005; Hamm et al., 2011; Kindermann, 2007; Kindermann & Skinner, 2009; Ryan, 2001; Wentzel, 1998). It has only been common for researchers to include multiple years of data in their designs when studying reciprocal effects of self-concept (Guay et al., 2003; Marsh, Hau et al., 2002; Marsh & O'Mara, 2008; Pinxten et al., 2010).

The current study was not only a two year longitudinal study, but it also employed a latent class repeated measures technique. LTA provides a method for examining latent class characteristics and predicting changes in these characteristics, as well as class

135

membership, over time. In addition, LTA can incorporate distal outcomes so class membership can be used to predict achievement scores and other outcome variables occurring at a later stage of development.

Methodological Approach. The present study utilized latent profile instead of latent class methodology to accommodate the continuous nature of latent factor scores. This is one of few studies reporting the use of latent profile analyses (for others see Dowdy et al., 2014; Pastor et al., 2007). Generalizability of LPA and LTA results were also confirmed by conducting cross-validation procedures. An independent hold-out sample was used to re-estimate the models and confirm the superiority of the final models selected as part of both the LPA and LTA procedures.

Limitations

There were problems with lack of clarity in the wording of the *SDQ-I* instrument. The All School Subjects questions asked about school subjects but did not instruct the child not to consider math or reading. This could influence the construct validity of this measure as students may be thinking of their interest and competence in either math or reading instead of addressing different content areas when they respond to these questions. In that case, the factor being represented by this scale would not be content areas other than reading and math.

It would be advisable to repeat the investigation into peer self-concept and its relationship to areas of academic self-concept with another more contemporary sample, as the ECLS-K data used for this study is dated from 2002 and 2004, more than ten years ago. The ECLS-K data also had more white respondents than minority. Even though the

data was weighted there could still be some cultural differences in the responses obtained from a primarily majority culture.

This study was also limited by the response scales and sample variables that were part of the ECLS-K dataset. In particular, the items assessing peer relations do not really get at feelings of belonging or social competence, aspects of peer relations shown to have significant effects on achievement (Deci & Ryan, 2008; Dowson & McInerney, 2001; Finn, 1989; Goodenow, 1993; Goodenow & Grady, 1993; Han et al., 2013; Martin & Dowson, 2009; Molloy et al., 2011; Oberle & Schonert-Reichl, 2013; Ryan & Deci 2000). Future research should supplement the *SDQ-I* measures with items that have a stronger relationship to feelings of belonging.

Additional research should also aim to supplement self-report items with a richer mix of qualitative data sources like interviews, classroom observations, and responses from significant others (e.g., peers, teachers, parents). Triangulation from these different sources of information will help to clarify the thought processes that underlie student perceptions and how these perceptions influence their academic motivation and performance.

Another limitation is the lack of significance testing available for differences between groups when using LCA and LTA. To test for the statistical significance of differences across classes, groups would need to be formed on variables not used in the creation of the latent classes. These analyses are beneficial in that they help to establish the validity of the resulting class structure. However, conducting separate analyses for verifying the significance of group differences was beyond the scope of the current study.

137

A couple of technical issues posed limitations when conducting data analyses. First, was the lack of convergence within some of the complex models being estimated as part of the latent transition analysis. Future research should investigate constraining some of the transitional parameters to equality. Secondly, the sampling weights necessary to make inferences to the larger population could not be accommodated with the Bootstrap Likelihood Ratio Test (BLRT) in Mplus, version 7.4.

Areas for Future Research

Person-oriented analyses. Understanding the complexity and interaction of forces impacting students in the classroom will help educators tailor educational experiences for individual students (i.e., peer coaching, establishing peer networks, peer-assisted leaning, and cooperative work groups). For educators it is as important to understand changes occurring within a child over the course of time as it is to understand the differences between students. Information about the complex array of factors impacting students can only be gained through the use of person-oriented research techniques. Based on the latent profiles and transitions conducted during the current research, it appears that students with high perceptions of math self-concept do not perform academically as well as students with lower math self-concept. Examining the underlying processes and direction of causation should be undertaken using person-oriented methodology within the reciprocal effects model framework (Arens et al., 2011; Marsh & Martin, 2008, 2011; Marsh et al., 2005, 2009, 2012; Pinxten et al., 2010).

138

Self-concept interventions. Further research should focus on understanding the relationship between self-perceptions of competence and interest and how these relate to actual academic performance. A number of specific patterns were found based on the latent transition results from this study:

- When students transitioned to the next higher self-concept class in fifth grade, their achievement test scores improved. This suggests that increases in self-perceptions of interest and competence may have some influence on achievement motivation.

- Students who had high self-concepts in grade three and stayed in the high self-concept class ("high SC") in grade five, had the lowest test scores in grade three. However, by the time they got to grade five they had the highest tests scores, with a gain of roughly 50 points in both reading and math. This may indicate that high self-perceptions of interest and competence bolstered student's achievement motivation. However, the transition probabilities for these students (staying in the high SC in grade three and grade five) are the lowest (.216). This indicates that most of the students that are in the high SC class in grade three are predicted to have lower self-concept in grade 5. This may mean that most student that are in the high SC class in grade three become more realistic about their specific abilities.

- For those in the low SC class in grade three, their achievement performance improves no matter what class they transition to in grade five, as long as they do not stay in the low self-concept group.

139

- For students who were in the average SC group in grade 3 as long as they do not decline in self-concept they maintain or improve achievement performance.

Further research should be conducted to determine the underlying causes for these relationships between self-concept and achievement. If these changes were better understood interventions could be implemented that would help students improve their self-perceptions of interest and competence across self-concept domains. Hopefully, the improvements in self-concept would reflect realistic assessments of competence so that improvements in achievement would follow.

Conclusion

The current study filled a gap in existing self-concept literature by examining peer self-concept and how it relates to other academic self-concept domains. Complex interactions between students and their environments were modeled through the use of person-oriented methodology. Latent profile and latent transition analyses provided insight into the way academic and peer self-concept interacted within elementary school students and how these may be involved in achievement motivation. Latent transition models revealed how self-concept and academic performance changed as children matured. Additional information about the complex relationships between the different domains of self-concept, individual characteristics, and achievement were obtained through the use of covariate analyses and distal outcomes. Both LPA and LTA are model-based ways to determine groups and assess changes in membership over time. The present study was the first to apply these methodologies to the study of peer self-concept.

140

REFERENCES

Adelson, J. L., McCoach, D. B., & Gavin, M. K. (2012). Examining the effects of gifted programming in mathematics and reading using the ECLS-K. *Gifted Child Quarterly*, *56*, 25-39.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52(3),* 317-332.

Aldenderfer, M.S., & Blashfield. R.K. (1984). *Cluster analysis.* Beverly Hills, CA: Sage.

Altermatt, E.R. (2011). Capitalizing on academic success: Students' interactions with friends as predictors of school adjustment. *Journal of Early Adolescence, 31*, 174-203.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261.271.

Andruff, H., Carraro, N., Thompson, A., Gaudreau, P., & Louvet, B. (2009). Latent class growth modelling: a tutorial. *Tutorials in Quantitative Methods for Psychology*, *5*(1), 11-24.

Arens, A.K., Yeung, A.S., Craven, R.G., & Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology, 103*(4), 970-981.

Atkins-Burnett, S., & Meisels, S. J. (2001). *Measures of socio-emotional development in middle childhood.* (Working Paper No. 2001-03). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Baker, J.A. (2006). Contributions of teacher-child relationships to positive school adjustment during elementary school. *Journal of School Psychology, 44*, 211-229.

Bandura, A. (1986). *Social foundation of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*(9), 1175-1184.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9-44.

Bergman, L.R. (1998). You can't classify all of the people all of the time. *Multivariate Behavioral Research*, *23*(4), 425-441.

Bergman, L. R., & El-Khouri, B. M. (1999). Studying individual patterns of development using I-states as objects analysis (ISOA). *Biometrical Journal*, *41*(6), 753-770.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, *9*(02)*,* 291-319.

Bergman, L. R., Magnusson, D., & El Khouri, B. M. (2003). *Studying individual development in an inter-individual context: A person-oriented approach*. Hove, United Kingdom: Psychology Press.

Bergman, L. R., & Trost, K. (2006). The person-oriented versus the variable-oriented approach: Are they complementary, opposites, or exploring different worlds? *Merrill-Palmer Quarterly*, *52*(3), 601-632.

Bergman, L. R., & Wangby, M. (2014). The person-oriented approach: A short theoretical and practical guide. *Estonian Journal of Education*, *2*, 29-49.

Berndt, T.J., & Bergy, L. (1996). Social self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept. Developmental, social, and clinical considerations* (pp. 171-209). New York, NY: John Wiley & Sons.

Biddle, B. (2014). *Social class, poverty and education*. London: Routledge.

Bong, M. (1998). Tests of the internal/external frames of reference model with subject-specific academic self-efficacy and frame-specific academic self-concepts. *Journal of Educational Psychology*, *90*(1), 102-110.

Bong, M., & Clark, R.E. (1999). Comparisons between self-concept and self-efficacy in academic motivation research. *Educational Psychologist, 34*(3), 139-153.

Bong, M., & Skaalvik, E.M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1-40.

Bouchey, H. A., & Harter, S. (2005). Reflected appraisals, academic self-perceptions, and math/science performance during early adolescence. *Journal of Educational Psychology*, *97*(4), 673-686.

Brookover, W., Thomas, S., & Paterson, A. (1964). Self-Concept of Ability and School Achievement. *Sociology of Education, 37*(3), 271-278.

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*(1), 108-132.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model

fit. *Sociological Methods & Research*, *21*(2)*,* 230-258.

Byrne, B. M. (1996). Academic self-concept: Its structure, measurement, and relation to

academic achievement. In B. A. Brackett (Ed*.), Handbook of self-concept:*

Developmental, social, and clinical considerations (pp. 287-316).

New York, NY: Wiley.

Byrne, B. M. (2002). Validating the measurement and structure of self-concept:

Snapshots of past, present, and future research. *American Psychologist,*

57, 897-909.

Byrne, B. M., & Shavelson, R. J. (1986). On the structure of adolescent self-

concept. *Journal of Educational Psychology*, *78*(6), 474-481.

Byrne, B. M., & Shavelson, R. J. (1996). On the structure of social self-concept for pre-,

early, and late adolescents: A test of the Shavelson, Hubner, and Stanton (1976)

model. *Journal of Personality and Social Psychology*, *70*(3)*,* 599-613

Calsyn, R. J.; & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of

others: Cause or effect of academic achievement? *Journal of Educational*

Psychology, 69(2), 136-145.

Cocks, R.J., & Watt, H.M.G. (2004). Relationships among perceived competence,

intrinsic value and mastery goal orientation in English and Maths. *The Australian*

Educational Researcher, 31(2)*,* 81-111.

Collins, L. M., Graham, J. W., Rousculp, S. S., & Hansen, W. B. (1997). Heavy caffeine

use and the beginning of the substance use onset process: An illustration of latent

transition analysis. In K. J. Bryant & M. Windle (Eds.), *The science of prevention:*

Methodological advances from alcohol and substance abuse research (pp. 79-99). Washington DC: American Psychological Association.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Hoboken, NJ: John Wiley & Sons.

Connolly, J. A., & Konarski, R. (1994). Peer self-concept in adolescence: Analysis of factor structure and of associations with peer experience. *Journal of Research on Adolescence*, *4*(3)*,* 385-403.

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin.

Cosden, M., Larsen, J. L., Donahue, M. T., & Nylund-Gibson, K. (2015). Trauma symptoms for men and women in substance abuse treatment: a latent transition analysis. *Journal of Substance Abuse Treatment*, *50*, 18-25.

Crocker, L., & Algina, J. (2008) *Introduction to classical and modern test theory*. Mason, OH: Cengage Learning.

Darensbourg, A.M., & Blake, J.J. (2014). Examining the academic achievement of black adolescents: Importance of peer and parental influences. *Journal of Black Psychology, 40*(2)*,* 191-212.

Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Deci, E.L., & Ryan, R.M. (2000).The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11,* 227-268.

Deci, E.L., & Ryan, R.M. (2008). Self-determination theory: A macro-theory of human motivation, development, and health. *Canadian Psychology, 49*(3)*, 182-185.

Dekovic, M., & Meeus, W. H. J. (1997). Peer relations in adolescence: Effects of parenting and adolescents' self-concept. *Journal of Adolescence, 20,* 163-176.

De Fraine, B., Van Damme, J., & Onghena, P. (2007). A longitudinal analysis of gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach. *Contemporary Educational Psychology*, *32*(1), 132-150.

Dewey, J. (1929). *The quest for certainty (Vol. 4, The Later Works: 1925-1953).* New York: Minton, Balch.

Diseth, Å., & Samdal, O. (2014). Autonomy support and achievement goals as predictors of perceived school performance and life satisfaction in the transition between lower and upper secondary school. *Social Psychology of Education*, *17*(2), 269-291.

DiStefano, C. (2012). Cluster analysis and latent class clustering techniques. In B. Larsen, T.D. Little, N.A. Card (eds.). *Handbook of developmental research methods* (pp.645-666). New York, NY: The Guilford Press.

DiStefano, C., & Kamphaus, R. W. (2006). Investigating subtypes of child development a *comparison* of cluster analysis and latent class cluster analysis in typology creation. *Educational and Psychological Measurement*, *66*(5)*, 778-794.

DiStefano, C. A., Kamphaus, R. W., & Mîndrilă, D. L. (2010). A typology of teacher-rated child *behavior*: Revisiting subgroups over 10 years later. *School Psychology Quarterly*, *25*(3)*, 152-163.

146

DiStefano C.A., & Mîndrilă, D.L. (2013). Cluster analysis. In T. Teo (Ed.) *Handbook of*
quantitative methods for educational research (pp. 103-122). Boston, MA: Sense
Publishers.

Distefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores:
considerations for the applied researcher. *Practical Assessment, Research &*
Evaluation, 14(20), 1–11.

Dowdy, E., Nylund-Gibson, K., Felix, E.D., Morovati, D., Carnazzo, K.W., & Dever,
B.D. (2014). Long-term screening for emotional and behavioral risk. *Educational*
and Psychological Measurement, 74(3), 453-472.

Dowson, M., & McInerney, M. (2001). Psychological parameters of students' social and
work avoidance goals: A qualitative investigation. *Journal of Educational*
Psychology, 93, 35-42.

Eccles, J. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.),
Achievement and achievement motivation (pp. 75–146). San Francisco, CA:
W. H. Freeman.

Eccles, J.S. (2005). Subjective task values and the Eccles et al. model of achievement
related choices. In A.J. Elliott & C.S. Dweck (Eds.), *Handbook of competence*
and motivation (pp.105-121). New York, NY: Guilford.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual*
Review of Psychology, 53, 109-132.

Ellis, L. A., Marsh, H. W., & Richards, G. E. (2002, August). A brief version of the Self-
Description Questionnaire II. In *Proceedings of the 2nd International Biennial*

147

Conference Self-Concept Research: Driving International Research Agenda.
Sydney: University of Western Sydney, SELF Research Centre.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling*, *8*(1)*,* 128-141.

Fan, W. (2011). Social influences, school motivation, and gender differences: an application of the expectancy-value theory. *Educational Psychology: An International Journal of Experimental Educational Psychology, 31*(2)*,* 157-175.

Fantuzzo, J. W., Davis, G. Y., & Ginsburg, M. D. (1995). Effects of parent involvement in isolation or in combination with peer tutoring on student self-concept and mathematics achievement. *Journal of Educational Psychology*, *87*(2)*,* 272-281.

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences, 19,* 499-505.

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

Finney, S. J., & DiStefano, C. (2013). Non-normal and categorical data in structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.) *Structural Equation Modeling: a Second Course* (pp.439-492).Charlotte, NC: Information Age Publishing, Inc.

Ford, M. E., & Nichols, C. W. (1991). Using goal assessments to identify motivational patterns and facilitate behavioral regulation and achievement. *Advances in Motivation and Achievement*, *7*, 51-84.

Freiberger, V., Steinmayr, R., & Spinath, B. (2012). Competence beliefs and perceived
 ability evaluations: How do they contribute to intrinsic motivation and
 achievement? *Learning and Individual Differences*, *22*(4), 518-522.

Froiland, J. M., & Oros, E. (2014). Intrinsic motivation, perceived competence and
 classroom engagement as longitudinal predictors of adolescent reading
 achievement. *Educational Psychology*, *34*(2), 119-132.

Gest, S. D., Rulison, K. L., Davidson, A. J., & Welsh, J. A. (2008). A reputation for
 success (or failure): the association of peer academic reputations with academic
 self-concept, effort, and performance across the upper elementary
 grades. *Developmental Psychology*, *44*(3), 625-636.

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic
 review of social, self-concept, and behavioral outcomes of peer-assisted
 learning. *Journal of Educational Psychology*, *98*(4), 732-749.

Goodenow, C. (1993). The psychological sense of school membership among
 adolescents: Scale development and educational correlates. *Psychology in the
 Schools, 30,* 79-90.

Goodenow, C., & Grady, K.E. (1993). The relationship of school belonging and friends'
 values to academic motivation among urban adolescent students. *Journal of
 Experimental Education, 62(1),* 60-71.

Grymes, J.M., & Lawler-Prince, D. (1993). Kindergartners' self-concepts and perceptions
 of peer interactions. *Paper presented at the Annual Meeting of the Mid-South
 Educational Research Association*. Retrieved from http://www.eric.ed.gov.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic

   achievement: Developmental perspectives on their causal ordering. *Journal of*

   Educational Psychology, *95*(1), 124-136.

Hamm, J.V., and Faircloth, B.S. (2005). Peer context of mathematics classroom

   belonging in early adolescence. *Journal of Early Adolescence, 25(3),* 345-366

Hamm, J.V., Schmid, L., Farmer, T.W., & Locke, B. (2011). Injunctive and descriptive

   peer group norms and the academic adjustment of rural early adolescents. *The*

   Journal of Early Adolescence, 31(1), 41-73.

Hamm J.V., & Zhang, L. (2010). The schooling context of adolescents' peer relations. In

   J. Meece & J. Eccles (Eds.), *The handbook of schooling effects on development*

   (pp. 518-554). Mahwah, NJ: Erlbaum.

Han, Q., Hu, W., Liu, J., Jia, X., & Adey, P. (2013). The influence of peer interaction on

   students' creative problem-finding ability. *Creativity Research Journal*, *25*(3),

   248-258.

Hancock, G. R., Harring, J. R., & Lawrence, F.R. (2013). Using latent growth modeling

   to evaluate longitudinal change. In G.R. Hancock & R.O Mueller (Eds.),

   Structural equation modeling: A second course (2nd edition, pp. 309-341).

   Charlotte, NC: Information Age Publishing.

Harter, S. (1982). The perceived competence scale for children. *Child Development*,

   53(1), 87-97.

Harter, S. (1983). Developmental perspectives on the self-system. In P. H. Mussen

   (Series Ed.) and E. M. Hetherington (Vol. Ed.), *Handbook of child psychology:*

150

Vol. 4: Socialization, personality, and social development (pp.275-385). New York, NY: Wiley.

Harter, S. (1985). *Manual for the self-perception profile for children.* Denver, CO: University of Denver.

Harter, S. (1988). Developmental processes in the construction of the self. In T. D. Yawkey & J. E. Johnson (Eds.), *Integrative processes and socialization: Early to middle childhood (pp.* 45-78). Hillsdale, NJ: Erlbaum.

Hay, I., & Ashman, A. F. (2003). The development of adolescents' emotional stability and general self-concept: The interplay of parents, peers, and gender. *International Journal of Disability, Development and Education*, *50*(1), 77-91.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, *70*(2), 151-179.

Higgins, E.T. (2007) Value. In A. W. Kruglanski, & E. Tory Higgins (Eds.), *Handbook of social psychology* (2nd edition, pp. 454-472). New York, NY: Guilford.

Hijzen, D., Boekaerts, M., & Vedder, P. (2006). The relationship between the quality of cooperative learning students' goal preferences, and perceptions of contextual factors in the classroom. *Scandinavian Journal of Psychology, 47*(1)*,* 9-21.

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

Ireson, J., & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction*, *19*(3), 201-213.

James, W. (1926). *The letters of William James.* Boston, MA: Little Brown.

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, 417-442.

Kaplan, A., & Midgley, C. (1997). The effect of achievement goals: does level of perceived academic competence make a difference? *Contemporary Educational Psychology*, *22*(4), 415-435.

Kim, H. Y., Schwartz, K., Cappella, E., & Seidman, E. (2014). Navigating middle grades: Role of social contexts in middle grade school climate. *American Journal of Community Psychology*, *54*(1-2), 28-45.

Kindermann, T. A. (1993). Natural peer groups as contexts for individual development: The case of children's motivation in school. *Developmental Psychology*, *29*(6), 970-977.

Kindermann, T. A. (2007). Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. *Child Development*, *78*(4), 1186-1203.

Kindermann, T. A., & Skinner, E. A. (2009). How do naturally existing peer groups shape children's academic development during sixth grade? *International Journal of Developmental Science*, *3*(1), 27-38.

Kirk, R.E. (2013). *Experimental design: Procedures for the behavioral sciences.* Los Angeles, CA: Sage.

152

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

Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: transitions in dating and sexual risk behavior. *Developmental Psychology*, *44*(2), 446-456.

Lawler-Prince, D., & Grymes, J.M. (1990). Public school kindergartners: Isolates, low self-concept, and family status. *ERIC Report ED 360 062.* Retrieved from http://www.eric.ed.gov.

Liem, A.D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology, 33*, 486-512.

Lim, H.J., & Kim, J. (2011). A longitudinal study of children's social behaviors and their causal relationship to reading growth. *Asia Pacific Education Review, 12* (2)*, 197-213.

Lindner-Muller, C., John, C., & Arnold K-H. (2012). Longitudinal assessment of elementary school students' social self-concept in relation to social preference. *Journal for Educational Research Online, 4*(1), 47-72.

Little, R.J., & Rubin D.B. (2002). *Statistical analysis with missing data.* New York, NY: Wiley.

Lo, Y., Mendell, N. R., & Rubin D.B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*(3), 767-778.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum Associates.

Marsh, H. W. (1984). Self-concept: the application of a frame of reference model to explain paradoxical results. *Australian Journal of Education, 28*, 165-181.

Marsh, H.W. (1987) The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analysis. *Journal of Educational Measurement*, *24,* 17-39

Marsh, H. W. (1988). *SDQ I manual & research monograph: Self-Description Questionnaire*. Psychological Corporation & Harcourt Brace Jovanovich.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, *81*(3), 417-430.

Marsh, H. W. (1990a). Causal ordering of academic self-concept and academic achievement: A multi-wave, longitudinal panel analysis. *Journal of Educational Psychology*, *82*(4), 646-656.

Marsh, H. W. (1990b). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology*, *82*(1), 107-116.

Marsh, H.W. (1990c). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2* (2), 77-172.

Marsh, H. W. (1990d). *Self-Description Questionnaire (SDQ) I: A theoretical and empirical basis for the* measurement *of multiple dimensions of preadolescent self-*

concept: A test manual and research monograph. New South Wales, Australia: Macarthur, University of Western Sydney.

Marsh, H. W. (1990e). *Self-Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and research monograph.* New South Wales, Australia: Macarthur, University of Western Sydney.

Marsh, H. W. (*1990*f). *Self-Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and research monograph.* New South Wales, Australia: Macarthur, University of Western Sydney.

Marsh, H. W. (1990g). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, *82*(4), 623-636.

Marsh, H. W. (1993). The multidimensional structure of academic self-concept: Invariance over gender and age. *American Educational Research Journal*, *30*(4), 841-860.

Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology*, *95*(4), 687-706.

Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology*, *76*(5), 940-956.

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, *80*(3), 366-380.

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and... *Educational Psychologist*, *34*(3), 155-167.

Marsh, H.W., and Craven, R. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment.* Orlando, FL: Academic Press.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*(2), 133-163.

Marsh, H. W., Craven, R. G., & Debus, R. (1991). Self-concepts of young children 5 to 8 years of age: Measurement and multidimensional structure. *Journal of Educational Psychology*, *83*(3), 377-392.

Marsh, H. W., Craven, R., & Debus, R. (1998). Structure, Stability, and Development of Young Children's Self-Concepts: A Multi-cohort–Multi-occasion Study. *Child Development*, *69*(4), 1030-1053.

Marsh, H. W., Debus, R., & Bornholt, L. (2005). Validating young children's self-concept responses: Methodological ways and means to understand their responses. In D.M. Teti (Ed.) *Handbook of research methods in developmental science* (pp.138-160). Hoboken, NJ: Wiley-Blackwell.

Marsh, H. W., Dowson, M., Pietsch, J., & Walker, R. (2004). Why Multicollinearity

    Matters: A Reexamination of Relations Between Self-Efficacy, Self-Concept, and

    Achievement. *Journal of Educational Psychology*, *96*(3), 518-522.

Marsh, H. W., Ellis, L. A., & Craven, R. G. (2002). How do preschool children feel about

    themselves? Unraveling measurement and multidimensional self-concept

    structure. *Developmental Psychology*, *38*(3), 376-393.

Marsh, H. W., Hau, K. T., & Kong, C. K. (2002). Multilevel causal ordering of academic

    self-concept and achievement: Influence of language of instruction (English

    compared with Chinese) for Hong Kong students. *American Educational*

    Research Journal, *39*(3), 727-763.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the

    study of self-concept: First-and higher order factor models and their invariance

    across groups. *Psychological Bulletin*, *97*(3), 562-582.

Marsh, H. W., & Holmes, I. W. M. (1990). Multidimensional self-concepts: Construct

    validation of responses by children. *American Educational Research*

    Journal, *27*(1), 89-117.

Marsh, H. W., Lüdtke, O., Trautwein, U., & Morin, A. J. (2009). Classical latent profile

    analysis of academic self-concept dimensions: Synergy of person-and variable-

    centered approaches to theoretical models of self-concept. *Structural Equation*

    Modeling, *16*(2), 191-225.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement:

    Relations and causal ordering. *British Journal of Educational Psychology*,

    81(1), 59-77.

157

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin*, *34*(4), 542-552.

Marsh, H. W., & O'Neill, R. (1984). *Self-Description Questionnaire III:* The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, *21*, 153–174.

Marsh, H. W., Parker, J., & Barnes, J. (1985). Multidimensional adolescent self-concepts: Their relationship to age, sex, and academic measures. *American Educational Research Journal*, *22*(3), 422-444.

Marsh, H. W., Parker, J. W., & Smith, I. D. (1983). Preadolescent self-concept: Its relation to self-concept as inferred by teachers and to academic ability. *British Journal of Educational Psychology*, *53*(1), 60-78.

Marsh, H. W., Relich, J. D., & Smith, I. D. (1983). Self-concept: The construct validity of interpretations based upon the *SDQ*. *Journal of Personality and Social Psychology*, *45*(1), 173-187.

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107-123.

Marsh, H. W., Smith, I. D., & Barnes, J. (1983). Multi-trait--multimethod analyses of the *Self-Description Questionnaire*: Student-teacher agreement on multidimensional ratings of student self-concept. *American Educational Research Journal*, *20*(3), 333-357.

Marsh, H. W., Smith, I. D., & Barnes, J. (1984). Multidimensional self-concepts: Relationships with inferred self-concepts and academic achievement. *Australian Journal of Psychology*, *36*(3), 367-386.

Marsh, H. W., Smith, I. D., Barnes, J., & Butler, S. (1983). Self-concept: Reliability, stability, dimensionality, validity, and the measurement of change. *Journal of Educational Psychology*, *75*(5), 772-790.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397-416.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology*, *16*(4), 331-345.

Marsh, H. W., Xu, M., & Martin, A. J. (2012). Self-concept: A synergy of theory, method, and application. In K.R. Harris, S. Graham, T. Urdan, (Eds.). *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues,* (pp. 427-458). Washington, DC, US: American Psychological Association.

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology*, *89*(1), 41-54.

Marsh, H. W., & Yeung, A. S. (1998a). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology*, *75*(2), 509-527.

Marsh, H. W., & Yeung, A. S. (1998b). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, *35*(4), 705-738.

Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research*, *79*(1), 327-365.

Masland, L.C., & Lease, A.M. (2013). Effects of achievement motivation, social identity, and peer group norms on academic conformity. *Social Psychology of Education, 16*, 661-681.

Molenaar, P.C.M. (2004). A manifesto on psychology as an idiographic science: Bringing the person back into scientific psychology—this time for ever. *Measurement: Interdisciplinary Research and Perspectives, 2*(1), 201-218.

Molloy, L.E., Gest, S.D., & Rulison, K.L. (2011). Peer influences on academic motivation: Exploring multiple methods of assessing youths' most "influential" peer relationships. *The Journal of Early Adolescence, 31*(1), 13-40.

Muthén, L.K. and Muthén, B.O. (1998-2015). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Muthén, L.K. and Muthén, B.O. (2005). *Mplus: Statistical analysis with latent variables: User's guide* (1998-2012). Los Angeles, CA: Muthén & Muthén.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological Methods*, *4*(2), 139-157.

160

Nagin, D.S.(2005). Overview of a semi-parametric, group-based approach for analyzing trajectories of development, *Proceedings of Statistics Canada Symposium 2002. Modelling Survey Data for Social and Economic Research*.

National Center for Education Statistics (2009) *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K).* [Data file and code book.] Retrieved from http://nces.ed.gov/ecls/dataproducts.asp.

Nichols, J.G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91(3),* 328-346.

Niehaus, K., & Adelson, J. L. (2013). Self-concept and native language background: A study of measurement invariance and cross-group comparisons in third grade. *Journal of Educational Psychology*, *105*(1), 226-240.

Niehaus, K., & Adelson, J. L. (2014). School support, parental involvement, and academic and social-emotional outcomes for English language learners. *American Educational Research Journal*, *51*(4), 810-844.

Nylund, K. L. (2007). *Latent transition analysis: Modeling extensions and an application to peer victimization.* Doctoral dissertation, University of California, Los Angeles.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4), 535-569.

Nylund, K. L., Bellmore, A., Nishina, A., & Graham, S. (2007). Subtypes, severity, and structural stability of peer victimization: what does latent class analysis say? *Child Development*, *78*(6), 1706-1722.

Nylund, K. L., Muthén, B., Nishina, A., Bellmore, A., & Graham, S. (2006). *Stability and instability of peer victimization during middle school: Using latent transition analysis with covariates, distal outcomes, and modeling extensions.* Unpublished manuscript.

Nylund-Gibson, K. L., Grimm, R., Quirk, M., & Furlong, M. (2014): A latent transition mixture model using the three-step specification. *Structural Equation Modeling: A Multidisciplinary Journal, 21*, 439-454.

Oberle, E., & Schonert-Reichl, K.A. (2013). Relations among peer acceptance, inhibitory control, and math achievement in early adolescence. *Journal of Applied Developmental Psychology, 34,* 45-51.

Obiakor, F.E., Stile, S.W., & Muller, D. (1987). Diminished self-concept of the visually impaired: Fact or fiction? *ERIC Report ED 900 942.* Retrieved from http://www.eric.ed.gov.

Osterman, K.F. (2001). Students' need for belonging in the school community. *Review of Educational Research, 70(3),* 323-367.

Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology*, *32*(1), 8-47.

Pinxten, M., Fraine, B.D., Van Damme, J., & D'Haenens, E. (2010). Causal ordering of academic self-concept and achievement: Effects of type of achievement measure. *British Journal of Educational Psychology, 80,* 689-709.

Pollack, J. M., Najarian, M., Rock, D. A., & Atkins-Burnett, S. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998- 99 (ECLS-K). Psychometric*

162

Report for the Fifth Grade. (NCES 2009-003). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Raufelder, D., Jagenow, D., Hoferichter, F., & Drury, K. M. (2013). The person-oriented approach in the field of educational psychology. *Problems of Psychology in the 21st Century*, *5*, 79-88.

Rinn, A.N. (2006). Effects of a summer program on the social self-concepts of gifted adolescents. *The Journal of Gifted Secondary Education, 17*(2), 65-75.

Ryan, A.M. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Development, 72*(4), 1135-1150.

Ryan, R.M., & Deci, E.L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54-67.

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement. In K.R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school*, (pp.171-195). New York, NY: Routledge.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*(3-4), 299-323.

Schmidt, M., & Cagran, B. (2008). Self-concept of students in inclusive settings. *International Journal of Special Education*, *23*(1), 8-17.

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). Estimating causal effects using experimental and observational designs. *Washington, DC: American Educational Research Association*.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323-338.

Schumacher, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd Edition), New York: Taylor & Francis Group.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, *57*(2), 149-174.

Schunk, D.H. (1995) Self-efficacy, and education, and instruction, In J.E. Maddus (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application.* New York, NY: Plenum Press.

Schunk, D. H., & Meece, J. L. (2006). Self-efficacy development in adolescence. *Self-Efficacy Beliefs of Adolescents*, *5*, 71-96.

Schunk, D. H., & Pajares, F. (2009). Self-efficacy theory. In K.R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school*, (pp.35-54). New York, NY: Routledge.

Schunk, D.H., & Zimmerman, B.J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32(4),* 195-208.

Schwartz, G. (1978). Estimating the dimension of a model, *Annals of Statistics, 6,* 461-464.

Sclore, L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52*, 333-343.

Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology*, *74*(1), 3-17.

164

Shavelson, R.J., Hubner, J.J., & Stanton, G.C. (1976). Self-concept: Validation of

    construct interpretations. *Review of Educational Research, 46(3),* 407-441.

Shavelson, R. J., & Marsh, H. W. (1986). On the structure of self-concept. In R.

    Schwarzer (Ed.), *Anxiety and cognitions.* Hillsdale, NJ: Erlbaum.

Silvia, P. J. (2005). Cognitive appraisals and interest in visual art: Exploring an appraisal

    theory of aesthetic emotions. *Empirical Studies of the Arts, 23,* 119-133.

Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school

    and whether I've got it: A process model of perceived control and children's

    engagement and achievement in school. *Journal of Educational Psychology*,

    82(1), 22-32.

Soares, L. M., & Soares, A. T. (1979). The Affective Perception Inventory (API),

    composite test manual. Trumbell, CT: SOARES Associates.

Steiger, J. H. (1990). Structural model evaluation and modification: an interval estimation

    approach. *Multivariate Behavioral Research, 25,* 173-180.

Steinberg, L., Brown, B. B., & Dornbusch, S. M. (1997). *Beyond the classroom: Why

    school reform has failed and what parents need to do.* New York, NY: Simon and

    Schuster.

Sterba, S. K., & Bauer, D. J. (2010). Matching method with theory in person-oriented

    developmental psychopathology research. *Development and

    Psychopathology*, *22*(02), 239-254.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. . In

    K.A. Bollen, & J.S. Long (eds.), *Testing structural equation models.* Newbury

    Park, CA: Sage.

165

Tashakkori A., & Kennedy, E. (1993). Measurement of self-concept in multicultural context: Psychometric properties of a modified *Self-Description Questionnaire*. *British Journal of Educational Psychology, 63,*337-348.

Tourangeau, K., Lê, T., Nord, C., Sorongon, A.G., & Chapman, C. (2009*). Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K): Eighth-Grade Methodology Report.* (NCES 2009-003). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Van de Schoot & Wong, (2012). Do delinquent young adults have a high or a low level of self-concept? *Self and Identity, 11,* 148-169.

Van Ryzin, M. J., Gravely, A. A., & Roseth, C. J. (2009). Autonomy, belongingness, and engagement in school as contributors to adolescent psychological well-being. *Journal of Youth Adolescence, 38,* 1-12.

Véronneau, M-H., & Dishion, T.J. (2011). Middle school friendships and academic achievement in early adolescence: A longitudinal analysis. *The Journal of Early Adolescence, 31* (1), 99-124.

von Eye, A., & Bogat, G. A. (2006). Person-oriented and variable-oriented research: Concepts, results, and development. *Merrill-Palmer Quarterly*, *52*(3), 390-420.

Wentzel, K. R. (1989). Adolescent classroom goals, standards for performance, and academic achievement: An interactionist perspective. *Journal of Educational Psychology, 81(2),* 131-142.

Wentzel, K. R. (1991a). Relations between social competence and academic achievement in early adolescence. *Child Development, 62(5),* 1066-1078.

Wentzel, K. R. (1991b). Social competence at school: Relation between social responsibility and academic achievement. *Review of Educational Research, 61(1),* 1-24.

Wentzel, K. R. (1993). Social and academic goals at school: Motivation and achievement in early adolescence. *Journal of Early Adolescence, 13*, 4-20.

Wentzel, K. R. (1996). Social and academic motivation in middle school: Concurrent and long-term relations to academic effort. *The Journal of Early Adolescence, 16,* 390-406.

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology, 89(3),* 411-419.

Wentzel, K. R. (1998). Social support and adjustment in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90,* 202-209.

Wentzel, K. R. (2003). Motivating students to behave in socially competent ways. *Theory into Practice, 42(4),* 319-326.

Wentzel, K. R. (2005). Peer relationships, motivation, and academic performance at school. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 279-296), New York, NY: Guilford Press.

Wentzel, K. R., (2009). Peers and academic functioning at school. In K.H., Rubin, W.M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups: Social, emotional, and personality development in context* (pp. 531.547). New York, NY: Guilford Press.

Wentzel, K.R., & Asher, S.R. (1995). Academic lives of neglected, rejected, popular, and controversial children. *Child Development, 66,* 754-763.

Wentzel, K.R., Baker, S.A., & Russell, S.L. (2012). Young adolescents' perceptions of
teachers' and peers' goals as predictors of social and academic goal pursuit.
Applied Psychology: An International Review, 61(4)*,* 605-633.

Wentzel, K. R., Barry, C. M., & Caldwell, K. A. (2004). Friendships in Middle School:
Influences on Motivation and School Adjustment. *Journal of Educational
Psychology*, *96*(2), 195-203.

Wentzel, K.R., Battle, A., Russell, S.L., & Looney, L.B. (2010). Social supports from
teachers and peers as predictors of academic and social motivation. *Contemporary
Educational Psychology, 35,* 193-202.

Wentzel, K. R., & Caldwell, K. (1997). Friendships, peer acceptance, and group
membership: Relations to academic achievement in middle school. *Child
Development*, *68*(6), 1198-1209.

Wentzel, K. R., & Watkins, D. E. (2002). Peer relationships and collaborative learning as
contexts for academic enablers. *School Psychology Review*, *31*(3), 366-377.

Wentzel, K.R., & Wigfield, A. (1998).Academic and social motivational influences on
students' academic performance. *Educational Psychology Review,
10(2),* 155-175.

Wentzel, K. R., & Wigfield, A. (Eds.). (2009). *Handbook of motivation at school*, New
York, NY: Routledge.

Wigfield, A., & Eccles, J.S. (2000). Expectancy-value theory of achievement motivation.
Contemporary Educational Psychology, 25, 68-81.

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs,
expectancies for success, and achievement values from childhood through

168

adolescence. In A. Wigfield & J.S. Eccles (Eds.). *Development of achievement motivation* (pp. 91-120). Cambridge, MA: Academic Press.

Wigfield, A., Tonks, S., & Klauda, S.L. (2009). Expectancy-value theory. In K.R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp.55-75), New York, NY: Taylor and Francis.

Williams, T., & Williams, K. (2010). Self-efficacy and performance in Mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology, 102*(2)*,* 453-466.

Wylie, R. C. (1961). *The self-concept: A critical survey of pertinent research literature.* Lincoln, NE: University of Nebraska Press.

Wylie, R. C. (1974). *The self-concept: A review of methodological considerations and measuring instruments.* Lincoln, NE: University of Nebraska Press.

Wylie, R. C. (1979). *The self-concept: Theory and research on selected topics (Vol. 2).* Lincoln, NE: University of Nebraska Press.

Zahra, A-T., Arif, M.H., & Yousuf, M.I. (2010). Relationship of academic, physical and social self-concepts of students with their academic achievement. *Contemporary Issues in Education Research, 3*(3)*,* 73-78.

APPENDIX A – *Self-Description Questionnaire-I* **Items Used in Grades 3 and 5**
(Four-point scale with anchors "Not at all true," "A little bit true," "Mostly true," and "Very true.")

Perceived Interest/Competence in Reading
1. I get good grades in reading
2. I like reading
3. Work in reading is easy for me
4. I am interested in reading
5. I cannot wait to read each day
6. I am good at reading
7. I like reading long chapter books
8. I enjoy doing work in reading

Perceived Interest/Competence in Math
1. I get good grades in math
2. I like math
3. Work in math is easy for me
4. I am interested in math
5. I cannot wait to do math each day
6. I am good at math
7. I can do very difficult problems in math
8. I enjoy doing work in math

Perceived Interest/Competence in All School Subjects
1. I am good at all school subjects
2. I enjoy work in all school subjects
3. Work in all school subjects is easy for me
4. I like all school subjects
5. I look forward to all school subjects
6. I get good grades in all school subjects

Perceived Interest/Competence in Peer Relations
1. I have lots of friends
2. I make friends easily
3. I get along with kids easily
4. I am easy to like
5. Other kids want me to be their friend
6. I have more friends than most other kids.

APPENDIX B – **Mplus Code for LPA and LTA**

Latent Profile Analysis Code for Grade 3 with Covariates and Free Variances
```
TITLE: LPA for gr 3;
DATA: file is gr3LPA.csv;
listwise=on;

VARIABLE:
Names are c56cw0 childid gender race w3ses w5ses gr3RIRT gr3MIRT gr5RIRT
gr5MIRT readgr3 read_SE mathgr3 math_SE allgr3 all_SE peergr3 peer_SE white black
Hispanic Asian Other;

Usevariables readgr3 mathgr3 allgr3 peergr3 gender w3ses black Hispanic Asian Other;

IDVARIABLE=CHILDID;
Missing =*;
Weight is c56cw0;
Classes=c(3);
MODEL:
%overall%
c#1 c#2 on gender w3ses black Hispanic Asian Other;
%c#1%
readgr3 mathgr3 allgr3 peergr3;
%c#2%
readgr3 mathgr3 allgr3 peergr3;
%c#3%
readgr3 mathgr3 allgr3 peergr3;
ANALYSIS: Estimator=MLR;
Type=mixture;
Starts=200 40;
OUTPUT: Tech1 Tech7 Tech11 sampstat;
savedata: file is gr3finals.sav;
save is cprobabilities;
PLOT: type is plot3;
Series=readgr3(1) mathgr3(1) allgr3(1) peergr3(1);
```

171

Latent Transition Analysis Code for Grade 3 with Distal Outcomes in Math
TITLE: latent transition file;
DATA: file is gr3_5LTA.csv;
 listwise=on;

VARIABLE: Names are read3 math3 all3 peer gender w3ses_3 black hisp asian other
c56cw0 childid cprob1_3 cprob2_3 cprob3_3 c3 read5 math5 all5 peer5 w5ses cprob1_5
cprob2_5 cprob3_5 cprob4_5 c5 gr3RIRT gr3MIRT gr5RIRT gr5MIRT;

Usevariables read3 math3 all3 peer3 gr5MIRT read5 math5 all5 peer5;

IDVARIABLE=CHILDID;
Missing are all (-9);
Weight is c56cw0;
Classes=c1(3) c2(4);
MODEL:
%overall%
c2 on c1;

MODEL c1:
%c1#1%
read3
math3
all3
peer3;

%c1#2%
read3
math3
all3
peer3;

%c1#3%
read3
math3
all3
peer3;

MODEL c2:
%c2#1%
read5
math5
all5
peer5;
[gr5MIRT] (p1);

172

www.manaraa.com

```
%c2#2%
read5
math5
all5
peer5;
[gr5MIRT] (p2);

%c2#3%
read5
math5
all5
peer5;
[gr5MIRT] (p3);

%c2#4%
read5
math5
all5
peer5;
[gr5MIRT] (p4);

model test: p1=p2;

ANALYSIS:
Estimator=MLR;
Type=mixture;
Starts=1000 160;
stiterations=20;

OUTPUT: sampstat STANDARDIZED TECH1 TECH15  Tech7 ;
```