

VALIDATION OF THE SCALE: A MEASURE OF MIDDLE SCHOOL STUDENTS'  
TRANSITION EXPERIENCES

Kathleen Megan Day Malley

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education (Educational Psychology, Measurement, and Evaluation).

Chapel Hill  
2014

Approved by:

Jill Hamm

Judith Meece

William Ware

Dana Griffin

Nisha Gottfredson

© 2014  
Kathleen Megan Day Malley  
ALL RIGHTS RESERVED

## **ABSTRACT**

Kathleen Megan Day Malley: A Validation of the SCALE: A Measure of Students' Middle School Transition Experiences  
(Under the direction of Jill Hamm)

The SCALE, a measure of students' middle school transition experiences, was validated for use with diverse, sixth-grade students attending metropolitan middle schools. Analyses were conducted to validate the SCALE using data from seven schools from both the fall and spring of sixth grade (N = 761; 52.30% female, 52.30% White, 48.10% free/reduced lunch status). Consistent with life course theory (Elder & Shanahan, 2006), three SCALE subscales were evident in the data, representing students' responses to the academic, procedural, and social demands of middle school. Overall, students' perceptions of their ability to meet the demands of middle school were positive, but academic demands were particularly difficult, specifically for male students. The SCALE exhibited full measurement invariance across gender and socioeconomic status, and partial invariance across race/ethnicity. SCALE subscales had good internal consistency in the fall and were stable across the sixth-grade year. Construct validity for all SCALE subscales was demonstrated through diverse students' experiences with transition demands early in sixth grade and significant concurrent and predictive relationships to indicators of students' adjustment in the fall and spring of 6th grade. After reviewing theoretical and statistical evidence gathered in the present study, the SCALE was determined to be suitable for use with diverse sixth-grade students.

## **ACKNOWLEDGEMENTS**

I am extremely thankful to those in my life who supported me throughout this journey. There are so many of you that I cannot possibly identify all of you by name without adding hundreds of pages to this dissertation. Thank you.

To my Committee, thank you for your support throughout this process. I am lucky to have committee members so knowledgeable and kind. Jill Hamm, thank you for all the opportunities you provided me and for your mentorship through the last few years. I learned from the best and I am eternally grateful. Bill Ware, thank you for allowing me to fall in love with statistics. Dana Griffin, thank you for reminding me why I love education so much and keeping me connected to my roots as a school counselor. Judith Meece, thank you for always encouraging and supporting my ideas. Nisha Gottfredson, thank you for all the positive support and always making the time for my questions.

To my husband, thank you for always believing in me and encouraging me to keep going until the end. To my mother, I would never have even made it here in the first place without your love, support, and sacrifice. To my Nana, thank you for being my inspiration. To my sister and brother, thank you for being my best friends and keeping me positive. Dad, thanks for always keeping me laughing. To the rest of my family: thank you for always listening to me ramble about education and statistics—and cheering me on. Thank you to my friends and colleagues, particularly Ashley, Jenni, Jessie, Mallory, Kerrylin, Liz, Catherine, Tonya, and Rachael, for always being there.

## TABLE OF CONTENTS

LIST OF TABLES .....	x
LIST OF FIGURES .....	xii
LIST OF ABBREVIATIONS AND SYMBOLS .....	xiii
CHAPTER 1: INTRODUCTION .....	1
The Current Study .....	4
CHAPTER 2: THEORETICAL FRAMEWORK .....	6
Conceptualizing the Middle School Transition .....	6
Middle School Situational Imperatives .....	10
Heightened academic expectations .....	10
Changes to procedural demands .....	11
Changing social relationships .....	11
Variability in students' experiences with middle school demands .....	14
Summary .....	15
Validity of Students' Adjustment Perceptions from the SCALE .....	15
Sources of evidence of construct validity .....	16
Summary .....	17
CHAPTER 3: LITERATURE REVIEW .....	19
The Transition to Middle School .....	19
Assessing Students' Experiences during the Middle School Transition .....	20
Students' Experiences and Adjustment during the Middle School Transition .....	21

Experiences with academic expectations.....	21
Changes to procedural expectations.....	24
Changes to social relationships.....	26
Establishing new relationships.....	26
Managing problematic peer situations .....	28
Summary .....	31
The Validity of Measures of Students' Transition Experiences .....	33
Compiling Theoretical Evidence .....	33
Compiling Statistical Evidence.....	34
Establishing the psychometric structure of existing scales.....	35
Establishing scale invariance in existing scales.....	36
Establishing scale reliability and stability of existing scales .....	38
Establishing construct validity of existing scales .....	40
The Present Study .....	48
CHAPTER 4: METHODS .....	50
Participants.....	50
Schools.....	50
Students.....	51
Procedures.....	51
Measures .....	52
Perceptions of Middle School Transition Experiences .....	52
Development of the SCALE .....	53
Content validity of the SCALE.....	53

Social, behavioral, and academic adjustment .....	53
Social adaptation .....	54
School behaviors .....	54
Academic behaviors .....	54
School belonging .....	55
School valuing .....	55
Academic achievement .....	56
Demographic variables .....	56
Plan of Analyses .....	56
<b>CHAPTER 5: RESULTS .....</b>	<b>59</b>
Missing Data .....	59
Descriptive Statistics .....	61
Univariate statistics .....	61
Bivariate statistics .....	63
Multivariate normality .....	64
Sampling adequacy .....	65
Summary .....	65
Power .....	66
Factor Analyses .....	68
Exploratory factor analysis .....	68
CFA-calibration .....	69
CFA-validation .....	71
Multiple Group CFAs .....	72

Gender.....	76
Free/reduced lunch status.....	77
Race/Ethnicity.....	78
Reliability and Stability .....	79
Internal consistency .....	80
Stability .....	80
Construct Validity .....	81
Concurrent and predictive relationships .....	82
Construct validity of the Academic subscale.....	82
Construct validity of the Procedural subscale.....	84
Construct validity of the Social subscale .....	85
Concurrent and predictive relationships across demographic groups.....	87
Correlations by gender.....	87
Correlations by socioeconomic status.....	90
Correlations by	
race/ethnicity	92.....
Latent mean	
differences	95.....
CHAPTER 6: DISCUSSION.....	98
Critical Transition Demands Captured by the SCALE.....	99
Variability in responding to middle school demands .....	99
Students' transition experiences and adjustment .....	103
Validity of Interpretations from the SCALE .....	107



Psychometric structure of the SCALE .....	108
Reliability and stability .....	109
Concurrent and predictive construct validity .....	110
Summary .....	114
Utility of the SCALE .....	114
Limitations .....	118
Conclusion .....	122
REFERENCES .....	151

## LIST OF TABLES

TABLE 1: Additional characteristics of participating schools. ....	123
TABLE 2: Items in the SCALE.....	124
TABLE 3: Cronbach’s Alpha Coefficients for the ICS-T Subscales Across Racial/Ethnic Groups .....	125
TABLE 4: Descriptive Statistics of SCALE Items - Fall of 6th Grade .....	126
TABLE 5: Descriptive Statistics of SCALE Items - Spring of 6th Grade.....	127
TABLE 6: Proportions of Response Values for Scale Items - Fall and Spring of 6th Grade.....	128
TABLE 7: Descriptive Statistics of Students’ Self-Reported Variables - Fall and Spring of 6th Grade .....	130
TABLE 8: Descriptive Statistics for Teacher-Rated Variables-Fall and Spring of 6th grade.....	131
TABLE 9: Polychoric Correlation Table for SCALE Items - Fall of 6th Grade .....	132
TABLE 10: Polychoric Correlation Table for SCALE Items - Spring 6th Grade.....	133
TABLE 11: Correlations between SCALE Items and Measures of Students’ Adjustment - Fall of 6th Grade.....	134
TABLE 12: Correlations between SCALE Items and Measures of Students’ Adjustment - Spring of 6th Grade .....	135
TABLE 13: Final EFA Model .....	136
TABLE 14: Final CFA Calibration and Validation Models.....	137
TABLE 15: Baseline Models Across Gender .....	138
TABLE 16: Measurement Invariance Across Gender: Comparisons of Model Fit.....	139
TABLE 17: Baseline Models Across Socioeconomic Status .....	140
TABLE 18: Measurement Invariance Across Socioeconomic Status: Comparisons of Model Fit... .....	141
TABLE 19: Baseline Models Across Race/Ethnicity .....	142
TABLE 20: Partial Measurement Invariance Across Race/Ethnicity: Comparisons of Model Fit ... .....	143

TABLE 21: Relationships Between SCALE Subscales and Student Adjustment Measures .....	144
TABLE 22: Relationships Between SCALE Subscales and Student Adjustment Measures by Gender.....	145
TABLE 23: Relationships Between SCALE Subscales and Adjustment Measures by Socioeconomic Status .....	146
TABLE 24: Relationships Between SCALE Subscales and Student Adjustment Measures by Race/Ethnicity.....	147

## **LIST OF FIGURES**

FIGURE 1: Scree plot including all SCALE variables.....	148
FIGURE 2: Final CFA calibration model for the SCALE.....	149
FIGURE 3: Gender differences on SCALE subscales in the fall of sixth grade .....	150

## LIST OF ABBREVIATIONS AND SYMBOLS

CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI [ , ]	Confidence interval
$df$	degrees of freedom
EFA	Exploratory Factor Analysis
$H_{\text{FORM}}$	Test for configural invariance
$H_{\Lambda x}$	Test for metric invariance
$H_{\Lambda, \Phi}$	Test for strict invariance
$H_{\Lambda, \Theta \delta}$	Test for equivalence of residual variances
ICC	Intraclass Correlation
KMO	Kaiser-Meyer-Olkin Measure of Sampling Adequacy
LGBTQ	Lesbian, gay, bisexual, transgender, or queer
$M$	Mean
MAR	Missing at Random
MCAR	Missing Completely at Random
MGCFA	Multiple Group Confirmatory Factor Analysis
MNAR	Missing Not Completely at Random
MSAs	Measures of Sampling Adequacy
$N/n$	Sample size
$p$	P-value
$r$	Pearson product correlation
$R^2$	Proportion of the variance

RMSEA	Root Mean Square Error of Approximation
$SE$	Standard error
SEsk	Standard error skewness
SEku	Standard error kurtosis
SEM	Structural Equation Modeling
$SD$	Standard deviation
TLI	Tucker-Lewis Index
x	item
$y^*$	Underlying response variable for y
$\leq$	Less than or equal to
$\geq$	Greater than or equal to
$\pm$	Plus or minus
$\alpha$	Cronbach's alpha
$\delta$	Delta
$\Delta \chi^2$	Chi-square difference test
$\lambda$	Lambda
$\tau$	Tau
$\theta$	Theta
$\phi$	Phi
$\xi$	Xi
$\chi^2$	Chi-squared
$\Lambda_x$	Capital lambda-x matrix
$\Phi$	Capital phi matrix

$\Theta_\delta$  Capital theta-delta matrix

## **Chapter 1: Introduction**

The middle school transition is a significant experience in education, one that students approach with feelings of both excitement and apprehension (Akos, 2002). During this transition, students' school context changes. Students physically change location as they leave elementary school and move to a new middle school (Juvonen, Kaganoff, Augustine, & Constant, 2004). Similar to other transitions, entering middle school also brings about new norms, expectations, and relationships for students. Students encounter more demanding academic standards and must adapt to new procedural rules (Anderson, Jacobs, Schramm, & Splittgerber; 2000; Farmer et al., 2013; Juvonen et al., 2004). They must also establish new relationships with teachers and peers, and manage problematic peer interactions that emerge following the introduction of new peers into a larger school context (Anderman & Mueller, 2010; Eccles & Roeser, 2009).

In the presence of so many changes and new expectations, the nature of students' adjustment depends upon their responses to contextual demands (Elder & Shanahan, 2006). Despite the efforts of educators and researchers to ease students' transition into their new middle school contexts over the past twenty to thirty years, many students have continued to demonstrate adjustment difficulties academically (Ryan, Shim, & Makara, 2013), socially (Karriker-Jaffe, Foshee, Ennett, & Suchindran, 2008; Niehaus, Rudasill, & Rakes, 2012), and behaviorally (Witherspoon & Ennett, 2011). Yet, other students transition into middle school well and many thrive in comparison to their adjustment in elementary school (e.g. Davidson, Gest, & Welsh, 2010; Wentzel, Battle, Russell, & Looney, 2010). Assessing students'



perceptions of their transition experiences provides a means of understanding how they respond to the changing demands, expectations, and relationships during the middle school transition and how this relates to their adjustment (Niehaus et al., 2012; Wang & Holcombe, 2010). However, there are few contemporary instruments that assess students' perceptions of their transition experiences; those available have not been adequately validated, particularly for use with students from diverse backgrounds. In one study, students' perceptions of the middle school transition in a rural school district were assessed with a single item, which limited the ability of researchers to determine the reliability of the measure and validity of their conclusions (i.e., Day, Hamm, Lambert, & Farmer, 2014; Hamm, Dadisman, Day, Agger, & Farmer, 2014). Other measures used in previous studies have not been assessed for reliability or validity (Akos, 2002; Akos & Galassi, 2004a, 2004b).

A measure by Elias, Ubriaco, Reese, Gara, Rothbaum, and Haviland (1992), the Survey of Adaptational Tasks of Middle School (SAT-MS), is one of the measures currently available that has been assessed for reliability and validity through multiple means. The authors consulted experts, empirical literature, and pilot tested the SAT-MS during the development of the scale. Furthermore, the researchers performed factor analytic procedures on the SAT-MS using data from two schools, one attended primarily by White, suburban students and the other attended by urban students from predominantly racial/ethnic minority groups. The researchers then compared the results of their factor analyses to assess the validity of the psychometric structure of the SAT-MS across schools. Mean differences on students' scale responses were examined for differences across gender and schools. Finally, the scale authors assessed the construct validity of the scale by correlating SAT-MS subscales with concurrent measures of students' adjustment, including assessments of self-concept, prosocial behavior, and their perceptions of

the school context. Relationships between subscales and other concurrent measures of students' adjustment were also analyzed for differences across gender and schools.

Despite these efforts, questions remain regarding the reliability of the scale and its validity for use with contemporary, diverse students. For example, the authors reported poor reliability for one of the subscales that assessed students' adaptation to the academic demands of middle school. In addition, the researchers did not take into account the ordinal nature of the item responses during the factor analytic procedures, which can contribute to biased parameters, standard error estimates (Schmitt, 2011), and underrated item correlations (Holgado-Tello, Chacón-Moscoso, Barbero-García, & Vila-Abad, 2010). These issues can lead to misguided conclusions about the psychometric structure of a scale.

The researchers found mean differences in students' responses on subscales of the SAT-MS across schools (Elias et al., 1992). In addition, during comparisons of the psychometric structure, researchers found inconsistencies as a few items within the SAT-MS loaded to different factors when used in geographically, economically, and racially dissimilar schools. Measurement bias at the item level may lead to misguided conclusions regarding the measurement properties of a scale, even in the presence of invariance in the relationship between overall scores of a measure and other criterion between groups (Millsap, 1997). However, no further analyses were conducted by Elias et al. (1992) to investigate the presence or absence of scale invariance by gender, socioeconomic status, or race/ethnicity to explain these inconsistencies. Therefore, items in the measure may have been culturally biased, leading students to respond in different ways depending on their demographic characteristics. Use of more rigorous tests of scale invariance, such as multiple group confirmatory factor analysis, would yield a more accurate assessment of the scale invariance of the SAT-MS across gender,

race/ethnicity, socioeconomic status, as well as identify potential areas of partial measurement invariance including at the item level (Cheung & Rensvold, 2002; Knight & Hill, 1998; Millsap & Kwok, 2004).

Finally, the items within the SAT-MS (Elias et al., 1992) are dated as the scale has not been revised to reflect current trends in the study of students' middle school transition experiences that may have changed during the last few decades. Current students may have different perceptions of the academic demands of middle school following the enactment of high-stakes testing policies such as the No Child Left Behind Act (2002) or perceptions of social challenges following the implementation of school-wide behavioral and procedural interventions, such as Positive Behavior Support (Sugai & Horner, 2002), that aim to improve teacher and peer relationships and decrease the number of discipline referrals. As a result, one of the more validated measures of students' perceptions of their transition experiences, the SAT-MS, may not be appropriate for use with contemporary, diverse students.

### **The Current Study**

Researchers and practitioners can use assessments to better understand middle school students' perceptions of their adjustment in the middle school transition (Day et al., 2014). However, available measures are dated and have questionable reliability and validity. In the current study, I address these concerns by validating the SCALE, a new measure of students' perceptions of their experiences with the academic, social, and procedural demands of the middle school transition recently administered in schools with diverse student bodies. The following goals will be addressed:

- Goal 1. Establish and confirm the factor structure of the SCALE through exploratory factor analysis and confirmatory factor analysis.

- Goal 2. Demonstrate the scale invariance of the SCALE across various student characteristics including gender, socioeconomic status, and race/ethnicity.
- Goal 3. Establish the reliability and stability of the SCALE.
- Goal 4. Demonstrate the construct and predictive validity of the SCALE.

## **Chapter 2: Theoretical Framework**

The current study is grounded in multiple theoretical frameworks. The combination of these theories provides a comprehensive theoretical basis for both the content of the SCALE and subsequent interpretations made from information gathered using the SCALE. To conceptualize how the items within the SCALE represent students' transition experiences, I integrate principles of stage-environment fit and life course theories. I use validity theory to conceptualize how theoretical and statistical information from the SCALE informs the reliability and validity of the measure and, subsequently, the usefulness of researchers' interpretations of students' perceptions of their experiences with the middle school transition from the SCALE.

### **Conceptualizing the Middle School Transition**

Leading theorists identify the relationship between students and the school context as influential to students' adjustment (Bronfenbrenner, 1975; Eccles et al., 1993; Elder & Shanahan, 2006). School contexts are composed of multiple microsystems that are characterized by physical features as well as by the policies and relationships among people within those systems (Bronfenbrenner, 1975). Physical features can include the layout of the building or the size of classrooms (Eccles & Roeser, 2011b). District-level and more proximal school-level and classroom-level policies and practices also set rules and expectations for student behavior (Eccles, 2004). Students, teachers, and staff, and the relationships among these individuals also add to the overall school context (Bronfenbrenner, 1994; Eccles & Roeser, 2011b). The physical, political, and social aspects of the school context create and regulate opportunities for students to interact with others and shape the relationships that form within the school (Eccles,

2004; Eccles & Roeser, 2011b). These relationships and interactions influence students' adjustment (Bronfenbrenner, 1994; Eccles, 2004).

During the course of their education, students experience comprehensive changes to their school context. The middle school transition, when most students move from k-5 elementary schools to middle schools enrolling students from grades 6 through 8, is an educational period in which students experience a vast contextual change (Juvonen et al., 2004). During the middle school transition, changes to students' school context include but are not limited to physical school location, rules, policies, as well as relationships between students and teachers (Eccles & Roeser, 2011b). This transition is a normative and planned occurrence in students' education; however, it has been regarded as a particularly difficult educational transition given reported declines in students' adjustment during this period (Eccles et al., 1993; Juvonen et al., 2004).

One prominent conceptualization of students' adjustment during the middle school transition has been stage-environment fit theory (Eccles et al., 1993). Early adolescents experience substantial change in many developmental domains at the same time that they transition to middle school (Eccles & Midgley, 1989; Wigfield, Byrnes, & Eccles, 2006; Wigfield, Lutz, & Wagner, 2005). Although these developmental changes coincide with the middle school transition, stage-environment fit theorists assert that the transition itself is not solely to blame for the adjustment declines commonly observed during this period (Eccles & Roeser, 2009; Midgley, Middleton, Gheen, & Kumar, 2002). Rather, a mismatch between characteristics of the school context and the developmental needs of early adolescents contributes to negative adjustment outcomes whereas alignment promotes positive outcomes (Eccles & Midgley, 1989; Eccles et al., 1993). Early adolescents' developmental needs may include the need for increased autonomy, mastery, and a focus on social relationships and are

often conceptualized by stage-environment theorists as consistent across individuals (Eccles et al., 1993). The academic, behavioral, and social declines that can occur during the middle school transition signify a mismatch between the demands of middle school and the developmental needs of early adolescents. Therefore, students may have difficulty meeting the demands required for their successful school adjustment following the transition (Eccles & Roeser, 2009; Midgley et al., 2002)

Life Course Theory is another influential developmental theory that scholars have drawn upon to conceptualize students' adjustment during both the transition into primary school (Entwisle, Alexander, & Olsen, 2003) and the high school transition (Benner, 2011). Life course theorists assert in the *principle of situational imperatives* that all transitions involve new behavioral and social demands within the context (Elder & Shanahan, 2006). Elder, Johnson, and Crosnoe (2003) stressed in a principle called *linked lives* that social relationships are particularly influential to individuals' adjustment. Although situational imperatives and social relationships are highly influential, they are not the only determining factors in individuals' adjustment (Elder, 1998; Elder & Shanahan, 2006). Rather, adjustment is shaped by the choices individuals make based on their experiences during transitions, also known as the *principle of human agency* (Elder, 1998). Yet, life course theorists have asserted that individuals from different groups may have differential experiences during transitions (Pallas, 2003). In particular, students' schooling opportunities and experiences differ by gender, racial/ethnic group membership, and/or socioeconomic status (Benner, 2011). Differential opportunities, particularly those that disadvantage some students in comparison to others, may affect how students are able to respond to situational imperatives and potentially contribute to school adjustment difficulties.

Comparisons of the principles of stage-environment fit theory and life course theory reveal several similarities. Both theories share an emphasis on the importance of the relationship between the context and the developmental characteristics of the individual to the individual's adjustment (Eccles, 2008). Also, both theories acknowledge that individuals have agency within their context that influences their development. Finally, both theories explain that, due to influences such as poverty, access to resources, or other social influence, students in different schools can be exposed to disparate opportunities that may influence the choices they make and their subsequent development (Eccles, 2008; Pallas, 2003). The presence of these similarities between the two theories is to be expected as a primary stage-environment fit theorist was influenced by the principles of life course theory (Eccles, 2008).

Despite the similarities between the two theories, there is a key difference in their conceptualizations of students' adjustment during the middle school transition. Stage-environment fit theory focuses on the match or mismatch between the characteristics of the context and the students' developmental needs, whereas life course theory focuses explicitly on the role of individuals' subjective experience of contextual characteristics on the course of their adjustment. Consequently, the principles of life course theory are a useful tool in organizing the conceptualization of students' perceptions of their middle school transition experiences.

The focus of the present study is on the SCALE, a measure that assesses students' subjective experiences and reactions, both positive and negative, to the middle school transition. Therefore, I conceptualize students' adjustment following the transition to middle school using the principles of life course theory, acknowledging that stage-environment fit theory also undergirds these conceptualizations. Specifically, in the following sections I focus on the academic, procedural, and social expectations and demands common to the middle school



transition experience in accordance with the *principle of situational imperatives*. I highlight the importance of social relationships to students' middle school adjustment, applying the *principle of linked lives*. Finally, I conceptualize the role of students' reactions to situational imperatives to their subsequent adjustment using the *principle of human agency*.

### **Middle School Situational Imperatives**

During the middle school transition, students encounter new academic and procedural expectations and norms, as well as social demands through changes to their relationships with both teachers and peers (Juvonen, 2007; Juvonen et al., 2004). Variability in transition experiences may occur as boys and girls and students from different racial/ethnic and socioeconomic status groups encounter differential expectations, demands, and norms in middle school (Farkas, 2003; Meece, Glienke, & Burg, 2006). According to the *principle of human agency* (Elder, 1998), students' experiences during the transition depends on their perceptions of situational imperatives and their responses to their experiences.

**Heightened academic expectations.** In middle school, teachers' academic expectations for students are characterized by a greater emphasis on academic performance (Schunk, Pintrich, & Meece, 2008; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). In addition, students' academic ability is strongly emphasized, and competition and social comparison between peers is encouraged (Eccles et al., 1993; Eccles & Roeser, 2009). At the same time, the difficulty of classwork and homework intensifies and the quantity increases (Anderson et al., 2000; Cauley & Jovanovich, 2006). Students also experience fewer opportunities for choice and decision-making in their classrooms, and receive more whole-class instruction and less personalized instruction than in elementary school (Schunk et al., 2008). Taken together, expectations for academic success in middle school require high levels of student performance

without as much assistance from teachers. The heightened focus on ability and achievement in tandem with less personalized instruction is detrimental to students' adjustment during the transition to middle school as it neglects early adolescents' needs for mastery, peer collaboration, and autonomy (Eccles et al., 1993; Eccles & Roeser, 2009).

**Changes to procedural demands.** During the transition, students move from an intimate, predictable elementary school structure in which they spend most of the day in one classroom with the same teacher and peers, to the middle school where they encounter a complex class schedule (Juvonen et al., 2004). Middle school students must move through multiple classrooms per day, each with different teachers and correspondingly, different expectations. Students also become responsible for using and organizing a locker between classes. Middle school teachers may expect students to fulfill the daily procedural demands without much help from adults; students' inability to meet expectations may result in disciplinary action (Anderson et al., 2000; Ellerbock & Kiefer, 2013).

Thus, students are expected to be more responsible and accountable for getting to each class on time with the appropriate materials and school work (Anderson et al., 2000; Cauley & Jovanovich, 2006). Early adolescents need more experiences that promote autonomy but also structure and assistance to develop decision-making and reasoning skills (Halpern-Felsher, 2009; Resnick, 2000). Consequently, adapting to the procedural demands of middle school may be more difficult for some students than others, particularly if teacher support is lacking (Anderson et al., 2000).

**Changing social relationships.** According to the *principle of linked lives* in life course theory, social relationships are particularly influential to the positive or negative nature of individuals' adjustment following a transition (Elder et al., 2003; Elder & Shanahan, 2006). A

greater focus on social relationships is also a defining developmental characteristic of early adolescence (Rodkin & Ryan, 2012). Considering the importance of social interactions to adjustment (Elder et al., 2003) and to early adolescence in general (Rodkin & Ryan, 2012), the social demands of middle school may be particularly salient for students and feature prominently in their transition experiences.

Scholars have identified two prominent social demands during the middle school transition: the demand to establish new teacher and peer relationships and the demand to manage problematic peer interactions. As students transition into middle school, their relationships with their elementary school teachers and peers are disrupted, and they must develop new relationships. Scholars have noted that characteristics of the middle school context may make establishing relationships with both teachers and peers especially difficult (Juvonen, 2007). Students have multiple teachers for short periods of time each day and classrooms have higher teacher-to-student ratios than in elementary school (Anderson et al., 2000; Juvonen, 2007). Middle school teachers are also more likely to use management techniques that discourage autonomous behavior and focus more on controlling students' behaviors (Collins & Steinberg, 2006; Eccles & Roeser, 2009). Furthermore, teachers are more likely to discipline students, even for minor misbehaviors (Anderson et al., 2000). The focus on controlling and disciplining behaviors in combination with less time with teachers during the school day makes establishing strong, supportive student-teacher relationships more challenging for students.

Aspects of the middle school transition also make establishing peer relationships more difficult for students. Students enter middle school with a combination of familiar elementary school peers and unfamiliar peers from other elementary schools (Juvonen, 2007). The transition allows for the reorganization of friendships and peer group affiliations, particularly in the

presence of numerous new peers (Farmer, Xie, Cairns, & Hutchins, 2007). However, shorter class periods with different groups of classmates throughout the day reduce students' opportunity to connect with their peers (Anderson et al., 2000; Juvonen, 2007). Subsequently, separation from familiar elementary school peers and fewer opportunities to form relationships with peers from other elementary schools makes peer relationships more challenging for students following the transition.

Additionally, although the presence of unfamiliar peers creates opportunities for new relationships, it disrupts students' social status within the larger peer network (Farmer et al., 2007). Social status can include students' acceptance as well as prestige with peers (Farmer et al., 2007; Rubin, Bukowski, & Parker, 2006). Disruptions to students' social status become problematic as early adolescents often use aggressive or bullying behaviors to dominate peers and to establish higher status (Espelage & Swearer, 2003; Pellegrini, 2002). Bullying, however, is not the only challenging peer situation that middle school students have to manage following the transition. The presence of older middle school peers can lead to exposure to antisocial behaviors such as fighting (Parker, Rubin, Erath, Wojskawowicz, & Buskirk, 2006). The social demands of the middle school transition can be mutually reinforcing; students who are unable to establish relationships or who have low status following the transition are more likely to become involved with peers engaged in disruptive behaviors, and are more likely to be victims of bullying or aggression from peers (Parker et al., 2006; Rubin et al., 2006). Thus, middle school students experience multiple social demands, and their ability both to build new relationships with teachers and peers and to manage difficult peer situations contributes to the quality of their transition experiences (Juvonen, 2007; Parker et al., 2006; Rubin et al., 2006).

**Variability in students' experiences with middle school demands.** Following the *principle of human agency* (Elder, 1998), the positive or negative nature of students' transition experiences is influenced by their responses to the demands imposed by middle school. However, the demands of middle school may be more challenging for some students than others because they conflict with their developmental needs (Eccles & Roeser, 2009; Resnick, 2000). Students in different gender, racial/ethnic, or socioeconomic status groups also may not experience the same opportunities (Elder et al., 2003; Pallas, 2003). As a result, some students may have limited choices for how to respond to middle school demands, which affects how they experience the transition.

For example, girls and boys may receive different messages regarding their achievement in certain school subjects (Meece et al., 2006). Teachers may consciously or unconsciously communicate to female students that math and science success is associated with being more masculine or to boys that language arts achievement is more appropriate for girls. Students from ethnic minority groups may also have different experiences during middle school (Garcia-Coll et al., 1996; Swanson et al., 2003). African American students may encounter racial/ethnic discrimination by peers and teachers, may be graded more stringently, or may be tracked into less advanced classes than their White peers (Farkas, 2003). African American and lower-income students are also more likely to attend schools that are lower-performing and less academically rigorous with fewer advanced course options than their White and economically advantaged peers. Differential opportunities, particularly those encountered by students based on their gender, race/ethnicity, and socioeconomic status may affect students' experiences during the transition (Farkas, 2003; Juvonen, 2007; Meece et al., 2006).

**Summary.** Transitions are critical points in individuals' adjustment (Elder et al., 2003; Entwisle et al., 2003). During the middle school transition, situational imperatives that arise from the new academic, procedural, and social demands and expectations of middle school present opportunities for changes to students' experiences. Changing social relationships, or *linked lives*, are especially influential to students' experiences (Elder et al., 2003). The positive or negative nature of students' transition experiences results from their responses to the academic, procedural, and social demands of middle school. Students may experience different opportunities during the middle school transition depending on their gender, race/ethnicity, or socioeconomic status (Garcia-Coll et al., 1996; Farkas, 2003; Meece et al., 2006; Swanson et al., 2003), which may affect their perceptions of the transition (Pallas, 2003). Accordingly, in the present study, I conceptualized students' experiences with the academic, procedural, and social demands they encountered during the transition to middle school with particular attention to differences in students' experiences based on their gender, race/ethnicity, and socioeconomic status.

### **Validity of Students' Adjustment Perceptions from the SCALE**

In the present study, I used current interpretations of validity theory to conceptualize how theoretical and statistical information gathered from the SCALE supported or refuted the validity of researchers' interpretations of students' responses on the SCALE. The *Standards for Educational and Psychological Testing* developed by the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) defined validity as "the degree to which evidence and theory support interpretations of test scores entailed by proposed uses of tests" (AERA, APA, & NCME, 1999, p. 7). Thus, according to contemporary scholars, validity is a representation of

researchers' interpretation of scale scores and not the scale itself (Kane, 2013). Moreover, the validity of scale interpretations is determined through the analysis of a collection of pertinent information that is used to support or refute the construct validity of the scale, or the notion that the scores actually represent the construct of interest (Cizek, 2012; Messick, 1989).

Information collected to support construct validity includes both theoretical and statistical evidence, such as information regarding the content and construct validity of the scale (AERA, APA, & NCME, 1999; Clark & Watson, 1995). This interpretation of validity theory is different from past conceptualizations, which separated information regarding aspects of the scale into different types of validity, e.g. content, external, or construct validity (Cronbach & Meehl, 1955; Downing, 2003). Therefore, according to current validity theory, all information that is used to determine the validity of a scale, including content, external, or construct validity, is gathered to present a unified representation of the construct validity of the interpretations drawn from a scale (AERA, APA, & NCME, 1999; Cizek, 2012; Messick, 1989).

**Sources of evidence of construct validity.** The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999) identify both theoretical and statistical evidence that can be gathered to judge the validity of scale interpretations. Theoretical evidence can include theoretical constructs, past research findings, as well as experts' opinions on the relationship between the measure and the construct of interest (AERA, APA, & NCME, 1999; Clark & Watson, 1995). Statistical evidence comprises information from a number of potential analyses of a scale, such as the internal structure of the scale, score comparability, and both the concurrent and predictive relationships of scale scores to variables external to the scale, as well as the reliability of the scale (AERA, APA, & NCME, 1999; Downing, 2003; Messick, 1995).

Analysis of the internal structure of the scale refers to the psychometric characteristics of the measure. The factor pattern of a scale involves information about the relationship of items to each other (Cizek, 2012; Downing, 2003). Scale scores are also associated with other measures of criteria that are hypothesized to assess the construct of interest or similar constructs (Messick, 1995). Scale scores may also be assessed for comparability across groups. The analysis of these relationships is used to support or refute whether or not the interpretations from the scale of interest are consistent with the intended construct (AERA, APA, & NCME, 1999; Messick, 1995). Assessing the predictive relationships of scale scores to other, future measures of relevant criteria enables researchers to assess how well the measure may relate to the construct of interest at a later time.

Finally, reliability refers to how well a scale performs in both consistent and predictable ways (DeVellis, 2012). Reliability has multiple components, including scale internal consistency and stability. The internal consistency of a scale measures how well items within the measure relate to other items in the same measure (Cronbach, 1951; Streiner, 2003). Scale stability, or test-retest reliability, measures how well a scale relates to itself over time and under different conditions (DeVellis, 2012).

**Summary.** Researchers can use the SCALE to interpret students' perceptions of their own adjustment to new situational imperatives associated with the middle school transition. The validity of researchers' interpretations from scale scores and, as a result, the utility of the SCALE relies upon characteristics of the measure itself (Cizek, 2012; Cizek, Rosenberg, & Koons, 2008). The validity of the interpretations made from scale scores is determined from a collection of theoretical and statistical evidence about the scale (AERA, APA, & NCME, 1999; Cabrera-Nyugen, 2010; Messick, 1989). Correspondingly, analyses of the construct validity and



reliability were used in the present study to determine the overall validity and utility of the SCALE for use with a diverse sample of sixth-grade students.

## **Chapter 3: Literature Review**

### **The Transition to Middle School**

The middle school transition is considered to be a difficult time in the course of students' education, associated with declines in adjustment (Anderman & Mueller, 2010; Eccles & Roeser, 2009; Juvonen et al., 2004). Studies have demonstrated that following the transition to middle school, students' academic achievement, motivation, and intrinsic value for school work declined between the 5th and 6th grades and throughout the remainder of 6th grade (Burchinal, Roberts, Zeisel, & Rowley, 2008; Lepper, Corpus, & Iyengar, 2005; Ryan et al., 2013). Discipline problems and incidences of school misbehavior also become more prevalent in middle school (Way, Reddy, & Rhodes, 2007; Witherspoon & Ennett, 2011) and students have reported greater difficulty with peers and teachers (Niehaus et al., 2012; Wentzel et al., 2010). In addition, adjustment differences across gender, socioeconomic, and racial/ethnic groups continue to be documented. Male students and African American students continue to exhibit achievement scores below female and White students in middle school (Akos, Rose, & Orthner, 2014; Balfanz, Herzog, & Mac Iver, 2007; Ryan et al., 2013). Economically disadvantaged students are more likely to experience lower academic achievement and lower academic growth in middle school, particularly in math, when compared to their economically advantaged peers (Akos et al., 2014). African American male and economically disadvantaged students, in particular, are also more likely to be perceived by teachers and peers as aggressive and to have a greater incidence of disciplinary referrals (Fenning & Rose, 2007; Graham, Bellmore, & Mize, 2006; Gregory, Skiba, & Noguera, 2010; Theriot & Dupper, 2009).

Nevertheless, not all students experience adjustment difficulties following the middle school transition and many exhibit positive outcomes (Davidson et al., 2010; Wentzel et al., 2010). Furthermore, researchers have not always found differences in students' adjustment outcomes between males and females or students from different racial/ethnic and socioeconomic status groups. For instance, academic motivation and intrinsic value for school work seem to decline for all students, regardless of their gender or race/ethnicity (Lepper et al., 2005; Ryan et al., 2013). In other studies, both boys and girls and students in different socioeconomic strata experienced difficulties with relationships with teachers and peers following the transition (Niehaus et al., 2012; Wang et al., 2009; Wentzel et al., 2010). Yet, the transition to middle school remains an important area of focus for both researchers and educators due to continued difficulties experienced by many students and particularly by male, African American, and economically disadvantaged students.

### **Assessing Students' Experiences during the Middle School Transition**

Assessing transition experiences provides an opportunity to understand further the variability in students' middle school adjustment (Day et al., 2014; Elias et al., 1992; Hamm et al., 2014). However, available measures of students' experiences may not be appropriate for use with contemporary, diverse students because scales are dated and have questionable reliability and validity. The intent of the present study is to validate the SCALE, a new instrument that assesses students' perceptions of their transition experiences. Both empirical and statistical information contribute essential evidence for scale validation (AERA, APA, & NCME, 1999). In this review, I present current research findings regarding students' transition experiences to provide empirical support for the content of the SCALE and to identify adjustment indicators to be used in the validation process. In addition, I review the procedures that have been used to

assess the validity of other available transition measures to identify statistical information significant for the validation of the SCALE. Together, the empirical and procedural information collected from these studies informs the validation of the SCALE.

### **Students' Experiences and Adjustment during the Middle School Transition**

Researchers have noted variability in students' overall transition experiences, as well as differences in the experiences of students in different gender, racial/ethnic, and socioeconomic status groups (Akos & Galassi, 2004b; Day et al., 2014; Elias et al., 1992; Hamm et al., 2014). For example, in many studies, most students rated the transition as being relatively easy (Akos & Galassi, 2004a, 2004b; Day et al., 2014; Hamm et al., 2014). Akos and Galassi (2004a) reported that girls perceived the transition to middle school to be more difficult for boys, but no differences were noted between White and African American students. However, Day et al. (2014) found that White male 6th-graders were the most likely and minority male 6th-graders were the least likely to perceive an easy transition into middle school. No studies were identified that measured differences in students' perceptions of the difficulty of their overall middle school transition experience between economically advantaged and disadvantaged students. Findings from other studies have indicated that the nature of students' middle school transition experiences were often related to their ease or difficulty in meeting the new academic, procedural, and social demands and expectations of middle school (Akos & Galassi, 2004a, 2004b; Barber & Olsen, 2004; Day et al., 2014; Deemer, McCotter, & Smith, 2003; Elias et al., 1992).

**Experiences with academic expectations.** Students have often reported anxiety about the level of difficulty of academic work they expect to experience in middle school (Akos, 2002; Akos & Galassi, 2004b). Parents have expressed academic concerns, anticipating that their

children would receive more difficult schoolwork in middle school and wondering if their students would be able to receive academic assistance if it was needed (Hamm et al., 2014). Following the transition, students' academic concerns may be realized as many perceived that both the difficulty and quantity of their schoolwork increased and reported that this was often the most challenging of aspect middle school (Akos & Galassi, 2004b; Deemer et al., 2003). Boys, in particular, have reported that they spent more time on homework in 6th grade than in elementary school and that the work was more difficult (Deemer et al., 2003).

In a number of studies, researchers have demonstrated that the expectations for academic success were different in middle school than elementary school. Teachers perceived that the students who were the most successful at adjusting after the middle school transition were more likely to be focused on school work and performed well in the classroom (Akos & Galassi, 2004b; Elias et al., 1992). Studies of both teachers' and students' perceptions have found that more so than elementary school teachers, middle school teachers expect students to complete their work autonomously and with less assistance (Deemer et al., 2003; Lane, Wehby, & Cooley, 2006). Moreover, middle school teachers were more likely to grade students' work more stringently than elementary school teachers (Randall & Engelhard, 2009). Following the transition, students perceived that their 6th grade teachers emphasized the importance of academic performance more often (Freeman & Anderman, 2005; Urdan & Midgley, 2003). Students also perceived a heightened emphasis on academic competition and social comparison between peers (Ryan & Patrick, 2001; Wang & Holcombe, 2010). Taken together, middle school students appear to be more responsible for increasing quantities of more challenging schoolwork without as much assistance from teachers. At the same time, students perceived that

the social and academic ramifications of academic performance became more significant, with achievement taking precedence over academic improvement.

The heightened expectations for academic performance and competition in middle school may be difficult for some students (Wang & Holcombe, 2010; Wigfield et al., 2006). Students who perceived that their teachers often endorsed performance were less likely to be academically motivated and engaged (Anderman, 2003; Urdan & Midgley, 2003). Furthermore, students who perceived more social comparison and competition in middle school had less favorable feelings of belonging and identification with school (Anderman, 2003; Wang & Holcombe, 2010). Conversely, when students perceived that their teachers valued academic improvement, they had greater feelings of school belonging and were more likely to be academically engaged and motivated (Anderman, 2003; Urdan & Midgley, 2003).

Yet, heightened academic expectations have not been consistently associated with maladjustment (Klem & Connell, 2004). In some studies, students were more academically engaged when their teachers had high academic expectations. However, students were more likely to be more academically engaged and to have higher levels of academic achievement if they perceived that their teachers not only had high expectations, but also provided academic support (Klem & Connell, 2004; Wang & Holcombe, 2010). Zimmer-Gembeck and Collins (2003) stressed that although early adolescents need more experiences with autonomy, they still require guidance and support from teachers. Support can include teachers' time, instructional assistance, feedback, goals endorsed in the classroom, as well as emotional support (Suldo et al., 2009; Turner, Gray, Anderman, Dawson & Anderman, 2013). Both boys and girls who perceived that their teachers had high academic expectations and felt supported by their teachers were more likely to feel positively about school and were more academically engaged (Suldo et

al., 2009; Turner & Patrick, 2004). Consequently, students' experiences with academic demands in middle school were influenced by the standards set by teachers as well as the available teacher support for attaining those standards (Klem & Connell, 2004; Wang & Holcombe, 2010).

Taken together, these findings indicate that the academic demands of middle school are significant to how students experience the transition. The findings from multiple studies suggest that the SCALE should assess students' perceptions of how well they are managing their schoolwork, performing academically, and their perceptions of teachers' academic expectations and support. Moreover, indicators of students' academic adjustment, engagement, and sense of school belonging should be included as constructs to validate students' perceptions of a successful academic transition.

**Changes to procedural expectations.** During the transition to middle school, students are expected to adapt to a number of organizational norms and expectations (Rudolph, Lambert, Clark, & Kurlakowsky, 2001). Students have reported that they are responsible for using lockers and combination locks between classes, arriving to multiple classes daily, on time and with the correct materials and homework (Deemer et al., 2003; Lane et al., 2006; Rudolph et al., 2001). Researchers have found that many teachers maintain different procedural expectations within their classrooms (Davis, 2006; Lane et al., 2006). Subsequently, after the transition, students tend to assume more responsibility for their movement throughout the school, autonomously organizing their materials for each class, and meeting additional expectations from teachers within each of their classrooms (McMullen, Shippen, & Dangel, 2007).

New procedural demands may not counter adolescents' developmental needs. Adolescents have a greater need for experiences that promote a sense of autonomy; but new procedural demands also require heightened self-control as well as decision-making capabilities

and forethought (Baker et al., 2001; Zimmer-Gembeck & Collins, 2003). Early adolescents are still developing decision-making capabilities; boys, especially, may develop decision-making skills more slowly than girls (Lane et al., 2006; McMullen et al., 2007; Wray-Lake, Crouter, & McHale, 2010). Teachers may also be more likely to enforce control over autonomy in schools with higher proportions of economically disadvantaged students (Lane et al., 2006). Thus, some students may find procedural changes difficult while others adapt easily.

Researchers have found that students' perceptions of their transition reflected the variability in their experiences with procedural demands (Akos & Galassi, 2004b; Deemer et al., 2003; Elias, 2001). Students frequently reported concerns about their abilities to remember their class schedule, to get to their classes on time, to remember the correct materials and homework, and to use their lockers properly (Akos & Galassi, 2004b; Deemer et al., 2003). Teachers and parents have corroborated students' procedural concerns, sharing their worries about their students' ability to navigate the schedule and keep up with materials during the school day (Akos & Galassi, 2004b). Yet, many students also looked forward to new procedural experiences, including the use of lockers and changing classes (Akos & Galassi, 2004b; Deemer et al., 2003). Moreover, despite apprehension, many reported that they felt that they had adapted successfully to the procedural demands of middle school.

Students who have difficulty adapting to procedural demands tend to experience negative adjustment outcomes, including higher levels of stress and increased incidences of disciplinary action (Lane et al., 2006; Rudolph et al., 2001). Thus, although procedural demands have received less empirical study than other areas of students' transition experiences, these types of demands are still important to their overall adjustment. According to the literature, the SCALE should assess students' perceptions of their ease in navigating their class schedule, organizing



class materials, and using lockers. Measures of students' academic achievement and behavioral adjustment, such as their compliance with school rules, should be used as comparable measures of adjustment to assess the validity of procedural items.

**Changes to social relationships.** The transition to middle school demands that students establish new relationships with teachers and peers and manage encounters with problematic peers (Cook, MacCoun, Muschkin, & Vigdor, 2008; Davidson et al., 2010). Students, as well as parents, have expressed particular concern about the social demands of middle school, reporting anxiety about students' abilities to establish new relationships (Akos & Galassi, 2004b). Parents expressed specific concerns that their children would establish supportive relationships, gain acceptance from peers, and avoid negative peer situations such as being pressured to engage in deviant behaviors by other students (Hamm et al., 2014). Therefore, both students and parents were concerned about both of the social demands that students would experience as they approached the transition.

***Establishing new relationships.*** Researchers have found that students often experience difficulties in establishing relationships with both teachers and peers following the middle school transition (Akos & Galassi, 2004b; Davidson et al., 2010; Deemer et al., 2003). Results from a number of studies suggest that students viewed their teachers as less supportive from the beginning of middle school and throughout the transition school year (Deemer et al., 2003; Klem & Connell, 2004; Lane, Pierson, & Givner, 2004; Niehaus et al., 2012). Girls, in particular, were more likely to feel greater teacher support than boys in 6th grade (Way et al., 2007; Wentzel et al., 2010). Results have been mixed regarding differing feelings of support between economically advantaged and disadvantaged students. Way et al. (2007) found that economically disadvantaged students reported lower feelings of teacher support, but no

differences were found in a study by Malecki and Demaray (2006). Students who had difficulty establishing supportive relationships with their teachers were more likely to experience other adjustment difficulties, such as lower grades, levels of intrinsic motivation, and more disciplinary referrals (Klem & Connell, 2004; Niehaus et al., 2012). Conversely, students who felt that they were supported by their teachers experienced higher grades and test scores and were more academically engaged, motivated, and oriented towards prosocial behaviors (Barber & Olsen, 2004; Brewster & Bowen, 2004; Malecki & Demaray, 2006; Niehaus et al., 2012; Wentzel et al., 2010). Perceptions of teacher support were particularly influential to the academic achievement of economically disadvantaged students (Malecki & Demaray, 2006).

Research findings suggest that students also have found it challenging to forge new relationships with peers. In one study, students reported that larger number of students in their middle school made it difficult to connect with new peers (Deemer et al., 2003). However, findings from other studies indicate that many students successfully established peer relationships following the transition (Davidson et al., 2010; Kingery & Erdley, 2007). Overall, students tend to perceive some difficulty in establishing peer relationships in middle school, but many seemed to successfully establish relationships. A number of studies report that students who successfully established peer relationships during the transition experienced more favorable adjustment, including higher levels of academic achievement and lower levels of loneliness in middle school (Davidson et al., 2010; Kingery & Erdley, 2007; Malecki & Demaray, 2006). Relatedly, middle school students who perceived that they were supported by peers were more likely to be motivated in school and to exhibit prosocial behaviors (Wentzel et al., 2010). Feeling supported by classmates was also found to positively impact students' reading achievement for economically disadvantaged students (Malecki & Demaray, 2006).

Furthermore, even if students did not form strong ties to their teachers, those who successfully established relationships with their peers had higher academic achievement and motivation (Davidson et al., 2010; Wentzel, Barry, & Caldwell, 2004). Thus, forming supportive peer relationships was especially important to a positive transition experience.

Students who had the most difficulty adjusting to the transition were those who were unable to form relationships with both teachers and peers (Davidson et al., 2010; Niehaus et al., 2012). These students were more likely to perform more poorly than their peers academically and were more likely to engage in problem behaviors and to associate with deviant peers (Davidson et al., 2010; Niehaus et al., 2012; Veronneau & Dishion, 2011). Researchers have also found gender differences in establishing relationships during the transition. Girls were thought to be more vulnerable to social difficulties due to their highly social nature (Roeser et al., 2008), but research findings indicate that boys experience social difficulties as well. Compared to girls, boys had lower feelings of teacher and peer support, but boys' feelings of teacher and peer support at the beginning of 6th grade declined at similar rates as did girls' during the transition year (Niehaus et al., 2012; Wentzel et al., 2010). Consequently, boys had lower feelings of support from teachers and peers following the transition, but both boys and girls experienced relationship difficulties. For all students, establishing supportive relationships with both teachers and peers had positive effects on their school adjustment following the transition to middle school.

***Managing problematic peer situations.*** Students reported that encounters with problematic peers, such as bullying and fighting, were one of the most difficult aspects of their transition experiences (Deemer et al., 2003; Elias et al., 1992; Ellerbock & Kiefer, 2013). Researchers have reported that the number of incidences of bullying, fighting, and other deviant

behaviors is higher in middle school than elementary school (Cillessen & Mayeux, 2007; Cook et al., 2008; Farmer, Hamm, Leung, Lambert, & Gravelle, 2011; Wang, Iannotti, & Nansel, 2009). In a study by Ellerbock and Kiefer (2013), students reported that the structure of middle school contributed to incidences of problematic peer behavior. Specifically, students perceived that the lack of supervision and presence of older peers during unstructured portions of the school day, such as before or after school or during lunch, provided opportunity for bullying and aggressive behaviors. Findings that some middle school teachers see bullying behaviors as normative and are reluctant to intervene (Yoon, Barton, & Taiariol, 2004) suggest that the availability of unstructured and unsupervised time contributes to occurrences of fighting and bullying (Ellerbock & Kiefer, 2013; Parault, Davis, & Pellegrini, 2007).

Students who were the victims of bullying in the 6th grade showed declines in school adjustment at the beginning and end of the transition year (Graham et al., 2006; Nansel, Haynie, & Simons-Morton, 2007). Students who bullied others were viewed as popular by their peers (Farmer, Estell, Leung et al., 2003; Galván, Spatzier, & Juvonen, 2011; Rodkin, Farmer, Pearl, & Van Acker, 2006). Consequently, bullying behaviors may be an effective strategy for establishing social status with peers following the middle school transition. The association between aggression and popularity strengthens in middle school (Galván et al., 2011). A study of middle school boys found that the perceived relationship between aggressive behaviors and popularity was higher for African American boys in comparison to White boys at the beginning of 6th grade, but that the association became equally as strong for White boys by the end of the 6th grade year (Xie, Dawes, Wurster, & Shi, 2013). Students' socioeconomic status, however, was not related to the association between popularity and aggression in this study. Yet, both bullies and bully-victims (i.e., students who both bullied others and were bullied themselves)

experienced higher levels of loneliness, poorer academic achievement and competence, and lower school-liking throughout the transition year (Erath, Flanagan, & Bierman, 2008; Graham et al., 2006). Economically disadvantaged students were more likely to be physically bullied, but economically advantaged students experienced bullying in other ways such as cyberbullying (Wang et al., 2009). Therefore, students may experience bullying or exhibit problematic behaviors regardless of their socioeconomic status. Males, especially minority males, were more likely to be identified as a bully or as a bully-victim and exhibited the most academic and behavioral difficulties in 6th grade (Graham et al., 2006; Wang et al., 2009). Students from racial/ethnic minority groups experienced discrimination from peers (Berkel et al., 2009), further increasing their risk of adjustment problems following the middle school transition (Burchinal et al., 2008).

Students with higher rates of deviant behaviors or discipline problems were more likely to suffer academically and be placed in less challenging classrooms (Farmer, Estell, Leung et al., 2003). Conversely, students who affiliated with positive, school-focused peers tended to experience favorable adjustment following the transition to middle school and throughout the remainder of the school year (Erath et al., 2008; Molloy, Gest, & Rulison, 2011; Veronneau & Dishion, 2011). Teachers also perceived that students who were more cooperative and had positive social skills were more likely to adjust easily to middle school (Akos & Galassi, 2004b; Elias et al., 1992). Results of a study by Fenning and Rose (2007) indicated that males and African American students engaged in more school misbehavior, although a study by Witherspoon and Ennett (2011) found that rates of misbehavior by White students increased at a higher rate during the transition year. Economically disadvantaged students were also more likely to exhibit school misbehaviors in 6<sup>th</sup> grade (Theriot & Dupper, 2009). Barber and Olsen

(2004) also found that girls reported more connections with deviant peers after the transition between 5th and 6th grades, whereas boys reported that their association with problematic peers declined during this time. Therefore, the findings are inconclusive regarding which students were more likely to have difficulty with resisting involvement in problematic behaviors.

Taken together, the social demands of middle school present challenges to students following the transition. Establishing relationships with peers and teachers, and managing problematic peer relations, such as bullying and engaging in deviant behaviors, are particularly salient experiences for students (Akos & Galassi, 2004b; Elias et al., 1992; Ellerbock & Kiefer, 2013). Although many students successfully established peer and teacher relationships and managed encounters with problematic peers, students who had difficulty with the social demands of middle school were more likely to experience negative adjustment outcomes including higher feelings of isolation, lower levels of school belonging, as well as academic difficulties (Niehaus et al., 2012; Way et al., 2007). The review of the literature suggests that the SCALE should contain items assessing both types of middle school social demands, such as students' ability to form relationships with peers and teachers and manage bullying/victimization experiences. Measures of students' social and behavioral adjustment, such as school belonging or engagement in aggressive behaviors, should be used to validate the social items in the SCALE.

**Summary.** In accordance with the principles of life course theory (Elder et al., 2003), the ease or difficulty of students' transition experiences tends to be related both to the demands that middle schools create for students, as well as students' responses to these demands. Studies of the middle school transition have reported that students, as well as parents and teachers, find that middle schools introduce a number of situational imperatives to 6th-graders including new academic, procedural, and social expectations and demands. In some cases students' experiences

with the academic, procedural, and social demands of middle school were differentiated by gender, race/ethnicity, or socioeconomic status (Berkel et al., 2009; Deemer et al., 2003; Graham et al., 2006; Malecki & Demaray, 2006; Niehaus et al., 2012; Theriot & Dupper, 2009; Wang et al., 2009; Wentzel et al., 2010). Students who were able to meet the demands of middle school were often more likely to exhibit positive indicators of adjustment, such as higher levels of academic achievement and engagement, feelings of school belonging, and were increasingly likely to exhibit prosocial behaviors (Barber & Olsen, 2004; Niehaus et al., 2012; Wang & Holcombe, 2010; Wentzel et al., 2010). However, not all students successfully meet the demands of middle school; those who do not tend to experience lower levels of academic engagement, feelings of school belonging, and higher incidences of problematic behaviors (Anderman, 2003; Rudolph et al., 2001; Urdan & Midgley, 2003).

The theoretical and empirical basis of the SCALE suggests that items should capture students' perceptions of their ease or difficulty in meeting the academic, procedural, and social demands of middle school. Measures used to validate the SCALE should include assessments of students' adjustment in these domains, such as their sense of school belonging, academic performance, engagement in bullying behaviors, and other comparable measures. Research findings have also suggested that students in different gender, socioeconomic status, and racial/ethnic groups may experience the academic, procedural, and social demands of middle school differently. Thus, the SCALE and its subscales should be examined for differences across gender, socioeconomic status, and racial/ethnic groups. In addition, relationships between the SCALE and SCALE subscales with concurrent and future measures of students' adjustment should be examined for differences across gender, socioeconomic status, and racial/ethnic groups.

## **The Validity of Measures of Students' Transition Experiences**

Past researchers have used a number of measures to assess students' perceptions of their experiences with the changes and demands they encountered during the middle school transition (Akos & Galassi, 2004b; Day et al., 2014; Elias et al., 1992). Assessments of students' perceptions of the middle school transition, as well as the perceptions of both parents and teachers, have been used to demonstrate the relationship between the ease or difficulty of students' transition experiences with both positive and negative indicators of their adjustment (Day et al., 2014; Hamm et al., 2014; Elias et al., 1992). Nevertheless, the reliability and validity of available measures of students' transition experiences are questionable.

In the following sections, I present the procedures researchers have used in past studies to assess the validity of measures of students' transition experiences. Following the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), I critique researchers' use of theoretical and statistical evidence in establishing the validity of available transition measures. I also critique researchers' use of specific types of statistical evidence including their assessments of the psychometric structure, reliability and stability of scales, their comparisons of scale scores, and comparisons of the relationships between scale scores to concurrent and later measures of students' transition adjustment. Finally, I combine the information regarding the empirical evidence of students' transition experiences with these critiques of the validation procedures used for available transition measures to guide the hypotheses for the validation of the SCALE in the present study.

**Compiling theoretical evidence.** A strength in the development of existing assessments of students' transition experiences is the use of theoretical, anecdotal, and empirical evidence to justify scale items. Though less common than the use of empirical research, some researchers



have conceptualized the validation of their scales using various theories including stage-environment fit theory (e.g., Akos, 2002) and Bronfenbrenner's ecological systems theory (e.g., Elias et al., 1992). More often, measures were constructed using information gathered from students, parents, and teachers, as well as practitioners and scholars familiar with the middle school transition (e.g. Akos, 2002; Akos & Galassi, 2004b; Elias et al., 1992). In addition, researchers used results of empirical research to construct scale items that reflected transition tasks or challenges commonly encountered by students (e.g. Deemer et al., 2003; Elias et al., 1992). Elias et al. (1992) also used empirical research to guide their decisions regarding the types of factors present within the SAT-MS following the use of factor analytic procedures. However, none of the current measures available have been revised since their inception to reflect current research findings or the opinions of current experts and are, therefore, potentially outdated.

SCALE items have been developed by experts and based on current empirical research as described more fully in the Methods. In the present study, interpretations of SCALE items are grounded in life course theory. Based on the review of relevant literature, it is anticipated that the SCALE will be composed of three subscales that reflect the academic, procedural, and social demands of the middle school transition.

**Compiling statistical evidence.** Many of the measures used in current studies of students' perceptions of their transition experiences have not been evaluated for validity using statistical evidence (e.g., Akos, 2002; Akos & Galassi, 2004b; Day et al., 2014; Hamm et al., 2014) or statistical evidence has not been reported (Akos & Galassi, 2004a; Deemer et al., 2003). Moreover, an assessment of rural students' transition experiences published in two studies (i.e., Day et al., 2014; Hamm et al., 2014), involved only a single item. Single-item scales have poor

measurement properties; no other information is available to assess whether or not the item is the best measure of the construct of interest and, consequently, assessments of construct validity are limited. Ideally, researchers should use at least three items per factor to represent a desired construct to ensure that the model is identified (Huck, 2012; Tabachnick & Fidell, 2007). Model identification is based on degrees of freedom (Bowen & Guo). *Degrees of freedom* are the difference between the number of known parameters and the unknown parameters to be estimated in a model (Bowen & Guo, 2012; Hoyle, 2012). Identified models have degrees of freedom greater than zero, indicating that they contain enough known information to allow for the estimation of parameters which are unknown and tests of model fit (Hoyle, 2012). Using at least three items per factor increases the likelihood that models are identified.

The Elias et al. (1992) transition instrument, the SAT-MS, included 28 items describing stressful tasks commonly associated with the middle school transition, such as “having to do harder schoolwork” or “forgetting your locker combination.” Students were asked to think about their experience with each task in the past month and rate their difficulty with each task on a 4-point, Likert-type scale ranging from *no problem* to *large problem*. The scale developers provided statistical evidence of the validity of interpretations made from the SAT-MS. For example, they established the psychometric structure of the scale, completed analyses of reliability, assessed mean score differences on latent factors, and correlated the SAT-MS with other measures of students’ adjustment. During the validation of the SCALE, these procedures and others recommended by the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999) will be used in the validation of the SCALE.

***Establishing the psychometric structure of existing scales.*** Elias et al. (1992) used factor analytic techniques to establish the psychometric structure of the SAT-MS. Nevertheless, a

weakness of their approach was failure to take the ordinal nature of response items into account. Failure to account for ordinal data in factor analytic procedures can potentially return biased parameters and estimates of standard errors (Schmitt, 2011). As a result, the likelihood of underrated item correlations and factor loadings is increased, which may contribute to errors when making determinations about the psychometric structure of scales (Holgado-Tello et al., 2010).

In the present study, both exploratory and factor analytic techniques will be used to establish and confirm the psychometric structure of the SCALE. These analyses will either support or refute the presence of three factors (i.e. the academic, procedural, and social demands of middle school) in the SCALE. The SCALE is ordinal in nature as each item contains a Likert-type scale with a small number of scale points designating ranked values, but the difference between values is unquantifiable (Bovaird & Koziol, 2012). In response, the present study will use factor analytic techniques that account for the ordinal nature of items present within the SCALE during both exploratory and confirmatory factor analytic procedures.

***Establishing scale invariance in existing scales.*** *Cultural bias* occurs when psychometric properties of a scale, such as scale language, influence groups to respond in certain ways (Knight & Hill, 1998). Even when scales relate to criterion similarly across groups, bias can exist by item (Millsap, 2007). Without the knowledge that cultural bias is present in a scale, researchers interpret group differences in scale responses as evidence of attitudinal differences between groups (Cheung & Rensvold, 2002; Knight & Hill, 1998). These interpretations from culturally biased scales are misguided as group differences are attributed to psychometric properties of the scale and not attitudinal differences between groups.

During validation procedures, researchers determine the presence or absence of cultural bias by testing for scale invariance. *Scale invariance* is present when the constructs assumed to be captured by a scale are the same for all participants, and all participants understand the scale constructs similarly, even if participants differ by gender, race/ethnicity, socioeconomic status or otherwise (Byrne, 2012; Kline, 2011; Knight & Hill, 1998). To establish scale invariance, researchers assess a number of scale properties, including the configural, metric, and scalar invariance of the scale (Cheung & Rensvold, 2002). *Configural invariance* is supported when the same number of factors is present within a scale and the same items load on the same factors for each group of interest (Byrne, 2012; Kline, 2011). If configural invariance is found, researchers can conclude that the same constructs are present for all groups. *Metric invariance* refers to consistency in the magnitude of loadings between items and factors across all groups of interest (Byrne, 2012). If metric invariance is present, researchers can conclude that the construct level of the scale is invariant, or that constructs are represented in the same way for each group (Kline, 2011).

Finally, *scalar invariance* also known as *strong factorial invariance* addresses the presence of invariance across measurement intercepts, or thresholds in the case of ordinal data, in addition to metric invariance (Millsap & Olivere-Aguilar, 2012). Finding scalar invariance allows researchers to be confident in their comparisons across groups' factor means such that groups have the same probability of attaining a mean score because the same factors are influencing all groups (Gregorich, 2006; Millsap & Olivere-Aguilar, 2012). Thus, differences across group means can be considered meaningful and not just the result of cultural differences affecting participants' responses on factor items (Gregorich, 2006). If analyses support configural, metric, and scalar invariance, the scale is considered to be invariant across groups

and researchers can be reasonably confident that a scale does not contain cultural bias (Byrne, 2012; Kline, 2011; Knight & Hill, 1998).

In the validation of the SAT-MS, the authors assessed the invariance of the scale structure by comparing the psychometric structure of the SAT-MS when it was distributed to two schools, one which served predominantly suburban, White students and the other which served urban students from predominantly minority racial and ethnic groups (Elias et al., 1992). This comparison indicated that there was some evidence of variance in the scale, as some items cross-loaded with different factors of SAT-MS across schools. These results suggest that there may have been cultural bias within the SAT-MS, but no further analyses of scale invariance were performed to assess sources of invariance and no changes were made to the scale. The use of more stringent statistical tests, like multiple-group confirmatory factor analysis, could have verified the presence of invariance and identified sources of cultural bias in the scale (Byrne, 2012; Kline, 2011; Knight & Hill, 1998). In the present study, multigroup confirmatory analysis procedures will be used to establish scale invariance for the SCALE, including configural, metric, and scalar invariance. Invariance assessments will be performed to assess the presence or absence of cultural bias across gender, socioeconomic status, as well as race/ethnicity as current empirical research indicates some differences in students' middle school transition experiences across groups (Berkel et al., 2009; Elias et al., 1992; Malecki & Demaray, 2006; Theriot & Dupper, 2009; Wang et al., 2009; Wentzel et al., 2010).

***Establishing scale reliability and stability of existing scales.*** The reliability of a scale describes how well scale scores consistently and predictably approach the true scores that represent a latent variable (DeVellis, 2012). There are multiple components to reliability, including internal consistency as well as stability. Internal consistency it is often assessed using

Cronbach's alpha, which measures how much a scale's variance is attributed to the relationships between scale items (Cronbach, 1951) and is commonly reported for measures (Iacobucci & Duhachek, 2003). With the exception of the SAT-MS, analysis of existing transition perceptions scales has not addressed scale reliability. By convention, Cronbach's alpha  $\geq .7$  is considered to indicate good reliability. For the SAT-MS, results of reliability analyses indicated good reliability for most of the subscales except for one, Academic Pressures, which was determined to have a Cronbach's alpha of .61 (Elias et al., 1992).

Stability refers to the correlation of the measure with itself over time (DeVellis, 2012). Pearson product correlations provide information regarding the linear association between scale scores from multiple test administrations (DeVellis, 2012; West & Finch, 1997). A critique of the use of Pearson's correlation in the measurement of scale stability is that it does not account for systematic error (DeVellis, 2012; Weir, 2005). *Intraclass correlations* (ICCs) are also used to measure scale stability. ICCs do account for systematic error in test administrations and can also account for different sources of systematic error, including measurement procedures, mean shifts, or rater error (McGraw & Wong, 1996; Shrout & Fleiss, 1979). ICCs for consistency assess the stability of scores, accounting for systematic error, between participants without attention to mean differences from the rater, similar to Pearson product correlations (Weir, 2005). ICCs for absolute agreement assume that mean differences are sources of systematic error (McGraw & Wong, 1996). Therefore, ICCs for absolute agreement measure score stability, accounting for mean shifts by the rater. Significant ICC coefficients and/or coefficients greater than .6 indicated that the SCALE is stable over time (Shoukri, Asyali, & Donner, 2004; Shrout, 1998; Weir, 2005).

Analyses of scale stability were not calculated or reported for the SAT-MS (Elias et al., 1992) or for other measures of students' transition perceptions. According to life course theory, transitions occur over time (Elder et al., 2003). The literature on the middle school transition suggests that many aspects of adjustment may decline between the fall and spring of 6th grade, e.g., school belonging or motivation (Anderman, 2003; Lepper et al., 2005; Niehaus et al., 2012). Thus, students' perceptions of their transition experiences may also be unstable but without findings from other studies, it is unclear whether or not the SCALE will remain stable over time.

In the present study, reliability will be assessed with both measures of internal consistency and stability. The internal consistency of the SCALE and its subscales will be calculated using Cronbach's alpha. Furthermore, stability will be assessed using the scores from the SCALE for both the fall and the spring of 6th grade. Given a lack of information regarding expected stability of the SCALE, Pearson product correlations and ICCs for consistency and absolute agreement will be conducted to explore the stability of the SCALE. The presence of invariance in the SCALE across gender, race/ethnicity, and socioeconomic status would also indicate that the SCALE was reliable across groups, but if scale variance is found, reliability analyses for each subscale using will be conducted for each gender, socioeconomic status, and racial/ethnic group (N. Bowen, personal communication, January 20, 2014; Millsap & Olivere-Aguilar, 2012).

***Establishing construct validity of existing scales.*** The relationship between scales and other comparable measures has been used to demonstrate construct validity. Validating researchers' interpretations about students' transition perceptions depends, in part, on the relationship between students' responses on perception assessments and other concurrent measures of their adjustment experiences (AERA, APA, & NCME, 1999; Messick, 1995).

Predictive validity is also established as the relationship between a scale and other comparable measures students' experiences at a later time (DeVellis, 2012). In addition, researchers hypothesize whether scores from the scale of interest will demonstrate strong relationships, also known as demonstrating convergent validity, or weak or negative relationships, known as divergent or discriminant validity, to other similar scales of interest (AERA, APA, & NCME, 1999; Campbell & Fiske, 1959; Downing, 2003; O'Leary-Kelly & Vorkurka, 1998).

Researchers can also use latent mean differences across groups as evidence of construct validity (Vandenburg & Lance, 2000). Latent mean differences across groups that are consistent with those found in past research may indicate that factors are assessing the intended constructs (Cronbach & Meehl, 1955). Equivalence among latent factor across groups may also indicate that constructs are operating consistently across groups (Byrne, 2012).

For the validation of the SAT-MS, researchers correlated the students' total scores to concurrent measures of their academic achievement, perceptions of the school climate, self-concept, as well as teacher-rated assessments of students' behaviors (Elias et al., 1992). The researchers also correlated the concurrent measures of students' adjustment with the individual subscales of the SAT-MS and examined the differences in relationships between subscales and other measures of students' adjustment across gender. Finally, researchers compared latent factor means across gender and across two schools located in economically diverse areas. Predictive validity was not examined for the SAT-MS. For both boys and girls, subscales and the scale as a whole correlated with measures of students' self-concept and perceptions of the school context but not with teachers' ratings of students' behavior. However, the strength of relationships between the subscales of the SAT-MS and measures of students' adjustment were stronger for boys than girls. Mean differences across gender on the SAT-MS as well as its



subscales were not found; however, mean differences were observed between schools. Students from an urban school serving more economically disadvantaged and racially/ethnically diverse students indicated increased difficulty on transition tasks than students in a suburban, blue-collar school. Researchers used these findings as evidence for the construct validity of the SAT-MS.

In studies that have included other assessments of students' perceptions of the transition, students' middle school transition experiences were associated with a number positive and negative of indicators of their adjustment following the transition to middle school. Students who perceived a more difficult transition to middle school were more likely to have lower feelings of school belonging (Day et al., 2014) and self-efficacy (Elias et al., 1992). Studies of parents' and teachers' perceptions of students' adjustment into middle school corroborate the association between students' transition experiences and indicators of their adjustment. Findings suggest that teachers perceived that students who adjusted successfully had better social skills, were cooperative with teachers, responsible, focused on school work, and academically successful (Akos & Galassi, 2004b; Elias et al., 1992). Furthermore, parents from one elementary school correctly predicted that students from that school would have more difficulty after the middle school transition; however, students from those schools displayed academic and behavioral adjustment difficulties, whereas parents anticipated social difficulties (Hamm et al., 2014).

In the present study, convergent and divergent relationships between the latent factors found for the SCALE and concurrent measures of students' adjustment will be assessed. Student adjustment measures will include of students' social, behavioral, and academic adjustment, social and behavioral characteristics and school behaviors, academic behaviors, and feelings of school belonging. In addition, measures of students' adjustment from the spring of 6th grade

will be used to establish the predictive validity of the SCALE. I predict that students' overall scores on the SCALE will demonstrate convergent validity with the popular, affiliative, and academic ICS-T subscales and students' overall adjustment from the ICS-T (Cairns, Leung, Buchanan, & Cairns, 1995). Students' scores on the SCALE will also demonstrate convergent validity with the Academic Effort scale (Gest, Rulison, Davidson, & Welsh, 2008), the Identification with School questionnaire (Voelkl, 1996), teachers' ratings of whether students are well-liked and are class leaders (Estell, Farmer, & Cairns, 2007), and students' school belonging as measured by the PSSM-B (Hagborg, 1998). The SCALE will demonstrate divergent validity with concurrent ratings of students' negative behaviors from the aggressive subscale of the ICS-T (Cairns et al., 1995), teachers' ratings of students being frequently bullied or bullying others (Estell et al., 2007), and the Defiance scale (Midgley et al., 2000).

In previous studies, students who experienced academic success and felt academically supported were more likely to be academically motivated, engaged, have a greater sense of school belonging, positive feelings about school; and higher levels of academic achievement (Klem & Connell, 2004; Suldo et al., 2009; Turner & Patrick, 2004; Wang & Holcombe, 2010). Subsequently, it is expected that students who perceive ease with the academic demands of middle school will also experience positive academic adjustment. Students' scores on the academic subscale of the SCALE in the present study are hypothesized to demonstrate convergent relationships with the academic ICS-T subscale, the overall adjustment scale from the ICS-T (Cairns et al., 1995), Identification with School (Voelkl, 1996), the Academic Effort scale (Gest et al., 2008), reading and math achievement scores, students' sense of school belonging (Hagborg, 1998), and teachers ratings of students as a class leader (Estell et al., 2007).

Results of a number of studies have indicated that students who were able to establish positive relationships with teachers and peers and manage problematic peer interactions were more likely to exhibit prosocial behaviors, be liked, and feel favorably towards school (Erath et al., 2008; Niehaus et al., 2012; Wentzel et al., 2010). Accordingly, in the present study, it is predicted that the social subscale of the SCALE will demonstrate positive relationships with sense of school belonging (Hagborg, 1998), being well-liked (Estell et al., 2007), overall adjustment, and teachers' ratings of student popularity and affiliative nature on the ICS-T subscales (Cairns et al., 1995). Divergent relationships will be present between the social subscale and teachers' ratings of students being aggressive (Cairns et al., 1995), frequently bullied or bullying others (Estell et al., 2007), and the Defiance scale (Midgley et al., 2000).

The academic and social subscales are expected to correlate with indicators of adjustment specific to each of their respective adjustment domains, as detailed above, and demonstrate discriminant validity through a lack of correlations with other adjustment measures. Specifically, the academic subscale would demonstrate a lack of relationships social and behavioral adjustment measures in the present study and the social subscale would also demonstrate a lack of relationships with academic measures. However, in past studies, students who were more academically adjusted were also more likely to be socially adjusted following the transition to middle school. For example, students who maintained positive relationships with peers and teachers and managed problematic interactions with peers were also more likely exhibit positive indicators of academic adjustment, such as academic engagement or achievement (Erath et al., 2008; Niehaus et al., 2012; Way et al., 2007) and those engaging in problematic behaviors were more likely to experience academic difficulty (Farmer, Estell, Leung et al., 2003). Therefore, it is also possible that the academic subscale will relate positively to

favorable indicators of students' social and behavioral adjustment, including popular and affiliative behaviors (Cairns et al., 1995) and being liked (Estell et al., 2007). Divergent relationships may also be present between the academic subscale and aggression (Cairns et al., 1995), bullying or being bullied by others (Estell et al., 2007), and the Defiance scale (Midgley et al., 2000). The social subscale may also relate positively associated with indicators of students' academic adjustment, such as the academic ICS-T subscale (Cairns et al., 1995), the Academic Effort scale (Gest et al., 2008), being a class leader (Estell et al., 2007), Identification with School (Voelkl, 1996), and math and reading achievement scores. Although these relationships are not anticipated, their presence would also support the construct validity of the academic and social subscales.

Finally, it is expected that the procedural subscale may not correlate with the other measures in the present study, as none of these measures assess students' organizational behaviors. However, given that many procedural demands involve organizing homework or classwork (Akos & Galassi, 2004b; Deemer et al., 2003); students' perceptions of their experiences with procedural demands may be related to indicators of their academic adjustment. Subsequently, it is anticipated that the procedural subscale of the SCALE will demonstrate convergent relationships with indicators of students' academic adjustment. Specifically, convergent relationships would be present between the procedural subscale and academic ICS-T subscale, the overall adjustment score from the ICS-T (Cairns et al., 1995), the Academic Effort scale (Gest et al., 2008), reading and math achievement scores, and students' feelings of school belonging (Hagborg, 1998), Identification with School (Voelkl, 1996), and teachers rating students as frequently a class leader (Estell et al., 2007).

Predictive validity will compare students' SCALE scores from the fall of 6th grade and to students' scores on the other measures of their adjustment from the spring of 6th grade. Positive and negative relationships between the SCALE and future measures of students' adjustment are expected to demonstrate the same convergent and divergent patterns as concurrent measures. Concurrent and predictive relationships between the SCALE and other measures of students' adjustment will also be completed for gender, socioeconomic status, and racial/ethnic groups. Elias et al. (1992) found that correlations between subscales of the SAT-MS and other measures of students' adjustment were stronger for boys. Correlations between measures of adjustment and SAT-MS subscales were more concentrated for girls as more adjustment measures correlated with the Peer Relationships subscale. The present study will assess whether similar patterns are present across boys and girls for the SCALE. Correlational differences between students' transition perceptions from the SAT-MS (Elias et al., 1992) and measures of students' adjustment were not assessed across socioeconomic status or racial/ethnic groups. The present study will examine these relationships across socioeconomic status and race/ethnicity, anticipating that relationships will be similar across these groups if the SCALE is determined to be invariant.

Finally, latent mean differences of SCALE factors will be compared across gender, socioeconomic status, and race/ethnicity. Findings of equivalence across groups will indicate that the underlying construct of each factor is operating consistently for each group (Byrne, 2012; Vandenburg & Lance, 2000). Differences in latent factor means that could contribute to evidence of construct validity would demonstrate agreement with results of past research. In terms of the academic factor, differences in means may indicate that female students may adapt to the academic demands of middle school more favorably as past research indicated that boys

perceived a more difficult time adapting to increased amounts of academic work (e.g. Deemer et al., 2003). Accordingly, I anticipate lower mean scores on the academic subscale for boys.

Although no studies can be found that specifically examined differences in students' perceptions of academic demands across socioeconomic status or race/ethnicity, economically disadvantaged students and African American students were more likely to experience lower academic achievement following the middle school transition (Akos et al., 2014). Consequently, I expect that economically disadvantaged students and African American students may also have lower mean scores on the academic subscale.

Girls may also have more favorable ratings on the social factor as research has reported that they experience greater feelings of teacher support in 6th grade (e.g., Malecki & Demaray, 2006; Way et al., 2007; Wentzel et al., 2010). However, other findings have indicated that both boys and girls experience social difficulties involving peers (e.g. Niehaus et al., 2012; Wentzel et al., 2010), so no differences may be found. Students from racial/ethnic minority groups may perform less favorably on the social factor as they are increasingly likely to experience bullying, discrimination, and behavioral difficulties (e.g., Berkel, et al., 2009; Graham et al., 2006; Wang et al., 2009). Differences in social adjustment between economically advantaged and disadvantaged groups have been mixed, with some researchers finding no differences in perceptions of teacher and peer support, social integration, or bullying experiences (e.g., Malecki & Demaray, 2006; Wang et al., 2009) and others noting social difficulties (e.g., Theriot & Dupper, 2009; Way et al., 2007). Therefore, no differences may be evident across socioeconomic groups on the social factor or economically disadvantaged students may exhibit a lower latent factor mean. Currently, no studies were identified that have indicated explicit

differences across groups with regards to procedural demands and, as a result, no specific relationships are anticipated.

### **The Present Study**

The transition to middle school is an important and historically difficult time in students' education (Anderman & Mueller, 2010; Eccles et al., 1993; Eccles & Roeser, 2009). The presence of new academic, procedural, and social demands that may not meet the developmental needs of early adolescents have been identified as a source of difficulty for students making the move to middle school (Eccles et al., 1993; Midgley et al., 2002; Zimmer-Gembeck & Collins, 2003). Studies of students' perceptions of the transition provide additional insight into the types of changes and demands that are difficult in the transition (Akos & Galassi, 2004b; Deemer et al., 2003; Elias et al., 1992; Hamm et al., 2014). Researchers have found variability in students' experiences with the demands of middle school and have associated their perceptions with both positive and negative indicators of their transition adjustment (Day et al., 2014; Elias et al., 1992; Hamm et al., 2014). Thus, assessments of students' transition experiences can be valuable to future studies of the middle school transition as they examine both students' perceptions of the demands of middle school and their ease or difficulty in responding to those demands.

The appropriateness of the measures presently available is questionable given the limited nature of scale validation. In the present study, I use contemporary techniques to establish or refute the validity of interpretations drawn from a new scale recently distributed to students from schools with diverse student bodies. During this study, the following goals and hypotheses will be addressed:

- Goal 1. Establish and confirm the factor structure of the SCALE through exploratory factor analysis and confirmatory factor analysis.

- Hypothesis 1: The SCALE will contain three subscales representing students' academic, procedural, and social adjustment.
- Goal 2. Demonstrate the scale invariance of the SCALE across gender, race/ethnicity, and socioeconomic status.
  - Hypothesis 2: Configural, metric, and scalar invariance, indicating the consistency of the overall scale structure, item loadings and thresholds, will be present across gender.
  - Hypothesis 3: Configural, metric, and scalar invariance, indicating the consistency of the overall scale structure and of item loadings, will be present across race/ethnicity for African American and White groups of students.
  - Hypothesis 4: Configural, metric, and scalar invariance, indicating the consistency of the overall scale structure, item loadings and thresholds, will be present across economically advantaged and disadvantaged groups of students.
- Goal 3. Establish the reliability and stability of the SCALE.
  - Hypothesis 5: If scale variance is found for gender, socioeconomic status, or race/ethnicity, reliability will be tested separately for each group.
  - Hypothesis 6: The total scale and subscales may demonstrate stability over time.
- Goal 4. Demonstrate the construct and predictive validity of the SCALE.
  - Hypothesis 7. The SCALE will demonstrate construct validity with concurrent measures of students' adjustment and mean score differences.
  - Hypothesis 8. The SCALE as completed at the beginning of the school year will demonstrate predictive validity with corresponding measures of students' adjustment at the spring of 6th grade.



## Chapter 4: Methods

The present study includes data from eleven control schools from a larger study, *SEALS: Supporting Early Adolescents' Learning and Social Success*. SEALS is a randomized controlled trial of the efficacy of the SEALS professional development program for teachers that sought to improve students' adjustment in the middle school transition. Within matched pairs of schools, schools were randomly assigned as either an intervention school that would receive the professional development program or a control school that carried out its usual daily activities. In the larger study, data were collected at four time points including the fall and spring of 6th grade and the fall and spring of 7th grade. As the intent of the current study is to validate a measure that assesses students' perceptions of their experiences during the transition year, data were used from the first two time points: fall and spring of 6th grade.

### Participants

**Schools.** Seven middle schools (grades 6 through 8) participated in SEALS as control schools. All schools were located in one state in the southern region of the United States. All schools were public, non-charter schools from metropolitan areas. Like other traditional middle schools which separate grade levels of students, sixth-grade students attended classes only with other sixth-grade students. Within grade levels, students were divided into teams, with a range from two to four teachers per team.

School size ranged from 595 to 1,052 students with an average size of 814 students ( $SD = 156.47$ ). The percentage of minority students per school ranged from 20.00% to 71.00% ( $M = 42.21\%$ ,  $SD = 17.77\%$ ). The largest racial/ethnic groups represented in schools were White,

African American, and Latino. The percentage of students on free and reduced lunch ranged from 21.80% to 61.50% between schools ( $M = 44.19\%$ ,  $SD = 15.12\%$ ). Proficiency rates in reading ranged from 61.00% to 88.10% ( $M = 71.21\%$ ,  $SD = 8.88\%$ ) and in math ranged from 73.90% to 94.10% ( $M = 83.77\%$ ,  $SD = 6.01\%$ ). See Table 1 for additional information on the schools in the present study.

**Students.** Participants were 6th-grade students ( $N = 761$ ). Three hundred ninety-eight students were female (52.30%). Gender information, obtained from school record data, was missing for 33 students (4.34%). Schools were relatively racially/ethnically diverse. In the resulting sample of participants, the majority of students were White ( $n = 398$ , 52.30%), 148 (19.44%) were African American, 134 (17.61%) Latino, 28 (3.68%) Multi-Racial, 17 (2.23%) Asian, 2 (.26%) Pacific Islander, and 1 (.13%) American Indian. Race/ethnicity, obtained from school record data, was missing for 33 students (4.34%). The majority of students were not identified as part of the exceptional students/ special education program ( $n = 659$ , 85.15%;  $n = 36$ , 4.73% missing), as indicated in school record data. School record data obtained for participants indicated that nearly half of the sample was eligible for free/reduced lunch ( $n = 366$ , 48.10%; 4.73% of participants missing).

## **Procedures**

Project staff visited classrooms in the fall of 6th grade year to explain the study to students and to distribute informed consent forms to be sent home to parents/guardians. Students who received consent from their parents/guardians and also assented to participate returned their consent forms to their homeroom teacher. A school liaison then returned all of students' consent forms to project staff.

In both the fall and the spring of 6th grade, trained research staff visited each school to administer student surveys to large groups of students. The trained staff members followed a research protocol and guided students through all the survey items. Students were able to ask questions, if needed. For their participation in each administration of the survey, students received a school supply item (e.g., a pencil). The response rate at the fall of sixth grade was 88.7%, with 761 out of 858 consented students completing the survey. Classroom teachers completed a set of surveys in which they rated the adjustment of each participating student at both the fall and the spring of 6th grade. Teachers were asked to complete surveys on their own time, and received financial compensation for their participation.

## **Measures**

Measures included in the present study are the SCALE and instruments used to validate the scale. Instruments used to validate the scale were chosen based on the literature on students' transition experiences and represent indicators of students' academic, social, and behavioral adjustment. Measures of students' adjustment were included to assess the concurrent and predictive construct validity of the SCALE and its subscales. Students' demographic characteristics were incorporated into many analyses to assess inconsistencies in the validity and reliability of the SCALE across gender, racial/ethnic, and economic groups.

**Perceptions of Middle School Transition Experiences.** The SCALE measures 6th-grade students' perceptions of their adaptation to their new middle school. The scale is composed of 20 items that measure academic, procedural, and social behaviors typically associated with students' transition into middle school. Example items include, "forgetting your locker combination" and "finding kids I can sit with at lunch" (see Table 2 for a complete list of

scale items). Students rate their perceptions about how much of a problem each task was for them on a four point, Likert-type scale from *no problem* to *large problem*.

***Development of the SCALE.*** The SEALS research staff created the SCALE using the SAT-MS (Elias et al., 1992) as a model. Seventeen out of the 28 items from the SAT-MS were used or slightly modified for the SCALE. The three remaining SCALE items were created using contemporary research literature about students' transition experiences and difficulties, and through consultation with experts on early adolescent development and the middle school transition.

***Content validity of the SCALE.*** During the development phase of the SCALE, two types of experts (psychologists and middle school students) were consulted to establish the content validity of the SCALE. Both groups were asked whether or not items reflected behaviors and experiences essential to middle school students' transition adjustment and identified key behaviors or experiences that were absent from the scale. Both groups were also consulted regarding the developmental appropriateness of the wording of scale items. Modifications were made to the SCALE where recommended and confirmed with both groups of experts prior to the administration of the SCALE to research participants.

***Social, behavioral, and academic adjustment.*** The Interpersonal Competence Scale-Teacher is an 18-item measure that addresses students' academic, behavioral, and social adjustment (ICS-T; Cairns et al., 1995). For example, teachers are asked to rate a student on items like [child] "argues" (behavior), "is friendly" (social), or "is good at math" (academic). Teachers rate each students' adjustment on each item on a seven-point, Likert-type scale from *never/very good at* to *always/not very good at* with a third anchor, *sometimes or so-so*, in the center of the scale. Reliability for the ICS-T for boys ( $\alpha = .87$ ) and girls ( $\alpha = .81$ ) has been very

good (Farmer & Hamm, 2010). The present study included four subscales of the ICS-T, including the Aggressive (argues, gets in trouble at school, fights), Popular (popular with boys, popular with girls, lots of friends), Affiliative (smiles, friendly), and Academic (good at math, good at spelling) subscales. Reliability for subscales has been very good when used with early adolescents ( $\alpha = .84$ ,  $\alpha = .83$ ,  $\alpha = .74$ ,  $\alpha = .80$  for each subscale, respectively; Farmer et al., 2012) as well as across racial/ethnic groups (Farmer & Hamm, 2010; see Table 3).

**Social adaptation.** Teachers were asked to rate students on seven items that assess students' social abilities on a 7-point, Likert-type scale (Estell et al., 2007). Four of the seven items were used in the present study. Items included *frequently class leader to never class leader*, *frequently bullied by peers to never bullied by peers*, *never bullies peers to frequently bullies peers*, and *liked by peers to not liked by peers*. A middle anchor of *sometimes* was present for each item. Three-week test-retest reliability for individual items has ranged from  $\alpha = .72$  -  $.93$  (Farmer, Estell, Bishop et al., 2003) and three-month test-retest reliability for individual items has ranged from  $\alpha = .60$  -  $.70$  (Farmer, Irvin, Sgamatto, Dadisman, & Thompson, 2009).

**School behaviors.** Midgley et al.'s (2000) Defiance scale contains five items that assess students' perceptions of their behavior in school. Students indicate their agreement with items such as, "I sometimes annoy my teacher during class", on a 5-point, Likert-type scale from 1 = *not at all true* to 5 = *very true*. Scores for all items are averaged and used as a composite score with higher values indicating greater behavioral adjustment difficulties. The scale has demonstrated high reliability in past studies with diverse adolescents (Cronbach's  $\alpha = .91$ , Gregory & Ripski, 2008).

**Academic behaviors.** The Academic Effort Scale (Gest et al., 2008) contains five items that assess teachers' perceptions of each student's academic effort. Scale items were derived

from other, well-validated scales like the Social Health Profile (Conduct Problems Prevention Research Group, 1999). Teachers rate their agreement on a 5-point, Likert-type scale on items like, “[student] works hard at school” or “does not try hard at school work” [reversed] from *very well* to *not at all well*. The Academic Effort Scale has shown high reliability ( $\alpha = .92$ , Gest et al., 2008).

**School belonging.** The Psychological Sense of School Membership-Brief (PSSM-B) (Hagborg, 1998) measures students’ sense of school belonging. The 11-item scale contains statements like, “I feel a real part of my school” or “I am included in lots of activities in my school”. Students indicate their agreement with each statement on a 5-point, Likert-type scale ranging from 1 = *completely false* to 5 = *completely true*. Across diverse samples of early adolescents, this scale has returned Cronbach’s alphas ranging between .71 - .88 (Hagborg, 1998; Hamm, Farmer, Dadisman, Gravelle & Murray, 2011).

**School valuing.** Students’ feelings of school involvement or valuing were assessed through seven items from the Identification with School questionnaire (Voelkl, 1996). The seven questions ask students to respond to prompts such as, “many of the things we learn in class are useless” or “dropping out of school would be a huge mistake for me” on a 5-point, Likert-type scale ranging from *strongly disagree* to *strongly agree*. Students’ scores on the seven items are averaged, with higher scores indicating a higher identification or valuing of school. This scale has shown strong construct validity through high correlations with class participation and student achievement measures (Finn & Frone, 2004; Voelkl, 1996). Cronbach’s alphas for this scale have ranged from .61 - .87 across diverse samples of early adolescents ( $\alpha = .78$  - .85 for White students and  $\alpha = .59$  - .70 for African American students; Farmer & Hamm, 2010).

**Academic achievement.** Students' percentile scores on-state level standardized tests were collected as representations of students' academic achievement. Reading and math percentile scores were collected from the spring of 5th and 6th grades. When students had multiple scores from one school year, e.g. in the case of a retest, the highest percentile score was retained for analyses.

**Demographic variables.** Students' gender, race/ethnicity, and free/reduced lunch status were taken from school record data. Gender and race/ethnicity were dummy coded so that 0 = female and White, respectively. Lunch pay status originally included free, reduced, and full-pay lunch status categories. Free and reduced lunch status students were combined into a single category. Lunch status was dummy coded so that 0 = full-pay lunch status and 1 = free/reduced lunch status.

### **Plan of Analyses**

The goal of the present study was to assess the validity of interpretations made from the SCALE, a measure of students' middle school transition experiences. During the validation of the SCALE, I addressed limitations of past assessments of similar measures of students' transition experiences. In particular, I focused on the ordinal nature of scale items and utilized stringent analytic procedures to assess the configural, metric, and scalar invariance of the SCALE as well as invariance of factor variances, covariances, and residuals across gender, racial/ethnic, and socioeconomic status groups.

Theoretical and empirical information was presented to provide support for the content of the SCALE. Further analyses provided statistical evidence for the validation of the SCALE. Analyses of missing data, descriptive statistics, and the Kaiser-Meyer-Olkin (KMO; Kaiser, 1970) measure of sampling adequacy were used to determine the suitability of the data for factor

analytic procedures. Analyses were conducted using SPSS (IBM Corporation, 2013), R 3.1.1 (R Core Team, 2014), and Mplus 7.11 programs (Muthén & Muthén, 2012). Power analyses were conducted for each of the factor analytic techniques conducted in the present study.

Factor analyses were conducted using MPlus software. Techniques were used to account for the ordinal and nested nature of the data in the present study. As recommended by Bowen & Guo (2012) and Muthén & Muthén (2012), a robust-least squares estimator that utilizes polychoric correlations, the WLSMV estimator, was used in analyses. The WLSMV estimator accounts for the ordinal nature of the data and is robust to violations of univariate and multivariate normality, a common characteristic of ordinal data (Bovaird & Koziol, 2012). In addition, when using ordinal factor analytic techniques in MPlus, it is important for bivariate combinations of all item values to be present. In other words, all bivariate combinations of items ratings must contain a value in each cell (Byrne, 2012; Muthén, Du Toit, & Spisic, 1997). For example, if a value of four on Item 1 is never present in combination with a rating of two on Item 2, this cell has no data, which causes problems in factor analytic procedures with ordinal data. Assessment of bivariate combinations of SCALE item values revealed a number of bivariate combinations that contained no data, particularly for ratings of three or four, which indicated difficulty on a transition element. Subsequently, values of three and four were collapsed into one rating category for fifteen items during analyses using structural equation modeling. Finally, students were also nested in schools. The school identification number was used as a cluster variable to account for the non-independence of observations based on school membership on parameter estimates and standard errors (Bowen & Guo, 2012; Muthén & Muthén, 2012).

The structure of the SCALE was determined using exploratory factor analysis (EFA). Eigenvalues greater than one, scree test analyses, and factor loadings of at least  $\pm.3$  were used to



determine the number of factors present in the SCALE (Cattell, 1966; Huck, 2012; Kline, 2011; Tabachnick & Fidell, 2007). Confirmatory factor analytic (CFA) procedures were utilized to validate the structure of the SCALE. The invariance of the SCALE across gender, socioeconomic status, and race/ethnicity was tested multiple group confirmatory factor analysis (MGCFAs). Model fit for all analyses was determined using goodness of fit indices, including chi-square ( $\chi^2$ ;  $p > .05$ ), root mean square error of approximation (RMSEA;  $\leq .05$  for close fit and between .05 - .08 for reasonable fit), comparative fit index (CFI;  $\geq .95$ ), and the Tucker-Lewis index (TLI;  $\geq .95$ ; Bollen, 1989; Bowen & Guo, 2012; West, Taylor, & Wu, 2012). For analyses using the WLSMV estimator, chi-square statistics are uninterpretable, but the corrected p-value is interpretable (Muthén & Muthén, 2012). A 90% confidence interval is also available with RMSEA in Mplus. The upper bound of the confidence interval is ideally  $\leq .08$  (Bowen & Guo, 2012).

Using SPSS, the reliability of the SCALE was established using Cronbach's alpha as a measure of internal consistency. The stability of the SCALE factors was assessed using Pearson product correlations and intraclass correlations (ICCs). Finally, concurrent and predictive construct validity of the SCALE was established by assessing the relationships between the SCALE and external measures of students' adjustment. Concurrent and predictive relationships were examined across gender, free/reduced lunch status, and racial/ethnic groups. Latent mean differences across gender, socioeconomic status, and race/ethnicity group were also analyzed as evidence of construct validity. Mplus was used to calculate both the relationships between factors and measures of students' adjustment as well as latent mean differences.

## Chapter 5: Results

### Missing Data

Seven hundred sixty-one students from eleven middle schools participated in the SEALS study during the sixth-grade year. Six hundred sixty-four students participated in during the spring of sixth grade, a difference of 12.61%. Of the 96 students no longer participating in the SEALS study in the spring of 6th grade, thirteen students (1.71%) were no longer consented to participate. Reasons for non-participation were unknown but could have included voluntary exit from the study, being sick the day of survey administration, or leaving the participating school.

Assessment of missing data can include examination of the percentage of missing values in a dataset, the number of cases with missing data, missing data within a case, and patterns of missing data (Tabachnick & Fidell, 2007). Across all waves of the study, the overall percentage of missing data values was 4.27%. During the first wave of the study, .001% of data values were missing. Only 59 participants (7.75%) were missing some values for the fall. The number of missing values on the SCALE for participants with missing data ranged from one to eleven, with the majority of participants only missing one or two values. During the second wave of the study in the spring of 6th grade, 7.78% of data values were missing. Excluding those students who did not participate in the spring of 6th grade, 61 cases (9.18%) had missing data. The number of missing data values on the SCALE per case ranged from one to ten, with most participants only missing one or two values.

Patterns of missing values can indicate group differences prior to a MGCFA (Byrne, 2012). This type of missingness is called *missing not completely at random (MNAR)* (Baraldi &

Enders, 2010; Kline, 2011). If data are MNAR, bias can enter the sample and may influence study conclusions. Data that are *missing completely at random (MCAR)* or *missing at random (MAR)* are able to be further analyzed. Little's MCAR test (Little, 1998) returned significant values for both the fall and the spring ( $\chi^2 = 115.62$ ,  $df = 89$ ,  $p = .03$ ;  $\chi^2 = 198.26$ ,  $df = 141$ ,  $p = .001$ , respectively), signifying that missing data were not MCAR. Therefore, additional analyses were completed to determine whether data were likely MAR or MNAR. Visual inspection of missing data patterns determined that missing data appeared random. The most common pattern of missing data for the fall was missing both reading and math standardized test scores, occurring in 50 cases. For the spring of sixth grade, the most common missing pattern was the lack of data for those who did not participate in the study during the spring of 6th grade. Correlations between missing data values demonstrated that missing data on teacher-rated scales were highly correlated, demonstrating that if teachers missed values for one scale, they were more likely to miss values on other scales. This was also evident for self-report scales in the spring; if students were missing any self-report scale scores, they were more likely to be missing all of the self-report items. Cross-tabulation tables and t-tests from the Missing Value Analysis function in SPSS were used to assess any significant relationships between missingness between SCALE items with values on other scales, school membership, gender, socioeconomic status, and race/ethnicity. No associations were found.

In summary, the amount of data missing for both waves of the SEALS study was low. In addition, although a significant chi-square from Little's MCAR test (Little, 1998) suggested that data were not missing completely at random, other characteristics of the missing data provided evidence that data were likely missing at random and were not MNAR for both the fall and the spring of 6th grade (Dong & Peng, 2013; Tabachnick & Fidell, 2007). When using Mplus and

the WLSMV estimator, the presence of some missing values is not problematic if data are not MNAR and if the amount of data per variable and per pair of variables, also known as *covariance coverage*, is greater than 90% during analyses (Muthén & Muthén, 2012). The amount of data per variable and covariance coverage was assessed for every subsequent analysis in the present study and data met these requirements. As a result, additional cases with missing data were not removed from the dataset and multiple imputation was not considered necessary.

### **Descriptive Statistics**

Data were screened for issues that would interfere with factor analytic procedures. Using SPSS software, means, variances, proportions of response values for each SCALE item, standard deviations, skewness, and kurtosis values were calculated for all SCALE items and for additional scales used in later analyses for both the fall and the spring of 6th grade. Bivariate polychoric correlations between SCALE items and ordinal measures of students' adjustment as well as polyserial correlations between SCALE items and continuous measures of students' adjustment were calculated. Mahalanobis distance values were analyzed to determine the multivariate normality of the data for both the fall and the spring. Finally, KMO analyses (Kaiser, 1970) for sampling adequacy were completed using Kerns' KMO function in R (Kerns, 2007).

**Univariate statistics.** Descriptive statistics for SCALE items for both the fall and the spring of 6th grade are presented in Table 4 and Table 5. Ranges were all within the expected limits. Means for SCALE items suggest that most students felt that they adjusted to the challenges of middle school favorably, ranging from 1.31 to 2.27, with a rating of one meaning *no problem*. Standard deviations ranged from .69 to 1.16, indicating some variables varied more than others. Spring means, variances, and standard deviations were similar to those in the fall of 6th grade.

Absolute values of skewness greater than  $\pm$  two or  $\pm$  three are considered extreme as are absolute values of kurtosis above seven (Bowen & Guo, 2012; Ware, Ferron, & Miller, 2013). The level of skewness and kurtosis for SCALE items were relatively high for a number of variables, particularly Items 2, 6, 19, and 20, but were not necessarily extreme. Skewness and kurtosis values were similar for SCALE items for the spring of students' sixth-grade year (see Table 5). Item 1, "getting lost and not being able to find your way around school" was slightly more skewed in spring than the fall and the mean was slightly lower, indicating that students reported a slightly more positive experience on this item at the end of the sixth-grade year. Skewness is typical for ordinal values, particularly if the number of scale points is low (Boviard & Koziol, 2012; Bowen & Guo, 2012; O'Connell, 2010) and for self-report measures as participants may sometimes inflate or under-report their performance (Kuncel, Credé, & Thomas, 2005; Zimmerman, Caldwell, & Bernat, 2002). Thus, this characteristic of the data was anticipated.

Proportions of response values for each SCALE item reflected the positive nature of participants' middle school experiences in both the fall and the spring of the 6th grade year (see Table 6). Items 1, 2, 6, 19, 20 particularly low ratios of participants indicating that they had a *large problem* with those items and high instances of *no problem* in the fall and the spring of sixth grade. Item 12 had a high proportion of students indicating favorable experiences in the spring of sixth grade. Items 4, "having a tough teacher;" 5, "having to do harder school work;" 9, "getting too much homework;" and 11, "not seeing your friends from elementary school enough" had larger proportions of students indicating difficulty on those items in the fall of sixth grade. Ratios between item response categories remained relatively consistent or only shifted slightly to indicate more favorable experiences from the fall to the spring for most SCALE items,

with the exception of Items 4, “having a tough teacher”; 7, “having an argument with a teacher”; 9, “getting too much homework”; 15, “not getting along with all of your teachers”; and 18, “teachers expecting too much of you”. The distribution of responses for these items indicated fewer participants as having *no problem* with these transition experiences and a greater percentage of students indicating more difficulty between the fall and the spring of 6th grade.

The descriptive statistics for additional self-report and teacher-rated adjustment scales were also examined for both the fall and the spring of sixth grade. The range of values was within expected limits for all scales. For self-report scales, means, variances, standard deviations, and skewness and kurtosis values were similar between the fall and the spring (see Table 7). Mean values indicated that participants rated themselves relatively favorably in terms of a lack of defiance towards their teachers, feelings of school belonging, and favorable feelings of school valuing. Scales varied as expected and skewness and kurtosis values were low.

As displayed in Table 8, values for teacher-rated scales were relatively consistent between the fall and the spring. Means for ICS-T Popular, Aggressive, Affiliative, Academic subscales and the ICS-T scale (Cairns et al., 1995) as a whole were all relatively close to the midpoint. Mean values for participants’ academic effort were moderately high. Mean values also indicated that teachers viewed most students as likely leaders, not likely to be bullied or bully others, and that most students were liked. Variance, standard deviation, skewness, and kurtosis values for these scales were all within acceptable ranges.

**Bivariate statistics.** Bivariate polychoric correlations for SCALE variables were calculated for both the fall (see Table 9) and the spring (see Table 10). SCALE item correlations for the fall ranged from .09 to .73 and ranged from .17 to .69 in the spring. Correlations were examined for signs of multicollinearity, indicated by items that are too highly correlated with one

another (Tabachnick & Fidell, 2007). None of the correlations were indicative of multicollinearity as all correlations were well below .90. Correlations between SCALE items and other adjustment scales were also calculated for both the spring and the fall of 6th grade as displayed in Tables 11 and 12, respectively. Correlations for both the fall and the spring of 6th grade were relatively low for many scales, but showed low moderate to moderate correlations between some SCALE items and other measures of students' adjustment.

**Multivariate normality.** Examining multivariate outliers contribute to evidence that data are multivariate normal (Bowen & Guo, 2012; Tabachnick & Fidell, 2007). Fifty cases were determined to be multivariate outliers using Mahalanobis distance values ( $p < .001$ ) for the fall of sixth grade. Calculations for the spring indicated 55 participants were multivariate outliers ( $p < .001$ ). Analysis of the variable values for these cases showed that many cases had values of four, *large problem*, on multiple SCALE items. Other participants that were not considered outliers also had values of four for some SCALE variables. Consequently, cases noted as being multivariate outliers were potentially those that were having a more difficult time on some combination of variables on the SCALE.

In summary, analyses using Mahalanobis distance values indicated that data were not multivariate normal. Mahalanobis distance values may be overly sensitive when used with ordinal data (Finch, 2012), but the presence of multivariate non-normality would not necessarily be unexpected, considering the skewed nature of individual SCALE items (Tabachnick & Fidell, 2007). The use of the WLSMV estimator in MPlus used in the present study is robust to violations of univariate and multivariate normality (Muthén & Muthén, 2012). Therefore, cases determined to be multivariate outliers were not removed from the overall dataset.

**Sampling adequacy.** Jay Kerns' Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (MSA) was implemented to determine the suitability of the data for factor analytic techniques (Kerns, 2007). Factor analytic techniques require a level of common variance between items and Kern's KMO function returns a score from *miserable* to *marvelous* that describes whether a dataset has the amount of common variance appropriate for factor analysis. In addition, Kern's KMO function returns measures of sampling adequacy (MSAs) for each variable, with an ideal  $MSA \geq .7$  (Kerns, 2007; Tabachnick & Fidell, 2007). Results of analyses for both the fall and the spring of 6th grade indicated that the degree of common variance was *marvelous* for both waves (= .91 and .90, respectively). All SCALE items for both waves had high MSA values, with the lowest being .85, signifying that data were adequate for factor analysis.

**Summary.** Analysis of individual SCALE items appeared to indicate that participants perceived their transition experiences to be positive overall and their perceptions remained relatively stable between the fall and the spring of 6th grade with a few exceptions. Correlations indicated adequate relationships between SCALE variables and no evidence of multicollinearity, critical for factor analytic procedures (Tabachnick & Fidell, 2007). KMO analyses also supported the adequacy of the data for further analysis using factor analytic techniques. Analyses of univariate and multivariate statistics indicated that there were deviations from normality and multivariate normality, a common occurrence in ordinal data (Bovaird & Koziol, 2012), justifying the use of polychoric correlations and the WLSMV estimator in Mplus for factor analyses (Flora & Curran, 2004; Lei & Wu, 2012; Muthén & Muthén, 2012).



## Power

Power is the probability of rejecting a false null hypothesis (Lee, Cai, & MacCallum, 2012). There are a number of expert recommendations for required sample size for factor analytic techniques, including a rule of 20 participants per parameter to be estimated, or a minimum of 200 participants (Bowen & Guo, 2012; Byrne, 2012; Kline, 2011). Fewer than 100 participants are not recommended (Bowen & Guo, 2012). MacCallum, Browne, and Sugarwara (1996) developed a framework for power analysis in SEM that uses desired effect size, level of statistical significance, and the model's degrees of freedom ( $df$ ) to determine the appropriate sample size. For factor analyses using ordinal data, the formula  $p(p-1)/2 + r$  is used to determine the maximum number of parameters to be estimated where  $p$  is equal to the number of variables and  $r$  is equal to the total number of available thresholds (Muthén & Asparouhov, 2002). The total of available thresholds is calculated as the number of response options minus one, multiplied by the total number of variables. The maximum number of parameters is also multiplied by the number of groups during multiple group analyses. Degrees of freedom are calculated as the maximum available parameters to be estimated minus the number of estimated parameters (Bowen & Guo, 2012).

Using the formulas from Muthén & Asparouhov (2002) and assuming that all 20 variables will load onto three factors, MacCallum et al.'s (1996) framework recommends a minimum sample size of approximately 200 for the exploratory factor analysis to achieve adequate power of .80 with  $\alpha = .05$  ( $df = 84$ ). One-third of the sample will be used for the EFA ( $n = 253$ ). For confirmatory factor analyses approximately 160 participants would be needed for adequate power ( $df = 127$ ). Two groups of 254 participants were used for the calibration and validation CFAs. For the MGCFAs, each group would also have to meet the recommended

sample size of 160 participants for adequate power. The MGCFA across gender and socioeconomic status met these requirements as gender groups included 264 females and 214 males and socioeconomic status groups included 245 participants who were designated as full-pay and 234 participants who were in the free/reduced pay lunch program. The MGCFA across race/ethnicity included 265 White and 99 African American participants. Accordingly, the African American student group did not meet the sample size given the requirements by MacCallum et al., (1996) but analyses were completed knowing that results would have to be interpreted with caution. Composites for gender with race/ethnicity and gender with socioeconomic status also created groups that were too small for MGCFA to be completed and were excluded from further analyses.

MacCallum et al.'s (1996) recommendations are based on continuous data and some researchers assert that larger sample sizes are needed for factor analytic techniques with ordinal data (e.g., Byrne, 2012; Muthén & Muthén, 2012). However, Flora & Curran (2004) found that when conducting factor analytic techniques with ordinal variables using robust weighted-least-squares estimators (e.g., WLSMV), polychoric correlations, and a small number of factors, samples as low as 200 provided adequate power. As cited by Byrne (2012), Jöreskog and Sörbom (1996) recommended a sample size of  $(p+1)(p+2)/2$  where  $p$  is the number of observed variables. Using this framework, each model would require a sample size of 231 participants and most models in the present study would meet this requirement. Models for males for the MGCFA across gender and African American students for the MGCFA across race/ethnicity were smaller than recommended.

Studies by MacCallum, Widaman, Zhang, and Hong (1999) and Preacher & MacCallum (2002) found that adequate factor structures can be retained from smaller sample sizes if high

communality between items is present and a few number of factors are estimated. MacCallum et al. (1999) recommended a high communality score of at least .80 for an item to be retained. Costello and Osborne (2005) noted that .40 to .70 is more common in social sciences research. Given the small sample sizes for the male and African American groups in the study, communality was calculated for SCALE items. Calculations revealed that Items 3, 8, and 20 had communalities of .38, .37, and .37, respectively. All other items had communalities of  $\geq .4$  although Item 11 was just over this requirement with a communality coefficient of .44. Communalities for other items ranged from .51 to .68. Items 3, 8, 20, and 11 were then flagged as potentially problematic items prior to conducting analyses, with the intent that they may be the first items to be excluded if problems developed during analyses.

## **Factor Analyses**

**Exploratory factor analysis.** EFA ( $n = 253$ ) was completed using both Geomin and Varimax rotations, representing oblique and orthogonal rotations, respectively. Analyses began with all variables included in the model. Models with one to four factors were considered. With all SCALE items in the model, eigenvalues greater than one indicated four factors present in the data. Scree plot analyses suggested two, three, or possibly four factors (see Figure 1). As problematic variables began to be removed, eigenvalues suggested three factors.

Loadings of at least  $\pm .3$  indicated that items were considered part of the factor (Kline, 2011; Tabachnick & Fidell, 2007). Initial analyses including all SCALE items returned many double-loaded items, regardless of the type of rotation or the number of factors. Ideally, individual items only load on one factor but researchers can provide a theoretical justification for why an item may be allowed to load on multiple factors (Bowen & Guo, 2012; Byrne, 2012;

Kline, 2011). However, theoretical justifications could not be made to allow all of the double-loaded items in the present study.

Models following the removal of items marked for communality concerns (items 3, “being treated more like a child”; 8, “not having the right books or supplies for class”; and 20, “understanding new rules”) improved fit, but continued to include many double-loaded items. Therefore, other items were causing poor fit in the model. Subsequent models excluded items without loadings that could be theoretically justified and those that would not load on a factor, until the model with the best fit was found. Items 3, “being treated more like a child;” 7, “having an argument with a teacher;” 10, “getting into fights;” 15, “not getting along with all your different teachers;” and 18, “teachers expecting too much of you;” were dropped from the final model.

Models that included more than one factor indicated moderate correlations between factors. Accordingly, an oblique rotation that allowed factors to correlate was the most appropriate rotation. Three-factor models consistently showed the best fit. The final model included fifteen items and three factors,  $\chi^2(63) = 82.73, p = .05$ ; RMSEA = .04, 90% CI [.00, .06], CFI = .98, and TLI = .97 (See Table 13). The three factors in the final model were congruent with the literature regarding students’ transition experiences and supported the hypotheses in the present study. The first factor reflected procedural experiences; the second factor described academic experiences, and the third factor contained items regarding social experiences with peers.

**CFA-calibration.** The first confirmatory factor analysis was completed to calibrate the structure of the SCALE ( $n = 254$ ). Similar to the EFA, it was expected that items would load on one factor, but double loaded items were considered (Bowen & Guo, 2012; Byrne, 2012; Kline,

2011). Factor loadings and fit indices were analyzed to assess model fit. In addition, Mplus provided  $R^2$  estimates that described the proportion of variance in the underlying continuous response variable,  $y^*$ , explained by the latent variable (Muthén, 2004). Low  $R^2$  estimates may also indicate that an item may not be well represented by a factor. The residual correlation matrix produced in Mplus was also examined to identify problems with the CFA models (Bowen & Guo, 2012; Muthén & Muthén, 2012). The residual correlation matrix indicates the difference between the observed and estimated correlations between variables in the model. Absolute values of correlations should remain small and  $\leq .10$  (Bowen & Guo, 2012; Kline, 2011). Residual correlations greater than .10 indicate a potential area of poor model fit, but it is unclear how many large correlations signify a significant problem in the model as a whole but the presence of few residual correlations greater than .10 is ideal (Kline, 2011).

The first model tested was the model suggested by the EFA results with three correlated latent factors from fifteen items representing procedural, academic, and social experiences. Fit statistics of the model showed good fit ( $\chi^2 (87) = 108.67, p = .06$ ), RMSEA = .03; 90% CI [.00, .05]; CFI = .97; TLI = .97). All item loadings were acceptable, with the lowest loading = .48. Examination of the residual matrix indicated a larger residual correlation between Items 6 and 13 (= -.22). However, that the majority of residual correlations were  $\leq .10$ . No modification indices were recommended. Three additional alternative models were also explored. A one factor model including all fifteen variables returned poor fit values ( $\chi^2 (90) = 193.42, p < .000$ ; RMSEA = .07, 90% CI [.05, .08]; CFI = .86; TLI = .85). A two-factor model that contained the Academic factor and a factor that combined the highly correlated Procedural and Social latent factors returned acceptable but not exceptional fit values ( $\chi^2 (89) = 135.42, p = .001$ ; RMSEA = .05, 90% CI [.02, .06]; CFI = .94; TLI = .93). However, the fit was not as strong as the fit for the

three factor model. Due to the high correlations between latent factors, a second-order factor model was estimated. This model also returned exceptional fit values ( $\chi^2(87) = 108.67$ ,  $p = .001$ ; RMSEA = .03, 90% CI [.00, .04]; CFI = .97; TLI = .97). All loadings were significant and of reasonable magnitude, with the lowest loading being .48. However, later multiple group tests of this model could not be estimated and, as a result, the three-factor model without a second-order factor was determined to be the final model (see Figure 2).

**CFA-validation.** Finally, a second CFA was completed ( $n = 254$ ) to validate the final fifteen-item, three-factor model. All parameters were significant and loaded on the appropriate latent factors. Fit statistics indicated a good fit to the model ( $\chi^2(87) = 104.12$ ,  $p = .10$ ; RMSEA = .03, 90% CI [.00, .05]; CFI = .98; TLI = .98). Examination of the residual matrix found few residuals  $\geq .10$ . Only two residual correlations were  $\geq .20$ . These correlations were between Items 14 and 16 ( $= .22$ ) and Items 1 and 8 ( $= -.21$ ). However, as most of the residuals were  $\leq .10$ , the residual correlation matrix did not seem to indicate a large issue with the model (Kline, 2005). No modification indices were suggested. Parameters were similar to the calibration sample; although with some slight differences, such as the magnitude of the correlation between the Procedural and Academic factors (see Table 14). Currently, there are no guidelines as to how similar parameter estimates need to be between calibration and validation analyses to indicate problems in the model (Bowen & Guo, 2012). Bowen and Guo (2012) asserted that if the fit of the model and parameters are adequate, that this indicates consistency between the calibration and validation models. As all parameters loaded successfully onto the specified factors and good overall model fit was obtained for both models, the model was determined to be adequately validated.

## Multiple Group CFAs

The purpose of MGCFA is to determine scale invariance. Matrices represent each aspect of the factor pattern (Bowen & Guo, 2012; Hoyle, 1995; Kline, 2011). Greek symbols are used to denote the parameters and their matrices, with the exception of items which are represented as  $x$ . Factors are represented with the symbol,  $\xi$  (xi), and the loading between  $x$  and  $\xi$  is represented by the symbol,  $\lambda$  (lambda). Measurement errors are symbolized by  $\delta$  (delta). The first matrix is  $\Lambda_x$  (capital lambda-x) and contains all the loadings between factors and their items. Second, is  $\Phi$  (capital phi) which represents the covariance matrix of latent variables and each element within the matrix is represented by  $\phi$  (phi). Finally, the  $\Theta_\delta$  (capital theta-delta) contains measurement errors and each error is represented by  $\theta$  (theta).

The MGCFA procedure involves a series of steps to test whether or not elements of the factor pattern remain invariant across groups (Bowen & Guo, 2012; Byrne, 2008; Byrne, 2012). Using the trimming method (Kline, 2011), each step is hierarchical in nature and the comparisons become increasingly stringent as more parameters are constrained, or forced, to be equal instead of freely estimated. The steps included comparing the baseline models, testing for configural invariance, testing for metric invariance, and testing for scalar invariance. Models tested within each step are judged for fit using goodness of fit indices. Then, to test whether or not the scale remains invariant from step to step, a  $\chi^2$  difference test is performed using the DIFFTEST function in Mplus (Muthén & Muthén, 2012; Muthén & Muthén, 2013). If the  $\chi^2$  difference test is not significant, the model fit is not significantly worse and invariance is assumed for that step (Kline, 2011). Unstandardized parameters are compared during MGCFA as standardized estimates can be influenced by the variance of the individual groups, making meaningful comparisons across groups difficult (Byrne, 2012; Muthén & Muthén, 2012).

The first step of a MGCFA, baseline, involves testing the fit of the predicted model for all groups separately (Byrne, 2008; Kline, 2011; Millsap & Yun-Tein, 2004). It is acceptable at this step for models to be slightly different (Byrne, 2012), but acceptable differences must be small, such as correlated errors in different groups. During this step, findings that each model for each group has the same number of factors and the same items are loaded on the same factors are required.

The second step is the test for configural invariance ( $H_{\text{FORM}}$ ; Bowen & Guo, 2012; Byrne, 2012; Cheung & Rensvold, 2002; Kline, 2011). The hypothesis for this analysis is that all groups, separately, have the same number of factors and that the same items load on the same factors when the models are tested at the same time (Byrne, 2012). All parameters are allowed to be freely estimated (Bowen & Guo, 2012; Byrne, 2012; Kline, 2011). If configural invariance is found, the same constructs are present for all groups and this model is used as the baseline for the following tests (Byrne, 2012; Kline, 2011).

If configural variance is found, the scale failed the most basic test and is not adequate for use with the groups of interest (Byrne, 2012; Kline, 2011). Groups may not understand the constructs of the scale in the same way, or constructs may need to include different factors for each group (Kline, 2011). If configural variance is found, different items can be used as scale reference variables and the model can be retested. In SEM, the loading for one item per factor, known as the scale reference variable, is set to one (Bowen & Guo, 2012; Kline, 2011). However, if the item chosen as the scale reference item is the only item that is invariant across groups, it could cause the configural invariance for the entire scale to be inaccurately rejected (Kline, 2011). In the case that configural variance is found regardless of which items are used as scale reference variables, the scale is determined to be variant across groups.



The test for metric invariance assesses the invariance of the magnitude of loadings between items and factors across groups. This process contains two steps. The first step constrains all loadings at one time ( $H_{\Lambda x}$ ) and the second constrains each loading separately ( $H_{\lambda}$  for each loading,  $1 \dots Q$ ) (Bowen & Guo, 2012; Cheung & Rensvold, 2002; Kline, 2011). If the hypothesis that all loadings are invariant ( $H_{\Lambda x}$ ) is retained, metric invariance is present across groups, meaning that constructs are represented in the same way for each group (Kline, 2011). Tests of  $H_{\lambda}$  indicate whether or not all loadings are invariant or only loadings for certain factors or items are invariant (Bowen & Guo, 2012; Cheung & Rensvold, 2002; Kline, 2011). The results of  $H_{\lambda}$  tests serve two purposes: to explain why  $H_{\Lambda x}$  may have been rejected and also as a double-checking measure to assure construct invariance.

Following a positive finding of metric invariance, tests for scalar or *strong factorial invariance* involve constraining the thresholds (signified by  $\tau$ ) (Millsap & Yun-Tein, 2004; Sass, 2011). In models with continuous data, intercepts are constrained but thresholds are constrained in the present study because data are ordinal. During tests of scalar invariance, thresholds are constrained equal across groups and one of the factor means is constrained to zero while all other factor means are permitted to be freely estimated (Millsap & Yun-Tein, 2004; Muthén & Muthén, 2013). If scalar invariance is found, measurement thresholds are consistent across groups and measured mean differences across groups are mostly likely due to the influence of common factors (Millsap & Olivere-Aguilar, 2012; Sass, 2011). Therefore, participants in one group would be expected to produce the same means for observed variables, given the same score on a latent construct (Milfont & Fischer, 2010). Making unbiased comparisons of latent factor means across groups requires that observed variable thresholds are invariant across groups (Byrne, 2008).

Assessment of configural, metric, and scalar invariance provide rigorous tests of the equivalence of the SCALE across groups (Bowen & Guo, 2012; Muthén & Muthén, 2012). Even more demanding tests of measurement equivalence are also available, including the invariance of variance and covariances of latent factors ( $H_{\Lambda, \Phi}$ ), also known as *strict factorial invariance*, and the equivalence of residual variances ( $H_{\Lambda, \Theta\delta}$ ). Invariance of latent variances and covariances provides evidence that the range of factor scores remains consistent and that the latent variables relate similarly to each other across groups (Byrne, Shavelson, & Muthén, 1989; Milfont & Fischer, 2010). Tests for residual invariance assess whether the same amount of measurement error is present for all groups (Bowen & Guo, 2012; Byrne, 2012; Byrne et al., 1989). Researchers disagree as to whether or not these more demanding tests of measurement invariance are necessary to provide evidence that a measure is without bias across groups (Bowen & Guo, 2012; Millsap & Olivere-Aguilar, 2012), with many stating that they are not (e.g., Byrne, 2012; Byrne et al., 1989; Millsap & Yun-Tein, 2004). In the present study, these tests of measurement invariance were evaluated with the understanding that results indicating variance would not mean that the SCALE provides inaccurate measurement of latent constructs across groups. Rather, the results for strict factorial invariance and residual invariance would only provide additional information that would be helpful to the interpretation of the measurement properties of SCALE.

If some model parameters are variant and others are not, this is called partial invariance (Byrne et al., 1989; Kline, 2011). Partially invariant scales include both invariant parameters that are constrained and variant parameters that are freely estimated across groups. Under certain conditions, some researchers find partial invariance to be an acceptable characteristic of a scale as invariance for all parameters is often extremely difficult and somewhat unlikely (Milfont

& Fischer, 2010). Findings of partial invariance may be insignificant enough that the partially invariant items will not cause the entire scale to insert bias into group comparisons (Byrne, 2012; Sass, 2011). Researchers have presented some guidelines for acceptable conditions of partial invariance. Configural invariance is required (Byrne, 2012; Milfont & Fischer, 2010) but manifestations of partial invariance during other tests of invariance may be acceptable if the number of variant parameters are the minority of parameters available and researchers can present theoretical or empirical evidence which would support the acceptability of partial invariance in that case (Byrne, 1989; Milfont & Fischer, 2010; Vandenburg & Lance, 2000).

For the present study, MGCFA were completed to assess the invariance of the SCALE across gender, socioeconomic status, and race/ethnicity. The three-factor model of the SCALE was tested for baseline, configural, metric, and scalar invariance as well as strict factorial invariance and residual invariance. The presence of partial invariance was considered using guidelines set forth by Byrne et al. (1989), Milfont and Fischer (2010), and Vandenburg and Lance (2000).

**Gender.** Two hundred sixty four females and 218 males were included in the MGCFA analysis. The baseline model for the female group indicated acceptable fit ( $\chi^2(87) = 127.26, p = .003$ ; RMSEA = .04, 90% CI [.02, .06]; CFI = .96; TLI = .95). A modification indice was recommended that allowed Item 11, “not seeing your friends in elementary school enough” to load on the Academic factor, but this was rejected for theoretical purposes. The baseline model for males was also acceptable ( $\chi^2(87) = 108.94, p = .03$ ; RMSEA = .03, 90% CI [.03, .05]; CFI = .96; TLI = .95). No modification indices were recommended. The magnitude of some parameters was slightly different; particularly the  $R^2$  values (see Table 15). All parameters were significant.

Tests of configural, metric, and scalar invariance all maintained good fit. Chi-square difference tests indicated that differences between models were not significant (see Table 16). Assessment of invariance isolating individual loadings also reflected invariance across groups. Consequently, the measurement model of the SCALE was determined to be invariant across gender. As reported in Table 16, tests for factor variance and covariance equivalence did not indicate good fit and returned a significant chi-square difference test. Thus, the range of factor scores was not consistent and factors related to each other differently for males and females. Residual variances were also variant across groups, indicating that the variance not accounted for by factors in the model was not the same across groups.

**Free/reduced lunch status.** The MGCFA across socioeconomic groups included 245 participants who were designated as full-pay and 234 participants who were in the free/reduced lunch program. Baseline analyses indicated good model fit for both full pay ( $\chi^2(87) = 114.87, p = .03$ ; RMSEA = .04, 90% CI [.01, .05]; CFI = .97; TLI = .96) and free/reduced lunch status groups ( $\chi^2(87) = 117.58, p = .02$ ; RMSEA = .04, 90% CI [.01, .06]; CFI = .97; TLI = .96). All parameters were significant and similar across models (See Table 17). No modification indices were recommended for either group.

Fit values for configural, metric, and scalar models across groups were good (see Table 18), with the exception of the chi-square statistics but this could have been due to the larger sample size. Assessment of individual loadings did not indicate any problems. Chi-square difference tests between models did not indicate a significant difference between configural, metric, and scalar models and the measurement model of the SCALE was determined to be invariant across socioeconomic status groups. Tests for equivalence across factor variances and covariances indicated equivalence across groups, indicating the range of factor scores and

relationships between factors were consistent across groups. Tests for residual invariance were significant indicating that the variance not accounted for by the factors was not consistent across groups.

**Race/ethnicity.** Two hundred sixty-five White and 99 African American participants were included in the MGCFA. The baseline model for White students indicated good fit ( $\chi^2$  (87) = 107.34,  $p$  = .07; RMSEA = .03, 90% CI [.00, .05]; CFI = .98; TLI = .97). No modification indices were suggested for the model. The baseline model for African American students indicated good fit ( $\chi^2$  (87) = 105.89,  $p$  = .08; RMSEA = .05, 90% CI [.00, .08]; CFI = .95; TLI = .94), although the confidence interval for the RMSEA value was quite wide, potentially due to the low sample size. As reported in Table 19, all parameters were significant and  $R^2$  values were reasonable, but it was noted that the  $R^2$  for Item 20 in the African American model was lower than most other values. No modification indices were recommended.

The configural model including both White and African American groups returned good fit indices. However, a significant difference in chi-square was found between the configural and metric models ( $\Delta\chi^2$  (12) = 25.34,  $p$  = .01) indicating that the loadings may not be consistent across groups. Analysis of individual item loadings found that Items 6, “finding kids I can sit with at lunch” and 12, “having trouble making new friends” were variant across groups. Allowing the loadings for both of these items on the Social factor to be freely estimated provided a good fit to the model and a non-significant difference in chi-square (see Table 20).

Partial invariance recommendations suggest that if the majority of items within a factor are not found to vary across groups that partial invariance may be acceptable (Byrne et al., 1989; Milfont & Fischer, 2010; Vandenburg & Lance, 2000). Differences in students’ difficulties with peers between racial/ethnic groups have been identified in several studies (e.g., Berkel et al.,

2009; Graham et al., 2006; Wang et al., 2009). Therefore, metric invariance for these items seemed probable and partial metric invariance was determined to be acceptable in the model.

The test for scalar invariance with the partially invariant models found no significant differences in chi-square, indicating the presence of scalar invariance. Thus, the measurement model across White and African American groups was determined to be partially invariant but acceptable. Tests for factor variance and covariance equivalence did not find significantly worse fit, indicating the relationships between factors and range of factor scores were similar across groups. The variance not accounted for by the latent factors was found to be different across groups as indicated by a significant chi-square difference test after constraining residuals.

### **Reliability and Stability**

Reliability is an assessment of the consistency or predictability of scores from a scale (Streiner, 2003). There are multiple components to reliability, including both the internal consistency of the scale as well as how a scale performs over time. Cronbach's alpha coefficients were calculated for each present factor, using the convention of .7 as an indicator of good reliability. The temporal stability of a scale determines how reliable a scale is from one time to another and is more commonly known as test-retest reliability (Multon, 2010). Pearson product correlations and intraclass correlations (ICCs) between factor scores from different time points provide evidence that a SCALE consistently assesses the constructs of interest over time (DeVellis, 2012). Pearson product correlations and ICCs were calculated for factors using scores from both the fall and the spring of sixth grade. Two-way random effects ICCs for both *consistency*, which does not take into account rater variability, and *absolute agreement*, which does account for systematic differences in rater variability, were evaluated (McGraw & Wong,

1996; Nichols, 1998). For ICCs, significant coefficients and values of .6 indicated that the SCALE was stable over time (Shoukri et al., 2004; Shrout, 1998; Weir, 2005).

**Internal consistency.** The internal consistency refers to how well items within a scale relate to each other and the confidence that these items are measuring constructs without error (Cronbach, 1951; Streiner, 2003). Internal consistency coefficients for the SCALE for the fall of 6th grade were  $\alpha = .70$  for the Procedural subscale,  $\alpha = .72$  for the Academic subscale, and  $\alpha = .79$  for the Social subscale. Measures of internal consistency were similar for the spring of sixth grade with Cronbach's alphas of .67, .71, and .81 for the Procedural, Academic, and Social subscales, respectively. According to convention, the internal consistency of the Procedural subscale for the spring of 6th grade was questionable. The Academic and Social subscales demonstrated good internal consistency across both time points. Measurement invariance implies comparable internal consistency across groups (Millsap & Olivere-Aguilar, 2012). Due to the findings of measurement invariance for the SCALE across gender, race/ethnicity, or socioeconomic status, Cronbach's alphas were not calculated for each group separately.

**Stability.** Scale scores were calculated for each latent factor by summing item scores in accordance with the methods used for the SAT-MS (Elias et al., 1992). Pearson product correlations between waves were significant and moderate (Procedural,  $r = .65$ ; Academic,  $r = .75$ , Social,  $r = .74$ ). ICCs for consistency, similar to Pearson product correlations, assessed the similarity between scores between the fall and the spring of sixth grade. Consistency results were  $ICC(646, 646) = .71, p < .001$  for the Procedural subscale,  $ICC(636, 636) = .65, p < .001$  for the Academic subscale, and  $ICC(615, 615) = .73, p < .001$  for the Social subscale. Results for absolute agreement, which accounts for changes in mean differences overtime, were almost identical;  $ICC(646, 646) = .71, p < .001$  for the Procedural subscale,  $ICC(636, 636) = .64, p <$

.001 for the Academic subscale, and  $ICC(615, 615) = .73, p < .001$  for the Social subscale. All were all greater than .6, used by some as a minimal standard of adequate stability for ICCs (Shoukri et al., 2004). ANOVAs for each subscale found significant changes in mean differences within individuals on each subscale,  $F(1) = 9.92, p = .002$  for the Procedural subscale,  $F(1) = 24.68, p < .001$  for the Academic subscale,  $F(1) = 8.47, p = .004$  for the Social subscale. However, differences in means across individuals did not seem to impact the stability of the SCALE over time.

### **Construct Validity**

Concurrent and predictive convergent and discriminant validity were assessed for the SCALE. Concurrent construct validity is the strength of the association between the measure of interest and other measures that assess the same or similar constructs and is determined by systemically relating scale scores to concurrent measures of students' adjustment (AERA, APA, & NCME, 1999; DeVellis, 2012). Predictive validity reflects how well a scale relates to future measures of the same or similar constructs as the measure of interest. Positive relationships between the scale of interest and either concurrent or future measures demonstrate convergent validity and negative or a lack of relationship between scales demonstrates divergent, or discriminant, validity (DeVellis, 2012; Downing, 2003; O'Leary-Kelly & Vorkurka, 1998). Correlations between the Academic, Procedural, and Social subscales and concurrent and future measures of students' adjustment were examined for evidence of convergent and divergent construct validity. Additional analyses examined concurrent and predictive relationships across gender, socioeconomic status, and race/ethnicity. Concurrent and predictive relationships were calculated using adjustment scales from the fall and spring of 6th grade.



Consistent with prior validity studies (e.g., Elias et al., 1992; Vandenburg & Lance, 2000) latent mean differences across gender, socioeconomic status, and racial/ethnic groups were also analyzed to provide evidence of construct validity. Differences in latent factor means indicate whether groups are divergent on the level of the underlying construct of the factor (Byrne, 2012). Finding equivalence across latent factor means may indicate that the underlying construct of the factor is operating in the same way across groups. Differences in latent factor means across groups that are similar to those found in past studies also provide evidence of the construct validity of a measurement (e.g., Elias et al., 1992).

**Concurrent and predictive relationships.** Overall, correlations between the Academic, Procedural, and Social subscales of the SCALE exhibited low moderate to moderate correlations with other measures of students' adjustment including measures of students' social, behavioral, and academic adjustment, social and behavioral characteristics, school behaviors, academic behaviors, and feelings of school belonging (see Table 21). This finding was similar to the strength of relationships found between individual items and measures of students' adjustment. Many of the correlations were negative due to higher scores on the SCALE representing increased difficulty and higher scores on many of the other adjustment scales indicating favorable adjustment.

**Construct validity of the Academic subscale.** It was predicted that the Academic subscale would demonstrate construct validity through positive correlations with measures of students' academic adjustment, engagement, and school belonging. Lower difficulty on the Academic subscale was significantly and positively associated with fall measures of students' school valuing, math and reading achievement, academic effort, school belonging, and teacher-rated academic and overall adjustment; demonstrating convergent validity. Concurrent

correlations were the strongest between the Academic subscale and feelings of school valuing and school belonging. Lower difficulty on the Academic subscale was also associated with measures of positive behavioral and social adjustment. A moderate, significant correlation was present between the Academic subscale and teacher-rated affiliative behavior. Significant, low moderate correlations were also found between the subscale and being liked and being bullied by peers. Low moderate, negative associations were also significant between the Academic subscale and bullying and defiant behaviors. A small correlation was present between the subscale and students' popularity, but this relationship was not significant.

Predictive correlations between the Academic subscale and measures of students' adjustment in the spring of 6th grade indicated that students who had favorable experiences with academic demands were also more likely to be academically, behaviorally, and socially adjusted in the spring. Convergent relationships between the subscale and measures of students' school valuing, math and reading achievement, school belonging academic effort, and academic and overall adjustment in the spring of 6th grade were significant and moderate. Divergent relationships between the Academic subscale and teacher-rated aggression, bullying, and being bullied and convergent relationships between the subscale and being a leader, being liked, affiliative behaviors, were low moderate and significant. No relationship was present between the Academic subscale and students' popularity.

Subsequently, evidence supported the concurrent and predictive validity of the Academic subscale through significant convergent relationships to fall and spring measures of students' academic adjustment, engagement, and school belonging. The Academic subscale was also correlated with almost all of the measures of students' behavioral and social adjustment during both the fall and the spring of 6th grade, with the exception of students' popularity. This was

anticipated as past studies have demonstrated relationships between academic adjustment and positive social and behavioral adjustment (e.g., Niehaus et al., 2012; Way et al., 2007; Wentzel & Caldwell, 1997). Thus, relationships between the Academic subscale and measures of students' social and behavioral adjustment were consistent with the findings of past studies, providing additional support for the construct validity of the Academic subscale.

***Construct validity of the Procedural subscale.*** No specific hypotheses were made regarding the relationship of the Procedural subscale to adjustment measures, but it was anticipated that the subscale may correlate with measures of students' academic and behavioral adjustment. During the fall of sixth grade, the Procedural subscale was significantly correlated with almost all measures of students' academic, behavioral, and social adjustment. The strongest concurrent relationship between the Procedural subscale and measures of students' adjustment was a significant correlation between the Procedural subscale and school belonging. The subscale was also moderately correlated with concurrent measures of students' math achievement, overall adjustment, being a leader, and being liked. Low moderate correlations were present between the subscale and students' reading achievement, academic effort, teacher-rated academic adjustment, and popularity in the fall. Therefore, students who had positive experiences with procedural demands were more likely to display favorable adjustment outcomes, such as higher achievement, belonging, academic effort, being liked, popular, and a leader. Low moderate, divergent correlations were present between the Procedural subscale and measures of students' aggression, defiance, bullying behaviors, and being bullied by peers in the fall of sixth grade. No significant relationship was present between the subscale and students' affiliative behaviors. Finally, a low moderate correlation was present between the Procedural subscale and school valuing, but the correlation was not statistically significant.

Predictive relationships between the Procedural subscale and measures of students' adjustment in the spring of sixth grade demonstrated that students who had positive procedural experiences were associated with favorable adjustment outcomes in the spring. Convergent relationships were significant and moderate between the subscale and students' math achievement, reading achievement, academic adjustment, overall adjustment, and belonging. Significant, low moderate correlations were also found between the subscale and school valuing, academic effort, and being seen as a leader. Significant, low moderate correlations were present between the subscale and measures of defiance and aggression; demonstrating divergent validity. Unlike the relationship with concurrent measures of students' adjustment, the Procedural subscale was not significantly related to students' bullying behaviors, popularity, being liked, or being bullied by others. There was also no correlation present between the subscale and students' affiliative behaviors.

Students who perceived favorable procedural experiences were more like to experience favorable adjustment outcomes in both the fall and the spring, lending evidence to the construct validity of the Procedural subscale. Relationships between the Procedural subscale and measures of students' academic, behavioral, and social adjustment were somewhat dissimilar from the fall to the spring as many of the spring social and behavioral measures were no longer correlated with the subscale. However, without prior knowledge regarding the anticipated convergent and divergent relationships between the Procedural subscale and other adjustment measures, it is unknown whether the lack of relationship with future measures of students' adjustment is indicative of stronger or weaker predictive validity for this subscale.

***Construct validity of the Social subscale.*** It was anticipated that the Social subscale would be negatively associated with antisocial behaviors, such as bullying or aggression, and

positively associated with measures of students' school belonging and other social characteristics, such as being liked or popular. The Social subscale demonstrated a significant, strong, convergent relationship with school belonging during both the fall and the spring. Moderate significant correlations were present between the subscale and fall and spring measures of being a leader and being liked. Significant, moderate concurrent and predictive relationships were present between the Social subscale and not being bullied by peers. The subscale also had low moderate, convergent relationships with fall and spring teacher-rated overall adjustment. No relationships were found between the Social subscale and fall and spring measures of students' affiliative behavior, academic effort, aggressive behavior, or bullying behaviors.

Correlations that were time-specific included a significant, low moderate correlation between the subscale and students' popularity in the fall. During the spring, the correlation between the subscale and students' popularity was strong and significant. Small correlations were present between the Social subscale and measures of students' school valuing, reading achievement, and academic adjustment in the fall, but correlations were not significant. Low moderate correlations were also significant between the Procedural subscale and spring measures of students' math achievement and academic adjustment. There were low moderate correlations between the subscale and spring measures of students' school valuing and reading achievement but correlations were not significant.

As predicted, the Social subscale demonstrated convergent validity through relationships with school belonging and positive social characteristics, such as being liked or popular. Students who adjusted to the demands of the Social subscale easily were more likely to experience favorable outcomes on these measures during both the fall and the spring of 6th grade. The Social subscale also demonstrated discriminant validity as few significant

relationships were present between the subscale and measures of students' academic adjustment. Furthermore, students who had less favorable experiences on the Social subscale were more likely to exhibit negative behavioral and social characteristics during both the fall and the spring of 6th grade. Therefore, the Social subscale was determined to demonstrate both concurrent and predictive construct validity.

***Concurrent and predictive relationships across demographic groups.*** Correlations between the SCALE and adjustment measures from both the fall and spring of 6th grade were compared across gender, socioeconomic status, and racial/ethnic groups. During analyses, patterns of significant concurrent and predictive correlations were compared across groups. The strength of correlations between SCALE subscales and some adjustment measures varied across groups. Large, unexpected differences in the magnitude of correlations across groups are discussed.

***Correlations by gender.*** During their study of students' transition perceptions, Elias et al. (1992) found stronger relationships for boys between measures of students' adjustment and subscales of the SAT-MS. Relationships between subscales and adjustment measures were also more varied for boys and, for girls, adjustment measures were more correlated with experiences with peer relationships. Table 22 displays the concurrent and predictive relationships between the SCALE and measures of students' adjustment across gender in the present study. Unlike Elias et al. (1992), the magnitude of significant correlations ranged from low moderate to strong for both boys and girls. Additionally, the Social subscale, composed of items concerning students' relationships with peers, was significantly related to more measures of students' adjustment for boys in the fall of sixth grade. In the spring of 6th grade, however, a number of adjustment measures were significantly related to the Social subscale for both boys and girls.

Significant, convergent relationships between the Academic subscale and concurrent measures of students' school valuing, math achievement, academic effort, and belonging were present for both groups. Fall reading achievement and teacher-rated academic adjustment and overall adjustment were significantly correlated to the Academic subscale for girls but not for boys. Divergent relationships between the Academic subscale and defiance and aggression were significant for both boys and girls in the fall. The Academic subscale was also significantly correlated with other measures of girls' adjustment, including a convergent relationships with being a leader, and divergent relationships with being not liked and bullied. Affiliative behaviors were significantly correlated with the Academic subscale for boys in the fall.

Predictive convergent relationships continued to be significant between the Academic subscale and measures of students' school valuing, academic effort, and belonging for both boys and girls. The relationship between the Academic subscale and math achievement was no longer significant for boys in the spring, but continued to be significant for girls. Reading achievement and being a leader in the spring was also significant for both boys and girls. A convergent, predictive relationship between affiliative behaviors and the Academic subscale was significant for boys. The Academic subscale displayed significant divergent relationships with defiance for both boys and girls in the spring. Significant, divergent relationships were also present between the Academic subscale and spring measures of girls' aggressive behaviors and being bullied. Bullying peers and not being liked were significantly and negatively correlated with the Academic subscale for boys in the spring.

Concurrent, convergent relationships between the Procedural subscale and math achievement, reading achievement, teacher-rated academic and overall adjustment, belonging, and being leader were significant for both boys and girls in the fall. A significant relationship

between school valuing and the Procedural subscale was also present for both groups, but the correlation was small for boys. Significant divergent relationships were also present between the subscale and aggressive behaviors and not being liked for both groups. The Procedural subscale also demonstrated significant, convergent relationships with academic effort and affiliative behaviors for girls and popularity for boys. The Procedural subscale was also significantly and divergently related to defiance, bullying behaviors, and being bullied for girls.

Convergent relationships between reading achievement, overall adjustment, and belonging continued to be significant for both groups in the spring. Relationships between the subscale and school valuing, academic effort, affiliative behaviors, and being a leader continued to be significant for girls and relationships with math achievement and popularity were significant for boys in the spring. A divergent, predictive relationship between the Procedural subscale and aggression continued to be significant for both groups. A significant, predictive relationship between defiance and the Procedural subscale was present for boys. For girls, there was also a significant, predictive correlation between the Procedural subscale and not being liked.

Significant, convergent concurrent and predictive correlations were present between feelings of belonging and the Social subscale for both boys and girls. However, significant concurrent relationships between the Social subscale and other measures of students' adjustment were lacking for girls. For boys, significant concurrent correlations were present with math and reading achievement, overall adjustment, being a leader, affiliative behaviors, being popular and divergent relationships included correlations with being bullied and not being liked.

A larger number of predictive relationships were significant for girls. The Social subscale was significantly related to spring measures of girls' overall adjustment, being a leader,



popularity, and not being bullied. Significant relationships were present between the Social subscale and spring measures of boys' math and reading achievement, overall adjustment, popularity, being liked and not being bullied.

Consequently, SCALE subscales demonstrated convergent and divergent relationships with both concurrent and predictive measures of students' adjustment. Although there were few concurrent correlations between the Social subscale and measures of students' adjustment for girls, the subscale did have a strong, significant correlation with girls' school belonging. The types of adjustment measures that were significantly correlated with the Academic and Procedural subscales were not always exactly the same across gender groups, but differences were did not indicate problems with the construct validity for either group.

*Correlations by socioeconomic status.* Results of analyses are displayed in Table 23. School valuing, math achievement, academic effort, overall adjustment, affiliative behaviors, and belonging were all significantly correlated with the Academic subscale for both full-pay and free/reduced lunch status students in the fall; demonstrating convergent validity. A divergent relationship between the Academic subscale and aggression was also present for both groups. For full-pay students, the Academic subscale was also significantly correlated with being a leader and a lack of defiance. The Academic subscale was significantly correlated with fall measures of students' reading achievement and academic adjustment for free-reduced lunch status students.

Predictive relationships between the Academic subscale and spring measures of students' math and reading achievement, academic effort, belonging, affiliative behaviors, and being a leader were significant for both groups. For full-pay students, the Academic subscale was also significantly correlated with overall adjustment, being liked, and not being bullied in the spring;

although, the correlation between the subscale and not being bullied was small. Predictive relationships were also significant between the subscale and school valuing, academic adjustment, and a lack of defiant behavior for free/reduced lunch status students.

For both groups, the Procedural subscale demonstrated concurrent, convergent relationships with math achievement, being a leader, popularity, and overall adjustment; but correlations were small for the free/reduced lunch status group, with the exception of the correlation with overall adjustment. In addition, the Procedural subscale was significantly correlated with bullying behaviors for both groups but the correlations were opposite in sign. Therefore, full-pay students who experienced increased difficulty with the procedural demands of middle school were less likely to bully others, but free/reduced lunch status students who had increased difficulties with procedural demands were more likely to bully others. For full-pay students, the Procedural subscale also demonstrated convergent validity with significant relationships to measures of school valuing, academic effort, belonging, and affiliative behaviors in the fall.

Correlations between the Procedural subscale and spring measures of students' reading achievement, belonging, and defiant behaviors were significant for both groups. Correlations were also stronger for the free/reduced lunch status students in the spring. The inverse relationship between bullying and the Procedural subscale was no longer significant and negligible for both groups. Predictive, convergent relationships between the Procedural subscale and measures of students' math achievement, academic effort, overall adjustment, popularity and being a leader were all significant for full-pay students. Predictive, divergent relationships were significant between the Procedural subscale and aggression, not being liked, and being bullied for full-pay students.

For both groups, significant convergent relationships were present between the Social subscale and fall and spring measures of students' school belonging. Significant concurrent, divergent relationships were also present for both groups; for the full-pay group, the Social subscale was negatively correlated with aggression and for the free/reduced lunch group the subscale was correlated with not being bullied. Significant concurrent relationships were also present between the Social subscale and measures of students' reading achievement, academic effort, overall adjustment, affiliative behaviors, being a leader, popular, and being liked for full-pay students. A significant predictive relationship was present for the free/reduced lunch group between the Social subscale and school valuing. Significant predictive relationships for the full-pay group included correlations with math achievement, overall adjustment, belonging, affiliative behaviors, being a leader, liked, popular, and not bullied.

A smaller number of significant concurrent and predictive relationships were found between measures of students' adjustment and students' experiences with procedural and social demands for the free/reduced lunch status students. Those relationships that were significant for the free/reduced lunch status groups were indicative of construct validity, with the exception of the correlation between the Procedural subscale and fall measures of bullying behaviors. However, due to the lack of studies on students' procedural adjustment, it is difficult to determine if this correlation is indicative of a violation of construct validity.

*Correlations by race/ethnicity.* Overall, fewer significant concurrent and predictive relationships were present between the Academic subscale and measures of students' adjustment for the African American student group (see Table 24). For both groups, the Academic subscale demonstrated convergent validity with fall measures of school valuing, math achievement, affiliative behaviors, and being a leader. Divergent relationships were also significant for both

groups between the subscale and fall measures of defiance. Additional significant correlations were present for White students between the Academic subscale and fall measures of reading achievement, academic effort, academic adjustment, overall adjustment, belonging, popularity, and being liked. Predictive relationships between the Academic subscale and spring measures of school valuing and reading achievement were significant for both groups. Additional significant predictive relationships were present for White students including correlations between the Academic subscale and measures of students' math achievement, academic effort, overall adjustment, belonging, affiliative behaviors, being a leader, being liked, not being bullied, and less aggression.

Significant concurrent relationships between adjustment measures and the Procedural subscale were present for both groups. Convergent relationships were present between procedural experiences and academic effort, academic adjustment, overall adjustment, being a leader, and being liked for both groups. For both groups, a significant, divergent relationship between procedural experiences and aggression was also present. For White students, the Procedural subscale was significantly correlated with school valuing and school belonging. A significant divergent relationship with between the Procedural subscale and fall measures of students' defiance was also significant for White students. For African American students, significant convergent relationships were present between the Procedural subscale and math achievement and affiliative behaviors and a significant divergent relationship was present with aggression.

A smaller number of significant predictive relationships between the Procedural subscale and spring adjustment measures were present for the African American group. Procedural experiences were significantly related to spring academic adjustment, overall adjustment, and

affiliative behaviors for African American and White students. Procedural experiences were also significantly correlated with spring measures of school valuing, math achievement, academic effort, belonging, being a leader, being liked, not being bullied, and not being aggressive for White students.

A smaller number of significant concurrent and predictive relationships were also found between adjustment measures and the Social subscale for the African American group. Significant, convergent relationships were found between concurrent measures of students' academic effort and belonging for both groups. A significant, divergent relationship was also found between the Social subscale and fall measures of defiance for both groups. A significant, concurrent relationship between social experiences and aggression was found for both groups, but the direction of the correlation was different between groups. For White students, increased difficulty with social experiences was related to higher teacher-rated aggression. For African American students, increased difficulty with social experiences was related to lower-teacher related aggression. For White students, the Social subscale was also significantly correlated with school valuing, reading achievement, academic adjustment, overall adjustment, affiliative behaviors, being a leader, being liked, being popular, and not being bullied.

The predictive relationship between the Social subscale and students' popularity was significant for both groups. In addition, for African American students, a significant, convergent relationship was found between positive social experiences and school valuing in the spring. In addition, the relationship between African American students' social experiences and teacher-rated aggression was significant and continued to demonstrate that African American students who had more difficulty with social experiences were rated as less aggressive by teachers. For White students, convergent, predictive relationships were significant between students' social

experiences and students' math and reading achievement, academic effort, academic adjustment, overall adjustment, belonging, affiliative behaviors, being liked, being a leader, and not being bullied.

Subsequently, for African American students, a smaller number of concurrent relationships were found between the Academic and Social subscales and measures of students' adjustment. Also, a smaller number of predictive relationships were found between measures of students' adjustment and all SCALE subscales. However, significant concurrent and predictive relationships between SCALE subscales and measures of African American students' adjustment provided evidence for the construct validity of the subscales, including the relationship between teacher-rated aggression and students' social experiences. In a recent study by Xie et al (2013), African American boys perceived a positive relationship between aggression and popularity. A study by Rodkin, Farmer, Pearl, and Van Acker (2000) also found that boys who were popular and displayed more antisocial behaviors were disproportionately more likely to be African American than White. Consequently, it is possible that African American students who were perceived as more aggressive by their teachers were experiencing social success in other areas and African American students perceived as less aggressive could have been experiencing other social difficulties. However, the African American sample in the present study was also much smaller than the White sample. Therefore, results should be interpreted with caution.

**Latent mean differences.** Mean differences were calculated across gender, socioeconomic status, and racial/ethnic groups. No significant differences were found between gender groups for the Procedural and Social subscales,  $p = .72$  and  $.87$  respectively. A significant difference was found for the Academic subscale, with males having a mean difference of  $.10$  ( $p = .006$ ), which indicated less favorable academic experiences for boys than girls (see

Figure 3). These findings are consistent with a study by Deemer et al. (2003) that found that males perceived more difficulty with academic tasks than females. As a result, mean differences suggested construct validity of the SCALE. No studies were found that explicitly examined mean differences across gender for procedural adjustment, so this finding was not indicative of a violation of construct validity.

No significant latent mean differences were found for any of the SCALE subscales across full-pay and free/reduced lunch status groups (Procedural,  $p = .19$ ; Academic,  $p = .32$ ; Social,  $p = .42$ ). Findings from past studies of students' social adjustment across socioeconomic groups have been inconsistent. In one study, economically disadvantaged students were more likely to engage in disruptive behaviors (Theriot & Dupper, 2009) but in other studies, results suggest that all students may experience social difficulties, such bullying, regardless of their socioeconomic status (Wang et al., 2009). Furthermore, no studies were identified that examined differences in perceptions of the difficulty of academic demands between different socioeconomic groups of students at the individual level. As a result, it was anticipated that differences in students' perceptions of their experiences with the academic, procedural, and social demands of middle school may not emerge. Non-significant differences in students' perceptions of their transition experiences were not considered to be indicative of a lack of construct validity of the SCALE.

Finally, latent means were compared across White and African American student groups. No significant differences were found across groups for any of the subscales (Procedural,  $p = .80$ ; Academic,  $p = .14$ ; Social  $p = .95$ ). Mean differences across race/ethnicity groups were anticipated for the Social subscale as researchers have found that African American students are more likely to experience bullying and discrimination from peers (e.g., Berkel et al., 2009; Graham et al., 2006; Wang et al., 2009). However, items from the SCALE assessed more

general social experiences with peers and may not have addressed difficulties with peer interactions that may be dissimilar across peers from different racial/ethnic groups. Thus, there did not seem to be any evidence that suggested a lack of construct validity for the Social subscale.



## Chapter 6: Discussion

Transitions are pervasive in life, embedded within each individual's development (Elder et al., 2003; Elder & Shanahan, 2006). The middle school transition is a particularly prominent educational period due the continued presence of declines in students' academic, behavioral, and social adjustment (e.g., Akos et al., 2014; Eccles et al., 1993; Niehaus et al., 2012, Ryan et al., 2013). During the middle school transition, students are confronted with a number of academic, procedural, and social expectations and demands. Students' responses to these contextual demands are important their adjustment following the transition (Elder, 1998; Elder et al., 2006). However, most available measures of middle school students' transition experiences are either significantly dated (i.e., Elias et al., 1992) or have not been subjected to rigorous validation (i.e., Akos, 2002; Akos & Galassi, 2004b; Day et al., 2014).

In the present study, a new measure of students' transition experiences, the SCALE, was validated. Following guidelines from the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999) and other validation studies, theoretical and statistical evidence was gathered in the present study to establish the construct validity of the SCALE. As discussed in the following sections, it was determined that theoretical and statistical properties of the SCALE supported the validity of interpretations drawn from the subscales regarding students' transition experiences. As a result, the SCALE is a suitable measure of diverse students' experiences with the academic, procedural, and social demands of middle school. Recommendations and limitations of the present study and the SCALE are discussed.

## **Critical Transition Demands Captured by the SCALE**

According to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), interpretations from validated scales should be supported by theory and empirical research. In accordance with the principles of life course theory, interpretations from the SCALE should represent students' responses to the contextual demands common to the middle school transition (Elder & Shanahan, 2006). Results of extensive statistical testing indicated three subscales in the SCALE representing students' experiences with academic, procedural, and social demands. These domains reflect past studies of students' middle school transition experiences which identified these areas as the most common demands experienced by students following the middle school transition (Akos, 2002; Akos et al., 2004b; Juvonen et al., 2004). Life course theory asserts that, during transitions, students encounter new behavioral and social demands and expectations that influence how they experience the context (Elder & Shanahan, 2006). Consistent with life course theory, the SCALE was able to identify the middle school demands commonly experienced by students following the transition.

**Variability in responding to middle school demands.** Life course theorists emphasize that students may perceive and respond to contextual demands differently (Elder, 1998). Consequently, it was expected that students' experiences with the demands of middle school would vary. In the present study, most students' perceptions of their experiences with the academic, procedural, and social demands were positive overall, but there was variability in individuals' experiences with demands. Results of the present study were consistent with the principles of life course theory as students' experiences with the demands of middle school varied (Elder, 1998; Elder & Shanahan, 2006). Furthermore, the results replicated findings from previous studies in which the majority of students did not perceive their experiences with the

transition to be negative or very difficult, but some students' experiences were more positive than others throughout the sixth-grade year (Akos & Galassi, 2004a, 2004b; Day et al., 2014).

In addition to differences in individuals' experiences, life course theorists assert that groups may also experience disparate opportunities within a context related to their gender, socioeconomic status, or race/ethnicity (Pallas, 2003). These disparate opportunities may contribute to group differences in students' responses to contextual demands and, as a result, their experiences of the middle school transition. In the present study, boys were more likely to have less favorable experiences with the academic demands of the middle school transition. Accordingly, males may have been subjected to differential academic experiences or opportunities following the transition which contributed to heightened perceptions of difficulty in meeting the academic demands of middle school. This is consistent with past research findings, in which boys were more likely than girls to report that they had trouble managing the increasing quantities of more challenging school work in middle school (Deemer et al., 2003). Research findings indicate that students excel academically in middle school when they are both challenged and supported by their teachers (Klem & Connell, 2004). A study by Deemer et al. (2003) recounted students' perceptions that meeting the academic demands of middle school were easiest when they received support from teachers and were on good terms behaviorally. In previous studies, boys were less likely than girls to feel supported by their teachers (Way et al., 2007; Wentzel et al., 2010). It is possible that males in the present study experienced less support from teachers than females, which contributed to increased difficulties in meeting the academic demands of middle school. Boys may also struggle more than girls with self-regulatory behaviors required in middle school which contribute to increased behavior problems in the classroom, making academic demands more challenging for boys to achieve (Lane et al.,

2006; McMullen et al., 2007; Wray-Lake et al., 2010). Future studies should assess aspects of middle school that may limit boys' ability to meet the academic demands of the transition.

In accordance with empirical research, the Social subscale represented two aspects of students' social experiences in middle school: the establishment of peer relationships and managing interactions with problematic peers (Akos & Galassi, 2004b; Cook et al., 2008; Davidson et al., 2010; Hamm et al., 2014). No significant differences were found in students' experiences with the social demands of middle school across gender, socioeconomic status, or race/ethnicity. However, none were anticipated across gender as both males and females have demonstrated social difficulties in past studies (Niehaus et al., 2012; Wentzel et al., 2010). Moreover, no differences were anticipated between economically advantaged and disadvantaged students because the results of past studies were inconsistent as to whether students experience problems with social relationships in relation to their socioeconomic status (Malecki & Demaray, 2006; Theriot & Dupper, 2009; Wang et al., 2009; Way et al., 2007).

Students' experiences with the procedural demands of middle school did not differ significantly across gender, socioeconomic status, and racial/ethnic groups. However, no group differences in students' experiences with procedural demands were anticipated because no previous studies were found that had assessed students' procedural experiences across any demographic group. Therefore, students had varying experiences with the social and procedural demands of middle school, but groups of students did not seem to experience differential treatment in relation to social or procedural demands (Pallas, 2003).

It was anticipated, however, that African American and economically disadvantaged would report more difficulty with the academic demands of middle school. In previous studies, economically disadvantaged and African American students were more likely experience

differential academic opportunities in school (Becker & Luthar, 2002; Farkas, 2003). For example, teachers may grade African American students more stringently (Farkas, 2003) and teachers may have lower academic expectations for African American or economically disadvantaged students (Becker & Luthar, 2002). African American and economically disadvantaged students also exhibited less favorable academic adjustment in comparison to their economically advantaged and White peers (Akos et al., 2014; Becker & Luthar, 2002). Yet, in the present study, students' perceptions of academic demands did not differ significantly by race/ethnicity or socioeconomic status. Given the results of previous studies, differential opportunities in middle school may contribute to academic performance disparities across racial/ethnic and socioeconomic status groups, but not necessarily to varying subjective experiences of the academic demands of middle school between groups.

It was also anticipated that African American students would perceive the social demands of middle school to be more challenging than their White peers, but this was not observed in the present study. In previous studies, African American students were more likely to experience negative social interactions, such as bullying and discrimination, which impacted their adjustment outcomes following the transition (Berkel, et al., 2009; Graham et al., 2006; Wang et al., 2009). However, social experiences specifically related to students' race or ethnicity, such as discrimination, were not specifically represented by items in the SCALE. Therefore, African American and White students may not have experienced differential opportunities during the transition that impacted their perceptions of their ability to meet the social demands assessed by the SCALE (Pallas, 2003), including establishing peer relationships and managing general problems with peers. Nonetheless, this does not imply that African American students do not experience other social challenges not assessed by the SCALE, specifically bullying or

discrimination related to students' race or ethnicity, which may impact their transition adjustment in other ways (Graham et al., 2006; Wang et al., 2009).

In summary, few group differences were evident in students' experiences of the academic, procedural, and social demands of the middle school transition. This may suggest that students from different racial/ethnic or socioeconomic groups did not experience pronounced differential opportunities that impacted their ability to meet the academic, procedural, and social demands of the middle school following the transition. The exception to this inference is that, consistent with the results of past studies (Klem & Connell, 2004; Lane et al., 2006), boys may have experienced opportunities within the middle school context that adversely affected the way they responded to the academic demands of middle school (Elder et al., 2006; Pallas, 2003). Given the results of the present study, in addition to assessing students' experiences with the critical demands of the middle school transition, the SCALE was useful in identifying groups' responses to transition experiences.

**Students' transition experiences and adjustment.** According to life course theory, students' experiences with contextual demands and expectations influence their adjustment following the transition (Elder & Shanahan, 2006). In order for the interpretations of SCALE subscales to be consistent with life course theory, experiences with the academic, procedural, and social demands of middle school would be related to positive and negative indicators of their adjustment (Elder, 1998; Elder & Shanahan, 2006). Results of the present study were consistent with life course theory (Elder, 1998; Elder & Shanahan, 2006) as students' success or difficulty on each SCALE subscale was significantly related to specific indicators of their adjustment at both the beginning and the end of the sixth-grade year. Students who adjusted favorably to the social, procedural, and academic demands of middle school were more likely to experience

positive academic, behavioral, and social adjustment outcomes at the beginning and end of the transition year regardless of their gender, socioeconomic status, or race/ethnicity. Those students with greater difficulty with the demands of middle school were also more likely to exhibit negative adjustment outcomes in all demographic groups analyzed in the present study.

In keeping with the findings of past studies (e.g., Anderman, 2003; Klem & Connell, 2004; Suldo et al., 2009; Wang & Holcombe, 2010), students who perceived that they adapted easily to the academic demands of middle school were more likely to report greater school valuing and belonging; were rated by teachers as expending greater academic effort and as well-adjusted overall at the beginning and end of sixth grade. Those students who easily adapted to the academic demands of middle school were also more likely to have higher standardized test scores in fifth grade and at the end of sixth grade. Furthermore, students who reported favorable experiences with the academic demands of middle school were also more likely to experience positive behavioral and social adjustment, such as being well-liked and exhibiting prosocial behaviors. Results of previous studies associated students' academic experiences and social adjustment, reporting that students who were higher achieving were more likely to be associated with prosocial behaviors, well-liked, and have positive relationships with peers (Davidson et al., 2010; Wentzel, 1993; Wentzel & Asher, 1995). The results of the present study reinforced existing studies that reported positive relationships between students' academic experiences and social adjustment.

Students who reported more favorable experiences on the Social subscale were likely to feel a sense of belonging to school, be viewed by teachers as well-adjusted overall, be popular, well-liked, and have positive social interactions with peers at both the beginning and end of sixth grade. Students with more favorable experiences adjusting to the social demands of middle

school were also likely to demonstrate academic benefits, including greater standardized test scores in math at the end of sixth grade. These findings are consistent with past studies in which students who established positive relationships with peers were more socially and academically adjusted following the transition to middle school (Davidson et al., 2010; Erath et al., 2008; Graham et al., 2006; Wentzel et al., 2004; Wentzel et al., 2010). However, one significant difference was found regarding how experiences with the social demands of middle school influenced White and African American students' adjustment. Unlike White students, African American students who felt that they responded to the social demands of middle school favorably were more likely to be perceived as aggressive by teachers throughout the sixth-grade year. According to the principles of life course theory, although individuals may experience similar contextual demands, demands may be interpreted differently; contributing to alternative adjustment outcomes (Elder, 1998; Elder & Shanahan, 2006). Furthermore, groups of students may experience contextual demands differently than other groups (Pallas, 2003). Thus, the findings of the present study, which indicate that African American students who experienced more favorable adjustment to the social demands of middle school were also more likely to be aggressive, are consistent with the principles of life course theory. In previous studies, African American males, in particular, were more likely to be seen as popular by their peers if they exhibited aggressive, or tough, behaviors (Rodkin et al., 2000). African American boys were also more likely to associate aggressive behaviors with popularity at the beginning of sixth grade (Graham, Taylor, & Hudley, 1998; Xie et al., 2013). Therefore, according to life course theory (Elder, 1998; Elder & Shanahan, 2006), African American students may be interpreting the social demands of middle school differently than White students, resulting in a positive



association between aggressive behaviors and perceived success in adapting to the social demands of middle school.

In the present study, students who easily adjusted to the procedural demands of middle school enjoyed more favorable academic, social, and behavioral adjustment. Students who had positive experiences with the procedural demands of middle school were less likely to exhibit antisocial behaviors and were well-liked at the beginning of the sixth-grade year. Students who perceived they adjusted to the procedural demands of middle school easily were also likely to have higher levels of academic effort and achievement, school belonging, likely to be perceived as a class leader, and well-adjusted overall at the beginning and end of the sixth-grade year. One exception to this pattern was that economically advantaged students who reported more negative experiences with procedural demands were less likely bully peers in the beginning of sixth grade, whereas economically disadvantaged students were more likely to bully others. However, this pattern was not significant for the spring of sixth grade.

Few studies have assessed the relationships between students' responses to procedural demands and their adjustment. In two studies, students who did not adapt easily to the procedural demands of middle school had higher levels of stress and were more likely to be disciplined by their classroom teachers (Lane et al., 2006; Rudolph et al., 2001). The present study adds to this limited literature, finding that, consistent with life course theory (Elder & Shanahan, 2006), students' responses to procedural demands were associated with their academic and social adjustment. The findings from the present study also add that negative behavioral adjustment outcomes that students may experience as a result of procedural difficulties vary across groups of students from different socioeconomic strata.

In summary, life course theorists assert that students' experiences with contextual demands are influential to their adjustment following a transition (Elder & Shanahan, 2006). In the present study, the SCALE captured students' experiences with academic, procedural, and social demands commonly associated with the middle school transition (Akos, 2002; Akos & Galassi, 2004a, 2004b). Consistent with other studies (e.g., Akos & Galassi 2004a, Day et al., 2014; Elias et al., 1992), students who were well-adjusted following the transition to middle school were likely to have favorable perceptions of their experiences with the academic, procedural, and social demands of middle school at the beginning and end of the sixth-grade year. Furthermore, consistent with life course theory (Elder & Shanahan, 2006), differences emerged among groups in terms of their responses to the demands of the middle school transition. Specifically, African American students' interpretations of social demands were different than White students' and contributed to differential adjustment outcomes between the two groups. Discrepant adjustment outcomes were also observed between economically disadvantaged and advantaged students during the fall of sixth grade, suggesting inconsistencies in groups' interpretations of the procedural demands of middle school. Taken together, it was determined that interpretations from the SCALE successfully captured the prominent demands of the middle school transition that were critical to students' adjustment, including those particular to students in different demographic groups.

### **Validity of Interpretations from the SCALE**

According to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), theoretical and statistical evidence is used to establish the construct validity of scale interpretations. It can be inferred from study findings that the SCALE subscales successfully represented students' experiences with academic, procedural, and social demands as

suggested by the principles of life course theory (Elder, 1998; Elder et al., 2003; Elder & Shanahan, 2006; Pallas, 2003) and past empirical research (e.g., Akos et al., 2004a, 2004b; Day et al., 2014; Elias et al., 1992). Thus, theoretical and empirical evidence supported the construct validity of the SCALE. Additionally, results of numerous, rigorous statistical tests provided evidence to support or refute the construct validity of the SCALE.

**Psychometric structure of the SCALE.** Three SCALE subscales representing students' experiences with the academic, procedural, and social demands of middle school were identified across multiple analyses. Factor analytic techniques were also appropriate for ordinal data (Muthén & Asparouhov, 2002; Byrne, 2012; Muthén & Muthén, 2013), which reduced the probability of bias in parameters and standard errors (Schmitt, 2011) and increased the likelihood that the psychometric structure of the SCALE was established without bias (Holgado-Tello et al., 2010). In addition, advanced techniques were used in the present study to assess the invariance of the SCALE across demographic groups, to assure that the interpretations of the SCALE were not culturally biased (Byrne, 2012; Knight & Hill, 1998).

Full configural, metric, and scalar invariance were found for the SCALE across gender and socioeconomic status. Partial invariance was found for the SCALE across White and African American groups as two items within the Social factor were metrically variant. Although some researchers may question the construct validity of partially invariant SCALES, findings of partial invariance are quite common (Vandenberg & Lance, 2000; Yang, 2008). Researchers have asserted that a small presence of partial invariance will not adversely affect latent means, particularly when modeled appropriately (Byrne et al., 1989; Kline, 2011; Millsap & Kwok, 2004; Yang, 2008). Sass (2011) also did not find evidence that partial invariance limited the ability of models to detect mean differences across groups. Therefore, for the

purposes SCALE, modeling partial invariance for the two Social subscale items during group comparisons using White and African American students would be necessary, but it is not anticipated that this will adversely affect the validity of the interpretations from the SCALE in a significant way (Byrne et al., 1989; Sass, 2011; Yang, 2008).

Results of factor analytic techniques in the present study provide support for the construct validity of the SCALE for use with diverse middle school students. The SCALE maintained its psychometric structure through progressively conservative tests. Due to findings of full and partial invariance, it was likely that the SCALE assessed students' experiences with the academic, procedural, and social demands of middle school without bias related to inappropriate statistical techniques (Holgado-Tello et al., 2010) or cultural bias related to students' gender, socioeconomic status, or race/ethnicity (Byrne, 2012; Kline, 2011; Knight & Hill, 1998).

**Reliability and stability.** Scales that are internally consistent are more likely to assess participants' true scores on a latent variable of interest, increasing the likelihood that scores are measured without error (Cronbach, 1951; DeVellis, 2012). Scales that are stable also increase the likelihood that scores are measured without error over time (DeVellis, 2012; McGraw & Wong, 1996; Weir, 2005). The internal consistency and stability for two of the SCALE subscales, Academic and Social, were acceptable. Cronbach's alpha for the Procedural subscale was acceptable when used with sixth-grade students at the beginning of sixth grade. Stability estimates were also acceptable for the Procedural subscale. However, Cronbach's alpha for the Procedural subscale was marginally questionable when used with sixth-grade students at the end of the sixth-grade year. As a result, both the Academic and Social subscales of the SCALE were determined to be reliable assessments of students' transition experiences at the beginning and end of the transition year, but the assessments of procedural experiences may offer more reliable

information about students' transition experiences if administered at the beginning of the transition year.

Students' ratings of their experiences with procedural items were very positive in the fall of sixth grade, but were extremely positive in the spring of sixth grade. Between 90% and 95% of participants reported having little to no problem with many procedural items by the end of the sixth-grade year. Cronbach's alpha can be sensitive to skewed data, particularly with smaller sample sizes (Sheng & Sheng, 2012). Subsequently, it is possible that the Procedural subscale was stable, in that students' assessments of their experiences were relatively consistent from the fall to the spring. However, enough students reported increasingly positive experiences with the procedural items by the end of the school year that skewness in the data contributed to a poor Cronbach's alpha estimate for the Procedural subscale.

Schools today may provide excellent intervention to students before or during the sixth grade transition that adequately prepares them to respond to the procedural demands of middle school assessed by the SCALE. Hamm et al. (2014) reported that, when asked about their students' preparedness for the procedural demands of middle school, parents recounted numerous intervention strategies used by schools to prepare students to meet these demands. In future studies of students in the middle school transition, other procedural tasks may be identified that are more demanding for today's students that can strengthen the internal consistency of the Procedural subscale across the transition year.

**Concurrent and predictive construct validity.** Finally, significant concurrent and predictive relationships between the Academic, Social, and Procedural subscales and indicators of students' academic, behavioral, and social adjustment in the fall and the spring of sixth grade were supported by past studies of students' adjustment experiences. Consistent with past studies,

students' experiences with academic demands were significantly related to their academic (Klem & Connell, 2004; Suldo et al., 2009; Turner & Patrick, 2004; Wang & Holcombe, 2010), social (Erath et al., 2008; Niehaus et al., 2012), and behavioral adjustment (Farmer, Estell, Leung et al., 2003). Students' social experiences were positively related to indicators of their social and behavioral adjustment (Erath et al., 2008; Niehaus et al., 2012; Wentzel et al., 2010). Finally, in the present study, students' procedural experiences were significantly correlated with indicators of students' academic, behavioral, and social adjustment in the fall and spring of sixth grade. These findings add to the overall literature on procedural demands and supported the findings from one previous study that found that students' procedural experiences contributed to their behavioral adjustment (Lane et al., 2006).

The vast majority of significant relationships between SCALE subscales and students' academic, behavioral, and social adjustment across gender, socioeconomic status, and racial/ethnic groups maintained the associations between the SCALE subscales and indicators of students' academic, behavioral, and social adjustment from the overall sample. An inverse relationship between students' bullying behavior early in the sixth-grade year and procedural adjustment was significantly different across socioeconomic groups. This finding could not be judged as a violation of construct validity as little information regarding the relationship between students' procedural experiences and indicators of adjustment is currently available. Another relationship was found to be significantly different across African American and White groups. A positive relationship between African American students' social experiences and higher levels of aggressive behavior was only found for the African American group, but this was supported by findings from previous studies (Graham et al., 1998; Rodkin et al., 2000; Xie et al., 2013). Consequently, relationships between students' sixth-grade adjustment and the transition

experiences of students from different gender, socioeconomic status, and racial ethnic groups was determined to support the concurrent and predictive construct validity of the SCALE.

However, for African American students, a smaller number of significant correlations were observed between Academic and Social subscales and fall measures of behavioral, social, and academic adjustment and between adjustment measures from the spring of sixth grade and all SCALE subscales. Fewer significant correlations were found between the Social and Procedural subscales and fall and spring measures of adjustment for economically disadvantaged students and a smaller number of significant correlations were also identified between females' social experiences on the SCALE and adjustment measures from the fall of sixth grade. The smaller number of significant correlations for these groups of students was not considered to be a violation of construct validity because, like other groups, significant correlations that were present were supported by empirical research.

Rather, the smaller number of significant relationships between students' transition experiences and measures of students' adjustment may have been related to the type of adjustment measure. Out of the fifteen adjustment measures used in the present study, three of these measures were self-report measures, two were standardized test scores, and the remaining ten were teacher-rated measures. In previous studies, teachers' and students' assessments of students' schooling experiences have not been aligned with one another (Neal, Cappella, Wagner, & Atkins, 2011; Pearl, Leung, Van Acker, Farmer, & Rodkin, 2007; Wienke Totura, Green, Karver, & Gesten, 2009). In one study (Neal et al., 2011), teachers with larger class-sizes or with classes with increased instances of aggressive behavior were more likely to view the social dynamics of their classrooms differently than students. As a consequence, some teachers

may be preoccupied with other classroom demands that inhibit their ability to perceive how each student is personally adjusting to middle school demands.

Furthermore, students struggling to adjust to middle school demands may not always exhibit adjustment problems in ways that are easily identifiable to teachers. For example, girls are more likely to display internalizing behaviors, such as anxiety or depression, which are not readily observable by definition (Barber & Olsen, 2004; Hogland, 2007). Furthermore, African American boys and girls have exhibited internalizing behaviors when experiencing discrimination in school (Brody et al., 2006). Therefore, there is the potential that some of the indicators of adjustment for African American, economically disadvantaged students, and females who are struggling with transition demands may manifest in ways not accounted for by the adjustment measures used in the present study. Therefore, the SCALE may prove to be an important assessment tool for future researchers and practitioners as it offers the potential to create a more holistic account of students' transition experiences when combined with teacher-rated adjustment measures. Future studies should continue to explore the relationship between students' and teachers' observations of students' adjustment to the demands of middle school, particularly for African American and economically disadvantaged students as those in these groups are more likely to experience significant adjustment difficulties following the middle school transition (Balfanz et al., 2007; Becker & Luthar, 2002; Graham et al., 2006; Ryan et al., 2013).

In addition, future studies may consider evaluating whether students' perceptions from the SCALE are correlated with other types of adjustment measures, such as peer-report measures. Peer-report measures, like teacher-rated measures, contribute distinctive information regarding students' schooling experiences and adjustment (Pearl et al., 2007). Comparing the



relationships between peer-report measures and the SCALE could inform how students' responses to demands of the middle school transition also contribute to other aspects of their adjustment, such as their social standing, through the perspective of their peers.

**Summary.** The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999) require that sufficient theoretical and statistical evidence be presented to support or refute the construct validity of a measure. The presence of three subscales assessing students' experiences with academic, procedural, and social demands was consistent with the principles of life course theory (Elder et al., 2003; Elder & Shanahan, 2006) and results of empirical research (Akos, 2002; Akos & Galassi 2004a, 2004b). Consequently, theoretical evidence supported the construct validity of the SCALE. Interpretations of analyses of the psychometric structure of the SCALE subscales, reliability analyses, and concurrent and predictive relationships of the subscales with measures of students' adjustment provided additional statistical support for the construct validity of the SCALE subscales when used with diverse sixth-grade students at the beginning of sixth grade. Moreover, Academic and Social subscales may also be appropriate for use with diverse sixth-grade students at the end of the transition year. As a result, acknowledging some limitations of the Procedural subscale, interpretations from the SCALE regarding students' experiences with the academic, procedural, and social demands of middle school were considered adequately validated and suitable for use with diverse sixth-grade students.

### **Utility of the SCALE**

In the present study, theoretical and empirical evidence supported the construct validity of the SCALE subscales. Thus, the SCALE is considered to be a suitable research tool for measuring students' experiences with the academic, procedural, and social demands of middle

school. Researchers and practitioners invested in intervention research may use the SCALE to assess the extent to which interventions change students' perceptions of the middle school transition. In the past twenty years, interventions aimed at improving the middle school transition have often incorporated strategies to assist students with becoming comfortable in their new school environment and with their new peers (Cauley & Jovanovich, 2006; Juvonen et al., 2004). Such strategies have included training fifth-grade students to use lockers or change classes, taking fifth-grade students on tours of their new middle school, or teaming in sixth grade (Hamm et al., 2014; Juvonen, 2007; Juvonen et al., 2004; Midgley & Edelin, 1998). In the present study, most students had particularly positive experiences with procedural and social demands, potentially due to a wider implementation of procedural- and social-oriented interventions in the last twenty years.

However, examination of students' responses on the items within subscales indicated that a larger percentage of students reported difficulty with items in the Academic subscale and the percentage of students perceiving that they had problems accomplishing academic demands also increased over the course of the sixth-grade year. This suggests that academic demands may be more challenging for students to master than procedural and social demands. Furthermore, demands were significantly more challenging for boys. Subsequently, in keeping with the results from the present study, there may be a need for targeted academic intervention during the middle school transition, particularly for boys, that eases students' experiences with academic demands in an effort to improve students' adjustment outcomes following the transition to middle school. In the present study, the Academic subscale was determined to be reliable and stable. Students' perceptions of their experiences with academic demands on the SCALE were also significantly related to their math and reading achievement scores at the beginning and end of sixth grade.

Therefore, the SCALE would be an appropriate measure to assess change in students' transition experiences over the course of the sixth-grade transition year as a result of intervention.

In addition, use of the SCALE in future studies may also assist in creating a more nuanced understanding of the gender differences in students' adjustment following the transition to middle school. In other studies, adjustment differences across gender groups have been somewhat mixed. For example, girls were more likely than boys to experience greater social adjustment in terms of feelings of teacher support (e.g., Malecki & Demaray, 2006; Way et al., 2007) but both boys and girls experienced social adjustment difficulties with peers in past studies (e.g., Neihaus et al., 2012; Wentzel et al., 2010). In the present study, no significant differences were found in males' and females' perceptions of their adjustment to the social demands of middle school. In the case of academic adjustment, gender differences have been inconsistent depending upon the indicator used in the study and the subject area. Males displayed lower standardized test scores in reading than females in middle school, but higher standardized test scores in math (Akos et al., 2014). However, females displayed more growth in math achievement from fifth to sixth grade. Studies have also found that females' math engagement was stronger than males' in sixth grade (Cleary & Chen, 2009). The subscales of the SCALE, which were found to be valid measures of students' social, academic, and procedural experiences across gender in the present study, could be used in the future studies to identify specific differences in the way that boys and girls experience different middle school demands. This information would then provide greater insight into specific types of transition experiences that may be different for boys and girls that contribute to adjustment differences, or a lack thereof, across gender groups following the transition to middle school.

In the future, practitioners may also use the SCALE as a screening tool to identify students and/or groups of students having trouble adapting to middle school demands. Practitioners could use the SCALE to identify specific domains of middle school adjustment that are problematic for students and/or groups of students. Identifying students experiencing difficulty following the transition to middle school may be challenging for teachers or other educators. Middle school teachers are often responsible for large numbers of students, which may make it harder for them to establish relationships strong enough to identify individual students struggling to adjust to the demands of middle school (Anderson et al., 2000; Juvonen, 2007). In previous studies, students with teachers who were more attuned to students' middle school experiences were more likely to experience positive adjustment outcomes (Farmer, Lines, & Hamm, 2011; Hamm et al., 2011). Intervention studies have demonstrated that teachers can learn how to be more attuned to the classroom experiences of their students, such as the social dynamics or the presence of bullying among classmates (Farmer, Hall, Petrin, Hamm, & Dadisman, 2010; Farmer et al., 2010; Hamm et al., 2011). The Social subscale was determined to be a reliable measure of students' social experiences. Accordingly, the SCALE may be a valuable tool for educators, aiding their identification of students who are experiencing problems adjusting to the social demands of middle school early in the transition year.

Furthermore, the SCALE may also be used to create a more comprehensive understanding of the middle school experiences of African American youth. African American youth, particularly males, have been more likely to be identified by teachers as aggressive in middle school (Graham et al., 2006; Skiba, Michael, Nardo, & Peterson, 2002; Wang et al., 2009). Relatedly, African American students and principally African American males have been more likely to experience disciplinary action as a result of aggressive behaviors, including school

suspension and expulsion (Farkas, 2003; Fenning & Rose, 2007). As a result, African American students are at an increased risk of negative outcomes, such as lower academic achievement or school drop-out. In the present study, African American students who perceived more success with the social demands of middle school were more likely to be perceived by their teachers as aggressive. This relationship, although documented in past studies (e.g., Graham et al., 1998; Xie et al., 2013), may indicate a difference between middle school students' and teachers' understanding of African American students' behavioral and social adjustment. Increasing the sensitivity of teachers to the needs of their students may contribute to a more positive middle school context for all students (Becker & Luther, 2002; Farmer et al., 2011; Hamm et al., 2011) and allow all students to respond to the adjustment demands of middle school in positive ways (Elder et al., 2006). The SCALE can be used to increase teachers' understanding of the personal transition experiences of African American students. Increasing teachers' knowledge of African American students' transition experiences may potentially contribute to a decrease in the over-identification of African American students, particularly males, as aggressive and encourage positive adjustment outcomes for students in the future.

### **Limitations**

Although rigorous techniques were used in the present study to validate the SCALE, this study is not without its limitations. Sample sizes for the majority of tests conducted in the present study were sufficient for adequate power. However, the sample sizes for male and African American groups were smaller than recommended for factor analytic techniques using ordinal data. The African American sample, in particular, was quite small. Thus, results from analyses in the present study for these groups should be interpreted with caution. Larger sample sizes in future studies of the SCALE would provide additional information regarding the

properties of the SCALE when used with these groups, including the presence of partial invariance across White and African American groups.

In future studies, researchers should also assess the measurement properties of the scale for additional ethnic groups. Latino students are one of the fastest growing groups in the United States (Eccles & Roeser, 2011a). Latino youth, like African American youth, may be more likely to be graded more stringently, be less highly represented in advanced classes, and also may experience discrimination from both teachers and peers (Farkas, 2003). Findings from past studies have indicated that Latino youth continue to demonstrate lower academic achievement than their peers and are more likely to be disciplined in school (Balfanz et al., 2007; Graham et al., 2006; Ryan et al., 2013). Akos and Galassi (2004a) also found that Latino students were significantly more likely to view the transition to middle school as more challenging than their White and African American peers.

Recent studies also indicate that upwards of 5% of students in middle school identify as lesbian, gay, bisexual, transgender, or queer (LGBTQ; Shields et al., 2013). Students identifying as LGBTQ may be more likely to experience bullying, discrimination, and harassment (GLSEN, 2009; Juvonen & Graham, 2014). Middle schools may also have fewer resources available for LGBTQ students (GLSEN, 2009). Therefore, LGBTQ youth may also have a harder time adjusting to the demands of middle school.

Additionally, students within groups may experience the middle school transition differently, based on other demographic characteristics (Pallas, 2003). For example, Day et al. (2014) found that white males were the most likely to perceive favorable transition experiences and African American males were the least likely. Group sizes were too small to complete invariance tests for composite groups in the present study. In future studies, researchers should

explore the measurement invariance of the SCALE including additional groups, such as Latino and LGBTQ youth as well as composites of students' demographic characteristics.

Five items from the SCALE as completed by students were not retained in the final version of the SCALE, including “being treated more like a child,” “getting into fights,” “teachers expecting too much of you,” “not getting along with all your different teachers,” and “having an argument with a teacher.” These items did not load in a predictable or interpretable manner. Three of these items specifically identified students' interactions with teachers. Life course theorists assert that relationships are particularly significant to students' adjustment following a transition (Elder et al., 2003). Studies have also reported that students who establish positive, supportive relationships with teachers following the transition to middle school were more likely to have higher levels of academic achievement, engagement, and motivation (Barber & Olsen, 2004; Niehaus et al., 2012; Wentzel et al., 2010). Students' difficulty or ease in establishing positive relationships with teachers during the transition may be particularly important to their adjustment in middle school. Future versions of the SCALE may consider revising items or adding new items that contribute to a factor that captures students' social experiences with teachers.

Additionally, in preliminary analyses of the SCALE, a second-order factor was found. All three subscales significantly loaded onto the second-order factor, suggesting that the second-order factor represented students' overall perceptions of their experiences with the demands of middle school. However, multiple group analyses of the second-order factor were unable to be completed, potentially due to low sample sizes in some groups or the need for additional factors for identification purposes (Byrne, 2012; Chen, Sousa, & West, 2005). Thus, the present study produced evidence that there was also a larger factor within the SCALE representing students'

overall transition experiences, but future studies with larger sample sizes will need to be completed to validate this factor for use with diverse sixth-grade students.

It should be noted that many of the items in the SCALE were drawn from the SAT-MS (Elias et al., 1992), including the five items that were not retained in the final version of the SCALE. However, subscales and the items that grouped together within each subscale were not the same between the SAT-MS and the SCALE. According to Elias et al. (1992), the SAT-MS contained subscales that represented psychosocial tasks that students experienced during the transition which contributed to their middle school adjustment. The SAT-MS included four subscales called Substance Abuse, Peer Relationships, Conflicts with Authority/Older Students, and Academic Pressures. In the SAT-MS, “having a tough teacher,” “forgetting your locker combination,” “getting lost and not being able to find your way around school,” and “having the wrong books and supplies for class” were grouped together in the Conflicts with Authority/Older Peers subscale with other items not shared between the SAT-MS and the SCALE. In contrast, the SCALE contains three subscales representing students’ experiences with the academic, procedural, and social demands of middle school. In the present study, “having a tough teacher” aligned with other academic items in the Academic subscale. “Forgetting your locker combination,” “getting lost and not being able to find your way around school,” and “having the wrong books and supplies for class” grouped together within the Procedural subscale of the SCALE. Many of the same items shared between the two scales were grouped together in a SAT-MS subscale called, Peer Relationships, and in the Social subscale of the SCALE. However, the Peer Relationships subscale contained some items not retained in the final version of the Social subscale of the SCALE. Therefore, even though the SCALE and the SAT-MS shared some of the same items, the items constructing the subscales and the constructs



characterized by subscales differed between the two scales. Consequently, the SCALE was determined to be a separate measure from the SAT-MS.

## **Conclusion**

Transitions, like the transition to middle school, are characterized by new demands and expectations (Elder & Shanahan, 2006). Students' responses to contextual demands influence their adjustment following a transition (Elder et al., 2003; Elder & Shanahan, 2006). In the present study, the SCALE was validated as a credible measure of students' transition experiences by demonstrating the capacity to differentiate students' perceptions of their abilities to adapt to the academic, procedural, and social demands of middle school and potential to predict later adjustment. The availability of a validated measure that assesses diverse students' transition experiences is beneficial to future researchers studying students' adjustment during the middle school transition. In the present study, using rigorous statistical techniques lacking in past studies of middle school transition perception measures, the SCALE was validated using a sample of diverse sixth-grade students and determined to be suitable for use. As a result, the SCALE is an acceptable tool available to researchers and practitioners invested in understanding and improving the transition experiences of modern, diverse middle school students.

Table 1

*Additional Characteristics of Participating Schools*

School	School Size	% Minority Students	% Free/Reduced Lunch	% Student Suspensions <sup>a</sup>	% Proficient in Reading	% Proficient in Math
A	1052	25.70%	21.80%	.07%	88.10%	94.10%
B	764	71.00%	61.50%	.21%	61.00%	73.90%
C	595	47.60%	58.70%	.49%	65.50%	82.40%
D	898	20.00%	27.20%	.29%	74.50%	85.40%
E	652	35.90%	43.10%	.16%	69.50%	83.50%
F	846	56.90%	52.60%	.31%	65.80%	81.60%
G	888	38.40%	44.40%	.18%	74.10%	85.50%

*Note.* Information provided by the National Center for Education Statistics (NCES) as well as

state-level school report card data.

<sup>a</sup>Average rate per 100 students.

Table 2

*Items in the SCALE*

Item
1. Getting lost and not being able to find your way around school.
2. Forgetting your locker combination.
3. Being treated more like a child.
4. Having a tough teacher.
5. Having to do harder school work.
6. Finding kids I can sit with at lunch.
7. Having an argument with a teacher
8. Not having the right books or supplies for class.
9. Getting too much homework.
10. Getting into fights
11. Not seeing your friends from elementary school enough.
12. Having trouble making new friends.
13. Kids trying to talk you into things you don't want to do.
14. Being bothered by the older kids.
15. Not getting along with all your different teachers.
16. Other kids teasing you.
17. Not being in the "in" group, like not being able to go around with the group of kids you'd like to hang around with.
18. Teachers expecting too much of you.
19. Having to change classes.
20. Understanding new rules.

Table 3

*Cronbach's Alpha Coefficients for the ICS-T Subscales Across Racial/Ethnic Groups*

<u>ICS-T Subscale</u>	<u>White Students</u>	<u>African American Students</u>
Aggressive	$\alpha = .85$	$\alpha = .83$
Popular	$\alpha = .85$	$\alpha = .83$
Affiliative	$\alpha = .72$	$\alpha = .67$
Academic	$\alpha = .78$	$\alpha = .75$

Table 4

*Descriptive Statistics of SCALE Items - Fall of 6th Grade*

Item	<i>M</i>	<i>SD</i>	Skewness	SEsk	Kurtosis	SEku
1	1.40	.69	1.73	.09	2.49	.18
2	1.32	.73	2.39	.09	4.87	.18
3	1.79	.97	.98	.09	-.17	.18
4	2.13	1.04	.44	.09	-1.02	.18
5	2.21	1.01	.38	.09	-.95	.18
6	1.31	.70	2.47	.09	5.63	.18
7	1.48	.92	1.87	.09	2.18	.18
8	1.67	.88	1.21	.09	.59	.18
9	2.25	1.10	.38	.09	-1.17	.18
10	1.56	1.02	1.60	.09	1.02	.18
11	2.27	1.16	.35	.09	-1.34	.18
12	1.47	.84	1.83	.09	2.49	.18
13	1.61	.94	1.45	.09	.96	.18
14	1.74	1.01	1.16	.09	.06	.18
15	1.55	.88	1.53	.09	1.28	.18
16	1.59	.96	1.49	.09	.94	.18
17	1.49	.87	1.72	.09	1.85	.18
18	1.85	1.00	.93	.09	-.32	.18
19	1.34	.75	2.29	.09	4.39	.18
20	1.38	.75	2.11	.09	3.90	.18

*Note.* SEsk = standard error skewness; SEku = standard error kurtosis.

Table 5

*Descriptive Statistics of SCALE Items - Spring of 6th Grade*

Item	<i>M</i>	<i>SD</i>	Skewness	SEsk	Kurtosis	SEku
1	1.29	.57	2.13	.09	4.61	.18
2	1.31	.69	2.50	.09	5.88	.18
3	1.83	.96	.89	.09	-.32	.18
4	2.38	1.06	.15	.09	-1.19	.18
5	2.30	1.01	.26	.09	-1.01	.18
6	1.33	.74	2.35	.09	4.67	.18
7	1.64	.98	1.35	.09	.54	.18
8	1.66	.89	1.22	.09	.55	.18
9	2.44	1.11	.13	.09	-1.33	.18
10	1.53	.95	1.66	.09	1.40	.18
11	2.17	1.13	.46	.09	-1.19	.18
12	1.42	.79	1.96	.09	3.10	.18
13	1.53	.90	1.60	.09	1.46	.18
14	1.68	.96	1.23	.09	.35	.18
15	1.70	.98	1.20	.09	.20	.18
16	1.52	.90	1.66	.09	1.59	.18
17	1.51	.86	1.63	.09	1.67	.18
18	1.99	1.04	.66	.09	-.83	.18
19	1.34	.73	2.26	.09	4.45	.18
20	1.33	.71	2.26	.09	4.50	.18

*Note.* SEsk = standard error skewness; SEku = standard error kurtosis.

Table 6

*Proportions of Response Values for Scale Items - Fall and Spring of 6th Grade*

Item	Response Value	n - Fall	% - Fall	n - Spring	% - Spring
1	1 - No Problem	532	70.00%	518	78.01%
	2	164	21.58%	116	17.47%
	3	52	6.84%	24	3.61%
	4 - Large Problem	12	1.58%	6	.90%
2	1 - No Problem	611	80.60%	528	80.00%
	2	78	10.29%	87	13.18%
	3	43	5.67%	25	3.79%
	4 - Large Problem	26	3.43%	20	3.03%
3	1 - No Problem	385	50.99%	314	47.72%
	2	205	27.15%	195	29.64%
	3	100	13.25%	97	14.74%
	4 - Large Problem	65	8.61%	52	7.90%
4	1 - No Problem	265	35.01%	166	25.34%
	2	227	29.99%	201	30.69%
	3	165	21.80%	166	25.34%
	4 - Large Problem	100	13.21%	122	18.63%
5	1 - No Problem	219	29.01%	161	24.58%
	2	264	34.97%	235	35.88%
	3	170	22.52%	161	24.58%
	4 - Large Problem	102	13.51%	98	14.96%
6	1 - No Problem	601	79.29%	527	79.97%
	2	102	13.46%	71	10.77%
	3	30	3.96%	38	5.77%
	4 - Large Problem	25	3.30%	23	3.49%
7	1 - No Problem	550	73.63%	418	63.33%
	2	100	13.37%	121	18.33%
	3	33	4.42%	57	8.64%
	4 - Large Problem	64	8.57%	64	9.70%
8	1 - No Problem	415	54.89%	378	57.01%
	2	220	29.10%	179	27.00%
	3	78	10.32%	68	10.26%
	4 - Large Problem	43	5.69%	38	5.73%
9	1 - No Problem	238	31.36%	170	25.72%
	2	241	31.75%	198	29.95%
	3	134	17.65%	135	20.42%
	4 - Large Problem	146	19.24%	158	23.90%
10	1 - No Problem	544	72.34%	474	71.82%
	2	80	10.64%	84	12.73%
	3	43	5.72%	47	7.12%
	4 - Large Problem	85	11.30%	55	8.33%

*(table continues)*

Item	Response Value	n - Fall	% - Fall	n - Spring	% - Spring
11	1 - No Problem	262	34.56%	246	37.27%
	2	205	27.04%	191	28.94%
	3	117	15.44%	98	14.85%
	4 - Large Problem	174	22.96%	125	18.94%
12	1 - No Problem	527	69.62%	479	73.02%
	2	146	19.29%	112	17.07%
	3	40	5.28%	39	5.95%
	4 - Large Problem	44	5.81%	26	3.96%
13	1 - No Problem	473	62.82%	451	68.54%
	2	165	21.91%	110	16.72%
	3	52	6.91%	54	8.21%
	4 - Large Problem	65	8.63%	43	6.53%
14	1 - No Problem	433	57.12%	388	58.70%
	2	171	22.56%	154	23.30%
	3	74	9.76%	65	9.83%
	4 - Large Problem	80	10.55%	54	8.17%
15	1 - No Problem	498	66.05%	388	59.24%
	2	144	19.10%	139	21.22%
	3	68	9.02%	69	10.53%
	4 - Large Problem	44	5.84%	59	9.01%
16	1 - No Problem	500	66.40%	457	69.56%
	2	128	17.00%	103	15.68%
	3	59	7.84%	50	7.61%
	4 - Large Problem	66	8.76%	47	7.15%
17	1 - No Problem	534	70.54%	443	67.02%
	2	119	15.72%	128	19.36%
	3	60	7.92%	52	7.87%
	4 - Large Problem	44	5.81%	38	5.74%
18	1 - No Problem	369	48.75%	275	41.67%
	2	213	28.14%	193	29.24%
	3	98	12.95%	112	16.97%
	4 - Large Problem	77	10.17%	80	12.12%
19	1 - No Problem	598	79.42%	524	79.03%
	2	83	11.02%	82	12.37%
	3	45	5.98%	34	5.13%
	4 - Large Problem	27	3.59%	23	3.47%
20	1 - No Problem	562	74.04%	521	78.58%
	2	132	17.39%	94	14.18%
	3	36	4.74%	33	4.98%
	4 - Large Problem	29	3.82%	15	4.37%



Table 7

*Descriptive Statistics of Students' Self-Reported Variables - Fall and Spring of 6th Grade*

	Min	Max	<i>M</i>	<i>SD</i>	Skewness	SEsk	Kurtosis	SEku
<u>Fall</u>								
Defiance	1.00	5.00	1.95	1.03	1.28	.09	.91	.18
Belonging	1.00	5.00	3.81	.69	-.75	.09	.36	.18
School Value	1.00	5.00	4.24	.63	-1.11	.09	1.56	.18
<u>Spring</u>								
Defiance	1.00	5.00	2.13	1.10	.96	.09	.09	.18
Belonging	1.00	5.00	3.71	.73	-.66	.09	.19	.18
School Value	2.00	5.00	4.20	.64	-.99	.09	.94	.18

*Note.* SEsk = standard error skewness; SEku = standard error kurtosis.

Table 8

*Descriptive Statistics for Teacher-Rated Variables-Fall and Spring of 6th grade*

	Min	Max	<i>M</i>	<i>SD</i>	Skewness	SEsk	Kurtosis	SEku
<u>Fall</u>								
Not Aggressive	1.00	7.00	5.78	1.38	-1.16	.09	.57	.18
Popular	1.00	7.00	4.72	1.22	-.16	.09	-.32	.18
Affiliative	2.00	7.00	5.53	1.11	-.65	.09	-.08	.18
Academic	1.00	7.00	4.76	1.49	-.40	.09	-.36	.18
Overall Adjustment	2.43	6.93	5.05	.86	-.34	.09	-.10	.18
Academic Effort	1.00	5.00	4.01	1.09	-.95	.09	-.05	.18
Never a Leader	1.00	7.00	3.81	1.93	.16	.09	-1.15	.18
Never Bullied	1.00	7.00	5.95	1.23	-1.51	.09	2.36	.18
Bullies Peers	1.00	7.00	2.07	1.55	1.62	.09	1.81	.18
Not Liked	1.00	7.00	2.39	1.31	.78	.09	-.11	.18
Reading Achievement	1.00	99.00	60.03	25.72	-.41	.09	-.71	.18
Math Achievement	1.00	99.00	66.19	25.58	-.52	.09	-.70	.18
<u>Spring</u>								
Not Aggressive	1.00	7.00	5.69	1.46	-1.08	.09	.36	.18
Popular	1.00	7.00	4.73	1.23	-.18	.09	-.43	.18
Affiliative	2.00	7.00	5.50	1.13	-.59	.09	-.20	.18
Academic	1.00	7.00	4.85	1.48	-.43	.09	-.40	.18
Overall Adjustment	2.00	7.00	5.04	.85	-.21	.09	-.44	.18
Academic Effort	1.00	5.00	3.94	1.15	-.90	.09	-.17	.18
Never a Leader	1.00	7.00	3.90	1.98	.08	.09	-1.27	.18
Never Bullied	1.00	7.00	5.91	1.31	-1.44	.09	1.67	.18
Bullies Peers	1.00	7.00	2.28	1.66	1.22	.09	.42	.18
Not Liked	1.00	7.00	2.49	1.35	.69	.09	-.23	.18
Reading Achievement	1.00	99.00	52.23	29.30	-.07	.09	-1.23	.18
Math Achievement	1.00	99.00	52.29	29.43	-.07	.09	-1.24	.18

*Note.* SEsk = standard error skewness; SEku = standard error kurtosis.

Table 9

*Polychoric Correlation Table for SCALE Items - Fall of 6th Grade*

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	.61																		
3	.26	.34																	
4	.29	.26	.33																
5	.21	.23	.12	.52															
6	.28	.30	.31	.26	.37														
7	.27	.57	.31	.39	.32	.09													
8	.40	.50	.31	.37	.20	.24	.47												
9	.16	.26	.23	.49	.60	.17	.34	.36											
10	.39	.55	.35	.30	.21	.22	.73	.47	.30										
11	.11	.14	.31	.23	.27	.27	.03	.12	.31	.16									
12	.41	.48	.49	.35	.29	.67	.26	.42	.24	.50	.46								
13	.42	.51	.47	.30	.29	.39	.48	.48	.33	.70	.21	.65							
14	.33	.55	.54	.30	.23	.24	.39	.34	.30	.54	.23	.58	.61						
15	.28	.43	.43	.54	.42	.31	.63	.47	.39	.51	.18	.39	.48	.43					
16	.31	.40	.42	.27	.23	.36	.34	.46	.32	.59	.26	.66	.71	.66	.33				
17	.26	.33	.47	.34	.27	.48	.23	.36	.24	.37	.36	.69	.58	.56	.36	.68			
18	.13	.41	.34	.47	.47	.28	.51	.40	.51	.47	.44	.39	.50	.37	.64	.47	.45		
19	.48	.50	.24	.31	.42	.54	.40	.45	.30	.48	.18	.44	.46	.30	.43	.32	.40	.43	
20	.33	.35	.14	.34	.32	.15	.31	.28	.30	.40	.09	.26	.38	.23	.32	.11	.27	.35	.55

Table 10

*Polychoric Correlation Table for SCALE Items - Spring 6th Grade*

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	.55																		
3	.34	.31																	
4	.25	.31	.48																
5	.24	.31	.31	.56															
6	.33	.34	.40	.26	.37														
7	.26	.48	.41	.43	.27	.21													
8	.36	.49	.31	.27	.34	.32	.54												
9	.23	.22	.26	.41	.59	.30	.35	.31											
10	.31	.49	.34	.32	.29	.26	.69	.58	.24										
11	.17	.27	.29	.29	.28	.30	.22	.29	.32	.22									
12	.35	.34	.31	.21	.35	.64	.18	.36	.26	.36	.44								
13	.38	.41	.46	.26	.26	.39	.47	.49	.20	.62	.35	.58							
14	.44	.36	.40	.30	.27	.35	.28	.33	.20	.37	.33	.50	.58						
15	.23	.36	.44	.53	.35	.22	.69	.40	.37	.51	.28	.17	.44	.35					
16	.30	.40	.42	.34	.30	.42	.40	.46	.21	.61	.30	.57	.64	.58	.36				
17	.28	.29	.37	.26	.30	.54	.28	.37	.28	.44	.47	.65	.59	.58	.33	.63			
18	.24	.34	.53	.57	.31	.30	.49	.36	.51	.44	.33	.36	.41	.39	.61	.36	.47		
19	.39	.40	.31	.27	.33	.38	.32	.33	.35	.46	.32	.46	.44	.43	.37	.46	.49	.40	
20	.32	.31	.30	.35	.31	.36	.48	.45	.35	.52	.29	.34	.50	.33	.45	.38	.38	.42	.56

Table 11

*Correlations between SCALE Items and Measures of Students' Adjustment - Fall of 6th Grade*

Item	Agg	Pop	Aff	Aca	Ovr	Def	Eff	Bel	Val	Read	Math	Lead	Blyd	Bllies	Like
1	-.01	-.09	-.03	-.03	-.06	.04	-.05	-.13	-.05	-.03	-.02	.22	-.11	-.06	.08
2	-.12	-.12	-.17	-.16	-.21	.09	-.14	-.17	-.08	-.24	-.24	.22	-.13	.05	.13
3	-.02	-.11	-.12	-.01	-.09	.13	-.09	-.31	-.10	.00	-.03	.09	.01	-.05	.06
4	-.14	-.05	-.16	-.08	-.15	.16	-.13	-.23	-.20	-.11	-.17	.09	-.07	.12	.09
5	-.05	-.11	-.16	-.14	-.17	.19	-.18	-.25	-.24	-.10	-.17	.20	-.14	.06	.12
6	-.06	-.17	-.12	-.04	-.15	-.02	-.07	-.35	-.09	-.08	-.06	.12	-.22	.01	.23
7	-.25	-.07	-.14	-.13	-.21	.42	-.19	-.17	-.22	-.16	-.16	.11	-.12	.21	.14
8	-.13	-.16	-.16	-.11	-.21	.13	-.15	-.18	-.08	-.09	-.11	.19	-.18	.09	.19
9	-.03	-.08	-.14	-.10	-.12	.14	-.09	-.23	-.23	-.07	-.10	.14	-.06	-.01	.08
10	-.14	-.11	-.16	-.15	-.20	.28	-.17	-.27	-.26	-.20	-.18	.17	-.13	.09	.18
11	.06	-.08	-.01	.01	-.03	.01	-.00	-.24	-.11	.09	.07	.03	-.03	-.06	-.01
12	.04	-.20	-.12	.01	-.12	-.07	-.05	-.45	-.11	-.01	-.03	.12	-.16	-.10	.18
13	-.07	-.08	-.15	-.12	-.16	.11	-.15	-.30	-.13	-.21	-.15	.14	-.11	.03	.17
14	.03	-.13	-.04	-.03	-.07	-.01	-.03	-.29	-.03	-.02	.01	.08	-.10	-.06	.10
15	-.26	-.11	-.19	-.14	-.25	.34	-.19	-.29	-.23	-.14	-.19	.18	-.11	.19	.17
16	-.08	-.17	-.11	-.13	-.19	-.01	-.13	-.42	-.10	-.11	-.10	.13	-.18	-.01	.19
17	.01	-.19	-.10	-.10	-.16	.01	-.09	-.43	-.09	-.10	-.09	.18	-.14	-.07	.17
18	-.19	-.17	-.19	-.17	-.25	.23	-.21	-.31	-.30	-.08	-.11	.17	-.13	.14	.19
19	-.05	-.06	-.10	-.10	-.13	.08	-.11	-.15	-.11	-.05	-.07	.15	-.12	.06	.15
20	-.21	-.10	-.16	-.14	-.23	.28	-.14	-.20	-.19	-.15	-.15	.14	-.16	.24	.16

*Note.* Agg = not aggressive; Pop. = popular; Aff = affiliative; Aca = academic; Ovr = overall adjustment; Def = defiance; Eff = academic effort; Bel = belonging; Val = school value; Read = reading achievement; Math = math achievement; Lead = never a leader; Blyd = never bullied; Bllies = bullies peers; Like = not liked.

Table 12

*Correlations between SCALE Items and Measures of Students' Adjustment - Spring of 6th Grade*

Item	Agg	Pop	Aff	Aca	Ovr	Defy	Eff	Bel	Val	Read	Math	Lead	Blyd	Bllies	Like
1	.05	-.14	-.04	-.05	-.09	-.05	-.01	-.12	-.05	.04	-.01	.15	-.11	-.10	.10
2	-.11	-.13	-.11	-.18	-.21	.10	-.14	-.13	-.04	-.15	-.14	.12	-.18	.11	.15
3	-.15	-.10	-.08	-.17	-.19	.14	-.11	-.34	-.15	.12	-.09	.05	-.08	-.17	.12
4	-.21	-.03	-.19	-.19	-.21	.17	-.24	-.28	-.16	-.10	-.10	.10	-.13	.24	.10
5	-.12	-.12	-.12	-.22	-.22	.13	-.22	-.24	-.17	-.11	-.11	.12	-.14	.09	.13
6	-.01	-.22	-.01	-.06	-.12	-.05	-.04	-.29	-.16	-.06	-.03	.06	-.19	.03	.20
7	-.43	-.14	-.22	-.32	-.40	.47	-.39	-.31	-.19	-.26	-.21	.15	-.17	.39	.21
8	-.18	-.22	-.18	-.22	-.29	.25	-.17	-.23	-.12	-.16	-.13	.19	-.16	.14	.17
9	-.10	-.12	-.13	-.07	-.14	.14	-.15	-.18	-.19	-.01	-.02	.09	-.14	-.11	.09
10	-.30	-.15	-.17	-.26	-.31	.27	-.30	-.32	-.17	-.32	-.26	.21	-.19	.26	.19
11	.08	-.07	-.01	-.05	-.02	.02	.04	-.17	-.03	.00	.03	.06	-.07	-.07	.01
12	.05	-.26	-.09	-.09	-.16	-.09	-.02	-.31	-.09	-.12	-.03	.16	-.22	-.10	.15
13	-.15	-.15	-.08	-.20	-.21	.11	-.17	-.28	-.10	-.25	-.17	.12	-.11	.12	.13
14	.02	-.11	-.01	-.08	-.06	.02	-.01	-.24	-.01	-.07	.01	.04	-.11	-.01	.09
15	-.32	-.10	-.17	-.22	-.29	.41	-.29	-.32	-.21	-.14	-.16	.12	-.12	.29	.19
16	-.13	-.24	-.13	-.17	-.25	.08	-.12	-.39	-.04	-.17	-.11	.08	-.25	.07	.17
17	-.01	-.19	-.07	-.12	-.15	-.01	-.12	-.37	-.09	-.09	-.06	.13	-.16	.02	.18
18	-.19	-.06	-.12	-.14	-.18	.25	-.23	-.36	-.30	-.08	-.06	.09	-.14	.19	.13
19	-.11	-.22	-.11	-.09	-.20	.12	-.18	-.22	-.13	-.17	-.10	.15	-.18	.07	.24
20	-.23	-.09	-.10	-.25	-.25	.29	-.25	-.29	-.17	-.26	-.22	.12	-.16	.12	.09

*Note.* Agg = not aggressive; Pop. = popular; Aff = affiliative; Aca = academic; Ovr = overall adjustment; Def = defiance; Eff = academic effort; Bel = belonging; Val = school value; Read = reading achievement; Math = math achievement; Lead = never a leader; Blyd = never bullied; Bllies = bullies peers; Like = not liked.

Table 13

*Final EFA Model*

<u>Item</u>	<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
1. Getting lost and not being able to find your way around school	<b>.70</b>	.01	-.02
2. Forgetting your locker combination	<b>.78</b>	-.06	.09
8. Not having the right books or supplies for class	<b>.38</b>	.20	.15
19. Having to change classes	<b>.66</b>	.23	.02
20. Understanding new rules	<b>.58</b>	.29	-.20
4. Having a tough teachers	.03	<b>.59</b>	.13
5. Having to do harder school work	.03	<b>.81</b>	-.01
9. Getting too much homework	-.04	<b>.69</b>	-.14
6. Finding kids I can sit with at lunch	.06	.23	<b>.47</b>
11. Not getting to see your friends from elementary school enough	-.20	.22	<b>.46</b>
12. Having trouble making new friends	.05	.08	<b>.79</b>
13. Kids trying to talk you into things you don't want to do	.29	.02	<b>.57</b>
14. Being bothered by older kids.	.21	-.11	<b>.63</b>
16. Other kids teasing you	.00	-.14	<b>.90</b>
17. Not being in the "in" group	-.10	.00	<b>.85</b>
<b>Factor Correlations</b>			
	<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
<u>Factor 1</u>			
<u>Factor 2</u>	.34		
<u>Factor 3</u>	.62	.40	

Table 14

*Final CFA Calibration and Validation Models*

<b>Factor</b>	<b>Item</b>	Calibration Unstandardized	Calibration Standardized	Validation Unstandardized	Validation Standardized
		Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
Procedural	1	1.00	.74 (.08)***	1.00	.64 (.08)***
	2	1.11 (.15)***	.82 (.07)***	1.06 (.17)***	.68 (.04)***
	8	.95 (.07)***	.70 (.05)***	1.26 (.19)***	.80 (.05)***
	19	.95 (.16)***	.70 (.06)***	.93 (.18)***	.59 (.07)***
	20	.83 (.12)***	.61 (.07)***	.95 (.16)***	.60 (.04)***
Academic	4	1.00	.61 (.05)***	1.00	.62 (.05)***
	5	1.39 (.15)***	.85 (.06)***	1.39 (.10)***	.86 (.04)***
	9	1.15 (.11)***	.70 (.07)***	1.25 (.10)***	.77 (.04)***
	6	1.00	.67 (.05)***	1.00	.74 (.03)***
	11	.73 (.08)***	.49 (.03)***	.56 (.06)***	.42 (.04)***
Social	12	1.22 (.14)***	.81 (.04)***	1.01 (.04)***	.82 (.04)***
	13	1.18 (.13)***	.79 (.05)***	1.11 (.09)***	.83 (.05)***
	14	1.13 (.10)***	.75 (.02)***	.74 (.05)***	.55 (.03)***
	16	1.16 (.09)***	.78 (.01)***	1.14 (.04)***	.85 (.02)***
	17	.97 (.12)***	.65 (.05)***	1.08 (.10)***	.80 (.06)***
<b>Factor Correlations</b>					
		Calibration		Validation	
<i>Procedural</i>		<i>Procedural</i>	<i>Academic</i>	<i>Procedural</i>	<i>Academic</i>
<i>Academic</i>		.48 (.07)***		.62 (.04)***	
<i>Social</i>		.70 (.05)***	.50 (.07)***	.71 (.05)***	.57 (.08)***
<b>R<sup>2</sup> Estimates</b>					
		Calibration		Validation	
Procedural	1	.55***		.40***	
	2	.68***		.46***	
	8	.38***		.67***	
	19	.49***		.35***	
	20	.37***		.36***	
Academic	4	.38***		.38***	
	5	.73***		.74***	
	9	.50***		.59***	
	6	.45***		.55***	
	11	.24***		.17***	
Social	12	.66***		.67***	
	13	.62***		.68***	
	14	.57***		.31***	
	16	.60***		.72***	
	17	.42***		.64***	

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$



Table 15

*Baseline Models Across Gender*

Factor	Item	Female	Male	Female	Male
		Estimate (SE)	Estimate (SE)	R <sup>2</sup>	R <sup>2</sup>
Procedural	1	1.00	1.00	.56	.46
	2	.95(.14)***	1.31 (.22)***	.60	.79
	8	.99 (.05)***	1.00 (.12)***	.55	.47
	19	.88 (.09)***	.76 (.18)***	.43	.27
	20	.90 (.05)***	.87 (.15)***	.45	.35
Academic	4	1.00	1.00	.52	.18
	5	1.22 (.07)***	2.10 (.55)***	.77	.81
	9	1.11 (.07)***	1.46 (.41)***	.64	.39
Social	6	1.00	1.00	.55	.34
	11	.68 (.04)***	.77 (.07)***	.25	.20
	12	1.14 (.06)***	1.44 (.20)***	.71	.71
	13	1.11 (.08)***	1.41 (.22)***	.67	.68
	14	1.08 (.04)***	1.08 (.21)***	.47	.39
	16	.97 (.09)***	1.32 (.18)***	.65	.59
	17	.83 (.10)***	.83 (.07)***	.50	.45
<b>Latent Factor Correlations</b>					
<i>Female</i>					
		<u>Procedural</u>	<u>Academic</u>		
<u>Procedural</u>					
<u>Academic</u>		.31 (.04)***			
<u>Social</u>		.37 (.06)***	.31 (.04)***		
<i>Male</i>					
		<u>Procedural</u>	<u>Academic</u>		
<u>Procedural</u>					
<u>Academic</u>		.10 (.03)**			
<u>Social</u>		.31 (.05)***	.11 (.05)**		

*Note.* Parameters are unstandardized.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Table 16

*Measurement Invariance Across Gender: Comparisons of Model Fit*

<b>Model</b>	<b><math>\chi^2 (df)</math></b>	<b><math>\Delta \chi^2 (df)</math></b>	<b>RMSEA [90 % CI]</b>	<b>CFI</b>	<b>TLI</b>
Configural	230.10 (174), $p = .003$	NA	.04 [.02 - .05]	.96	.95
Metric	241.63 (186), $p = .004$	18.93 (12), $p = .09$	.04 [.02 - .05]	.96	.96
Scalar	255.14 (202), $p = .007$	20.42 (16), $p = .20$	.03 [.01 - .05]	.96	.96
Strict	270.78 (211), $p = .003$	21.32 (9), $p = .01$	.03 [.02 - .05]	.96	.96
Residual Variances	279.03 (219), $p = .004$	28.80 (14), $p = .01$	.03 [.02 - .05]	.96	.96

Table 17

*Baseline Models Across Socioeconomic Status*

Factor	Item	Full Estimate (SE)	Free/Reduced Estimate (SE)	Full R <sup>2</sup>	Free/Reduced R <sup>2</sup>
Procedural	1	1.00	1.00	.47	.48
	2	.92 (.15)***	1.06 (.17)***	.40	.53
	8	.95 (.15)***	1.13 (.12)***	.42	.61
	19	.85 (.14)***	.93 (.09)***	.34	.42
	20	1.09 (.14)***	.63 (.10)***	.57	.19
Academic	4	1.00	1.00	.45	.35
	5	1.25 (.12)***	1.49 (.16)***	.71	.78
	9	1.12 (.12)***	1.21 (.17)***	.57	.51
Social	6	1.00	1.00	.56	.43
	11	.61 (.05)***	.73 (.15)***	.21	.23
	12	1.16 (.07)***	1.24 (.12)***	.75	.67
	13	1.04 (.09)***	1.21 (.12)***	.60	.63
	14	.87 (.09)***	.99 (.11)***	.42	.42
	16	1.09 (.06)***	1.15 (.11)***	.66	.57
	17	.78 (.08)***	1.16 (.11)***	.44	.58

**Latent Factor Correlations***Full Pay*

	<u>Procedural</u>	<u>Academic</u>
<u>Procedural</u>		
<u>Academic</u>	.29 (.02)***	
<u>Social</u>	.38 (.04)***	.28 (.05)***

*Free/Reduced Pay*

	<u>Procedural</u>	<u>Academic</u>
<u>Procedural</u>		
<u>Academic</u>	.18 (.03)***	
<u>Social</u>	.33 (.04)***	.19 (.04)***

*Note.* Parameters are unstandardized.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Table 18

*Measurement Invariance Across Socioeconomic Status: Comparisons of Model Fit*

<b>Model</b>	<b><math>\chi^2 (df)</math></b>	<b><math>\Delta \chi^2 (df)</math></b>	<b>RMSEA [90% CI]</b>	<b>CFI</b>	<b>TLI</b>
Configural	227.08 (174), $p = .004$	NA	.03 [.02, .05]	.97	.96
Metric	240.56 (186), $p = .004$	20.51 (12), $p = .06$	.04 [.02, .05]	.97	.96
Scalar	258.34 (202), $p = .005$	26.61 (16), $p = .05$	.03 [.02, .05]	.97	.97
Strict	267.09 (211), $p = .01$	15.79 (9), $p = .07$	.03 [.01, .05]	.97	.97
Residual Variances	278.00 (219), $p = .004$	28.79 (14), $p = .01$	.03 [.02, .05]	.97	.97

Table 19

*Baseline Models Across Race/Ethnicity*

Factor	Item	White	African American	White	African American
		Estimate (SE)	Estimate (SE)	R <sup>2</sup>	R <sup>2</sup>
Procedural	1	1.00	1.00	.56	.37
	2	1.15 (.21)***	1.06 (.16)***	.43	.41
	8	1.06 (.15)***	1.01 (.13)***	.51	.37
	19	.77 (.11)***	.91(.13)***	.26	.30
	20	.54 (.05)***	.62 (.18)***	.43	.14
Academic	4	1.00	1.00	.66	.41
	5	1.48 (.13)***	1.30 (.25)***	.26	.70
	9	1.38 (.17)***	1.15 (.25)***	.65	.55
Social	6	1.00	1.00	.47	.66
	11	.75 (.05)***	.56 (.07)***	.30	.21
	12	1.11 (.06)***	1.05 (.10)***	.65	.72
	13	1.00 (.11)***	1.11 (.15)***	.53	.81
	14	.88 (.04)***	.83 (.06)***	.41	.46
	16	1.06 (.05)***	1.13 (.09)***	.59	.83
	17	.96 (.05)***	.78 (.09)***	.49	.40
<b>Latent Factor Correlations</b>					
<i>White</i>					
		<u>Procedural</u>	<u>Academic</u>		
<u>Procedural</u>					
<u>Academic</u>		.23 (.11)***			
<u>Social</u>		.34 (.06)***	.25 (.06)***		
<i>African American</i>					
		<u>Procedural</u>	<u>Academic</u>		
<u>Procedural</u>					
<u>Academic</u>		.21 (.11)**			
<u>Social</u>		.36 (.14)***	.23 (.08)**		

*Note.* Parameters are unstandardized.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Table 20

*Partial Measurement Invariance Across Race/Ethnicity: Comparisons of Model Fit*

<b>Model</b>	<b><math>\chi^2</math> (df)</b>	<b><math>\Delta \chi^2</math> (df)</b>	<b>RMSEA [90% CI]</b>	<b>CFI</b>	<b>TLI</b>
Configural	207.97 (174), $p = .04$	NA	.03 [.00, .05]	.97	.96
Metric (Partial)	219.96 (184), $p = .04$	17.83 (10), $p = .06$	.03 [.01, .05]	.97	.96
Scalar	236.65 (200), $p = .04$	22.67 (16), $p = .12$	.03 [.01, .05]	.97	.97
Strict	245.86 (209), $p = .04$	13.80 (9), $p = .13$	.03 [.01, .05]	.97	.97
Residual Variances	256.51 (217), $p = .03$	27.30 (14), $p = .01$	.03 [.01, .05]	.97	.97

Table 21

*Relationships Between SCALE Subscales and Student Adjustment Measures*

<u>Measure</u>	<b>Concurrent Relationships (Fall)</b>			<b>Predictive Relationships (Spring)</b>		
	<u>Academic</u>	<u>Procedural</u>	<u>Social</u>	<u>Academic</u>	<u>Procedural</u>	<u>Social</u>
School Value	-.33***	-.14	-.14	-.20***	-.10*	-.09
Math Achievement	-.24***	-.21***	-.01	-.21***	-.24***	-.10*
Reading Achievement	-.15***	-.19*	-.10	-.23***	-.24**	-.14
Academic Effort	-.23***	-.17***	-.05	-.22***	-.17**	-.07
Academic	-.20***	-.17**	-.09	-.25***	-.24***	-.13**
Overall Adjustment	-.23***	-.22***	-.12*	-.23**	-.20*	-.19**
Belonging	-.32***	-.27***	-.49***	-.31***	-.22***	-.35***
Affiliative	-.20***	-.07	-.10	-.15**	-.06	-.08
Defiance	.18***	.12***	.03	.19***	.13**	.00
Not Aggressive	-.21***	-.16**	-.02	-.19**	-.12**	-.02
Bullies Peers	.15*	.12***	.02	.18*	.09	.03
Never a Leader	.17***	.22***	.11**	.19*	.18*	.22***
Not Liked	.17**	.21**	.24***	.15**	.10	.22***
Never Bullied	-.15*	-.19***	-.23***	-.16*	-.13	-.23***
Popular	-.11	-.14*	-.19***	-.02	-.14	-.30***

*Note.* Correlations are standardized. Ratings on the SCALE range from 1, *no problem*, to 4,

*large problem.*

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Table 22

*Relationships Between SCALE Subscales and Student Adjustment Measures by Gender*

Measure	Concurrent Relationships (Fall)			Predictive Relationships (Spring)		
	Academic	Procedural	Social	Academic	Procedural	Social
<i>Female</i>						
School Value	-.33***	-.15*	-.15	-.12*	-.10**	-.07
Math Achievement	-.28***	-.19***	.02	-.25***	-.08	-.04
Reading Achievement	-.15***	-.20*	-.08	-.24***	-.20*	-.08
Academic Effort	-.24***	-.23***	.00	-.27***	-.16***	-.05
Academic	-.31***	-.11*	-.06	-.29***	-.11	-.08
Overall Adjustment	-.28***	-.15***	-.10	-.26**	-.18**	-.11**
Belonging	-.31***	-.31***	-.49***	-.35***	-.22***	-.36***
Affiliative	-.19*	-.16**	-.08	-.20	-.20***	-.07
Defiance	.27***	.15***	.09	.14*	.04	.00
Not Aggressive	-.15**	-.17***	.04	-.18***	-.10*	.02
Bullies Peers	.11	.17***	.01	.13	.06	.06
Never a Leader	.19***	.19***	.05	.21***	.15**	.15**
Not Liked	.16***	.21***	.14	.10	.17**	.10
Never Bullied	-.15**	-.15***	-.02	-.17***	-.04	-.17***
Popular	-.05	-.03	-.10	.01	-.08	-.15**
<i>Male</i>						
School Value	-.34***	-.08*	-.06	-.28***	.00	-.07
Math Achievement	-.14**	-.19***	-.13**	-.16	-.28***	-.26***
Reading Achievement	-.11	-.24**	-.18**	-.21**	-.29**	-.22**
Academic Effort	-.22**	-.12	-.14	-.22*	-.15	-.10
Academic	-.01	-.18*	-.09	-.17***	-.27***	-.13
Overall Adjustment	-.17	-.28**	-.22*	-.20*	-.24*	-.26*
Belonging	-.33***	-.20*	-.49***	-.25***	-.19***	-.40***
Affiliative	-.29***	-.10	-.15*	-.26**	.00	-.12
Defiance	.25***	.06	-.05	.23*	.20**	.05
Not Aggressive	-.31***	-.23***	-.14	-.11	-.14*	-.09
Bullies Peers	.13	.08	.01	.23*	.07	.00
Never a Leader	.13	.27*	.22**	.22**	.14	.20
Not Liked	.08	.16*	.26***	.17*	.13	.32***
Never Bullied	-.06	-.11	-.27**	-.09	-.11	-.22***
Popular	-.10	-.25***	-.25*	-.08	-.24*	-.38***

*Note.* Correlations are standardized. Ratings on the SCALE range from 1, *no problem*, to 4,

*large problem*.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .



Table 23

*Relationships Between SCALE Subscales and Adjustment Measures by Socioeconomic Status*

Measure	Concurrent Relationships (Fall)			Predictive Relationships (Spring)		
	Academic	Procedural	Social	Academic	Procedural	Social
<i>Full-Pay Lunch Status</i>						
School Value	-.31***	-.19**	-.11	-.13	-.05	.00
Math Achievement	-.15*	-.23**	-.11	-.19***	-.26***	-.18
Reading Achievement	-.08	-.09*	-.20**	-.23***	-.18**	-.21**
Academic Effort	-.14***	-.23***	-.18***	-.13**	-.23**	-.04
Academic	-.03	-.08	-.08	-.05	-.20***	-.07
Overall Adjustment	-.17*	-.19***	-.25***	-.14*	-.26**	-.22**
Belonging	-.37***	-.34***	-.57***	-.29***	-.20***	-.36***
Affiliative	-.19**	-.11*	-.23***	-.15**	-.13	-.11*
Defiance	.17*	.17*	.12	.05	-.01	-.03
Not Aggressive	-.13**	-.11***	-.17***	-.04	-.14***	.01
Bullies Peers	.04	-.11***	.05	.06	.05	.09
Never a Leader	.28**	.34*	.32*	.17***	.25**	.18*
Not Liked	.15	.28***	.32***	.14**	.23***	.26**
Never Bullied	-.11	-.09	-.15	-.08*	-.18***	-.21**
Popular	-.12	-.17*	-.24**	-.04	-.24***	-.31**
<i>Free/Reduced Lunch Status</i>						
School Value	-.33***	.00	-.08	-.23***	-.08	-.15***
Math Achievement	-.27***	-.09**	-.02	-.22**	-.04	-.03
Reading Achievement	-.15*	-.15	-.05	-.26**	-.14*	-.09
Academic Effort	-.25***	-.10	.07	-.29*	-.08	-.04
Academic	-.25*	-.09*	.00	-.33**	-.09	-.09
Overall Adjustment	-.24*	-.13*	-.02	-.28	-.07	-.12
Belonging	-.24***	-.13	-.45***	-.33**	-.20***	-.34***
Affiliative	-.26**	-.05	-.02	-.31*	-.03	-.01
Defiance	.31	.16	.04	.23***	.18**	-.02
Not Aggressive	-.18*	-.09	.05	-.20	-.04	.02
Bullies Peers	.10	.18***	.00	.16	.04	-.04
Never a Leader	.24	.16**	.06	.24*	.02	.18
Not Liked	.08	.06*	.10	.13	-.05	.07
Never Bullied	-.07	-.03	-.14***	-.19	.02	-.19
Popular	.00	.08*	-.09	-.02	.06	-.19

*Note.* Correlations are standardized. Ratings on the SCALE range from 1, *no problem*, to 4,

*large problem*.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Table 24

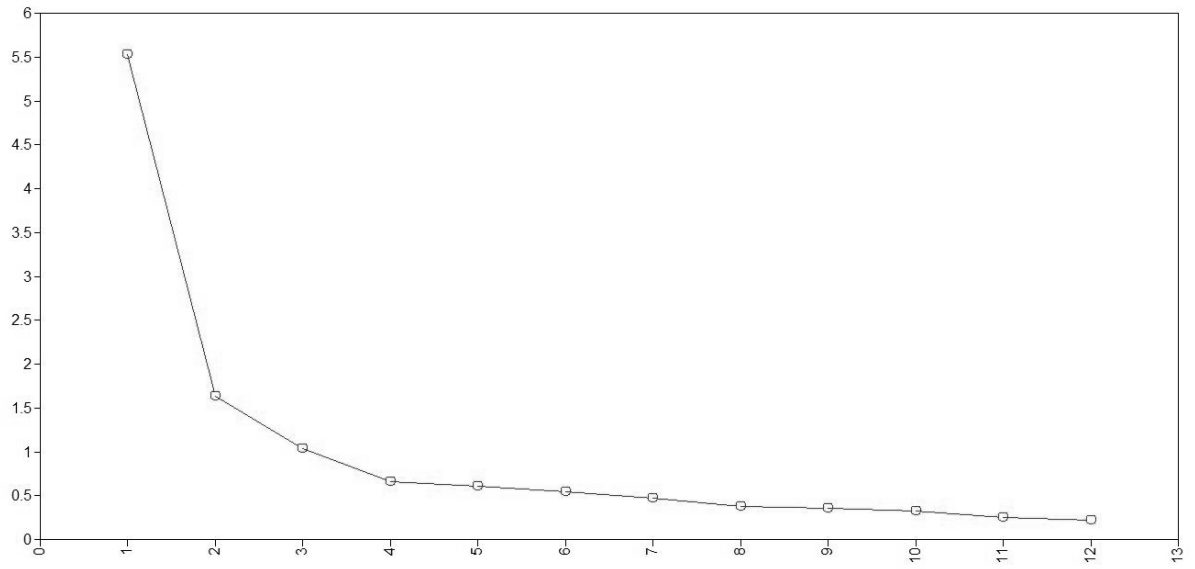
*Relationships Between SCALE Subscales and Student Adjustment Measures by Race/Ethnicity*

Measure	Concurrent Relationships (Fall)			Predictive Relationships (Spring)		
	Academic	Procedural	Social	Academic	Procedural	Social
<i>White</i>						
School Value	-.30***	-.24***	-.13**	-.15**	-.08*	-.02
Math Achievement	-.16***	-.14	-.11	-.20***	-.23*	-.20**
Reading Achievement	-.12***	-.09	-.16*	-.22***	-.19	-.18**
Academic Effort	-.20**	-.24***	-.24***	-.21**	-.29***	-.22**
Academic	-.19***	-.12*	-.16***	-.19**	-.19*	-.19***
Overall Adjustment	-.24**	-.22*	-.31***	-.23**	-.26*	-.33***
Belonging	-.48***	-.28***	-.55***	-.37***	-.23***	-.49***
Affiliative	-.22**	-.15	-.23***	-.27**	-.28*	-.27***
Defiance	.19*	.17*	.16*	.07	.08	.04
Not Aggressive	-.16	-.20**	-.20**	-.18*	-.24***	-.13
Bullies Peers	.11	.12	.11	.21	.11	.07
Never a Leader	.20***	.23*	.27**	.20***	.20**	.30**
Not Liked	.15*	.21**	.29***	.19**	.26*	.31***
Never Bullied	-.13	-.18	-.22*	-.21**	-.21***	-.31***
Popular	-.14**	-.10	-.22***	-.02	-.16	-.31***
<i>African American</i>						
School Value	-.34***	-.04	-.16	-.23***	-.06	-.21**
Math Achievement	-.30**	-.32*	-.13	-.04	-.05	-.15
Reading Achievement	-.05	-.19	-.14	-.20**	-.24	-.19
Academic Effort	-.10	-.22**	-.28**	-.07	-.12	.09
Academic	-.16	-.26*	-.14	-.13	-.20*	-.07
Overall Adjustment	-.12	-.33**	-.01	-.05	-.23**	-.21
Belonging	-.04	-.08	-.34***	-.10	-.18	.09
Affiliative	-.28*	-.35**	-.03	-.07	-.17**	-.04
Defiance	.26***	.15	-.24*	-.02	.11	-.08
Not Aggressive	.11	-.24*	.20***	.16	-.11	.30***
Bullies Peers	-.11	.25*	-.13	-.04	.22	-.07
Never a Leader	.16***	.26*	.02	.14	-.05	.09
Not Liked	.01	.24***	.12	-.07	.11	.17
Never Bullied	-.08	-.02	-.12	.05	.03	-.17
Popular	-.11	-.14	-.07	-.04	-.04	-.40***

*Note.* Correlations are standardized. Ratings on the SCALE range from 1, *no problem*, to 4,

*large problem*.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .



*Figure 1.* Scree plot including all SCALE variables. Eigenvalues are represented on the vertical axis and factors on the horizontal axis. Large changes in slope indicate the potential number of factors within the data (Tabachnick & Fidell, 2007).

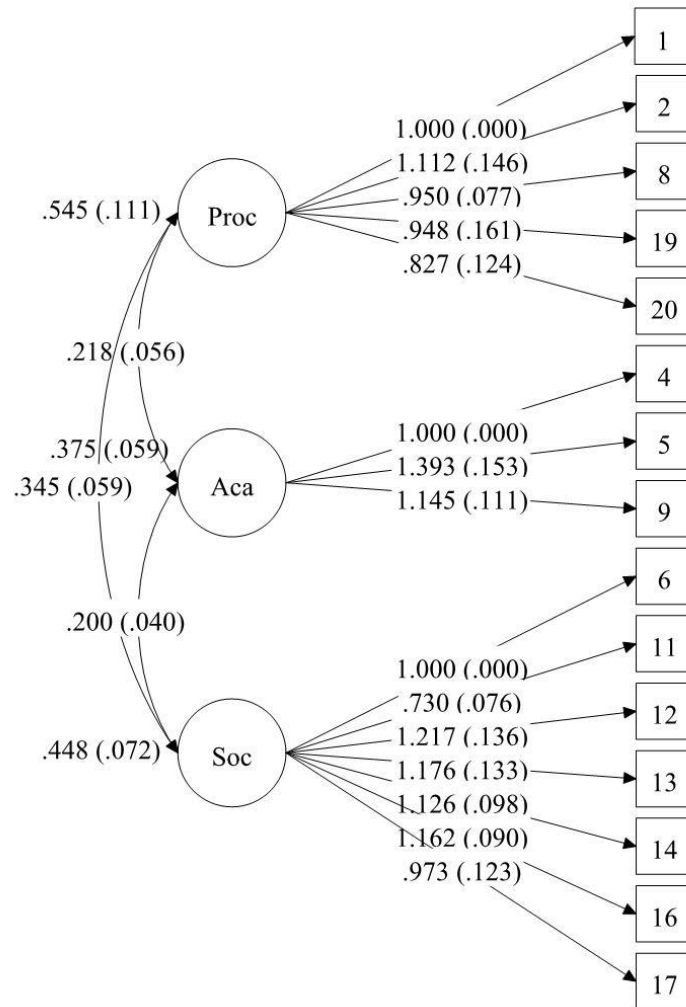


Figure 2. Final CFA calibration model for the SCALE. *Proc* represents the Procedural subscale, *Aca* represents the Academic subscale, and *Soc* represents the Social subscale. Parameters are unstandardized.

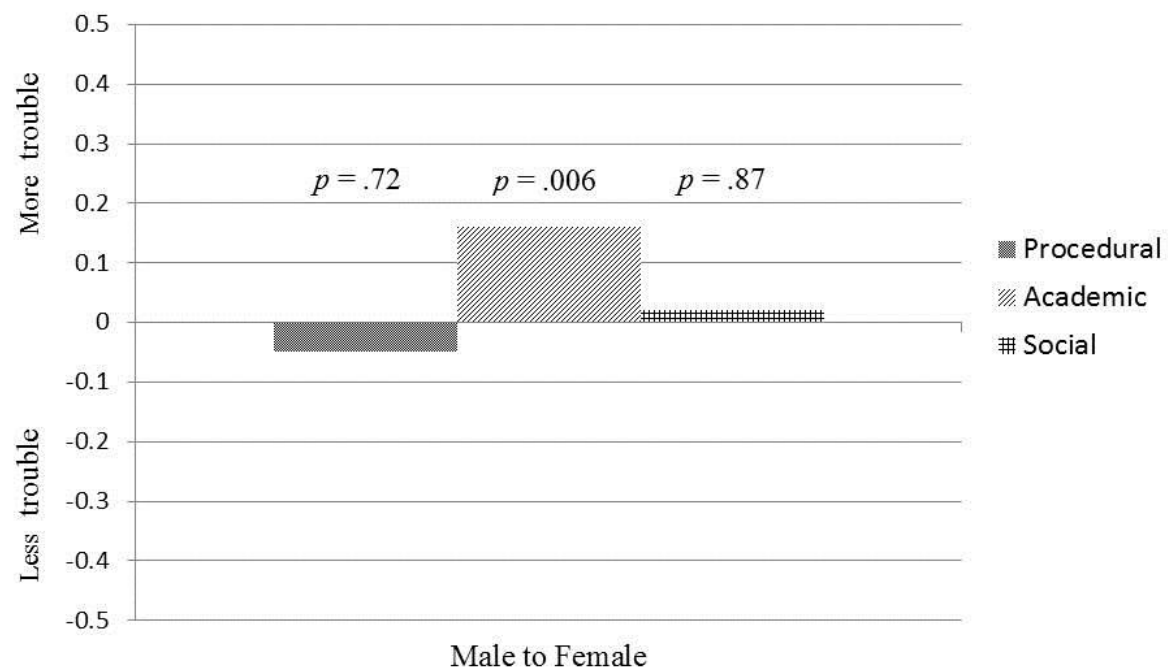


Figure 3. Gender differences on SCALE subscales in the fall of sixth grade. Females were held to zero and males' scores were compared to the female group.

## REFERENCES

- Akos, P. (2002). Student perceptions of the middle school transition from elementary to middle school. *Professional School Counseling, 5*, 229-345.
- Akos, P., & Galassi, J. P. (2004a). Gender and race as variables in psychosocial adjustment to middle and high school. *The Journal of Educational Research, 98*, 102-108. doi:10.3200/JOER/98.2.102-108
- Akos, P., & Galassi, J. P. (2004b). Middle and high school transitions as viewed by students, parents, and teachers. *Professional School Counseling, 7*, 212-221.
- Akos, P., Rose, R. A., & Orthner, D. (2014). Sociodemographic moderators of middle school transition effects on academic achievement. *The Journal of Early Adolescence*. Advance online publication. doi:10.1177/0272431614529367
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA/APA/NCME]. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anderman, L. H. (2003). Academic and social perceptions as predictors of change in middle school students' sense of school belonging. *Journal of Experimental Education, 72*, 5-22. doi:10.1080/00220970309600877
- Anderman, L. H., & Mueller, C. (2010). Middle school transitions and adolescent development. In J. Meece & J. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 198-215). New York: Routledge.
- Anderson, L. W., Jacobs, J., Schramm, S., & Splittgerber, F. (2000). School transitions: Beginning of the end or a new beginning? *International Journal of Educational Research 33*, 325-339. doi:10.1016/S0883-0355(00)00020-3
- Baker, J. A., Derrer, R. D., Davis, S. M., Dinklage-Travis, H. E., Linder, D. S., & Nicholson, M. D. (2001). The flip side of the coin: Understanding the school's contribution to dropout and completion. *School Psychology Quarterly, 16*, 406-426. doi:10.1521/scpq.16.4.406.19901
- Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist, 42*, 223-235. doi:10.1080/00461520701621079
- Baraldi, A. N. & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48*, 5-37. doi:10.1016/j.jsp.2009.10.001

- Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle school and high school. *Journal of Adolescent Research, 19*, 3-30. doi:10.1177/0743558403258113
- Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychologist, 37*, 197-214. doi: 10.1207/S15326985EP3704\_1
- Benner, A. D. (2011). The transition to high school: Current knowledge, future directions. *Educational Psychology Review, 23*, 299–328. doi:10.1007/s10648-011-9152-0
- Berkel, C., Murry, V. M., Hurt, T. R., Chen, Y. F., Brody, G. H., Simons, R.L., Cutrona, C., & Gibbons, F. X. (2009). It takes a village: Protecting rural African American youth in the context of racism. *Journal of Youth and Adolescence, 38*, 175-188. doi:10.1007/s10964-008-9346-z
- Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.
- Bovaird, J. A., & Koziol, N. A. (2012). Measurement models for ordered-categorical indicators. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp.495-511). New York: The Guilford Press.
- Bowen, N. K., & Guo, S. (2012). *Structural equation modeling. Pocket guides to social work research methods*. New York: Oxford University Press.
- Brewster, A. B., & Bowen, G. L. (2004). Teacher support and the school engagement of Latino middle and high school students at risk of school failure. *Child and Adolescent Social Work Journal, 21*, 47-67. doi:10.1023/B:CASW.0000012348.83939.6b
- Brody, G. H., Chen, Y., Murry, V. M., Ge, X., Simons, R. L., Gibbons, F. X., Gerrand, M., & Cutrona, C. E. (2006). Perceived discrimination and the adjustment of African American youths: A five-year longitudinal analysis with contextual moderation effects. *Child Development, 77*, 1170-1189. doi: 10.1111/j.1467-8624.2006.00927.x
- Bronfenbrenner, U. (1975, July). The ecology of human development in retrospect and prospect. Paper presented at the Conference on Ecological Factors in Human Development by the International Society for the Study of Behavioral Development, Guilford, England.
- Bronfenbrenner, U. (1994). Ecological models of human development. In *International encyclopedia of education* (Vol. 3, 2<sup>nd</sup> ed., pp. 1643-1647). Oxford, England: Elsevier.
- Burchinal, M. R., Roberts, J. E., Zeisel, S. A., & Rowley, S. (2008). Social risk and protective factors for African American children's academic achievement and adjustment during the transition to middle school. *Developmental Psychology, 44*, 286–292. doi:10.1037/0012-1649.44.1.286

- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20, 872-882.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge, Taylor, and Francis Group.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456-466. doi:10.1037/0033-2909.105.3.456
- Cabrera-Nyugen, P. (2010). Author guidelines for reporting scale development and validation results. *Journal of the Society for Social Work and Research*, 1, 99-103. doi:10.5243/jsswr.2010.8
- Cairns, R. B., Leung, M. C., Buchanan, L., & Cairns, B. D. (1995). Friendships and social networks in childhood and adolescence: Fluidity, reliability, and interrelations. *Child Development*, 66, 1330-1345. doi:10.1111/1467-8624.ep9510075266
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105. doi:10.1037/h0046016
- Cattell, R. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276. doi:10.1207/s15327906mbr0102\_10
- Cauley, K. M., & Jovanovich, D. (2006). Developing an effective transition program for students entering middle school or high school. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 80, 15-25. doi:10.3200/TCHS.80.1.15-25
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, 12, 471-492. doi:10.1207/s15328007sem1203\_7
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255. doi:10.1207/S15328007SEM0902\_5
- Cillessen, A. H. N., & Mayeux, L. (2007). Expectations and perceptions at school transitions: The role of peer status and aggression. *Journal of School Psychology*, 45, 567-586. doi:10.1016/j.jsp.2007.05.004
- Cizek, G. J. (2012). Defining and distinguishing validity: Interpretations of score meaning and justifications of test use. *Psychological Methods*, 17, 31-43. doi:10.1037/a0026975
- Cizek, G. J., Rosenberg, S. L., & Koons, H. H. (2008). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement*, 68, 397-412. doi:10.1177/0013164407310130



- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in scale development. *Psychological Assessment*, 7, 309-319. doi: 10.1037/1040-3590.7.3.309
- Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, 47, 291-314. doi: 10.1016/j.jsp.2009.04.002
- Collins, W. A., & Steinberg, L. (2006). Adolescent development in interpersonal context. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional and personality development* (Vol. 3, 6th ed., pp. 1003-1068). Hoboken, NJ: John Wiley and Sons.
- Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology*, 67, 648-657. doi:10.1037/0022-006X.67.5.648
- Cook, P. J., MacCoun, R., Muschkin, C., & Vigdor, J. (2008). The negative impacts of starting middle school in sixth grade. *Journal of Policy Analysis and Management*, 27, 104-121. doi:10.1002/pam.20369
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10, 1-9. Retrieved from <http://pareonline.net/getvn.asp?v=10&n=7>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. doi: 10.1007/BF02310555
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302. doi:10.1037/h0040957
- Davidson, A. J., Gest, S. D., & Welsh, J. A. (2010). Relatedness with teachers and peers during early adolescence: An integrated variable-oriented and person-oriented approach. *Journal of School Psychology*, 48, 483-510. doi:10.1016/j.jsp.2010.08.002
- Davis, H. A. (2006). Exploring the contexts of relationship quality between middle school students and teachers. *The Elementary School Journal*, 106, 193-223. doi:10.1086/501483
- Day, K. M., Hamm, J. V., Lambert, K., & Farmer, T. W. (2014). Social integration and students' perceptions of the transition to middle school. *Journal of Educational and Developmental Psychology*, 4(1), 10-26. doi:10.5539/jedp.v4n1p10
- Deemer, S., McCotter, S., & Smith, N. (2003, April). Sharing the responsibility: A university-school district investigation of the middle school transition. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

- DeVellis, R. F. (2012). *Scale development: theory and applications* (3rd ed). Thousand Oaks, CA: Sage Publications, Inc.
- Dong, Y., & Peng, C. J. (2013). Principles missing data methods for researchers. *SpringerPlus*, 2(222), 1-17. doi:10.1186/2193-1801-2-222
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37, 830-837. doi:10.1046/j.1365-2923.2003.01594.x
- Eccles, J. S. (2004). Schools, motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.). *Handbook of adolescent psychology* (2nd ed., pp. 125-153). Hoboken, NJ: John Wiley & Sons.
- Eccles, J. S. (2008). Agency and structure on human development. *Research in Human Development*, 5, 231-243. doi: 10.1080/15427600802493973
- Eccles, J. S., & Midgley, C. (1989). Stage–environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Eds.), *Research on motivation in education: Vol. 3. Goals and cognitions* (pp. 13–44). New York: Academic Press.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & MacIver, D. (1993). Development during adolescence: The impact of stage–environment fit on adolescents’ experiences in schools and families. *American Psychologist*, 48, 90-101. doi:10.1037/0003-066X.48.2.90
- Eccles, J. S., & Roeser, R. W. (2009). Schools, academic-motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 404-434). Hoboken, NJ: John Wiley & Sons, Inc. doi: 10.1002/9780470479193.adlpsy00101
- Eccles, J. S., & Roeser, R. W. (2011a). School and community influences on human development. In M. H. Bornstein & M. H. Lamb (Eds.), *Developmental science: An advanced textbook* (6th ed., pp. 571-643). New York: Psychology Press.
- Eccles, J. S., & Roeser, R. W. (2011b). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21, 225-241. doi: 10.1111/j.1532-7795.2010.00725.x
- Elder, G. H. (1998). The life course as developmental theory. *Child Development*, 69, 1-12. doi:10.1111/j.1467-8624.1998.tb06128.x
- Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 3-19). New York: Kluwer Academic Publishers.

- Elder, G. H., & Shanahan, M. J. (2006). The life course and human development. In R. M. Lerner & W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (6th ed., Vol. 1, pp. 665-715). Hoboken, NJ: John Wiley & Sons.
- Elias, M. J. (2001). Easing transitions with social-emotional learning. *Principal Leadership*, 1, 20-25.
- Elias, M. J., Ubriaco, M., Reese, A. M., Gara, M. A., Rothbaum, P. A., & Haviland, M. (1992). A measure of adaptation to problematic academic and interpersonal tasks of middle school. *Journal of School Psychology*, 30, 41-57. doi:10.1016/0022-4405(92)90019-2
- Ellerbock, C. R., & Kiefer, S. M. (2013). The interplay between adolescent needs and secondary school structures: Fostering developmentally responsive middle and high school environments across the transition. *The High School Journal*, 96, 170-194. doi:10.1353/hsj.2013.0007
- Entwisle, D. R., Alexander, K. L., & Olson, L. A. (2003). The first grade transition in life course perspective. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 229-250). New York: Kluwer Academic/Plenum.
- Erath, S. A., Flanagan, K. S., & Bierman, K. L. (2008). Early adolescent school adjustment: Associations with friendship and peer victimization. *Social Development*, 17, 853-870. doi:10.1111/j.1467-9507.2008.00458.x
- Espelage, D. L., & Swearer, S. M. (2003). Research on bullying and victimization: What have we learned and where do we go from here? *School Psychology Review*, 32, 365-383.
- Estell, D. B., Farmer, T. W., & Cairns, B. D. (2007). Bullies and victims in rural African American youth: Behavioral characteristics and social network placement. *Aggressive Behavior*, 33, 145-159. doi: 10.1002/ab.20176
- Farkas, G. (2003). Racial disparities and discrimination in education: What do we know, how do we know it, and what do we need to know? *Teachers College Record*, 105, 1119-1146. doi:10.1111/1467-9620.00279
- Farmer, T. W., Estell, D. B., Bishop, J. L., O'Neal, K. K., & Cairns, B. D. (2003). Rejected bullies or popular leaders? The social relations of aggressive subtypes of rural African American early adolescents. *Developmental Psychology*, 39, 992-1003. doi:10.1037/0012-1649.39.6.992
- Farmer, T. W., Estell, D. B., Leung, M. C., Trott, H., Bishop, J., & Cairns, B. D. (2003). Individual characteristics, early adolescent peer affiliations, and school dropout: An examination of aggressive and popular group types. *Journal of School Psychology*, 41, 217-232. doi:10.1016/S0022-4405(03)00046-3

- Farmer, T. W., Hall, C. M., Petrin, R. A., Hamm, J. V., & Dadisman, K. (2010). Evaluating the impact of a multicomponent intervention model on teachers' awareness of social networks at the beginning of middle school in rural communities. *School Psychology Quarterly*, 25, 94-106. doi:10.1037/a0020147
- Farmer, T. W., & Hamm, J. V. (2010). Project REAL: Teacher and student measures data dictionary. Unpublished manuscript.
- Farmer, T. W., Hamm, J. V., Lane, K. L., Lee, D., Sutherland, K. S., Hall, C. M., & Murray, R. A. (2013). Conceptual foundations and components of a contextual intervention to promote student engagement during early adolescence: The Supporting Early Adolescent Learning and Social Success (SEALS) model. *Journal of Educational and Psychological Consultation*, 23, 115-139. doi:10.1080/10474412.2013.785181
- Farmer, T. W., Hamm, J. V., Leung, M. C., Lambert, K., & Gravelle, M. (2011). Early adolescent peer ecologies in rural communities: Bullying in schools that do and do not have a transition during the middle grades. *Journal of Youth and Adolescence*, 40, 1106-1117. doi: 10.1007/s10964-011-9684-0
- Farmer, T. W., Hamm, J. V., Petrin, R. A., Robertson, D. R., Murray, R. A., Meece, J., & Brooks, D. S. (2010). Supporting early adolescent learning and social strengths: Promoting productive contexts for students at-risk for EBD during the transition to middle school. *Exceptionality*, 18, 94-106. doi:10.1080/09362831003673192
- Farmer, T. W., Irvin, M. J., Sgamatto, A. N., Dadisman, K., & Thompson, J. H. (2009). Interpersonal competence configurations in rural Appalachian fifth graders: Academic achievement and associated adjustment factors. *Elementary School Journal*, 109, 301-321. doi:10.1086/592309
- Farmer, T. W., Lines, M. M., & Hamm, J. V. (2011). Revealing the invisible hand: The role of teachers in children's peer experiences. (Introduction to the special issue). *Journal of Applied Developmental Psychology*, 32, 247-256. doi:10.1016/j.appdev.2011.04.006
- Farmer, T. W., Petrin, R., Brooks, D. S., Hamm, J. V., Lambert, K., & Gravelle, M. (2012). Bullying involvement and the school adjustment of rural students with and without disabilities. *Journal of Emotional and Behavioral Disorders*, 20, 19-37. doi:10.1177/1063426610392039
- Farmer, T. W., Xie, H., Cairns, B. D., & Hutchins, B. C. (2007). Social synchrony, peer networks, and aggression in school. In P. Hawley, T. Little, & P. C. Rodkin (Eds.), *Aggression and adaptation: The bright side to bad behavior* (209-234). Mahwah, NJ: Lawrence Erlbaum Associates.
- Fenning, P., & Rose, J. (2007). The overrepresentation of African American students in exclusionary discipline: The role of school policy. *Urban Education*, 42, 536-559. doi:10.1177/0042085907305039

- Finch, W. H. (2012). Modern methods for the detection of multivariate outliers. *Frontiers in Psychology*, 3(211), 1-12. doi:10.3389/fpsyg.2012.002115
- Finn, K. V., & Frone, M. R. (2004). Academic performance and cheating: Moderating role of school identification and self-efficacy. *Journal of Educational Research*, 97, 115-123. doi:10.3200/JOER.97.3.115-121
- Flora, D. B. & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9, 466-491. doi:10.1037/1082-989X.9.4.466
- Freeman, T. M., & Anderman, L. H. (2005). Changes in mastery goals in urban and rural middle school students. *Journal of Research in Rural Education*, 20, 1-12.
- Galván, A., Spatzier, A., & Juvonen, J. (2011). Perceived norms and social values to capture school culture in elementary and middle school. *Journal of Applied Developmental Psychology*, 32, 346-353. doi:10.1016/j.appdev.2011.08.005
- Garcia-Coll, C., Lamberty, G., Jenkins, R., McAdoo, H. P., Crnic, K., Wasik, B. H., & Garcia, H. V. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, 67, 1891-1914. doi: 10.2307/1131600
- 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, 625-636. doi: 10.1037/0012-1649.44.3.625
- GLSEN (2009). *The experiences of lesbian, gay, bisexual and transgender middle school students* (GLSEN Research Brief). New York: Gay, Lesbian and Straight Education Network.
- Graham, S., Bellmore, A. D., & Mize, J. (2006). Peer victimization, aggression, and their co-occurrence in middle school: Pathways to adjustment problems. *Journal of Abnormal Child Psychology*, 34, 363-378. doi:10.1007/s10802-006-9030-2
- Graham, S., Taylor, A. Z., & Hudley, C. (1998). Exploring achievement values among ethnic minority early adolescents. *Journal of Educational Psychology*, 90, 606-620. doi: 10.1037/0022-0663.90.4.606
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, 44, S78-S94. doi:10.1097/01.mlr.0000245454.12228.8f

- Gregory, A., & Ripski, M. B. (2008). Adolescent trust in teachers: Implications for behavior in the high school classroom. *School Psychology Review, 37*, 337-353.
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher, 39*, 59-68.  
doi:10.3102/0013189X09357621
- Hagborg, W. J. (1998). School membership among students with learning disabilities and nondisabled students in a semirural high school. *Psychology in the Schools, 35*, 183-188.  
doi:10.1002/(SICI)1520-6807(199804)35:2<183::AID-PITS10>3.0.CO;2-8
- Halpern-Felsher, B. (2009). Adolescent decision-making: An overview. *The Prevention Researcher, 2*, 3-7.
- Hamm, J. V., Dadisman, K. A., Day, K. M., Agger, C. A., & Farmer, T. W. (2014). The move to middle school: Parents' expectations and early adolescents' adjustment in rural communities. *Journal of Educational and Developmental Psychology, 4*, 46-65.  
doi:10.5539/jedp.v4n1p46
- Hamm, J. V., Farmer, T. W., Dadisman, K. A., Gravelle, M., & Murray, A. R. (2011). Teachers' attunement to students' peer group affiliations as a source of improved student experiences of the school social-affective context following the middle school transition. *Journal of Applied Developmental Psychology, 32*, 267-277.  
doi:10.1016/j.appdev.2010.06.003
- Hogland, W. L. G. (2007). School functioning in early adolescence: Gender-linked responses to peer victimization. *Journal of Educational Psychology, 99*, 683-699. doi: 0.1037/0022-0663.99.4.683
- Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity, 44*, 153-166. doi:10.1007/s11135-008-9190-y
- Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural equation modeling* (pp. 1- 15). Thousand Oaks, CA: Sage Publications.
- Hoyle, R. H. (2012). Model specification in structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp.181-194). New York: The Guilford Press.
- Huck, S. W. (2012). *Reading statistics and research* (6th ed.). Boston, MA: Pearson Education, Inc.
- Iacobucci, D. & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology, 13*, 478-487. doi:10.1207/S15327663JCP1304\_14

- IBM Corporation. (2013). IBM SPSS Statistics for Windows, Version 22.0 [Computer Software]. Armonk, NY: IBM Corp.
- Juvonen, J. (2007). Reforming middle schools: Focus on continuity, social connectedness, and engagement. *Educational Psychologist*, 42, 197-208. doi:10.1080/00461520701621046
- Juvonen, J., & Graham, S. (2014). Bullying in schools: The power of bullies and the plight of victims. *Annual Review of Psychology*, 65, 159-185. doi:10.1146/annurev-psych-010213-115030
- Juvonen, J., Le, V., Kaganoff, T., Augustine, C., & Constant, L. (2004). *Focus on the wonder years: Challenges facing the American middle school*. Santa Monica, CA: RAND.
- Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35, 401-415. doi:10.1007/BF02291817
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1-73. doi:10.1111/jedm.12000
- Karriker-Jaffe, K. J., Foshee, V. A., Ennett, S. T., & Suchindran, C. (2008). The development of aggression during adolescence: Sex differences in trajectories of physical and social aggression among youth in rural areas. *Journal of Abnormal Child Psychology*, 36, 1227-1236. doi:10.1007/s10802-008-9245-5
- Kerns, G. K. (2007) R function: KMO sampling adequacy. <https://stat.ethz.ch/pipermail/r-help/2007-August/138049.html>
- Kingery, J. N., & Erdley, C. A. (2007). Peer experiences as predictors of adjustment across the middle school transition. *Education and Treatment of Children*, 30, 73-88. doi:10.1353/etc.2007.0007
- Klem, A. M., & Connell, J. P (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, 74, 262-273. doi:10.1111/j.1746-1561.2004.tb08283.x
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: The Guilford Press.
- Knight, G. P., & Hill, N. E. (1998). Measurement equivalence in research involving minority adolescents. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: conceptual, methodological, and theoretical issues* (pp. 183-210). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75, 63-82. doi:10.3102/00346543075001063

- Lane, K. L., Pierson, M., & Givner, C. C. (2004). Secondary teachers' views on social competence: Skills essential for success. *Journal of Special Education, 38*, 174-186. doi:10.1177/00224669040380030401
- Lane, K. L., Wehby, J. H., & Cooley, C. (2006). Teacher expectations of students' classroom behavior across the grade span: Which social skills are necessary for success? *Exceptional Children, 72*, 153-167. doi: 10.1177/001440290607200202
- Lee, T., Cai, L., & MacCallum, R. C. (2012). Power analysis for tests of structural equation models. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp.181-194). New York: The Guilford Press.
- Lei, P-W., & Wu, Q. (2012). Estimation in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 164-179). New York, NY: Guilford Press.
- Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*, 184–196. doi:10.1037/0022-0663.97.2.184
- Little, R. J. A. (1998). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198 – 1202. doi:10.1080/01621459.1988.10478722
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149. doi: 10.1037/1082-989X.1.2.130
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99. doi:10.1037/1082-989X.4.1.84
- Malecki, C. K. & Demaray, M. K. (2006). Social support as a buffer in the relationship between socioeconomic status and academic performance. *School Psychology Quarterly, 21*, 375-395. doi:10.1037/h0084129
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30-46. doi:10.1037/1082-989X.1.1.30
- McMullen, R. C., Shippen, M. E., & Dangel, H. L. (2007). Middle school teachers' expectations of organizational behaviors of students with learning disabilities. *Journal of Instructional Psychology, 34*, 75-80.
- Meece, J., Glienke, B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology, 44*, 351-374. doi:10.1016/j.jsp.2006.04.004



- Messick, S. (1989). Validity. In R. L. Linn (Ed.) *Educational Measurement*, (3rd ed., pp.13–103). New York: Macmillan.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741–749. doi: 10.1037/0003-066X.50.9.741
- Midgley, C., & Edelin, K. C. (1998). Middle school reform and early adolescent well-being: The good news and the bad. *Educational Psychologist*, 33, 195-206. doi:10.1207/s15326985ep3304\_4
- Midgley, C., Maehr, M. L., Hruda, L., Anderman, E. M., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor: University of Michigan.
- Midgley, C., Middleton, M. J., Gheen, M. H., & Kumar, R. (2002). Stage-environment fit revisited: A goal theory approach to examining school transitions. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 109-142). Mahwah, NJ: Lawrence Erlbaum Associates.
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research* (Special Issue on Current Issues in Exploratory and Confirmatory Data Analysis), 3, 112-131.
- Millsap, R. E. (1997). Invariance in measurement and prediction: The relationship in the single-factor case. *Psychological Methods*, 2, 248-260. doi:10.1037/1082-989X.2.3.248
- Millsap, R. E. (2007). Invariance in measurement and prediction revisited. *Psychometrika*, 72, 461-473. doi:10.1007/S11336-007-9039-7
- Millsap, R. E. & Kwok, O. M. (2004). Evaluating the impact of partial invariance on selection in two populations. *Psychological Methods*, 9, 93-115. doi:10.1037/1082-989X.9.1.93
- Millsap, R. E., & Olivere-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 380-392). New York: The Guilford Press.
- Millsap, R. E., & Yun-Tien, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39, 479-515. doi:10.1207/S15327906MBR3903\_4
- 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, 13-40. doi:10.1177/0272431610384487

- Multon, K. (2010). Test–retest reliability. In N. Salkind (Ed.), *Encyclopedia of research design* (pp. 1496-1499). Thousand Oaks, CA: SAGE Publications, Inc.  
doi:10.4135/9781412961288.n457
- Muthén, B. O. (1998 – 2004). *Mplus Technical Appendices*. Los Angeles: Muthén & Muthén.
- Muthén, B., & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. Mplus Web Notes, No. 4, Version 5.  
Retrieved from <http://www.statmodel.com/download/webnotes/CatMGLong.pdf>
- Muthén, B. O., Du Toit, S. H. C., Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous variables. Retrieved from  
[http://www.statmodel.com/bmuthen/articles/Article\\_075.pdf](http://www.statmodel.com/bmuthen/articles/Article_075.pdf)
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Muthén, L. K., and Muthén, B. O. (2013). *Version 7.1 Mplus Language Addendum*. Retrieved from <https://www.statmodel.com/download/Version7.1xLanguage.pdf>
- Nansel T. R, Haynie, D. L., & Simons-Morton, B. G. (2007). The association of bullying and victimization with middle school adjustment. In J. E. Zins, M. J. Elias, & C. A. Maher, (Eds.), *Bullying and peer harassment: A handbook of prevention and intervention* (pp. 49-59). New York: Hawthorne Press.
- Neal, J. W., Cappella, E., Wagner, C., & Atkins, M. S. (2011). Seeing eye to eye: Predicting teacher-student agreement on classroom social networks. *Social Development*, 20, 376-393. doi:10.1111/j.1467-9507.2010.00582.x
- Nichols, D. P. (1998). Choosing an intraclass correlation coefficient. *SPSS Keywords*, 67.  
Retrieved from <http://www.utexas.edu/its/rc/answers/spss/spss4.html>.
- Niehaus, K., Rudasill, K. M., & Rakes, C. R. (2012). A longitudinal study of school connectedness and academic outcomes across sixth grade. *Journal of School Psychology*, 50, 443-460. doi:10.1016/j.jsp.2012.03.002
- No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).
- O'Connell, A. A. (2010). An illustration of multilevel models for ordinal response data. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society*. Proceedings of the Eighth International Conference on Teaching Statistics, Ljubljana, Slovenia. Voorburg: The Netherlands: International Statistical Institute.  
[www.stat.auckland.ac.nz/~iase/publications.php](http://www.stat.auckland.ac.nz/~iase/publications.php)

- O'Leary-Kelly, S. W. and Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of Operations Management*, 16, 387-405. doi:10.1016/S0272-6963(98)00020-5
- Pallas, A. M. (2003). Educational transitions, trajectories, and pathways. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 165-182). New York: Kluwer Academic Publishers.
- Parault, S., Davis, H. A., & Pellegrini, A. (2007). Exploring the social contexts of bullying and victimization. *Journal of Early Adolescence*, 27, 1-30. doi:10.1177/0272431606294831
- Parker, J. G., Rubin, K. H., Erath, S. A., Wojskawowicz, J. C., & Buskirk, A. A. (2006). Peer relationships, child development, and adjustment: A developmental psychopathology perspective. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol. 1: Theory and method* (2nd ed., 419-491). Hoboken, NJ: John Wiley & Sons.
- Pearl, R., Leung, M. C., Van Acker, R., Farmer, T. W., & Rodkin, P. C. (2007). Fourth- and fifth-grade teachers' awareness of their classrooms' social networks. *The Elementary School Journal*, 108, 25-39. doi:10.1086/522384
- Pellegrini, A. D. (2002). Bullying, victimization, and sexual harassment during the transition to middle school. *Educational Psychologist*, 37, 151-163. doi:10.1207/S15326985EP3703\_2
- Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. *Behavior Genetics*, 32, 153-161. doi:10.1023/A:1015210025234
- R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Randall, J., & Engelhard, G. (2009). Differences between teachers' grading practices in elementary and middle Schools. *The Journal of Educational Research*, 102, 175-186, doi:10.3200/JOER.102.3.175-186
- Resnick, M. D. (2000). Protective factors, resiliency, and healthy youth development. *Adolescent Medicine: State of the Art Reviews*, 11, 157-164.
- Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2000). Heterogeneity of popular boys: Antisocial and prosocial configurations. *Developmental Psychology*, 36, 14-24. doi:10.1037/0012-1649.36.1.14
- Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2006). They're cool: Social status and peer group supports for aggressive boys and girls. *Social Development*, 15, 175-204. doi:10.1046/j.1467-9507.2006.00336.x

- Rodkin, P. C., & Ryan, A. M. (2012). Child and adolescent peer relations in educational context. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA education psychology handbook. Vol. 2: Individual differences and cultural and contextual factors* (pp. 363–389). Washington, DC: American Psychological Association. doi:10.1037/13274-015
- Roeser, R. W., Galloway, M., Casey-Cannon, S., Watson, C., Keller, L., & Tan, E. (2008). Identity representations in patterns of school achievement and well-being among early adolescent girls: Variable- and person-centered approaches. *Journal of Early Adolescence*, 28, 115-152. doi:10.1177/0272431607308676
- Rubin, K. H., Bukowski, W., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional and personality development* (6th ed., Vol. 3, pp. 571-645). Hoboken, NJ: John Wiley and Sons.
- Rudolph, K. D., Lambert, S. F., Clark, A. G., & Kurlakowsky, K. D. (2001). Negotiating the transition to middle school: The role of self-regulatory processes. *Child Development*, 72, 929-946. doi: 10.1111/1467-8624.00325
- Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, 38, 437-460. doi:10.3102/00028312038002437
- Ryan, A. M., Shim, S. S., & Makara, K. A. (2013). Changes in academic achievement and relational self-worth across the transition to middle school. *Journal of Youth and Adolescence*, 42, 1372-1384. doi:10.1007/s10964-013-9984-7
- Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, 29, 347-363. doi:10.1177/0734282911406661
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29, 304-321, doi:10.1177/0734282911406653
- Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). Classroom and school influences. In *Motivation in education: Theory, research, and applications* (3rd ed., pp. 334-358). Upper Saddle River, NJ: Pearson Education, Inc.
- Sheng, Y. & Sheng, Z. (2012). Is coefficient alpha robust to non-normal data? *Frontiers in Psychology*, 3(34), 1-13. doi:10.3389/fpsyg.2012.00034
- Shields, J. P., Cohen, R., Glassman, J. R., Whitaker, K., Franks, H., & Bertolini, I. (2013). Estimating population size and demographic characteristics of Lesbian, Gay, Bisexual, and Transgender youth in middle school. *Journal of Adolescent Health*, 52, 248-250. doi:10.1016/j.jadohealth.2012.06.016

- Shoukri, M. M., Asyali, M. H., & Donner, A. (2004). Sample size requirements for the design of a reliability study: Review and new results. *Statistical Methods in Medical Research*, 13, 251-271. doi:10.1191/0962280204sm365ra
- Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research*, 7, 301-317. doi:10.1191/096228098672090967
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 36, 420-428. doi:10.1037/0033-2909.86.2.420
- Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: sources of racial and gender disproportionality in school punishment. *The Urban Review*, 34, 317-342. doi: 10.1023/A:1021320817372
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80, 99-103. doi:10.1207/S15327752JPA8001\_18
- Sugai, G., & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Behavior Psychology in the Schools*, 24, 23-50.
- Suldo, S. M., Friedrich, A. A., White, T., Farmer, J., Minch, D., & Michalowski, J. (2009). Teacher support and adolescents' subjective well-being: A mixed-methods investigation. *School Psychology Review*, 38, 68-85.
- Swanson, D. P., Spencer, M. B., Harpalani, V., Dupree, D., Noll, E., Ginzburg, S., & Seaton, G. (2003). Psychosocial development in racially and ethnically diverse youth: Conceptual and methodological challenges in the 21st century. *Development and Psychopathology*, 15, 711-743. doi:10.1017.S0954579403000361
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn and Bacon.
- Theriot, M. T., & Dupper, D. R. (2009). Student discipline problems and the transition from elementary to middle school. *Education and Urban Society*, 42, 205-222. doi:10.1177/0013124509349583
- Turner, J. C., Gray, D. L., Anderman, L. H., Dawson, H. S., & Anderman, E. M. (2013). Getting to know my teacher: Does the relation between perceived mastery goal structures and perceived teacher support change across the school year? *Contemporary Educational Psychology*, 38, 316-327. doi:10.1016/j.cedpsych.2013.06.003
- Turner, J. C., & Patrick, H. (2004). Motivational influences on student participation in classroom learning activities. *Teachers College Record*, 106, 1759-1785. doi:10.1111/j.1467-9620.2004.00404.x

- Urdan, T., & Midgley, C. (2003). Changes in the perceived classroom goal structure and pattern of adaptive learning during early adolescence. *Contemporary Educational Psychology*, 28, 524–551. doi:10.1016/S0361-476X(02)00060-7
- Vandenburg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-70, doi:10.1177/109442810031002
- Veronneau, M., & Dishion, T. J. (2011). Middle school friendships and academic achievement in early adolescence: A longitudinal analysis. *Journal of Early Adolescence*, 31, 99-124. doi:10.1177/0272431610384485
- Voelkl, K. E. (1996). Measuring students' identification with school. *Educational and Psychological Measurement*, 56, 760-770. doi:10.1177/0013164496056005003
- Wang, J., Iannotti, R., & Nansel, T. (2009). School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *Journal of Adolescent Health*, 45, 368–375. doi:10.1016/j.jadohealth.2009.03.021
- Wang, M., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, 47, 633-662. doi:10.3102/0002831209361209
- Ware, W. B., Ferron, J. M., & Miller, B. M. (2013). *Introductory statistics: A conceptual approach using R*. New York, NY: Routledge.
- Way, N., Reddy, R., & Rhodes, J. (2007). Students' perceptions of school climate during the middle school years: Associations with trajectories of psychological and behavioral adjustment. *American Journal of Community Psychology*, 40, 194-213. doi:10.1007/s10464-007-9143-y
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19, 231-240.
- Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology*, 85, 357-364. doi:10.1037/0022-0663.85.2.357
- Wentzel, K. R., & Asher, S. R. (1995). The academic lives of neglected, rejected, popular, and controversial children. *Child Development*, 66, 754-763. doi:10.2307/1131948
- 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, 195-203. doi:10.1037/0022-0663.96.2.195

- 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. doi:10.1016/j.cedpsych.2010.03.002
- Wentzel, K. R., & Caldwell, K. A. (1997). Friendships, peer acceptance, and group membership: Relations to academic achievement in middle school. *Child Development*, 68, 1198-1209. doi:10.2307/1132301
- West, S. G., & Finch, J. F. (1997). Personality measurement: Reliability and validity issues. In R. Hogan, B. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 143-165). San Diego, CA: Academic Press.
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209-231). New York: The Guilford Press.
- Wienke Tatura, C. M., Green, A. E., Karver, M. S., & Gesten, E. L., (2009). Multiple informants in the assessment of psychological, behavioral, and academic correlates of bullying and victimization in middle school. *Journal of Adolescence*, 32, 193-211. doi:10.1016/j.adolescence.2008.04.055
- Wigfield, A., Byrnes, J. B., & Eccles, J. S. (2006). Adolescent development. In P. A. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 87-113). Mahwah, NJ: Erlbaum.
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional and personality development* (Vol. 3, 6th ed., pp. 933-1002). Hoboken, NJ: John Wiley and Sons.
- Wigfield, A., Lutz, S. L., & Wagner, A. L. (2005). Early adolescents' development across the middle school years: Implications for school counselors. *Professional School Counseling*, 9, 112-119.
- Witherspoon, D., & Ennett, S. (2011). Stability and change in rural youth's educational outcomes through the middle and high school years. *Journal of Youth and Adolescence*, 40, 1077-1090. doi:10.1007/s10964-010-9614-6
- Wray-Lake, L., Crouter, A. C., & McHale, S. M. (2010). Developmental patterns in decision-making autonomy across middle childhood and adolescence: European American parents' perspectives. *Child Development*, 81, 636-651. doi:10.1111/j.1467-8624.2009.01420.x
- Xie, H., Dawes, M., Wurster, T. J., & Shi, B. (2013). Aggression, academic behaviors, and popularity perceptions among boys of color during the transition to middle school. *American Journal of Orthopsychiatry*, 83, 265-277. doi:10.1111/ajop.12039

- Yang, Y. (2008). *Partial invariance in loadings and intercepts---their interplays and implications for latent mean comparisons*. (Doctoral dissertation, The University of Nebraska - Lincoln). *ProQuest Dissertations and Theses*, 201. Retrieved from <http://search.proquest.com/docview/304530413?accountid=14244>. (304530413).
- Yoon, J. S., Barton, E., & Taiariol, J. (2004). Relational aggression in middle school: Educational implications of developmental research. *The Journal of Early Adolescence*, 24, 303-318. doi:10.1177/0272431604265681
- Zimmer-Gembeck, M. J., & Collins, W. A. (2003). Autonomy development during adolescence. In G. R. Adams & M. Berzonsky (Eds.), *Blackwell handbook of adolescence* (pp. 175–204). Oxford, UK: Blackwell.
- Zimmerman, M. A., Caldwell, C. H., & Bernat, D. H. (2002). Discrepancy between self-report and school-record grade point average: Correlates with psychosocial outcomes among African American adolescents. *Journal of Applied Social Psychology*, 32, 86-109. doi:10.1111/j.1559-1816.2002.tb01421.x