University of Massachusetts Amherst

ScholarWorks@UMass Amherst

Doctoral Dissertations

Dissertations and Theses

November 2017

Social-Emotional Development Assessment: Scale Development for Kindergarten through Second Grade Youth Universal Screening

James F.M. Brenchley

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Educational Assessment, Evaluation, and Research Commons, School Psychology Commons, and the Student Counseling and Personnel Services Commons

Social-Emotional Development Assessment:

Scale Development for Kindergarten through Second Grade Youth Universal Screening

A Dissertation Presented

by

JAMES F. M. BRENCHLEY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2017

College of Education

Social-Emotional Development Assessment:

Scale Development for Kindergarten through Second Grade Youth Universal Screening

A Dissertation Presented

By

JAMES F. M. BRENCHLEY

Approved as to style and content by:

_____
Sara A. Whitcomb, Chair

_____
John M. Hintze, Member

_____
Craig S. Wells, Member

_____
Gloria T. DiFulvio, Member

                          _____
                          Joseph B. Berger, Senior Associate Dean
                          College of Education

Kim. I am so blessed to have her as my partner. She provides me with great balance and

support, but also is able to challenge me to accomplish more than I ever think I am able.

ABSTRACT

SOCIAL-EMOTIONAL DEVELOPMENT ASSESSMENT:

SCALE DEVELOPMENT FOR KINDERGARTEN THROUGH SECOND GRADE

YOUTH UNIVERSAL SCREENING

SEPTEMBER 2017

JAMES F.M. BRENCHLEY, B.A., LE MOYNE COLLEGE

M.Ed., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Sara Whitcomb

The goal of this project was to address a significant gap in the research literature with regard to available screening tools that allow young children to self-identify needs related to their social-emotional development. A review of current evidence-based social-emotional tools led to the identification of seven domains most frequently utilized: self-regulation, emotional regulation, social skills, self-concept, school connectedness, social responsibility, and optimism/positivity. To accomplish this endeavor, two studies were conducted to develop a screening measure that demonstrated adequate psychometric properties, but also minimized cost related to time for implementation. The first study was a review of 105 pilot scale items by kindergarten through second grade teachers ($n = 12$). Teachers ranked items based on importance that students demonstrate these skills at this developmental stage and also the readability of items. Rankings were then compared to that of the principal investigator. A balance of items from the seven domains with the highest rankings were taken to develop a 30-item pilot survey. The second study was a pilot of a self-report survey completed by kindergarten through second grade students (n

= 384) from two different districts in the Northeast. This survey was completed by students via group administration. Special consideration was given to the survey design to maximize the likelihood that students would remain engaged and provide reliable information. Students were presented with visuals and additional prompts to aid in administration of the tool. Classic item analysis approaches found one item that was an outlier and was removed from analysis. The remaining 29 items were reviewed by an exploratory factor analysis. It was found that this scale presents with unidimensionality and explained 30.5% of the variance. Items were then compared utilizing a graded response model of item response theory. Following this review, 12 items were identified for future research from five of the seven original domains (self-regulation, social skills, school connectedness, social responsibility, and optimism). This scale will require future assessments to further validate the measure, but marks an important step in the potential development of a scale which allows young students to self-report social-emotional needs and receive early intervention supports.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

STATEMENT OF THE PROBLEM

According the Collaborative for Academic, Social, and Emotional Learning (CASEL), the passing of the Every Student Succeeds Act (ESSA) in December 2015 could substantially increase the availability of evidence-based and "result-driven" instruction in the areas of social-emotional learning for students (CASEL, 2017a). ESSA broadens the understanding of student success to allow for the inclusion of indicators beyond traditional academic scores. The impact of this federal policy could have significant implications for the field of social-emotional learning and the types of programs available.

Current research on social-emotional learning indicates that certain programs, when implemented correctly, can have dramatic effects in areas of mental health and general well-being for youth. Social-emotional learning has been defined as "the process through which we learn to recognize and manage emotions, care about others, make good decisions, behave ethically and responsibly, develop positive relationships, and avoid negative behavior" (Zins, Weissberg, Wang, & Walberg, 2004, p. 4). A recent meta-analysis of 213 studies on social-emotional learning found that programs in this field have been shown to be effective both in and outside of school settings, effective for students from many different racial, ethnic, and economic backgrounds, and across grade levels (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). In addition, these researchers found that not only do social-emotional programs lead to improved social-emotional skills, improved pro-social behavior, connectedness, and reduction of undesirable behavior, but the acquisition of these skills can also lead to improved academic achievement and improved acceptance by peers and their teachers.

Despite the promising results of these programs, their availability and the structure through which they are implemented seems to indicate there are still significant gaps in the identification and service delivery of programs in the area of social-emotional development. Greenberg et al. (2003) estimated that 20 percent of youth are experiencing some form of psychological problem and of these students nearly 80 percent do not receive the appropriate services and supports to treat or prevent its occurrence. Public health researchers have shown that the majority of time in research and prevention work is spent on intervention and education; however, the largest effects in change come from policy and infrastructure building. Yet, these types of prevention strategies receive the least amount of time in terms of research and practice (Frieden, 2010). As social-emotional learning interventions continue to develop, having better assessment tools to understand how programs are effective and to build service delivery models will be paramount.

Through the years, the expected role of schools in the identification and treatment of student mental health concerns is one that has continued to grow. Schools are not just seen as academic centers for learning, but are also charged with providing education and support in the area of social, emotional, and behavioral growth for the youngest members of our community. School Psychologist training and ethical guidelines place an emphasis on the ability to adequately screen and link appropriate interventions in not only academics, but also in areas of behavior and social-emotional development (Ysseldyke, et al., 2006). To meet the demand of providing free and appropriate services to all youth, schools are beginning to shift to a tiered model of identification and service delivery for academic and behavior programs. Prevention programs and interventions are categorized

as universal (delivered to all students), selective (targeted programs delivered to a small, at-risk subset of the population, approximately 5-15%), or indicated (intense, specific programs often delivered individually to students of highest risk, less than 5 percent).

Within this model, there is a need not only for programs and curricula that fit within these levels, but for assessment tools that can be used for both screening at early levels and diagnostic purposes at later levels. Assuming instruction is effective for most students, an appropriate screening tool within this model would be one that can be administered on a universal level and accurately identifies a subset of the population that are not making effective progress with the universal instruction as compared to a majority of their developmentally equivalent peers. Screening tools, unlike more in-depth diagnostic instruments, should be practical and easily administered. Administration and scoring of these assessments needs to be relatively brief and should be an instrument that classroom teachers could administer. However, in the area of social-emotional development, few universal and feasibly administered screening tools exist.

Recent social-emotional assessment reviews have identified over 100 instruments used for the screening and diagnosing of mental health disorders and social- emotional functioning (Humphrey et al., 2011; Williams, 2008). In reviewing these measures researchers identified significant gaps in early child and primary-aged screening tools, options for youth self-reports, tools designed for screening as opposed to diagnostic assessment, and assessments that measure strengths in addition to deficits.

These systematic searches revealed a significantly higher number of screening and assessment tools for adolescents than pre-adolescent children. Considering the fundamental goal of screening and early intervention, it is important to recognize many

mental health disorders are already beginning to or have emerged during primary grades (Kessler et al., 2005). In addition, self-reports are often neglected in screening and assessment tools until later childhood or adolescence. This is most likely due to the difficulty of obtaining a reliable measure from younger children on social-emotional development and the costs associated with this type of screening (Kamphaus, 2012).

However, the absence of self-reports in assessments for youth creates a different systemic problem. Internalizing disorders are often overlooked in education settings. It is much easier for teachers and staff to observe challenging externalizing behaviors than it is to observe the presence of more internalizing concerns (Walker & Severson, 1992). A system that does not afford youth the opportunity to self-reflect and give their impression to staff around their current functioning may be a contributing cause to the large number of underserved youth. Systematic self-report screening would also assist support staff in identifying youth who have socially and emotionally developed as typically developing children should, but are currently experiencing a temporary setback and may benefit from temporary supports. In sum, without obtaining direct reports from youth, a sizable portion of data is missing to make appropriate service decisions during critical developmental periods.

The aforementioned reviews also found that many current assessment tools are impractical for schools to use as universal screeners. In their search, they found that the total time for completion of instruments ranged from ten minutes to more than an hour for each child. When factoring in additional time for scoring, this would be extremely time consuming to conduct assessments with these methods at a universal level for one benchmark period, let alone as a systematic data collection procedure with multiple

benchmarks within schools. Ettelson and Laurent (2002) note that multi-stage screenings tend to be a cost-effective and time-saving way to deliver services in a number of areas addressed in schools. However, there is a need for a screening tool that can be used to help streamline this process, and allow for these more thorough diagnostic tools to be used for youth who have the potential of being at elevated risk.

Another important consideration of selecting an appropriate assessment tool for systematic screening would be the population with which it was normed. In particular, when considering behaviors associated with appropriate social-emotional development, it is important that both the questions and the sample its norms are based on are reflective of the overall social and cultural norms of that community. In the review by Humphrey and colleagues (2011), they found that normative data in diverse samples were missing for most measures, including such factors as race, socio-economic status, and urbanism, which may impact the likelihood of a child being identified as at-risk.

Lastly, many of the items within instruments that are used to assess social-emotional development often focus on deficits as opposed to the presence of strengths. When considering how to best target interventions, how to evaluate which components of interventions have been successful, and how to develop appropriate treatment plans, there are great benefits to having a more strength-based approach to assessment. Researchers have shown that through a strength-based approach, the focus on the positive is more likely to increase motivation for the child to grow in these strengths, gaps in scores are seen as skills to learn as opposed to deficiencies, and they increase client involvement, among other benefits (Epstein, Dakan, Oswald, & Yoe, 2001). However, a majority of instruments reviewed reflected a model that focuses on the weaknesses of youth for

identification as opposed to the instructional areas of strength to build upon (Humphrey et al., 2011; Williams, 2008).

In response to the low reliability and predictive validity of youth self-reports, some researchers have begun to address the problem of self-report reliability of younger children. Watkins (2008) reviewed various formats for self-reports in youth to address the concern of under identification of youth with more internalizing concerns. Watkins surveyed teachers regarding the types of symbols and forms of expression children were most familiar with and then compared various self-report choices for youth such as thumbs up/thumbs down methods, happy and sad faces, glass full or empty, and so on. Watkins then used the method that led to the highest acceptance by teachers and had the best potential for reliability among youth to implement a self-report.

Although the instrument Watkins developed did not lead to concurrent validity with teacher measures, moderate reliability was found for kindergarten and 1st grade youth. The format was particularly reliable at measuring items that were more behaviorally-anchored and featured somatic complaints than items that were emotional in nature. This may indicate that developmentally, children this age do have the potential to be consistent, accurate reporters when asking about behavior as opposed to emotional and internal states. This suggests that an effective screening tool that examines whether behaviors are associated with positive social-emotional development would have the highest likelihood of reliable responses for youth.

The purpose of this study was to develop a list of behaviors associated with domain areas of social-emotional development that kindergarten through second grade teachers in various settings endorse as being developmentally appropriate for the youth

they work with. These behaviors were then used to develop a screening tool that was administered to kindergarten through second grade youth in multiple settings. These data were analyzed to determine whether there are items that are reliably endorsed and predictive of appropriate social-emotional development. The final product aims to be a strength-based, social emotional, universal screening tool that is practical, takes into account diversity in its sample, is developmentally appropriate in the behaviors it reflects, and can be reliably administered to youth at a critical time in social-emotional development. Specifically the overall research questions of this study were:

1. Do primary grade teachers agree upon the domains and behaviors associated with positive social-emotional development?

2. Of the newly developed measure based on teacher ratings, what factors emerge as being predictive of the overall construct of social-emotional competency across all kindergarten-second grade youth?

3. What is the internal consistency of the newly developed measure of children's social-emotional competence?

4. Which pilot items explain the greatest amount of the variance based on an individual's overall proficiency score?

5. Are there significant differences in ratings of children from various grades and districts?

CHAPTER II

REVIEW OF THE LITERATURE

In recent decades, the school's role has expanded to include children's interpersonal and life skill development in addition to being a center for academic skill growth. This has led to a greater need for research and policy development related to how schools can best address this need, such as the field of social-emotional learning. This chapter will begin with an introduction to the history of social-emotional learning and the benefits that have been seen in its programming. Next, this chapter will discuss Response to Intervention (RTI), a useful framework for guiding the implementation of social-emotional learning. When applied to social-emotional learning, this framework highlights the need for measurement tools that enable early identification of students at risk for negative mental health outcomes. This chapter will then consider current social-emotional learning screening instruments and their limitations. The domains most frequently assessed by these instruments will be identified and discussed in order to gain an operational understanding of each domain. This chapter will conclude with the proposal of a new measurement tool that can fill a critical void in the implementation structure of social-emotional learning in elementary school settings.

Social-Emotional Learning

Social-Emotional Learning (SEL) began to quickly emerge as a field of research in the 1990s. In 1994, the Collaborative for Academic, Social, and Emotional Learning (CASEL) was formed. The goal of this collaborative was to advance the science and policy of SEL to ensure proper implementation of programming for children from preschool through high school (CASEL, 2017b). Although many definitions of SEL and understandings of SEL are noted, one widely used definition describes SEL as "the

process through which we learn to recognize and manage emotions, care about others, make good decisions, behave ethically and responsibly, develop positive relationships, and avoid negative behavior" (Zins et al., 2004, p. 4). Unlike academic instruction, many believe social and emotional skills are not something that are seen as requiring direct instruction, but can occur naturally over time. However, a core tenant in this field is that although the perception is that these skills occur organically over time, these are also skills that can and should be explicitly taught and reinforced. Weissberg, Durlak, Domitrovich, and Gullotta (2015) note that in its development, CASEL included the word "learning" with "social and emotional" because the development of these skills is a process and schools can serve as a primary place where this learning can take place.

In 1997, CASEL partnered with the Association for Supervision and Curriculum Development (ACSD) and published *Promoting Social and Emotional Learning: Guidelines for Educators* (CASEL, 2017b). This was the first publication to provide a framework for the implementation of SEL programming within schools and addressed preschool through 12th grade. Over the next 20 years, SEL saw a dramatic increase in the volume of policy development, programs developed, and research being conducted. In 2004, Illinois became the first state to develop learning standards related to SEL for preschool through high school-aged students. By 2015, all 50 states had SEL standards for at least preschool-aged children (Weissberg et al., 2015).

Research has demonstrated positive effects of SEL programs both in behavioral and academic outcomes. SEL programs have been effective at reducing disruptive behaviors and increasing classroom engagement through targeted interventions for at-risk students (O'Connor, McCormick, Cappella, McClowry, & Society for Research on

Educational Effectiveness, 2014). Universal SEL approaches have been shown to reduce aggressive behaviors and increase social competence during a three year longitudinal study (Bierman et al., 2010). When comparing SEL to a control condition, academic gains in early reading skills have been demonstrated (Ashdown & Bernard, 2012). Moreover, a study conducted in Israel found that incorporating SEL as a part of literature instruction led to significant increases in content knowledge. In this study, the researchers interwove story and character lessons. This integrated approach increased the students' motivation to learn, relationship development (cohesion), and positive behaviors in these classrooms as compared to a control condition (Shechtman & Abu Yaman, 2012).

Benefits of SEL approaches appear to extend beyond the humanities. A three-year study of 24 Mid-Atlantic schools found a relationship between teachers' use of an SEL classroom strategy and students' overall achievement on mathematic assessments (Griggs, Rimm-Kaufman, Merritt, & Patton, 2013). Jones, Brown, and Aber (2011) conducted a two-year study on schools that implemented a combined SEL and literacy intervention. They not only found gains in areas related to literacy, but also significant improvements in mathematic scores for these students as compared to their peers.

Another important finding in the study by Jones and colleagues was a reduction in symptomology related to Attention Deficit Hyperactivity Disorder (ADHD) and depression for students who had undergone the SEL and literacy intervention. The idea that SEL programming could potentially lead to better mental health outcomes has been a major factor behind the push for increased SEL programming in schools. Mental illness rates are a growing concern in the United States, and school-aged children are no exception. The most recent Centers for Disease Control (CDC) surveillance report found

that from 2005 to 2011 there was an increase in the rates of mental health disorders identified for adolescents. Approximately 13-20 percent of children in the United States experience a mental disorder in a given year, a rate that has consistently been increasing since 1994 (CDC, 2013). This concerning trend underscores the need for early intervention strategies. The predominant approach of intervening when mental health symptoms are already present appears to be ineffective and ultimately an unsustainable model given these trends.

The more proactive approach of early SEL intervention has benefits that can be seen in a recent review of suicide prevention programs. Researchers found that a SEL universal program, The Good Behavior Game (GBG), was an especially effective prevention tool for decreasing suicide ideation and the likelihood and age of onset of suicide attempts (Wilcox et al., 2008, Katz et al., 2013). In fact, in a grading system by the Oxford Centre for Evidence-Based Medicine Criteria, GBG received the highest grade among programs that have been shown to have positive outcomes in regards to suicide ideation and prevention (Katz et al., 2013). The GBG was a classroom management program that divided the classrooms into teams with points given when a team exhibited negative behaviors during a specific period. The team with the lowest point total for a set period of time would win some form of reinforcement. This early SEL intervention had a greater impact on ideation and attempt rates for early adults than did programs that directly targeted suicide prevention for adolescents (i.e., Sources of Strength; CARE [Care, Assess, Respond, Empower]/CAST [Coping and Support Training]). A recent study found that a potential mechanism for the success of GBG to be positive childhood peer formation, particularly for the more disruptive, aggressive

students (Newcomer et al., 2015). These researchers hypothesized that due to difficulty with social adaption, these students are at greater risk of being "non-preferred," or a student others would endorse as not liking. By cultivating this protective factor while students are in first and second grade, it may have the potential to have a larger impact downstream than more universal and targeted interventions that occur later in development, during adolescent years.

Research has focused not only on the effectiveness of SEL interventions; but on the best practices for their implementation. CASEL recommends that in order for SEL programs to maximize their efficacy that schools and researchers prescribe to the SAFE (sequenced, active, focused, and explicit) approach. With this approach, success at improving the social and emotional health of students is maximized when programs have a planned progression and are connected (sequenced), are designed to maintain engagement and energy from students (active), program components are clear in their emphasis (focused), and curricula targets specific areas or subskills (explicit). This approach challenges the notion that effective SEL gains can be made organically and without thoughtful planning. In a meta-analysis of 213 SEL universal programs, Durlak and colleagues (2011) found that adherence to the SAFE criteria served as an important factor for the overall efficacy of an SEL intervention. When interventions met all four SAFE criteria, they were more likely to be successful.

Although SEL interventions appear to be promising, public health researchers have shown that the largest effects in change come from policy and infrastructure building. Yet, these macro-level prevention efforts often receive the least amount of time in terms of research and practice (Frieden, 2010). One macro-level framework for

prevention, Response to Intervention (RTI), began as a way to address the unsustainable rate of growth in number of students requiring special education instruction. Rather than wait until students fall behind, RTI utilizes universal screening to identify students at risk. Similarly, a Multi-Tiered System of Support (MTSS) approach like RTI could be a fruitful avenue to coordinate the efforts of SEL intervention planning and to meet the needs of all students. Rather than wait until students experience adverse outcomes such as poorer mental health, universal screening approaches for SEL could identify students at risk.

Response to Intervention

RTI could be a useful framework to guide the structure of SEL within schools. Although RTI methods have recently expanded to include behavioral components, early in its development, RTI methods focused primarily on prevention of academic failure. The National Center on Response to Intervention (NCRTI) lists the four components of RTI: a school-wide, multi-level instructional and behavioral system for preventing school failure; screening; progress monitoring; and data-based decision making for instruction, movement between levels, and disability identification (NCRTI, 2010). A key feature of RTI is the use of dynamic assessment, which includes formative assessment practices that guide core instruction for all students. In contrast, the traditional model has focused on more summative assessment approaches and has been described as a "wait-to-fail" approach.

RTI saw a very rapid rise in its use and research following the most recent reauthorization of Individuals with Disabilities Education Act (IDEA) in 2004. One of the most significant changes in IDEA regarded the procedures of identification of

students with learning disabilities. The law reflected a response to concerns related to over-identification of students with disabilities and a "wait-to-fail" model within special education. Within a traditional model, students received special education support and remediation only after they received a low standardized achievement score – most commonly below the 16th percentile – on a formal, standardized assessment. The level of achievement would also need to represent a significant discrepancy from the student's underlying cognitive abilities, most often assessed separately by an intelligence test. This "wait-to-fail" model also was associated with large disparities with regard to prevalence rates among states, socioeconomic groups, and race/ethnicities. Some argued that the system led to funds being disproportionately applied to middle class children with "dubious disability" identification as opposed to students from lower socioeconomic groups that lacked a discrepancy score, but were low achieving (Fuchs and Fuchs, 2006).

In addition to identification practices, the law also permitted local educational agency special education departments to use up to 15 percent of their budget towards early intervention programming. Many districts were struggling to meet special education needs prior to the reauthorization of IDEA. In particular, districts with underperforming students were faced with steep costs related to special education supports and disproportionality with regard to levels of identification among various demographics. At the time, learning disabilities (LD) had been on the rise in terms of the percentage of individuals identified as LD since its first inclusion as a disability area in 1975 (U.S. Office of Education, 1977). The percentage of children identified as having a Specific Learning Disability had tripled from 2% in 1976 to 6% in 2000. Following policy and the rise in the use of RTI principles within academics, this number began to level off and has

shown some evidence of decreasing. A National Center for Learning Disabilities study (2011) found the national percentage of students with an LD to be 5.2%. Additionally, IDEA special education service utilization data showed that rates of special education use for LD decreased by approximately 19% from 2004 to 2013 (U.S. Department of Education, NCES, 2016). Thus, an RTI model became an alternative for the process of identifying students requiring additional supports and potential referrals for special education supports.

Early intervention is key to the RTI approach. Early researchers noted students who did not learn to read as early as first grade were at significant risk of remaining weak readers throughout elementary school (Juel, 1988). This was termed the "Matthew Effect" by Stanovich (1986), whereby students with foundational reading skills would continue to engage and seek out opportunities for growth in reading abilities, whereas struggling readers would tend to become avoidant and lose opportunities for reading growth. Fuchs and Fuchs (1998) similarly noted the dual discrepancy that exists in struggling readers whereby they start out at a lower point than their peers, but also learn at a slower rate. This leads to a larger gap between the struggling reader and the average peer as time goes on. Thus, the RTI model's tenets of early identification and intervention are crucial in this context.

Various multi-tiered models exist for an RTI approach to learning, though a common conceptualization is the three-tiered model adopted from public health research. This breaks service delivery into three tiers: universal, targeted/secondary, intensive/tertiary. Within this model, all students are exposed to the preventive and proactive curriculum and supports. SEL supports at this tier would include things such as

school-wide prevention curriculum or approaches to promote a positive school climate. Through regular screenings, students are identified when they begin to struggle and receive additional supports and instruction in specific areas of need. These students also receive an increase in the frequency of progress monitoring to determine the effectiveness of that intervention approach. Data are used to drive the decision as to the effectiveness of the initial intervention and whether a different or more intensive intervention is needed. Students in this tier may receive a specialized skills training, such as social skills, in a small group over a fixed period. The last step is more intensive, individualized intervention approaches for students who have undergone several less intensive interventions and have failed to close the learning gap. These students may require supports such as individualized counseling or even special education supports. Researchers propose that for costs related to time and resources, this model is predicated on getting to a place where approximately 80 percent of a population is adequately served by universal approaches, 15 percent by targeted, tier two supports, and only 5 percent require more intensive interventions (Sugai & Horner, 2006).

As previously mentioned, dynamic assessment is an important feature of RTI. Both the American Psychological Association (APA) and National Association of School Psychologists (NASP) have called for improved social-emotional assessment tools that can work within a positive behavioral support or behavioral RTI model. In the APA Task Force on Evidence-Based Practice for Children and Adolescents (2008), a call was made for clinicians to be utilizing assessments that inform diagnosis, treatment planning, and outcome goals. The assessments must be able to provide accurate feedback as to the appropriate treatment for a youth and whether an evidence-based program is effectively

meeting the needs of that youth. Similarly, the NASP Practice Model (NASP, 2017)

domains place as much emphasis, if not more, on assessment related to social and life

skills as these domains do with regard to academic achievement. Schools' primary

responsibility is the development of cognitive and academic skills, but effective learning

is influenced significantly beyond what occurs within the school setting during academic

instruction. If students' basic health – both physical and mental – is not adequately

addressed, then optimal learning cannot occur. Thus, school psychologists aren't charged

with simply treating conditions, but with promoting wellness and resiliency. In order to

accomplish this objective, there needs to be improved tools that allow for the early

identification of students at-risk for adverse mental health outcomes that would benefit

from SEL intervention. RTI provides a useful framework for this early identification

through universal screening.

Universal Screening

One form of assessment that has been described as an essential first step to initiate

and direct early intervention supports and services is universal screening (Severson,

Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). Within an RTI framework,

universal screenings are brief assessments that focus on a specific skill, yet are strongly

associated with later future outcomes (Jenkins, 2003). Fuchs and Fuchs (2006) caution

that universal screening is meant to measure a portion of your population that is

"potentially" at-risk for later poorer outcomes. These are not necessarily direct measures

of an overall learning outcome, but identification of students that are at heightened risk.

This can lead to false positives, or students who are classified as at-risk, but are in fact on

an average trajectory for skills development. The alternative is false negatives, or

17

students that are at heightened risk, but are missed by the screening and not identified as needing additional support services.

The concept of universal screening in schools for areas of social-emotional needs is not a new phenomenon. In their paper reviewing an 11-item, teacher-report screening tool, Cowen (1973) referenced the prior decade as a shift from "repairing rooted dysfunction" to developing interventions that look to "prevent disorder and build health." This philosophic shift and importance being placed on greater support and review with the "young child" and the child's early social institution (school), is now over 50 years in the making. However, with an emphasis on early prevention and screening tools being developed a half-century ago, why don't more schools participate in regular screenings for warning signs or resiliency skill needs during early childhood?

One main reason why there has not been more progress made in this research area is that the resources for conducting a thorough screening often can be too daunting for schools with finite resources. Glovers and Albers (2007) proposed that screening tools be evaluated more so in terms of their practicality than other psychometric properties. They argued that those selecting tools should also consider cost, feasibility of test administration and data analysis, ability to generate stakeholder buy-in, appropriateness for the intended sample, and the ability of the data to inform treatment and prevention efforts.

Kamphaus (2012) focused on the single aspect of screening cost. With regard to the hour it takes to administer a student self-report and the various professionals involved, the estimated cost in terms of labor was $1,680. However, this researcher further delved into this figure as it only references the one time cost of the screening

itself. When professional time is considered in the selecting of the screening tool, training of staff to collaborate and coordinate the screening, training of staff to administer the measures, obtaining passive or active parental consent, collection of materials, cleaning and entering data, running the analysis and comparing and merging any prior datasets, and the follow-up meetings with stakeholders to review data and make programming decisions, the costs can add up to be quite a significant investment on the part of a school district.

Although challenges related to universal screening need to be further addressed, current approaches such as utilizing record review data seem to be ineffective at accurately identifying students at risk for mental health concerns. Kuo, Vander Stoep, Herting, Grupp, and McCauley (2013) researched the use of various systematic record reviews to identify students at heightened risk. They derived algorithms based on grades, attendance, suspensions, and basic demographic information and compared it to the screening data they had obtained for middle school students. When using the best model derived from this information, there was still a false positive rate of over 50% and an estimated 50-75% of students showing early signs of depression that were not identified. From a purely economic model, this common approach also comes with negative resource ramifications, even when schools implement this less systematically. There are significant intervention costs with any targeted intervention with regard to staff time, training, materials, and coordinating with teachers and parents. The study by Kuo and colleagues indicates that record review approaches, even when done systematically, can lead to false positives at a rate of over 20%. This leads to intervention supports and staffing being directed to students who aren't at-risk. Additionally, students are being

19

missed through this approach that may later develop more severe needs such as depressive symptoms that can lead to more clinical and long-term mental health conditions.

Despite some initial costs and challenges, universal screening with self-reports have been conducted successfully with students that are in late elementary and middle school. Vander Stoep and colleagues (2005) conducted a study on implementing a universal screening of all incoming middle school students. Items were taken from the Mood and Feeling Questionnaire (MFQ; Angold, Costello, Messer, & Pickles, 1995) and the Youth Self-Report (YSR; Achenbach & Rescorla, 2001). Of the sixth graders screened, 15% were noted as having an elevated score and were referred for a follow-up assessment. Of those students, 71% were noted as experiencing "significant distress." Staff then worked with students and families to link the student to school counselors, tutors, or after-school programs.

Eklund and colleagues (2009) compared the method of universal screening versus traditional teacher-referral approaches to identify students as needing additional social-emotional supports. Their study compared third and fourth grade students who were indicated based on the use of a screening instrument with a teacher report and traditional methods (student referred to child-study teams, students referred for special education testing, or receipt of special education or building level support services and counseling). Their study found that traditional methods only accounted for just under half of the students that were found to be at-risk based on the universal screening measure. These researchers concluded that universal screening enhanced the potential for early intervention as there was increased sensitivity than more traditional referral methods.

The most effective screening tools are also practical and simple because they can easily be used for repeated use over time. McCarty and colleagues (2013) found that tracking patterns in growth related to depressive symptoms and conduct was more predictive of later substance abuse than any point-in-time datum level measured. These researchers found that these symptoms are quite dynamic over time in adolescence and thus the need for multiple screenings and tracking of student progress over time becomes a critical step in effectively using universal screenings as a prevention tool.

In regard to the most effective screening tools in the field of SEL, Humphrey and colleagues (2011) and Williams (2008) reviewed hundreds of assessment tools based on systematic search criteria and later filtering of results with different, but specific, inclusion and exclusion criteria. They generated tables and reviews of the top instruments in this field that have had extensive research with regard to their reliability and validity. As part of their review, Humphrey and colleagues had concluded that one area of particular gap with current tools available was in early childhood, self-report measures. With all the efforts to push towards earlier intervention for youth, and the internalizing nature of many mental health disorders, this void is particularly troubling. However, it may be in part due to the difficulty with having children serve as self-reporters of emotional health.

Youth Self-Reports

Research has been mixed on the validity and utility of obtaining self-report information from elementary school aged children. However, this is a period of time that is incredibly important in emotional and behavioral development. Developmental researchers have referred to the ages of 6-12 as middle childhood (Collins, 1984).

Authors summarize the research in this period by noting that although there is continuity to the developmental processes that were occurring in early periods, this is the period in which children begin to consolidate their default responses to regularly occurring stimuli. Children's ability to monitor their own behavior and engage in personal goal setting significantly increases during this period (Brown et al, 1983). Researchers noted school as the primary vehicle for knowledge gains to occur in both academics and behavior, and also cultural norms and values (Epps & Smith, 1984). However, as important as this time is, Watkins (2008) noted there are very few tools available to reliably obtain this information from youth. This is despite Reynolds' (1992) recommendation that internalizing disorders are best measured through self-report and interviews with children due to the subjective and indirect nature of behavior for outside observers (parents, teachers).

However, concerns have been noted with regard to the reliability of children to serve as self-reports of their behavior. Research has shown instability in self-report measures such as those aiming to measure self-competence for younger children. Jacobs, Lanza, Osgood, Eccles, and Wigfield (2002) found in longitudinal research that children's reports of self-competence don't remain constant, but levels decrease on average as children move to later elementary school ages. Some have theorized that this is in part due to children having an inflated sense of self-abilities at younger ages and thus aren't the best self-raters. Parsons and Ruble (1977) conducted a study where even when given an impossible task, young children would persist and still predict success. Dweck and Elliott (1983) found that young children can have high expectations even in the face of failure. Research has also found that children can have difficulty with self-

ratings of emotions, as there are concerns related to the stability of these items as younger children can be more state driven than trait driven in their response style (Watkins, 2008). Thus, some have concluded that children are not accurate perceivers nor self-reporters of their skills during this time in development.

More recently there has been some challenge to this perspective as more research suggests that children as young as preschool age are able to engage in accurate self-reflection activities on items that are behavioral in nature (Muller & Kerns, 2015). Others have also pointed out that people in general tend to be overly positive in self-ratings (Alicke & Govorun, 2005) and better measures need to be made to control for this phenomenon when considering the positive response style of children (Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007).

Studies have shown significant associations between children's self-reports of internalizing concerns as early as the first grade and their mental health in adolescence. Ialongo and colleagues used a modified version of the Children's Depression Inventory (CDI) in a longitudinal study and found that children's self-ratings via semi-structured interviews were significant predictors of academic functioning, use of mental health services, suicidal ideation, and diagnosis of major depressive disorder by age 14 (Ialongo, Edelsohn, Wethamer-Larsson, Crockett, & Kellam, 1995; Ialongo, Edelsohn, & Kellam, 2001). Similarly, these researchers found relationships between self-reported anxiety and later academic function and anxious symptomology. These studies demonstrate strong potential related to youths being able to serve as valid self-reporters, but their approach makes it difficult for replication as these data were collected via a semi-structured

interview and limits the ease and practicality for use as a consistent universal screening tool.

In the medical world, emerging research is also demonstrating that although prior approaches indicated that eight was the recommend age to reliably begin giving self-report measures in pediatrics, it is now being recommended, with adjustments to the scales given, that children as young as five can be reliable and valid raters on a health-related quality of life measure (Varni, Limbers, & Burwinkle, 2007). One of the necessary shifts came from research on likert scale use with this population. In this study they limited the number of response options to three instead of five for older children and adults. This was based on Weng's (2004) research that found fewer response categories led to lower test-retest reliability. However, scales where all responses are clearly defined increase test-retest reliability as opposed to where there are just anchor labels, making it more advantageous on their scale to limit the response choices to three for children at that developmental period.

Despite concerns related to the reliability of children as self-reporters, it is clear that current approaches are lacking. Although there have been increases in social-emotional programming and policy in schools, the number of individuals experiencing mental health disorders is still increasing (CDC, 2013). In addition, approximately half of all mental health illness cases begin by the age of fourteen (Kessler, Chiu, Demler, & Walters, 2005). Therefore, despite risks related to false positives, it appears greater efforts are needed to identify ways of allowing students to indicate when they are beginning to experience forms of mental health duress, as it is clear students are being missed.

In addition, there is a growing body of research showing that when students are more involved in the identification and treatment planning phases, the more impactful the intervention. A recent study on the implementation of the Second Step curriculum found that adherence to the set script was not a significant predictor of the overall effect of the program on positive outcome measures, but student engagement in the lesson did have a significant effect when comparing classes with low versus high engagement (Low, Smolkowski, & Cook, 2016). These researchers make the argument that competency or quality of delivery is just as important as content. This followed a study conducted by Reyes, Brackett, Rivers, Elbertson, and Salovey (2012), where analyses found significant differences in outcomes based on program teachers' quality of implementation (low-quality vs. moderate- or high-quality). Connecting with students and their own engagement in the lessons can be a significant predictor of program success.

Research has shown that mental health treatment has improved results when steps to build engagement, such as motivational interviewing, are implemented during the pre-treatment phase. Dean, Britt, Bell, Stanley, and Collings (2016) conducted a study with adolescents receiving treatment related to anxiety and mood disorders and found that when motivational interviewing techniques were utilized, participants attended more sessions, had increased treatment initiation, and self-ratings with regard to readiness for treatment. Currently, research is lacking with regard to more child-centered approaches to assessment and selection of treatment goals (Connors, Prout, Kozlowski, & Stephan, 2016). However, a recent review of 11 studies, with eight meeting the criteria as having "best evidence," showed promise in regards to motivational interviewing (Snape & Atkinson, 2016). In eight separate studies where motivational interviewing activities were

used, seven of the eight had significant positive effects. The remaining study was classified as neutral. Overall, though still understudied, research on student-focused approaches to identification and intervention seems promising for improved mental health outcomes in school.

<div align="center">**<u>Strength-Based Assessment</u>**</div>

Over the past two decades there has been an effort to increase the use of a more person-centered approach to assessment (Climie & Henley, 2016). This approach to using measurement tools to review a student's strengths, abilities, and positive qualities has been termed strength-based assessment (Epstein & Sharma, 1998). Epstein (2000) noted that this approach is not novel as diagnosticians and counselors often gather this information through interviews and observation, but usually this practice was done informally and without standardized instruments. This has led to a need in the development of standardized assessment tools, which review the appropriate skills necessary to serve as buffers for later psychopathology.

Strength-based assessment is founded in the theories of positive psychology. Epstein (2000) highlights the basic principles guiding the development of these assessment tools: (1) every child has unique individual strengths, (2) children are influenced by the ways others respond and motivation increases when strengths are highlighted, (3) when a strength is not exhibited it is due to a lack of necessary experiences to develop mastery and not a deficit, and (4) when goals are developed by schools, a strengths-based approach is more likely to engage families and lead to a more likely achievement of those goals.

This shift in focus of assessments on the assets a child possesses is in part due to concerns related to limitations in current deficit-oriented assessment. An assessment approach which focuses primarily on deficits can lead to negative implications for a child's self-confidence, reduction in motivation, reflecting on past failures and prompting negative expectations, stigma or stereotypes for these youth, and decreased sense of belonging and alienation of the youth measured (Laursen, 2003). Although these tools can provide an understanding of these children in comparison to their normative peer group, these potential pitfalls can have quite negative ramifications for the child and their potential benefits from treatment.

Despite these limitations, Merrell (2008) notes the significantly higher number of traditional-deficit oriented assessment tools available and frequency of their use in comparison to strength-based tools. This is in part due to what Rashid and Ostermann (2009) noted in their review of strength-based assessments. Although there is a clear need for more strength-based approaches for treatment, deficit-oriented measures are not without their merit. These researchers highlight work that has shown that negative emotions, interactions, and critiques have greater impact on the psyche than positive. By understanding these negative emotions and thinking patterns in individuals, it can become a central part of effective treatment approaches. In addition, these tools allow for a process by which specific types of psychoses can be identified and understood. These measured deficits are what is often needed in order to review Diagnostic Statistical Manual (DSM) criteria for different mental health impairments. Under our current health care model, this label at times is what may allow for receipt of appropriate treatment and community-based mental health services.

Therefore, strength-based assessments do not necessarily replace more traditional assessment methods, but can work in combination to provide a more complete picture of youth. Suldo and Shaffer (2008) explored a dual-factor model in understanding youth. These researchers contend that understanding a child's level of illness (psychopathology) in combination with positive indicators of well-being provide the best predictor of overall mental, physical health, and achievement in adolescents. Research has also examined treatment outcomes when utilizing strength-based assessment tools in the pre-treatment phase for the youth and the adults providing support. Cox (2006) found that when incorporating a strength-based assessment in treatment for youth with considerable emotional and behavioral disorders, it did not have significant effects with regard to youths' overall functioning at post-test. However, there was an increase in parent satisfaction and a reduction in missed appointments. Furthermore, Climie and Henley (2016) outline the benefits of strength-based approaches as providing a more complete representation of the student. They further discuss that these measures encapsulate a preventative focus and they provide a more positive and optimistic view of children for the key stakeholders charged with assisting a child's development. When strength-based measures are used in conjunction with prevention-oriented programs, it is more likely for interventionists and families to observe and reinforce significant gains in strengths as opposed to simply looking for decreases in deficits.

SEL Domains

Given the importance of SEL and its potential to improve mental health later in development, the lack of self-report universal screening tools, measuring SEL domains in young children, leaves a significant void. Thus, the current study focused on development

of a tool that addresses the self-report needs currently present for SEL universal screening. A critical task in this endeavor was to draw from the theories that have been driving the understanding of the key domains that underlie SEL. Moreover, it was imperative to identify developmentally appropriate behaviors within these domains that would lend themselves to early identification of more extreme behavioral and social-emotional concern later in development. To identify appropriate subscales of overall SEL, the current study drew upon systematic literature reviews of SEL assessment tools conducted by Humphrey and colleagues (2011) and Williams (2008). Many tools that were reviewed have demonstrated reliability and validity, yet lack a self-report component or practicality as a universal screener. The current research used the tools that showed the most utility in order to identify the skills that were most frequently assessed. This process led to the identification of seven primary domains that were included this study: self-regulation, emotional regulation, social skills, self-concept, school connectedness, social responsibility, and optimism.

Regulation Skills

Self-regulation is an umbrella term, which refers to management of attention, emotion, and stress response that is both more overtly controlled by individuals and also more nonvolitional in nature. Blair, Ursache, Greenberg, and Vernon-Feagans (2015) outline self-regulation as a combination of two main subcategories, executive functioning and effortful control. Executive functioning is the ability to maintain focus and attention to a specific task or demand, even when there are outside elements competing for the individual's attention. These researchers refer to working memory and the ability to purposefully and smoothly shift attention as primary elements of this domain of self-

regulation. Effortful control, on the other hand, refers more to an individual's temperament and disposition, making it more difficult to regulate or control.

Emerging research on self-regulation has led some to argue that instruction and opportunities for growth in these areas should be as important as academic instruction (Blair & Diamond, 2008). These authors note that if students lack the regulation skills to sustain and focus their attention, hold new information in their short-term memory, and relate new information to prior knowledge, they will struggle to acquire new content.

As students begin kindergarten, there is already quite a bit of demand on students to meet regular demands on self-regulation (Becker, Miao, Duncan, & McClelland, 2014). Researchers in this field describe the challenges young children face integrating these skills to move between tasks (in some cases desired and undesired). Children must attend to directions while engaged in a desired task, inhibit the desire to continue with play, understand and hold onto direction while other demands are placed on attention, and then shift to the academic task required by the teacher. This is occurring in the wake of a period of increased development on average for students. Posner, Rothbart, Sheese, and Voelker (2012) note that during the preschool-aged years of children (ages 2-5), there is significant growth in the prefrontal cortex, which is responsible for much of the quality of self-regulation that students are able achieve. Becker and colleagues posit that this period through kindergarten is an ideal time to measure and observe the development of self-regulation for children.

Self-regulation is a specific point of concern to teachers. In a national survey of kindergarten teachers in 2001, teachers reported that roughly half of their kindergarten students enter kindergarten with at least one specific problem area that negatively impacts

their school readiness (Rimm-Kaufman, Pianta, & Cox, 2000). The number one area of concern was children's ability to follow directions (46%). The next closest concern was academics, but at just 36%. These early regulation difficulties and its impact on behavior can have significant consequences. Gilliam (2005) found in the first national study on expulsion rates in preschool settings that expulsion occurs at a rate of three times that of school-aged children.

Of further concern is the research on the rise of pharmaceutical interventions with young children. At the turn of the last century, research found that stimulant medications being prescribed to preschoolers nearly tripled during the early 1990's (Zito, Safer, Gardner, Boles, & Lynch, 2000). Although this could be due to a number of potential factors according to the researchers (expanding diagnostic considerations for ADHD, more recognition of girls experiencing ADHD symptoms, less stigma, and expansion in early detection in school settings), it still marks an area of significant concern as students enter kindergarten.

On the positive, growth in skills associated with self-regulation has been shown to have a number of positive outcomes for young children. Fuhs, Farran, and Nesbitt (2013) examined the effects of preschool teaching and environment on self-regulation development. These researchers assessed classrooms' overall emotional climate and degree of quality instruction at predicting self-regulation development. Classrooms that had a heightened degree of reinforcement for positive behaviors and quality of instruction had students with improved cognitive self-regulation. These children's engagement in academic tasks increased and gains were seen in mathematics and early literacy for these classrooms. Sawyer and colleagues (2015) also found that when improvements are made

in children's task attentiveness and emotional regulation longitudinally from ages 2-3 to 6-7, these gains meant significant improvements related to math and literacy development. These researchers recommend that efforts to promote the particular regulation skill of "task attentiveness" has the potential to significantly improve children's academic achievement.

Positive relationships between self-regulation skills and academic gains aren't just limited to measures of attending and engagement. Self-regulation skills have historically been strongly associated with math skill development. Bull, Espy, Wiebe, Sheffield, and Nelson (2011) used a confirmatory factor analysis model to examine the relationship between executive control and early math achievement. These researchers found a fairly robust relationship, and mediation data indicated that this linkage was distinct from more crystallized functioning. They argued that although some other areas of self-regulation (working memory and executive functioning) may be more associated with later, more complex math skill development, effects of self-regulation are already seen in early math skill development.

The link between early literacy and self-regulation development has been less direct in research. Blair and colleagues (2015) suggested that reading is less demanding on brain development and is more dependent on crystallized intelligence and the accessing of prior knowledge. Lonigan, Allan, and Phillips (2017) theorized that it may be that researchers need to further breakdown literacy into its smaller subcomponents to understand the link between self-regulation and literacy skill development. Specifically, print knowledge skills and areas such as vocabulary may be less strongly associated with self-regulation, whereas meaning-related skills and phonological processing skills may be

more strongly associated with self-regulation. In their initial study, they noted that executive functioning did not seem to have unique features to explain literacy development, but they found a relationship between teacher-rated attention and acquisition of skills. They concluded that due to the complexity of both self-regulation and literacy, better understanding of their subcomponents' linkages needs to be further understood in order to more appropriately target intervention efforts when educational challenges arise.

Social Skills

Gresham and Elliott (1993), creators of the Social Skills Rating System and Social Skills Improvement System, refer to social skills as learnable behaviors that lead to the ability to have positive interactions with others and to avoid engaging in behaviors that could result in undesired social outcomes. Providing instruction and opportunities for students to learn how to get along with others, become active in their learning, and engage in self-care are three of the most important goals that schools should strive to accomplish (DiPerna, Bailey, & Anthony, 2014).

There have been a number of reviews on whether social skills can be directly instructed. Gresham, Cook, Crews, and Kern (2004) conducted a mega-analysis that looked at six meta-analyses on social skills training for students with or at risk for emotional and behavioral disorders. In their review of effect sizes, they found that nearly 65% of individuals who were in social skills training treatment conditions saw an improvement in skills as compared to roughly 35% in control conditions. The sample was substantial as it included 338 studies and more than 25,000 children.

Some researchers have looked to further understand the relationship between social skills and achievement, beyond its correlation. In a longitudinal study over a nine year period (kindergarten through eighth grade), Caemmerer and Keith (2015) examined whether social skill competency positively influences academic performance or whether improved academic achievement led to improved social standing and skills in this domain. Their findings indicated a bi-directional relationship between the two overall. They noted that social skills were consistently improved by improved academic standing. However, social skill improvements seemed to lead to improved academic performance in two different periods over the course of the evaluation (spring kindergarten to first grade and fifth through eighth grade).

In contrast, one study on the importance of social skill development has shown that when controlling for the effects of prosocial behavior, the relationship between third and eighth grade achievement becomes insignificant (Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000). These researchers found the best predictor of eighth grade achievement to be an assessment of prosocial behavior that was taken by collection of self-assessment data, peer and teacher ratings in areas such as degree of helpfulness, sharing, kindness, and cooperativeness. This study was somewhat unique in its use of all three sources to gather information on individual social competency.

DiPerna, Volpe, and Elliott (2002) proposed that social skills are part of a continuum of skills referred to as "academic enablers." The first step when a student presents with academic difficulties is to further assess their current skill level for that academic area. However, these academic enablers become the second area to explore. Gaps in an enabler can contribute to that learning difficulty and a targeted goal to assist in

the remediation of that skill may ultimately lead to academic improvement. In addition to motivation and study skills, this model lists social skills as an enabler. This model recommends that as part of the problem solving process, questions are reviewed related to what social skills are important and necessary for this student to be successful in the classroom. Has the student demonstrated the skill (either currently or in the past)? And are these skills utilized at an expected rate and proficiency?

In addition to research showing that improvements in social skills are related to improved academic scores, early social skill development in the elementary school setting is predictive of a number of major health and academic outcomes. Buhs, Ladd, and Herald (2006) found that peer exclusion in kindergarten increased risk for peer maltreatment in later grades. Moreover, student maltreatment was able to predict later school disengagement. Building positive relationships with peers even as early as kindergarten can have lasting impact for a student's future learning outcomes and acceptance within the school environment. These researchers noted Coie's model (1990), which stated peer rejection can lead to chronic exclusionary concerns. These students can begin to devalue relationships as their position within the group becomes more tenuous, perhaps as a protective factor. Gresham, Vance, Chenier, and Hunter (2013) asserted that the importance of screening and identifying with gaps in social skills is imperative.

Self-Concept

Historically, various components and definitions have been outlined for self-concept. According to Carl Rogers, self-concept includes three areas: self-image, self-esteem, and ideal self (Hall & Lindzey, 1957). Baumeister (1999) provides a definition that indicates that self-concept is comprised of knowledge of attributes and also as to

what and who the self is. Lewis (1990) wrote about self-concept development and includes an initial step referring to the recognition of a separate self from others. The next component becomes categorization of the self as to the attributes and features of the self that can group the self or separate the self from others. It begins with more concrete identifications such as gender, age, and physical features. However, it begins to shift in school-aged children to relate to psychological traits and analysis of abilities as compared to others.

Prior experience and perceived competence can dictate an individual's own perception of ability to be successful. Self-concept can also be impacted by the ways an individual compares themselves to others in the environment when presented with similar tasks. Self-concept refers to an individual's perception of abilities and likelihood of success in a domain based on a combination of internal and external factors. This leads to the component referenced by Rogers, which is self-esteem. Self-esteem is related to affect and the degree of positive or negative perceptions people hold about their own ability. Self-concept is the cognitive process by which this determination is made over the course of development. Argyle (2008) outlines four factors of concept formation that influence self-esteem: perception of how others react to the individual, how the individual feels they compare to others, the social roles the individual is assigned, and what an individual's identification is in relation to the role in which they are assigned.

In schools, researchers have looked specifically at academic self-concept and its relationship with academic and behavioral development in schools. The Shavelson model was introduced in the 1970's as a hierarchical model to understand self-concept (Brunner et al., 2010). At the most general level, this model emphasizes the aspects of self-concept

that are most associated with self-esteem. The model then moves to more domain specific

areas: academic, social, emotional, and physical. This model reflects that a momentary

presentation of an individual's self-concept is understood both by the general self and the

current demand and environment. Individual's self-concept is not static, but dynamic and

its level of positivity can vary based on setting.

A Marsh/Shavelson model was developed to further explore self-concept

development in the area of academics (Marsh, 1990). Just as the previous model had a

general self, this model has an academic self-concept that influences all academic areas.

However, despite the presence of a general academic self-concept that will correlate with

both verbal and mathematical domains, the individual domains don't necessarily correlate

with each other (Marsh, Byrne, & Shavelson, 1988).

School Connectedness/Belonging

Although school connectedness has been associated with many positive outcomes

related to mental health and school performance for youth, researchers have noted that

operationalizing the definition as to what school connectedness is has been a challenge.

In fact, Whitlock, Wyman, and Moore (2014) found in their review that connectedness is

linked to nine different research fields: attachment theory, social support theory, bio-

ecological models of human development, resilience frameworks, stage-environment fit

theories, social development and learning theories, and social capital theories.

With regard to school connectedness, Barber & Schluterman (2008) in their

review of conceptualizations and understandings of connectedness outlined three areas

that seem to emerge in the literature related to school connectedness. The first component

is a youth's perception of quality of the relationships to the people of the school. Studies

on effects of connectedness typically focus on whether students feel cared for and supported by their peers and adults. The second area of school connectedness is students' relationship to the greater school community. Students who are connected feel a part of the larger school environment and identify as a part of that community. Lastly, school connectedness refers to the youth's feelings around the importance of school. This reflects research related to attendance and self-identifying as trying their best when at school as a positive predictor of connectedness.

The CDC (2009) outlined that improved school connectedness is associated with improved school attendance, higher academic achievement as measured by report card grades and test scores, decreased likelihood to use drugs or drink alcohol, reduced violent behavior, and less likelihood of emotional difficulties such as suicidal ideation or eating disorders. In a study on connectedness with nearly 2,000 students, Thorpe (2003) found an indirect effect on student achievement. Thorpe theorized that school-initiated connectedness efforts have an indirect effect on achievement by taking efforts to minimize student alienation before it emerges in the seventh grade. Specifically, alienation, which has harmful effects on learning outcomes, is reduced through increases in students' integration in the school setting, ratings of engagement in leaning, and sense of connection and quality of relationship with peers and adults. This further supports Barber and Schluterman (2008)'s operational understanding of connectedness and the positive outcomes associated with its promotion on multiple levels.

Social Responsibility

Social responsibility has been characterized as "a reflection of concern for the greater good and welfare of others that extends beyond personal wants, needs, or gains"

(Wray-Lake, Syvertsen, & Flanagan, 2016, p. 130). Researchers have examined people's

stability with regard to value formation based on Schwartz's (1992) theory of human

values. Bardi, Lee, Hofmann-Towfigh, and Soutar (2009) found that value formation and

rank order of these values related to both self-enhancement (self-directed goals and

desires) and self-transcendence (concern related to other's welfare) remain relatively

stable in adulthood. Wray-Lake and colleagues found in their research that social

responsibility values decline during adolescence. Levels measured in children at age nine

decrease steadily until leveling off at age 16. Schools can serve as a buffer: School

solidarity, community connections, trusted friendships, and volunteer opportunities had a

positive association with social responsibility values. Further, these researchers argue that

most value formation literature focuses on its stability in adulthood, but not during

periods of more plasticity. Their work focused on a critical period of adolescence, but a

similar argument can be made for the need for further understanding during the period of

childhood.

Thapa, Cohen, Guffey, and Higgins-D'Alessandro (2013) conducted a large

review of school climate research. Reviewing more than 200 studies, they found that

improved school climate leads to many significant improvements in health and

achievement for students (i.e. achievement, mental health, motivation to learn, drop-out

rates). Importantly, they found that the teaching and learning related to social-moral

emotional learning and civic responsibilities was a main contributor to overall improved

school climate.

Social responsibility research has also been associated with improvements in

children's levels of aggression, victimization, and emotionality (Leadbeater, Thompson,

& Sukhawathanakul, 2016). These researchers argue that this protective factor acts in such a way that negative behaviors are incompatible with these more positive behaviors. When a student is demonstrating prosocial skills such as initiating and enacted helping and caring behaviors, combined with teachers providing opportunities and reinforcing these skills, it minimizes the likelihood of these other negative behaviors to occur.

Positive SEL skills related to social responsibility, such as empathy and perspective-taking, begin to emerge early in development, in four and five year olds. These skills have been found in longitudinal research to be a predictor of personality characteristics and social and emotional health in later adolescence and early adulthood (Eisenberg et al., 2002). Over 25 years, children that engaged in "spontaneous sharing" during preschool and early childhood ages were later found to be more prosocial in early adulthood.

A recent study on a mindfulness SEL intervention with fourth and fifth grade students showed significant gains between self-report pre- and posttests for treatment condition students when given scales related to empathy, optimism, mindfulness, and social responsibility (Schonert-Reichl et al., 2015). Students' gains in the treatment condition as compared to the control group reflected other gains measured by behavioral assessments completed by teachers, peer acceptance, and academic scores from school records.

Optimism/Positivity

Karen Reivich (2010), a positive psychologist, outlines an understanding of optimism with two components. It is the combination of being able to see having a hopeful outlook and a belief that world is the best possible version of itself. In terms of

intervention, Reivich focuses on the portion of optimism that refers to the way people explain why things occur in their life, or the concept of explanatory style (Seligman, 1991). Dr. Reivich in her work refers to optimistic and pessimistic thinking, not optimistic and pessimistic people. Positive psychologists argue that style of thinking is not static, but there is control in the ability to challenge and shape one's way of thinking. When people are optimistic they think of their problems or struggles as temporary, they don't overgeneralize the problem beyond its specific domain, and they attribute it to external causes.

Carol Dweck (2006) has written about two type of mindsets: fixed and growth. Individuals with a fixed mindset tend to avoid challenges and view them as more permanent deficits. Individuals with a growth mindset tend to remain engaged and persist during challenges. Researchers in this field have studied the role of implicit theories of intelligence and outcomes based on the attitudes one has related whether intelligence is stable and deterministic of overall ability. An incremental theory of intelligence is one in which it is malleable, and the other is an entity (fixed) theory.

Longitudinal studies have found positive outcomes when students develop more of an incremental theory with regard to their own intelligence even when controlling for pre-treatment achievement scores. Blackwell, Trzesniewski, and Dweck (2007) found that teaching incremental theory to 7th graders led to increases in motivation and achievement. Good, Aronson, and Inzlicht (2003) found that this approach can be especially helpful at challenging stereotype threat for individuals in marginalized groups. Their research showed improved achievement for females in mathematics as compared to control groups and improvements in reading scores for minority and low-income

41

adolescents as compared to control groups. Bevel and Mitchell (2012) found that academic optimism predicted and accounted for more variance than poverty when looking at reading achievement.

Early research on optimism in schools focused on personality as a way of understanding optimism and its positive outcomes for academics. Gough (1953) found that in utilizing a personality measure to compare higher achieving high school students from lower achievers, one of the most significant predictors was having "optimistic self-confidence." Positive psychologists argue this is in part due to the relationship between optimism and resiliency. Having a mindset that sees positivity and opportunity in situations, foster one's resiliency and persistence despite challenge. People who believe they can reach a desired goal are going to work harder to achieve and be less apt to give up (Wise & Rosqvist, 2006).

Boman, Furlong, Shochet, Lilles, and Jones (2009) reviewed a number of the more promising programs in the instruction of optimism within the school setting. Through meta-analysis they concluded that the most successful programs have been ones that target the preadolescent group, and are more targeted as opposed to universal. However, they caution without adequate screening information to make the program information targeted, this approach was less sustainable.

One universal program with a core component targeting optimism is the Penn Resiliency Program (PRP). In a review of the program, which has been implemented in a number of settings in multiple countries, Seligman, Ernst, Gillham, Reivich, and Linkins (2009) found that direct instruction in this area can lead to a number of positive outcomes. One training set is designed for students aged nine to 14 years old. PRP can be

implemented in 12 90-minute lessons or 18-24 60-minute lessons. PRP has been found to reduce and prevent depression, reduce hopelessness, prevent clinical depression and anxiety, and reduce behavioral problems (such as aggressive acts and delinquency).

Current Study

The outlined research indicates an increasing need for SEL programming to be effectively implemented in schools. With increases in available SEL evidence-based programs, the need for a framework to coordinate the implementation of various programs and review their effectiveness is needed. A RTI model provides a useful starting place to conceptualize how to address this need; it highlights the need for measurement tools that will provide feedback as to the effectiveness of universal programming efforts and areas to target for intervention, while also obtaining data to identify students at-risk for later negative outcomes. While there are evidence-based screening tools available to schools, there are fewer tools available for the early elementary school years, and very few that include the self-reports of young children. Based on the internalizing nature of many mental health concerns in adolescence, the need to obtain student data including behaviorally anchored perceptions of self-competencies and social-emotional skills development seems like a large void. However, obtaining this information has presented with challenge, particularly for more universal approaches. Reliable self-ratings from youth at this stage that are related to later health outcomes is a great need that this study aims to address.

Based on available tools, seven domain areas have been selected that have shown to be related to later more positive health outcomes. Steps have been taken to design a tool that may maximize the likelihood of obtaining reliable self-ratings from youth in

kindergarten through 2nd grade. The end product will be one that could lead to the creation of a brief, universal screening tool that could be administered within a whole class setting, with simple scoring procedures that balances predictive properties with practicality. This would allow for the potential of early identification of students at-risk and would target SEL programming that builds skills and may serve as a way to close gaps before the onset of a later potential mental illness or significant school-related difficulties.

CHAPTER III

METHODOLOGY

The purpose of this study was to develop a social-emotional screening instrument that could be administered to kindergarten through second grade students. Unlike more diagnostic instruments that can be given one-on-one and with more time constraints, a balanced approach was taken to ensure the measure had adequate psychometric properties, but also was feasible with regard to administration and scoring. In order for this to be a tool that could be used multiple times within a school year, the instrument should ultimately take less than 15 minutes to administer, around one minute to score for each student, and have a delivery and protocol that will allow students to provide reliable responses within a large-group setting.

To accomplish this goal, the study included a screening tool that was informed by a literature review and two data collection phases. The literature review provided the theoretical basis for the initial pool of items, which covered the seven domain areas that had been identified as being associated with positive social-emotional development. The first phase of data collection was a teacher survey conducted to review the initial item list. Teachers ranked items based on the behaviors that they believed to be the most important and that students were demonstrating by that age level, and by how well students would understand the language of the item. Based on the results of the teacher survey and guided by theory, the screening tool was then condensed. The second phase of data collection was a pilot of the revised screening tool within kindergarten through second grade classrooms. Items from this pilot survey were then reviewed with regard to

their reliability and their ability to provide useful information based on their relationship with the student's overall proficiency on the measure.

Settings and Participants

Kindergarten through second grade teachers from a district in the Northeast were invited to participate in the initial phase of this study. The district has three primary buildings and is located in a suburban setting. The district's most recent Basic Education Data System information reported the district to be relatively homogenous with regard to race and ethnicity and economically above state averages. With regard to race and ethnicity, the district's Caucasian students account for 84% of its population as compared to the state average of 45%. Students regarded as being economically disadvantaged account for 12% of its population as compared to the state average of 54% for that year. The district is a higher performing school: on the state Common Core assessment for the last year, the district had an overall proficiency percentage that was approximately twice the statewide rate on the grade 3-8 Mathematics and English Language Arts assessments.

For the student pilot study, the same district (District 1) was used as with the teacher survey, but with an additional district (District 2) also providing students to the sample. District 2 is also homogenous with regard to race and ethnicity, but is a rural district with a more economically disadvantaged population. This district's Caucasian population accounted for 90% of its population. However, 42% of its population are regarded as economically disadvantaged. This district is a lower performing district when using Common Core assessment performance as an indicator. The district has been identified by the state as a school in need of developing a Local Assistance Plan based on underperformance of various subgroups on the state Common Core assessments. This

district has performed just below state averages with regard to percentage of students achieving scores of proficiency.

Sample Size

The initial portion of the sample was recruited from District 1, which employed 23 kindergarten through second grade teachers at the time of this study. All teachers were invited to take part in the teacher survey. A response rate of 57% was achieved with 12 fully completing the survey. Within this district, the student pilot survey was conducted in one of the three elementary schools. This school had the greatest level of diversity with regard to socioeconomic status and race/ethnicity. The students who qualify as economically disadvantaged make up 14% of the population within that school. With regard to race and ethnicity, Caucasian students account for approximately 80% of the overall population. All 11 kindergarten through second grade teachers in this school were invited to have their classrooms participate in the study. Every class consented to participate. However, one first grade class had to drop out due to a last minute scheduling conflict. The resulting student pilot sample size in this district was 207 kindergarten through second grade children. The second district to participate in the student pilot study had 17 kindergarten through second grade classrooms. Overall, 10 classes from this district gave consent for data to be collected and a comparable sample size was collected within this district ($n = 177$). This provided an overall student sample of 384 students with over 100 students within each grade level assessed.

Recruitment Method

Participants of the district were invited to participate through email and at meetings. The primary researcher was the school psychologist who worked for both

districts that participated. The first phase was the teacher survey. The teacher survey was made available electronically through the Qualtrics website. Teachers were also given the option of a paper and pencil version, but all teachers that participated completed their survey online. Active teacher consent was obtained through a consent letter that outlined the purpose of the survey, approximate time for completion, incentives for participation, and statement on the confidentiality of their responses (see Appendix A).

The second phase was the student pilot study. This was conducted at elementary schools via a paper and pencil survey filled out by children of the district. Consent for this portion was obtained through a passive method with a letter sent home to parents informing them of the upcoming survey. The letter outlined the purpose of the survey, steps to ensure confidentiality, and a form to be mailed or sent with their child that allows their child be withdrawn from the study (see Appendix B). No letters were returned and all kindergarten through second grade students present on the day of data collection participated in the study. Student item responses were then used to analyze the reliability of the measure and determine the items that accounted for the highest variance of the social-emotional competency construct.

Incentives for Participation

Incentives were offered to increase participation of classroom teachers and districts. For the teacher survey, the first ten teachers to complete the survey were given a $5 gift card to Dunkin' Donuts or Starbucks and all teachers completing the survey were entered to win one of two $25 Barnes and Noble gift cards.

Administrators of participating schools were offered a presentation of final results. In addition, if future research establishes the screening tool as valid,

administrators will receive a copy of the final instrument and a scoring sheet for their districts.

<div align="center">**Procedure**</div>

<div align="center">**Teacher Survey**</div>

The goal of the teacher survey was to obtain ratings from teachers on the domain areas and behaviors that they believe most represent social-emotional competency within their classrooms. Teachers were asked to rate behaviors that they feel are exhibited by most students of their classroom.

Items were developed using literature reviews on social-emotional competency measures (Humphrey et al., 2011; Williams, 2008). All scale domain areas were identified by the principal investigator and based on an extensive literature review. Similar domain areas such as connectedness and belonging were counted as one. Domain areas were then tallied for frequency at which they were measured in existing social-emotional diagnostic and screening tools. The domain areas identified with greatest frequency were: social skills, emotional regulation, self-regulation, self-concept, social responsibility, school connectedness/belonging, and optimism/positivity.

A review of the literature for each domain area was conducted to gain a theoretical understanding of these constructs and the potential markers of positive development within each domain. The principal investigator then developed survey items based on this review with an emphasis on behaviors that may be observed in kindergarten through second grade. Participants would give their recommendation of whether an item, such as "I take turns," should be included in the student survey. Teachers ranked items on a five-point scale from "do not include" to "absolutely include." Participants were asked

to give consideration to both the appropriateness of the behavior for these developmental ages and also the readability of the items. An equal number of questions were developed for each domain and can be reviewed in Appendix C. Behaviors in these domain areas were then ranked based on the results of this survey and used to develop items for the youth self-rating scale. The survey also include demographic questions including the school and grade the participants taught. This survey can be reviewed in Appendix D.

Pilot Testing of Youth Scale

The principal investigator administered the youth surveys in twenty different kindergarten through second grade classrooms across two separate districts in the Northeast. Administration occurred in a group setting within each classroom. All students of the classroom were asked to participate and surveys remained anonymous. The classroom teacher, and in some cases a classroom aide, assisted by monitoring students' progress on the survey. In addition, after the administration of the survey, classroom teachers were asked if they felt they could give the measure after seeing it administered. This served as a brief measure of feasibility.

The survey consisted of 30 items with ten items per page. The pages were printed in color and each item had a number ranging from 1-10 with each number having its own color. Additional consideration was given as to which response option visuals would produce the most reliable information from students. Rosenberg and Bryant (2001) compared how young children responded to similar items based on different survey response icons. Although younger children tended to have a more positive response style overall, the distribution was particularly skewed when facial expression icons were used: Children may be less willing to select a frowning face option. In contrast, children

50

showed more discrimination when a simple yes/no or thumbs up/down format was used. The response choices for each item were "usually or always true" as indicated by a thumbs up clip art, "not true or rarely true" as indicated by a thumbs down clip art, or "sometimes true" as indicated by a sideways thumb clip art. The layout of the page was the item number, followed by response choices, and lastly the item statement. This was done to minimize the amount of tracking students would need to do to go from the item number to their selection. A light pink background color was given to the even numbered items to provide contrast as the student progress from item to item and aid their tracking as they progressed through each page. The survey had a total of three pages. Each page had a different color shape (red square, green circle, or blue triangle) visible in the corner to assist students and the investigator ensure they were on the right page. This allowed the item numbers to never exceed the number ten to maximize student ease, yet allow all students to respond to all 30 pilot items. Surveys were also counterbalanced by rotating the order of the survey pages. This resulted in three versions of the surveys, with each set of ten items being presented first one-third of the time that the survey was administered.

The survey was introduced to the students with directions relating to the purpose of the survey. Students were also given an opportunity to practice the response choices of endorsing whether they felt an item was "usually or always true," "not true or rarely true," or "sometimes true." Practice items were given by asking students to physically indicate their response choice by giving a thumbs up, thumbs down, or sideways thumb in order to mirror the choices on the survey. Students were asked to respond to practice statements that would be nearly universally endorsed in the same way for all students. For instance, students would give a thumbs up to the statement, "my school has a roof."

51

Students were encouraged to be honest with their answers and leave their name off their paper.

Cards were held up showing the number and color of the item that was currently being read and with a verbal prompt as to the item number and color before the reading of each item. The principal investigator and the classroom teacher would circle the room and make sure students were on the correct item and that each student was ready to move on to the next item before proceeding. At the end of the first ten items, students were then prompted to turn the page and make sure they had the right symbol in the bottom corner of the page before administering the next ten items.

Once the 30 items were completed, students would then turn back to the front page and the investigator would collect the class items. Each survey had a number in the top corner to track the class identification and a participant number for each student, but no names were ever attached to any items. The class identification also included grade level, which served to provide data for between-group comparisons by grade. Total administration time ranged from approximately 15 to 20 minutes with kindergarten classrooms tending to take the longest time. Scoring of each survey took approximately one minute per student.

Data Analytic Plan

Teacher Survey

Teacher surveys were analyzed by computing descriptive statistics, including mean, median, mode, skewness, and kurtosis to assess the distribution of scores for each item. Frequency tables were then examined and items were ranked based on the level that teachers had endorsed as being social-emotional developmental behaviors that they were

52

most apt to observe. These items were examined to determine the social-emotional

domain areas that teachers found to be most representative of positive social-emotional

development for that age group. Highly rated questions within these domains were then

selected to be used as part of the screening tool administered to youth. The goal was to

have approximately 30 items that could be used as part of the initial youth screening tool.

This was based on the assumption that following an item analysis, only half to one-third

of the questions would prove to be reliable and account for a significant amount of the

variance of social-emotional development.

<div align="center">**Initial Screening Pilot**</div>

After the initial youth screening was conducted, an initial analysis of items was

conducted to assess for the distribution of responses: mean, median, mode, skewness, and

kurtosis. These descriptives, along with frequency tables, were reviewed to determine

whether any items should be removed as outliers both in terms of mean and standard

deviation using item analysis methods. Internal reliability was assessed using Cronbach's

alpha.

An exploratory factor analysis was then used with the remaining items to

determine dimensionality and which questions were most related to the overall measure

of social-emotional development. A Scree plot was examined for information related to

the unidimensionality of the items and the degree of variance explained by the first

factor. Based on this review if one factor were to emerge as explaining a significant

proportion of variance, an item response theory approach would be taken to review items

that had at least a moderate factor loading. Items with low factor loadings were removed

from further analyses. Criteria for appropriateness of factor loadings was influenced by

guidelines outlined by Ford, MacCallum, and Tait (1986). Although, a general heuristic is that factor loadings have a value of at least 0.3, it is important to compare these values within the context of the study to determine appropriateness as opposed to one specific cut-score.

In the event of unidimensionality, Item Response Theory (IRT) was selected to determine which items are most useful to include in future studies and if any items can be eliminated from the pilot instrument without significantly impacting the reliability of the overall measure. The fewer the questions needed, while still maintaining an accurate representation of the construct (social-emotional development), the more likely the scale will allow for future administration that is reliable and practical.

IRT was initially developed to review dichotomous variables, but later models were developed to review polytomous variables (Penfield, 2014). This approach allows for the review of individual items to determine the relationship between performance on the item and the overall measure based on an individual's proficiency. An IRT approach allowed for determination of which items were most related to the overall construct being measured, but also for maximizing the amount of insight into items that provide variance for students that performed lower on the measure.

The goal of the scale was to identify students who are less proficient and who are more likely to endorse positively-worded items as only rarely or sometimes occurring for them. The items being reviewed were polytomous with three response options (thumbs up, down, or sideways). Based on items being categorical in nature, similar to a likert scale, the IRT model that was used was a graded response model. Once a final scale was determined, its items were then analyzed to determine any between-group differences

(grade level, school type). These differences were assessed using a sum score and

calculating a *t*-test.

CHAPTER IV

RESULTS

Teacher Survey

The goal of the teacher rating scale was to help narrow the initial item pool of 105 items to 30 items for the student pilot study. Seven districts were contacted to participate in this portion of the data analysis with one district giving consent. Two elementary schools from this district participated. Out of a potential pool of 24 kindergarten through second grade teachers, 12 completed the survey. This sample size was lower than the desired goal for this portion of the study.

Teachers were given the option to complete the survey online or via pen and paper. All teacher participants elected to complete their survey online. Table 1 provides the means for the 30 most highly rated items endorsed by respondents in descending order.

Table 1: Item Means from Teacher Survey (continued onto next page)

Item	Mean
I take turns	4.81
I like playing games even when I lose	4.81
I like coming to school	4.81
People at school care about me	4.81
I like to learn	4.81
I listen carefully to the teacher	4.72
I get my work done when I'm supposed to	4.72
My family cares about me	4.72
I work hard at school	4.72
I do my best when I work	4.72
I like myself	4.72
I raise my hand when I have a question	4.63
I stay in my seat when I'm supposed to	4.63
I invite kids to play with me	4.63
Kids at school like me	4.63
I am quiet in the hallways	4.54

I don't poke other kids	4.54
I don't hit other kids	4.54
I can sit and listen to a whole story without getting up	4.54
I like to share my toys	4.54
Kids want to be my friend	4.54
I like other kids	4.54
I don't get upset when I lose	4.54
Other kids will let me play with them if I ask	4.54
I wait my turn in line	4.45
There are many people I can talk to if I have a problem	4.45
My teacher notices when I do my best work	4.45
I don't yell at people	4.41
I like talking with kids in my class	4.36
I can join in games other kids are playing	4.36

Items that were the most highly rated by teacher respondents were concentrated from four domains in particular: connectedness, positivity, self-regulation, and social skills. Of the 30 top-rated items, only one each came from the item pools for emotional regulation, self-concept, and social responsibility.

In conjunction with the results from the teacher survey, a qualitative theoretical approach was implemented (Creswell, 2013). The primary investigator rated his top four items for each domain area and then cross referenced it with the top thirty items based on mean score from the teacher survey. Items that were indicated by both were the first to be selected for the initial pilot. This approach led to 15 items being immediately selected for the student pilot survey. In the four domains where the majority of the highly rated items occurred in the teacher survey (connectedness, positivity, self-regulation, and social skills), 14 items selected by the investigator aligned with the top-16 rating for these four domains from the teacher survey. Also, the bottom ten items from the teacher survey were examined and none of these items were ones that were selected by the principal investigator. Items from the other three domains (emotional regulation, self-concept, and

social responsibility) were selected by the principal investigator to ensure adequate

representation from each domain in the initial pilot survey. These items can be seen in

Appendix E.

Four of the original items were then selected to be written to describe negative,

undesired behaviors. These items were:

I cry when it is time to come to school
I yell at people
I get upset when I lose
I poke other kids

This was done to serve as a check of children's ability to reliably complete the measure.

These negatively worded items should not be positively correlated with the other 26

items which reflect more positive, desired behaviors.

Student Pilot Rating Scale

One goal of this pilot study was to determine whether students could reliably

complete this scale through a whole-group administration. In addition, this study

reviewed which items provided the most useful information and if the scale length could

be reduced from the original 30 items. Students in kindergarten through second grade

classrooms in two different schools within different districts is New York State

contributed to the overall study sample. The districts were diverse with regard to

socioeconomic status (12% vs. 42% economically disadvantaged) and region type (588

vs. 118 people per square mile). Consent was obtained from building principals and

district administration. Consent letters were sent home and no parents elected to opt their

child out of the research study. One first grade classroom in the suburban district could

not participate in the study due to a scheduling conflict. All other kindergarten through

second grade classrooms in the two districts participated. The data collection for this

school was scheduled and collected in June. The data for the rural district was collected

in the month of February.

Student Demographics

The student pilot study had a total initial sample of 385 students. One student

presented as a significant outlier with regard to response selection and overall proficiency

score. This student's data was removed from further analysis resulting in a final sample

of 384. Table 2 outlines the sample characteristics by school and grade:

Table 2: Number of Student Participants by Category

Grade	School Type		
	Suburban	Rural	Total
Kindergarten	79	49	128
First	38	69	107
Second	89	60	149
Total	206	178	384

Item Descriptive Statistics

All items were reviewed using classical item analysis techniques. Table 3 outlines

the response characteristics for students on all 30 items:

Table 3: Item Descriptive Statistics (continued onto next page)

		Percent In Each Category					Classical
Item	N	1	2	3	Mean	SD	Discrimination
1	383	11.7	29.8	58.5	2.47	.697	.329
2	383	4.7	21.4	73.9	2.69	.555	.437
3	383	6.8	17.5	75.7	2.69	.592	.346
4	383	58.0	30.3	11.7	1.54	.696	.269
5	384	2.1	24.7	73.2	2.71	.498	.222
6	382	7.6	26.2	66.2	2.59	.629	.412
7	384	1.3	11.2	87.5	2.86	.381	.337
8	382	8.9	31.7	59.4	2.51	.655	.353
9	384	5.5	22.9	71.6	2.66	.578	.440
10	384	0.8	16.7	82.6	2.82	.406	.398
11	384	6.5	19.5	74.0	2.67	.592	.385
12	381	1.3	18.6	80.1	2.79	.441	.381

13	383	82.0	16.2	1.8	1.20	.443	.226
14	384	3.6	24.2	72.1	2.68	.538	.344
15	379	5.3	42.7	52.0	2.47	.596	.128
16	384	5.7	28.4	65.9	2.60	.596	.461
17	384	2.1	13.8	84.1	2.82	.435	.299
18	383	1.3	18.8	79.9	2.79	.441	.395
19	380	91.1	6.3	2.6	1.12	.394	.256
20	383	5.5	58.2	36.3	2.31	.569	.401
21	384	0.5	17.2	82.3	2.82	.400	.312
22	380	11.3	32.4	56.3	2.45	.689	.372
23	384	8.6	41.7	49.7	2.41	.644	.311
24	384	4.2	9.4	86.5	2.82	.479	.313
25	383	10.7	25.8	63.4	2.53	.682	.397
26	384	3.1	17.7	79.2	2.76	.495	.429
27	383	1.0	18.0	80.9	2.80	.427	.353
28	384	1.0	6.0	93.0	2.92	.309	.308
29	382	83.0	10.2	6.8	1.24	.564	.256
30	384	7.6	17.7	74.7	2.67	.610	.452

Note. Categories: 1 = thumbs down (never); 2 = sideways thumb (sometimes); 3 = thumbs up (almost always)

Based on the means and frequencies within the three response options, students displayed a pattern of responding affirmatively to each item with the exception of the four negatively worded items (items 4, 13, 19, and 29). The negatively worded items were not positively correlated with the overall score. These negatively worded items were then reverse scored for all further analysis. With regard to their distribution, items were negatively skewed overall. Internal consistency was calculated by Cronbach's alpha and was found to be adequate (0.83). Correct item-total correlations were calculated as a measure of classical discrimination. A general heuristic is for this value to be 0.3 or larger in order to demonstrate an adequate discriminating relationship between performance on an item and the overall total score (Nunnally & Bernstein, 1994). As can be seen in Table 3, 22 of the 30 items met criteria of being 0.3 or larger. Nunnally and Bernstein (1994) indicate that this is not a fixed cutscore, but do state that any item with

an item-total correlation below 0.2 would be described as very poor in their ability to discriminate. Only one item fell below the 0.2 threshold.

Exploratory Factor Analysis

An exploratory factor analysis (EFA) was conducted using principal axis factoring to determine the factor structure of positive social-emotional development. Determining the number of appropriate factors in EFA can at times be problematic. Kaiser (1960) proposed a general practice of only keeping factors with eigenvalues over one. However, researchers have argued against this practice based in part because of the arbitrary nature of this value and whether factors with eigenvalues of 1.01 versus 0.99 are truly different in their acceptability (Ledesma & Valero-Mora, 2007). Alternatively, Cattrell's scree plot can provide a visual to review eigenvalues in descending order to determine where the last significant drop occurs and value descent begins to level off and creates the look of an elbow. Thus, a scree plot was generated to gain insight into the number of factors that emerged from the initial screening (Figure 1).

Figure 1. Scree plot representing eigenvalues.
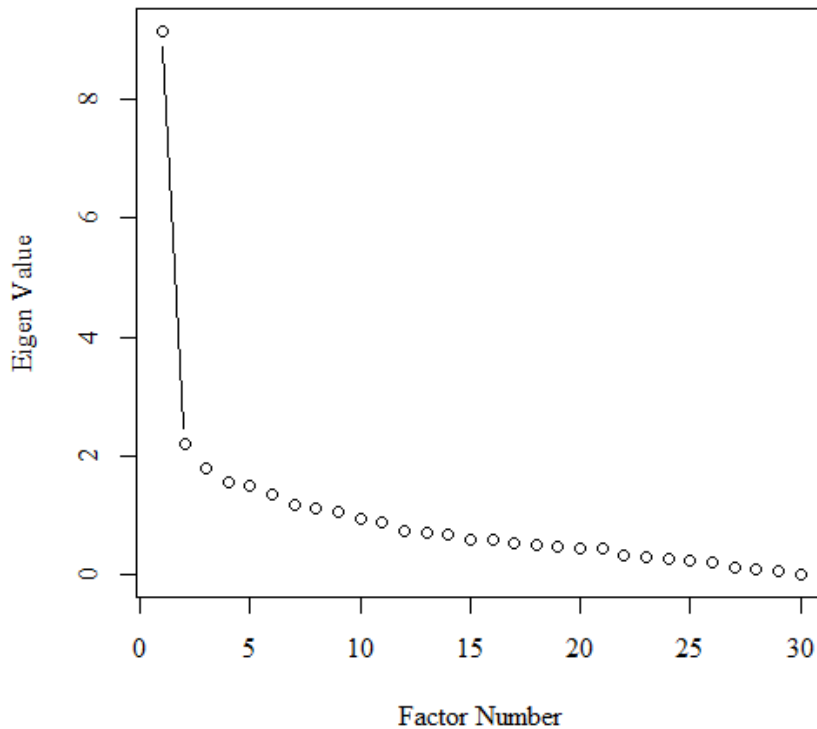
Although the scree plot approach has received some criticism due to the subjective nature of making a determination, in the current study the difference in the drop from factor one to two as compared to two to three is quite striking. Therefore, it appears there is one predominant factor in this model. Moreover, the first factor accounted for 30.5 percent of the total variance. When determining unidimensionality, having the first factor explain at least 20 percent of the variance is recommended (Reise & Revicki, 2015). Based on the review of the scree plot and the amount of the variance explained by the first factor, the results of the factor analysis supported unidimensionality. While items represented seven domains, together they appear to reflect a broader construct of positive social-emotional development. Individual item factor loadings can be seen in Table 4:

Table 4: Factor Loadings from Principal Axis Factoring

Item Number	Factor Loading
1	0.429
2	0.583
3	0.572
4	0.325
5	0.414
6	0.547
7	0.555
8	0.511
9	0.594
10	0.709
11	0.522
12	0.539
13	0.454
14	0.544
15	0.154
16	0.613
17	0.635
18	0.594
19	0.548
20	0.543
21	0.537
22	0.509
23	0.447
24	0.557
25	0.586
26	0.634
27	0.587
28	0.685
29	0.419
30	0.632

Various researchers have given guidelines as to appropriate factor loading scores. Typically, recommendations range from 0.3 to 0.4 as acceptable minimum thresholds for item inclusions in further analysis. Hair, Anderson, Tatham, and Black (1998) provided recommendations based on overall sample size. With this sample having over 350 participants, these researchers suggest factor loadings greater than 0.3 as an acceptable gauge for item inclusion. All items meet the 0.3 criteria with the exception of one item.

Item 15, "I can do a lot of things without help from adults," had a very low factor loading and was removed from further analyses. The item was intended to reflected positive self-concept. However, it appears the wording "without help from adults" made the question confusing to students. In some cases, students with higher scores on the overall measure were less likely to endorse this item. Students may still actively seek adult support at this age and their level of independence may not make this as useful an item for this age group. Another item, "I get upset when I lose," had a factor loading just above 0.3. Although its loading was close to the minimum threshold, it was not immediately removed from additional analyses because it was not as discrepant from other items.

Item Response Theory

One assumption underlying the use of Samejima's GRM is that covariation of items is measuring a singular dimension. Based on the unidimensionality of the data following the initial exploratory factor analysis, individual items were analyzed utilizing an item response theory (IRT) approach. Samejima's GRM was selected based on the items' response options being polytomous (having more than two response choices) and ordered. The response options in the current study were similar to a 3-point Likert scale with options including a thumbs down representing never, a sideways thumb representing sometimes, and thumbs up representing almost always. This analysis gave information for each item based on an individual's overall proficiency.

Table 5 contains the results of the GRM item parameter estimates.

Table 5: GRM Parameter Estimates (continued onto next page)

Item	b_1 (SE)	b_2 (SE)	a (SE)
1	-2.95 (0.50)	-0.52 (0.22)	0.75 (0.13)
2	-2.60 (0.33)	-0.97 (0.26)	1.41 (0.21)

3	-2.57 (0.36)	-1.18 (0.28)	1.21 (0.19)
4	-3.68 (0.77)	-0.62 (0.27)	0.59 (0.13)
5	-4.56 (0.80)	-1.29 (0.43)	0.90 (0.16)
6	-2.50 (0.33)	-0.71 (0.21)	1.18 (0.18)
7	-3.76 (0.59)	-1.86 (0.51)	1.32 (0.23)
8	-2.72 (0.38)	-0.48 (0.20)	0.98 (0.15)
9	-2.64 (0.33)	-0.93 (0.25)	1.31 (0.19)
10	-3.37 (0.45)	-1.27 (0.43)	1.82 (0.27)
11	-2.89 (0.41)	-1.17 (0.29)	1.08 (0.17)
12	-4.00 (0.62)	-1.37 (0.46)	1.27 (0.21)
13	-4.55 (0.83)	-1.87 (0.55)	0.94 (0.18)
14	-3.25 (0.44)	-1.03 (0.30)	1.17 (0.18)
16	-2.55 (0.30)	-0.67 (0.21)	1.36 (0.18)
17	-3.19 (0.43)	-1.52 (0.39)	1.47 (0.23)
18	-3.57 (0.50)	-1.27 (0.41)	1.42 (0.22)
19	-3.35 (0.57)	-2.25 (0.58)	1.25 (0.26)
20	-2.98 (0.39)	0.64 (0.15)	1.10 (0.15)
21	-4.75 (0.81)	-1.61 (0.61)	1.17 (0.20)
22	-2.38 (0.33)	-0.30 (0.16)	1.00 (0.15)
23	-3.11 (0.48)	0.01 (0.17)	0.84 (0.13)
24	-3.05 (0.46)	-1.89 (0.44)	1.21 (0.22)
25	-2.14 (0.27)	-0.59 (0.18)	1.21 (0.17)
26	-2.86 (0.36)	-1.20 (0.32)	1.50 (0.23)
27	-3.75 (0.54)	-1.35 (0.44)	1.41 (0.22)
28	-3.16 (0.45)	-1.96 (0.54)	1.87 (0.35)
29	-3.61 (0.73)	-2.23 (0.58)	0.80 (0.18)
30	-2.14 (0.25)	-0.99 (0.23)	1.53 (0.22)

These results provide information as to the relationship between a student's overall proficiency and the likelihood of their endorsement of the three different response choices. The initial threshold parameter estimate (b_1) is based on proficiency and provides the estimate at which at student would have a 50% chance of endorsing a sideways (sometimes) or thumbs up (almost always) for that particular item. The lower the estimate, the lower the proficiency at which a student would be likely to indicate a response other than thumbs down (never). For many of these items this value was relatively low ranging from -4.746 to -2.135. This indicates that it was relatively easy to score above a thumbs down (never), and this was not a frequently selected response for that item. The second proficiency value (b_2) represents the likelihood of scoring above a two, meaning students selecting the thumbs up option. These scores ranged from -2.247 to 0.638. The last value is the parameter estimate (a) which provides the discrimination for each item. In general, if an item's discrimination score is above 0.8 it is considered an item that provides useful information and a range of scores consistent with the overall proficiency rating (Ayala, 2009). Despite limitations related to a three point scale, nearly all items achieved this cutscore.

Item Category Response Functions

Item category response function graphs were also generated to assist in the selection of the best items on this scale for future study. An example is shown in Figure 2, which represents graph for Item 2, "There are a lot of people I can talk to if I have a problem."

Figure 2. Item category response function graph for item 2.
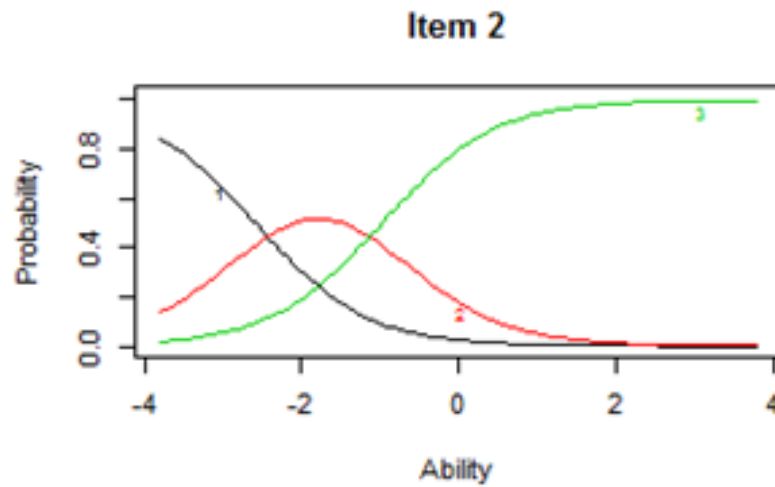
This graph demonstrates the probability that a responder will select a particular response choice as a function of the overall proficiency (ability). Line 1 represents the likelihood a respondent for this item would select the "thumbs down" choice, Line 2 a "sideways thumb," and Line 3 a "thumbs up." This item was one that was measured to have an appropriate degree of discrimination based on its parameter estimate ($a = 1.409$). It demonstrates that as ability increases it becomes more likely a respondent would select with a thumbs up or sideways thumb for that particular item. Alternatively, for students with low proficiency it shows it is unlikely they would select thumbs up for their response choice option. This item also shows an increase in the probability that a student would select the sometimes option for this item when they fall closer to the mean level of proficiency. Conversely, Figure 3 below represents the graph for item 4: "I get upset when I lose."

Figure 3. Item category response function graph for item 4.
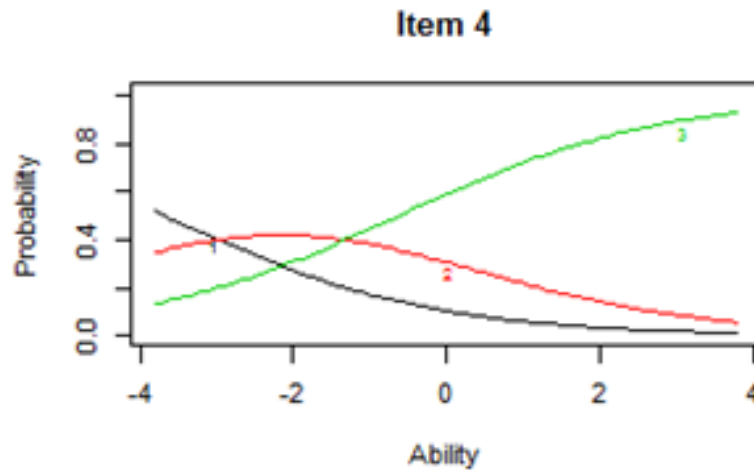
This item shows a similar pattern with regard to the relationship between ability and probability of a particular response choice. However, there is much less discrimination as compared to the prior item, particularly between the sometimes option and almost always selection choice. This item is one that would not provide as much useful information, statistically, as its distribution of response categories does not discriminate consistent with an individual's overall proficiency to the level of other items.

All item response category graphs can be seen in Appendix F.

Item Information Curves

Another area reviewed as part of the IRT analysis was the item information curves. These curves provided a visual as to the amount of information (i.e., precision) that is obtained across the different levels of proficiency. In examining these curves, it is important to consider the size of the peak and its location. Similar to the response function graphs, item information is plotted against ability with both representing continuous variables. Item information curves for Item 2 and Item 4 are displayed below (Figure 4) as a representation of an item that presents with more useful item information

(Item 2) and one that presents with less useful information (Item 4) for the overall

proficiency score.

Figure 4. Item information curves for items 2 and 4.
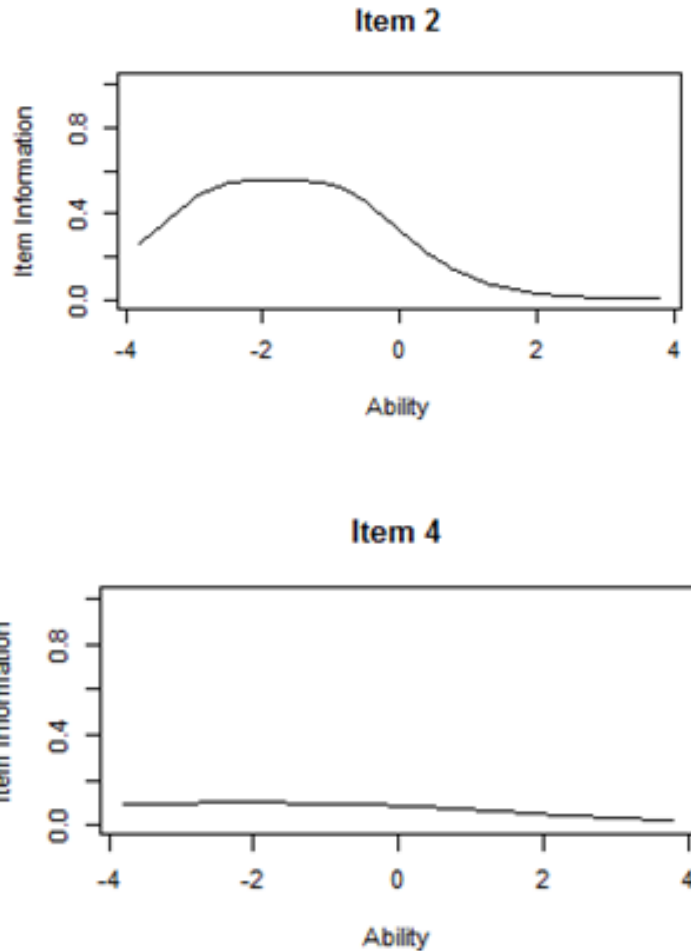
The *y*-axis in Figure 4 represents item information with larger values indicating more

prevision at the respective ability level. The *x*-axis represents the ability. In reviewing

these two items based on information, the obvious difference is in the height of the curve

for Item 2 as compared to Item 4. Item 2 represents an item that provides greater

information as compared to Item 4. An additional aspect to consider is the location of the

curve. The curve for Item 2 occurs predominately between -4 to 0. This means that the

information obtained through this item is most useful at identifying lower to moderate performers. This is particularly useful for the current study as the goal is to use these items to create a universal screening measure.

All item response category graphs can be seen in Appendix G.

Domain Area Review

The last consideration with regard to final item selection was a review of the items by domain area to determine whether items in all domains are appropriate for the final screening tool. Even domains that are associated with positive social-emotional development may not necessarily be developmentally appropriate for this age group, and it is possible that the behaviors used to develop the items were not written in a way that allowed for accurate self-rating for this age group. Examining information on items' usefulness within each domain may provide some insight into these considerations. Table 6 shows items grouped by domain and information related to their overall utility based on rater proficiency. The discrimination value classification was determined based on the guidelines outlined by Baker (2001). Baker states these labels can provide a useful way of conveying meaning to these numeric values. Excluding extreme values of none or perfect classification, this system consists of five descriptors ranging from very low to very high. Only one item did not meet classification of providing at least a moderate level of discrimination.

Table 6: Item Utility by Domain (continued onto next page)

Domain Area	Item	Discrimination (a)	Factor Loading	a classification
Self-Regulation	I listen carefully to the teacher	1.17	0.54	Moderate
	I get my work done when I'm supposed to	0.90	0.41	Moderate

	I raise my hand when I have a question	1.27	0.54	Moderate
	I poke other kids	1.25	0.55	Moderate
	I wait my turn in line	1.87	0.69	Very High
Emotional Regulation	I yell at people	0.94	0.45	Moderate
	I cry when it's time to come to school	0.80	0.42	Moderate
	I use my words to tell someone if I'm angry	1.00	0.51	Moderate
	I can tell people how I am feeling	1.18	0.55	Moderate
Social Skills	I take turns	1.32	0.56	Moderate
	I like playing games even when I lose	1.53	0.63	High
	I invite kids to play with me	1.17	0.54	Moderate
	I can join in games other kids are playing	0.84	0.45	Moderate
	When I ask kids to play with me they say yes	1.10	0.54	Moderate
Self-Concept	My teacher cares about me even when I make a mistake	1.21	0.56	Moderate
	I can do a lot of things without help from adults	--	0.15	--
	Other kids like me even if we sometimes argue	0.75	0.43	Moderate
	I feel included by my friends during recess	0.98	0.51	Moderate
School Connectedness	I like coming to school	1.21	0.59	Moderate
	People at school care about me	1.36	0.61	High
	Kids at school like me	1.31	0.59	Moderate
	There are many people I can talk to if I have a problem	1.41	0.58	High
Social Responsibility	I like to learn	1.21	0.57	Moderate
	I like to help other kids at school	1.82	0.71	Very High
	I like to help kids when they are sad	1.42	0.64	High
	I like to help my teacher	1.50	0.63	High
Optimism	I do my best when I work	1.41	0.59	High
	I like myself	1.08	0.52	Moderate
	I get upset when I lose	0.59	0.33	Low

My teacher notices when I do my best work	1.42	0.60	High

Note. "--" indicates that data are not available because the item was removed prior to this analysis.

Based on a review of items, two domain areas had lower discrimination scores for their items: Emotional Regulation & Self-Concept. One item in Self-Concept had a discrimination value above a 1: "My teacher cares about me even when I make a mistake." However, other items provided less discrimination, especially in comparison to other domains. This may be in part due to the wording of these items being more complex than the items in different domains. Each item in this domain contained a qualifier, such as the item, "Other kids like me even if we sometimes argue." Also, the one item that was higher in discrimination may be influenced by a student's connectedness and optimism, domains which tended to have items with a greater degree of discrimination value. Similarly, the emotional regulation domain also contained only one item with a discrimination value over 1: "I can tell people how I am feeling." There are other items that reflect aspects of this statement such as, "There are many people I can talk to if I have a problem." Thus, neither domain was included in item selection for the final scale.

Items for further analysis were then taken from the remaining domain areas: self-regulation, social skills, school connectedness, social responsibility, and optimism. Twelve items were selected overall with three items each being contributed by the optimism and connectedness domains, and two each from self-regulation, social skills, and social responsibility. Social responsibility did have three items that received a classification of at least "high," but two of those items were quite similar ("I like to help other kids at school" and "I like to help kids when they are sad"). The first of those two

72

items was selected to be included in the final analysis as well as the following eleven

items:

I raise my hand when I have a question
I wait my turn in line
I like playing games even when I lose
I invite other kids to play with me
People at school care about me
Kids at school like me
There are many people I can talk to when I have a problem
I like to help my teacher
I do my best when I work
I like myself
My teacher notices when I do my best work

An analysis of the internal consistency was conducted. After reducing the scale

from the original 30 items down to these 12 items, Cronbach's alpha reduced from 0.83

to 0.74. Given a commonly accepted minimum value of 0.7 for reliable tests, the

reliability of the new 12-item scale remained adequate. Kline (2000) highlights this

criteria and notes that ten is the minimum number of test items for a reliable measure. It

is notable that the reliability of the 12-item scale is still acceptable given that there were

only 12 items with a three-point scale, conditions under which an adequate reliability is

difficult to achieve.

Behavior Group Comparisons

A proficiency score for the final 12-item screening tool was calculated by

summing the items, with each thumbs down response scored as 1, sideways as 2, and

thumbs up as 3. Thus, total possible scores ranged from 12-36. The mean proficiency

score for the overall sample was 32.81 ($SD = 3.19$). A 2 (school: suburban or rural) x 3

(grade: K, 1, 2) ANOVA was conducted to determine whether there were any differences

between these groups on their screening proficiency score. There was no significant main

effect for school type, $F(1,378) = 0.09$, $p = .77$, $\eta^2_p < .01$, grade, $F(2,378) = 0.02$, $p = .98$,

$\eta^2_p < .01$, or interaction between school and grade, $F(2,378) = 0.13$, $p = .15$, $\eta^2_p = .01$.

This suggests that there was no difference in performance between these different

populations on this measure of social-emotional wellness.

CHAPTER V

DISCUSSION

The goal of this research study was to begin development of a strength-oriented screening self-report tool that could be used to identify students at-risk for poor social-emotional development. Based on research showing increases in mental health concerns in adolescents, it is important to engage in efforts upstream to try to identify students at increased risk earlier in their development. However, previous studies have had difficulties in obtaining reliable information from younger children on self-report measures. An additional concern has been the investment that universal screenings can take with regard to resources, both financial and time for staff. In order for a screening measure to be a viable option it would have to demonstrate adequate psychometric properties, but also minimize the investment of time and resources for districts in order to be practical. Thus, the current study sought to fill these gaps by developing a social-emotional screening tool that could be feasibly administered to and completed by young children within schools.

Summary of Study Findings

The study began with a review of current literature on social-emotional rating scales that are currently available, with a particular emphasis on those that have demonstrated adequate psychometric properties. This review led to an identification of seven domain areas that are most frequently cited as being associated with positive social-emotional development. These domains were further explored and developmentally appropriate potential behaviors were identified that may be associated with these areas. The target developmental stage was kindergarten through second grade

75

students. Overall, 15 items were developed for each of the seven domains, leading to an initial total of 105 items for review. These items were created to be positively worded and behavioral in nature, as research has indicated that children may improve in the self-report rating when reporting on behaviors versus emotions (Watkins, 2008). These 105 items were then reviewed by kindergarten through second grade teachers who participated in a survey. These teachers gave input on the developmental importance of demonstrating these behaviors in school and also the likelihood that students would understand these questions if they were read aloud to them.

Thirty items were selected for the student pilot survey based on the feedback of teachers and a theoretically-guided review conducted by the principal investigator. These items were then presented to students in a format that attempted to maximize the ability to get reliable, accurate ratings from kindergarten through second grade youth. The survey form was designed to minimize visual tracking demands and the frequency of missed items for these younger students. Directions with practice items were included to aid student comprehension. There were also visual supports displayed in the classroom to assist student with their ability to accurately follow along as items were read. These data were collected through administration to entire classes, with class sizes ranging from 14 to 25 students across two schools in two different districts.

Items from this student pilot screening were analyzed using classical item analysis techniques, an exploratory factor analysis, and item response theory approaches. The top 12 items were then selected for between-group comparisons and to be utilized in future research studies to further develop the screening instrument.

Teacher Survey

The teacher survey provided information that allowed for the reduction of the initial item pool from 105 to 30 items. The sample size ($n = 12$) was much smaller than initially desired due to participating districts needing to withdraw from this portion of the study. This sample size did not allow for advanced analysis or any between-group comparisons related to how positive social-emotional development may vary based on region of the country, type of population served, or grade of students served.

Teachers ranked each item on how important they felt it was to include on a social-emotional screening tool. The top 30 ranked items tended to represent four of the seven domains: self-regulation, social skills, connectedness, and optimism. Items from the bottom 30 items disproportionately came from the domains of self-concept, emotional regulation, and social responsibility. The category with the highest ranking was self-regulation. This finding was consistent with past research such as that conducted by Rimm-Kaufman and colleagues (2000). In their national study, these researchers found that kindergarten teachers ranked the ability to follow directions as the number one priority area with regard to school readiness. This regulation skill was higher ranked than academic skills and self-concept items related to self-viewed proficiency in mathematics and reading, which is consistent with the findings of this study as well.

Although self-regulation items were highly rated by teachers, it was not the only domain area that contributed a large number of positively endorsed items. Social skills, optimism, and connectedness also had a number of highly endorsed items on the teacher survey, with few items in these domains receiving a rating of less than either "important" or "absolutely include." This is notable because it reflects a perception by teachers that

positive social-emotional development is more than just behaviors related to compliance. Instead, the results suggest that teachers prioritize students' feelings and relationships when it comes to their social-emotional growth. Furthermore, in support of the research outlined in the literature review for these other domains, although regulation skill development is important to engage in the learning process, these other areas are also key contributors to appropriate social, emotional, and academic development at these ages. Teacher item ratings were then compared to the principal investigator's ranking of items within each domain and a sample of the top-rated items from each domain was selected for the student pilot survey.

Student Pilot Survey

The goals of the student pilot survey were to examine the psychometric properties of the scale and its items, to further reduce the number of items, and to evaluate the scale's feasibility in terms of its administration and quality of youth self-reports.

In terms of the psychometric properties of the scale and its items, the reliability of the scale was assessed through Cronbach's alpha and was found to be adequate. The item-total correlation estimates demonstrated that most items were appropriately related to the overall proficiency score. However, one item was removed after this portion of the analysis and appeared to be a poorly written item. The remaining 29 items were included as part of an exploratory factor analysis. This analysis revealed that there was unidimensionality within the items and this factor was able to adequately explain 30.5% of the variance. This was despite the psychometric challenge of a Likert scale with only three response options.

An IRT approach was taken to review the items. Because items had polytomous and ordered response options, a graded response model was utilized. Parameter estimates and graphs were reviewed to evaluate individual items. These analyses revealed that items were adequate at a minimum within all domains, though the areas with the highest concentration of quality items were the domains of social responsibility, school connectedness, and optimism/positivity. Domains with the lowest number of quality items were self-concept and emotional regulation. These items provided less understanding and correlation to the total score than items from the other domains. Compared with the results from the teacher survey, the domains with the highest rated items were fairly consistent between the student and teacher surveys. The one exception was the social responsibility items, which performed better in the student pilot as compared to the teacher survey. Based on these results, final items for additional analysis were selected predominately from the social responsibility, connectedness, optimism, self-regulation, and social skills domains.

The self-concept domain contained items with lower ratings on both the teacher and student rating scales. In part, this may have been due to the wording of the items. Self-concept items tended to have qualifiers and/or were perhaps too complex. One such item was, "Other kids like me even if we sometime argue." For some students, they may have had a tough time conceptualizing that arguments happen at times, even between friends. However, students' lack of endorsement on this item may also be related to the item being influenced by at least two factors. Students who don't positively endorse this statement may feel as though they aren't well received by peers in general. Also, some students may be hesitant, as they don't view themselves as being argumentative in the

first place. Flahive, Chuang, and Li (2015) note that children's ability to make comparisons regarding interpersonal forms of self-concept appears to emerge around ages 8 to 11. Although some of the items for this scale were written based on an attempt to gather information related to self-esteem even in challenging situations, these types of items might not have been as developmentally appropriate.

Emotional regulation items tended to be simpler in their wording than self-concept items, but were also lower rated by both teachers and students. In part, this may be due to the behaviors within this domain occurring less frequently than items within the self-regulation domain. The average student may be willing to identify that "waiting their turn in line" or "raising their hand when they have a question" are areas for improvement, which are items in the self-regulation domain. However, they may be less willing to state that they engage in poking behaviors or cry as a reaction to different situations, items that were in the emotional regulation domain. Given that there was also a higher concentration of items in the emotional regulation domain that were reflective of negative behaviors, this may have affected the utility of the items in this domain. Moreover, there were also items within this domain that assessed information related to emotional state across different situations. Watkins (2008) found that emotions were more difficult to assess than behaviors in youth self-reports. Not surprisingly, items in the current study that were more emotional in nature were less reliable and provided less information related to proficiency than the positively-worded, behavioral items.

Based on these results, items from self-concept and emotional regulation domains were not part of the 12 items selected for final analysis. Items from the other five domains were reviewed with an emphasis on performance (factor loading, discrimination

value, visual inspection) and avoiding overly redundant items. The internal consistency of the final 12 items was then evaluated. Although reliability decreased after reducing the scale from 30 to 12 items, the reliability achieved on the 12-item scale is still considered to be adequate (Kline, 2000). With only 12 items and the utilization of a three-point scale for students this age, this still marked a positive step toward development of a useful screening tool.

This 12-point scale was also utilized to conduct between-group comparisons. Comparisons between grades (K vs. 1 vs. 2) and between schools (suburban vs. rural) were conducted based on overall proficiency scores on the screening tool. None of these comparisons demonstrated significant differences in any of the areas assessed. This was a positive finding with regard to the generalizability of the tool. It demonstrates continuity between two different schools with different types of populations, which is particularly notable given that norms for behavioral expectations may differ between schools. Moreover, when considering item selection this consistency in scoring from kindergarten through second grade lends support that this same scale would be appropriate to administer to all three of the grades assessed on this evaluation. Thus, this was an important first step in establishing generalizability for the scale.

Regarding the feasibility of the screening tool for implementation in early childhood classrooms, student survey data results were encouraging. Students in kindergarten through second grade demonstrated the ability to participate in a screening measure of this nature during a whole-group administration. Out of nearly 400 students, only one student's missing data rendered their score unusable. Otherwise, missing data were relatively rare and students were able to engage in the process of completing this

screening measure even in kindergarten classrooms. In one classroom, there was a student with significant disabilities and limited verbal communication who had a one-to-one aide. Notably, he was able to complete the survey with minimal prompts. The aide was overheard making a comment to the classroom teacher that the student was being very honest with his responses. Overall, with two adults typically in the room (classroom teacher and principal investigator), students were able to successfully engage in the completion of all items. With occasional prompts and repetition of items, students did reasonably well with keeping up. For older students in this survey, there were even times when they would rush ahead to complete the items. These students would be asked to wait for the rest of the group to ensure they did not misread an item. Thus, administration of the screening tool within kindergarten through second grade classrooms was considered a success.

Students also reacted positively to the screening tool. Students were observed to make statements such as, "I like this test," "when can we do this again," and "why can't I put my name on it, I want people to know how I feel." This seemed to further support the argument that there is a real need to attempt to capture the feelings and perspective of students, even at this younger age. These younger students were engaged, but also enjoyed being able to complete a survey of this nature. Current practices leave little opportunity to provide a structured avenue for students to self-reflect and indicate a need for additional support in the areas of social-emotional development. The importance of improving behaviors that may impact learning, building positive relationships with peers and adults, and feeling a sense of purpose and importance at school does not start at third

grade. However, a majority of self-reporting tools start around third grade at the earliest. It appears that current approaches are missing out on a critical intervention point.

The task of developing appropriate universal screening tools to utilize in early social institutions (i.e., schools) has been a goal in psychology research for more than 50 years (Cowen, 1973). These researchers recognized the need to "repair rooted dysfunction" in young children in order to prevent later disorder from developing. However, due to challenges related to cost (Kampaus, 2012) both in terms of time and materials, there has not been adequate progress at addressing this significant gap. Eklund and colleagues (2009) found that at least 50% of students in their study that would have self-identified as needing additional social-emotional supports were missed using traditional referral approaches. Students at these early grades seem to be seeking the chance to share how they feel about school and their own development, but due to costs and lack of appropriate tools, schools are falling well short of addressing this need. The screening tool in the current study may be a useful avenue for schools to bridge this gap.

Limitations and Future Directions

This study marked an important first step in determining whether children in kindergarten through second grade could appropriately engage and provide reliable information on a whole-group administered survey of social-emotional development. However, as a first step, a limitation of the current study is that it could not establish validity of the screening tool. The logical next step is to systematically collect and review information related to the validity of the final items. Establishing validity is particularly important as this is a primary concern of research on the accuracy of children as self-raters.

In considering validity concerns there are a few areas that would be important to consider moving forward. Messick (1995) outlines six considerations for educational and psychological test development with regard to construct validity. In this model of validity review, Messick posits that validity comes not just from measures of statistical properties, but also the social values associated with the measure of interest. His model asserts that consideration should be given to these areas: content, substantive, structural, generalizability, external, and consequential.

Content refers to the ability of a measure to be representative of all construct domains. In addition to content, the substantive aspect of validity is the gathering of information related to testing consistencies in response and whether the test is designed to keep participants engaged. The structural aspect of validity refers to the scoring and rubric development for a test and whether it is appropriate for that particular construct domain. Generalizability refers to scores and whether the construct is applicable across settings and populations. External validity is related to criterion validity and whether the tool is convergent with appropriate measures and discriminant from opposing measures. Lastly, the consequential aspect of validity assesses what low or high scores on a test may be associated with. This can be either positive or adverse consequences.

The substantive aspect of validity for the screening tool would be an especially fruitful avenue to explore. A review of test-retest reliability to see the stability of these ratings for students over time would be a helpful next step. Younger children have been noted to potentially be more state-specific in their self-ratings as opposed to trait-specific. For instance, a student's self-reflection on social skills may be more driven by their last period of recess as compared to a general assessment of recess periods overall. Being

able to compare scores over a period of a few weeks to determine stability without the implementation of an intervention would be an important measure to give information related to this concern area.

Another aspect of validity that would be important to assess for this screening tool in the future is criterion validity. A study could be conducted to assess how a subset of students completing the self-report measure compares to a teacher or parent rating of those students. Measures to assess consequential validity will also be important to examine the relationship between the tool and various outcome measures (attendance, discipline, grades, teacher ratings of friendship formation, etc.). Students could be given this screening toward the beginning of the year and a review could occur at the end of the year to determine whether a lower performance on the screening tool predicts negative outcomes for those students. Alternatively, for students that demonstrate a higher proficiency, demonstrating a relationship with more positive health and school outcome measures would provide useful support for this area of validity. An added advantage of assessing teacher observations of students' social-emotional learning is that it could be compared to the scores on the screening tool for purposes of assessing the sensitivity of the screening tool. Specifically, it could be determined whether the screening scores differentiate students across the spectrum of social-emotional learning or whether the tool is most useful in differentiating the most high-risk students from typically developing social-emotional learners. This would then further support the argument that the tool provides useful screening information that could allow for potential early intervention.

Another limitation of this study was the lack of diversity in the overall sample; which negatively impacted its generalizability. This group was rather homogenous with

regard to race/ethnicity. Schools' Caucasian populations accounted for approximately 90% of the sample in one district and 84% of the sample for the other district. A particular group that was underrepresented was African-American youth. Another limitation for generalizability was the lack of diversity of geographic regions. In the area of social-emotional learning, the importance of reflecting cultural norms and expectations is quite important. Identifying behaviors and positive qualities that are universal and valued as important in development across a more diverse sample is paramount. The current study could not accomplish that aim and is something that needs to be addressed in order to make any recommendations related to the generalizability of this study's findings.

Applications in the Field

A positive contribution of this scale is the potential to identify students that are at heightened risk for later internalizing concerns. In both the teacher and student surveys, items of value came from multiple domain areas. Although teachers have noted the ability to follow directions is important for school readiness (Rimm-Kaufman et al., 2000), early compliance behaviors are not the whole picture of positive development. Similar to this research, a heightened portion of the top ranked items by teachers tended to come from the self-regulation item pool. Although regulation skill items ultimately did provide useful information for this scale, this area should not be the sole or even primary focus of a measure like this one based on the student responses. In fact, these external behaviors are already considered to be important by teachers and appear to be an area of observational focus. In thinking of this from an RTI perspective and what unique information a measure like this may provide, this contribution is important.

Results from the student pilot data found items from areas such as connectedness, social responsibility, and optimism to have the greatest number of items that provided high or very high discrimination values. This may be a key value of this survey, if the final measure is ultimately able to identify students with particular struggle in these areas such as feeling positive connections, a sense of purpose, and a more positive cognitive framework when responding to challenges. In thinking about identifying students with internalizing concerns (depression, anxiety, future suicide ideation/attempts, etc.), these items may provide early insight as to protective factors that if addressed at these young ages may potentially prevent the occurrence of these more significant mental health problems. This more proactive approach seems to be emerging in the literature as being more likely to have an impact on preventing these significant mental health concerns than even more targeted interventions during adolescence (Wilcox et al., 2008). These ratings would provide potential insight into areas that are more likely to be missed on current tools used in screening (teacher rating scales, office discipline referral information, etc.) as they are less observable, particularly at these early stages.

Another potential strength of this survey refers to the ability to guide treatment decisions. The screening instrument showed unidimensionality and thus it would not be prudent for this particular tool to be broken down into domain subcategories. However, when a student is identified as at-risk based on their overall proficiency, a review of items may give an indication of areas of potential concern. A follow up interview with the student or an additional self-rating measure with an expanded item pool may be a next step. With items coming from different domain areas, particular areas of weakness can provide useful information as to which interventions may be most effective in building

social-emotional health. Intervention strategies for young children with regulation deficits should look different than those that are struggling with difficulties stemming from an area such as connectedness.

From the practicality of implementation standpoint, this tool was quite promising with regard to its cost for a district. Within this study, first and second grade classrooms were able to complete this 30-item survey in approximately 10-15 minutes. Kindergarten classrooms took slightly longer with administration time taking up to 20 minutes. However, if the survey items were to be reduced to the suggested 12 items, it is conceivable that completion time for kindergarten students would also be under 15 minutes. In addition, students would become more familiar with this process over time if used as a repeated screening measure and thus directions and time for needed for clarification may also reduce this time.

NCES (2001) estimated the average primary school (pre-k through 3rd grade) in the United States to be 446 students. This works out to around 100 students per grade. Scoring of the pilot survey was roughly one minute per survey. After some initial set up of a data system, most schools would be able to collect kindergarten through second grade screening data with one staff member (psychologist, counselor, social worker, etc.) and enter that data within one day if this staff member was given release time. This would not incur nearly the cost that has been estimated in past research on the investment for screenings within social-emotional domain (Kamphaus, 2012).

Although a strength of the potential screening tool is its ease of implementation and practicality, there are some potential avenues for further improvement. The ability to utilize technology to answer questions could prove to be quite valuable. In one of the

schools in the study, there had been a recent one-to-one technology initiative. Projections

have shown a rapid increase in the frequency of students and staff in public schools that

are assigned their own personal computing device. There was an increase from 23 percent

in 2012 to 54 percent by 2016 (Molnar, 2015). With more student familiarity, the ability

to put this survey online that then linked to a data management and analysis program

would alleviate one of the primary costs of the survey with regard to scoring time. This

would all be predicated, however, on the ability of students to represent a similar degree

of reliability on their ability to answer these same questions online, using an appropriate

polling app or software.

<u>Summary</u>

Overall, this study presents as an important first step in addressing a current void

in social-emotional screening assessment of youth. This study demonstrates that children

can engage in a whole class screening assessment of social-emotional development on a

measure lasting fewer than 20 minutes. Nearly all items administered were found to have

adequate measures of reliability related to the overall test score. Areas such as

connectedness, social responsibility, and optimism appear to be areas that provide useful

information and assists in the discrimination between students that demonstrate higher

levels of proficiency on this measure. This screening tool may provide useful treatment

information for early intervention approaches that may reduce risks of later mental health

concerns, particularly those that are more internalizing in nature. Yet, much work is still

needed before this tool can be endorsed as appropriate at meeting these goals. A re-

administration of this scale and steps to ensure the final measure's validity and

generalizability will be critical prior to implementation recommendations.

TEACHER SURVEY CONSENT

Study of Social-Emotional Development of Kindergarten-2nd Grade Students
Consent for Voluntary Participation (Teacher Survey)

Facts about this project:

The is Part 1 of a two-part study

Purpose: The purpose of this study is to develop a brief rating scale that kindergarten through second grade students could complete that assesses their social-emotional development. This teacher survey will assist in the development of the items for this screening tool based on the behaviors you see in your classroom and whether you feel the item's language could be understood by your students.

Incentives: Participating school districts will receive a copy of the results of the study and a copy of the rating scale and an excel scoring sheet (if one of value is developed) based on teacher and student input of this district and one other. The rating scale may be helpful in screening of students in need of additional supports or interventions in developing appropriate social-emotional skills. In addition, all teachers completing the survey within your district will be entered to win one of two gift cards of $25 to Barnes and Noble. Also, the first ten teachers to complete this survey will receive a $5 dollar gift card to Barnes and Noble.

Survey format: The teacher survey that you are being asked to complete can be filled out online, using qualtrics.com, or a paper-and-pencil version is available. The survey asks teachers to rate potential rating scale items on how important it is that items be included in a screening tool for social-emotional development.

Confidentiality: No names or identifying information will be reported on the survey. Email addresses will be compiled only to select winners for raffle prizes. The only demographic information that will be collected is the grade and district in which you work to allow for between group comparisons.

You have the right to withdraw from part or all of the study at any time. Your participation is voluntary and a decision not to participate will have no negative consequences for you.

Your informed consent to participate in the study under the conditions described above is assumed by your completing the survey and submitting it to the researcher. Do not complete the survey or submit it if you do not understand or agree to these conditions.

If you have any questions about the project, please contact me at:
James Brenchley
(607) 244-0407
jbrenchl@educ.umass.edu

You can also contact my dissertation chair, Dr. Sara Whitcomb, at
swhitcomb@educ.umass.edu.

APPENDIX B

PARENT CONSENT FORM

Parent Consent Form
Requesting Student Participation in a Youth Survey

Dear Parent:

We are asking permission for your child to participate in a survey that will be administered in <insert date>. All students in the kindergarten through second grade throughout the district are being invited to participate.

The purpose of the survey is to assist in the development of a screening tool that would allow students to self-report on their social-emotional development. This tool may assist the district in being able to identify students or groups of students who may feel they need additional support or instruction in this area. The school will receive a report presenting the results of the survey that can be used to examine current social-emotional development of students and provide a potentially useful tool to assist student service delivery.

The survey is entirely anonymous. Students will not put their names or any other identifying information on the survey booklet. All results from the study will be presented only in group summary form, like many opinion polls. There is a copy of the questionnaire in the principal's office, if you wish to review it.

Your child's participation in the survey is completely voluntary. There are no costs or risks to your child in completing the questionnaire. Each child will be given the option of leaving blank any question that he or she prefers not to answer. You may decline to have your child participate, if you wish. If you do decline, your child will be allowed to read or participate in an alternative activity while the survey is being administered.

The survey is being conducted by James Brenchley, a doctoral student from UMass Amherst. If you have any questions regarding the study, you may email him (jbrenchl@educ.umass.edu) or call (607) 244-0407.

Please check the box below if you DO NOT want your child to participate in the study and send the letter back to the school.

☐ I DO <u>NOT</u> want my child to participate in the study.

_____   _____
Parent's Name   Child's Name

INITIAL ITEM POOL

Self-Regulation	Emotional Regulation
1. I listen carefully to the teacher	1. I use my words to tell someone if I'm angry
2.I get my work done when I'm supposed to	2. I don't cry in class
3.My work is not messy	3. I tell people that I'm happy
4. I wait my turn in line	4. I don't cry at recess
5. I am quiet in the hallways	5. I use my words to tell someone if I'm upset
6. I don't poke other kids	6. I smile a lot
7. I don't hit other kids	7. I laugh a lot
8. I raise my hand when I have a question	8. I don't cry when it's time to come to school
9. I stay in my seat when I'm supposed to	9. I can tell people how I am feeling
10. I can sit and listen to a whole story without getting up	10. I don't break toys when I get angry
11. I can want a toy or game without grabbing it from others	11. I don't hit or kick the wall or desks when I get upset
12. I sit in my seat when I'm on the bus	12. I can tell someone I'm upset without yelling
13. The work I give to teachers is always my best work	13. I don't yell at people
14. If other kids are talking when they aren't supposed I can still get my work done	14. I only get sad for a little bit of time
15. If an activity gets cancelled I don't complain	15. If the class doesn't get the full time for recess, I am not angry
Social Skills	**Self-Concept**
1.I like to share my toys	1. My teacher cares about me even when I make a mistake
2.I like talking with kids in my class	2. My family cares about me
3.I want more friends	3. Other kids like me even if we sometimes argue
4.I can join in games other kids are playing	4. People like me even when I'm having a bad day
5. I like the kids I sit with at lunch	5.I do as well as other kids on my work
6.I like learning about kids in my class	6. I feel included by my friends during recess
7. I take turns	7. I am a good reader
8.I like playing games even when I lose	8. I am as good a friend as other kids my age
9. I let other kids pick the games we played during recess	9. I am good at math
10. I invite kids to play with me	10. I do a lot more good things than bad things
11.When I ask kids to play with me they say yes	11. I like to learn new games even if they seem hard at first

12.Other kids ask me to play with them	12. I am not great at every game I try
13. I like to come up with new games to play at recess	13. I like how I look
14.I notice when other kids are getting upset	14. I can do a lot of things without help from adults
15. I usually know why kids are upset in school	15. When other kids are playing a game with me, I want them to do their best
School Connectedness/Belonging	**Social Responsibility**
1. My teacher likes me	1. I like to learn
2. I like coming to school	2. I want to make school better
3. I feel important at school	3. I like to help my teacher
4. Kids at school like me	4.I like to help other kids at school
5. People at school care about me	5. I clean up any games or toys after recess without the teacher telling me to
6. I fit in at school	6. I clean up after lunch
7.I have lots of fun at recess	7. I like to help kids when they are sad
8. Teachers are always saying good job to me	8. I like to help other kids if they are angry
9. Kids think I do a good job at things	9. I get a teacher if kids are arguing
10. Other Kids don't try to hurt my feelings	10. I get an adult if kids are fighting
11. Kids want to be my friend	11. I like it when the teacher gives me jobs
12. I have enough friends	12. I like to think about how to make school better for everyone, not just me
13. There are many people I can talk to if I have a problem	13. I am a good listener to other kids
14. School is wonderful place	14. I can be friend with a kid that others say they don't like
15. People are happy at school	15. I ask kids to play with me who look different than me
Optimism/Positivity	
1. I work hard at school	
2. I do my best when I work	
3. I am a good kid	
4. I am special	
5. I am smart	
6. Good things happen to me	
7.Teachers are helpful	
8. I like other kids	
9. I like myself	
10. I don't get upset when I lose	
11. I am kind	
12. Other kids want me to do well in school	
13. Other kids will let me play with them if I ask	
14. My teacher notices when I do my best work	
15. Kids in my class are fair when we play games	

APPENDIX D

TEACHER SURVEY

Please review the following potential items for a social-emotional development survey that students in kindergarten through second grade will complete. Questions need to be reflective of behaviors that you feel are most associated with positive social-emotional development in the students you work with and are also items that your students would understand when read to them. Please review the following statements and rank them on a scale of 1 (Do not include) to 5 (Absolutely Include) on how important you feel they are to include in a youth self-report survey.

	Do not include	Little Importance	Somewhat Important	Important	Absolutely Include
Student can tell people how they are feeling	1	2	3	4	5
Student agrees they like other kids	1	2	3	4	5
Student take turns	1	2	3	4	5
Student doesn't hit other kids	1	2	3	4	5
Student works hard at school	1	2	3	4	5
Student thinks school is wonderful place	1	2	3	4	5
Student agrees they like themselves	1	2	3	4	5
Student feels included by friends during recess	1	2	3	4	5
Student likes playing games even when I lose	1	2	3	4	5
Student feels the kids at school like me	1	2	3	4	5
The student feels the kids at school care about the student	1	2	3	4	5
Student notices when other kids are getting upset	1	2	3	4	5
Student likes learning about kids in class	1	2	3	4	5
Student doesn't hit or kick the wall or desks when they get upset	1	2	3	4	5
Student likes talking with kids in my class	1	2	3	4	5
Student think people are happy at school	1	2	3	4	5
Student feels they are doing as well as other kids on their work	1	2	3	4	5
The student feels they fit in at school	1	2	3	4	5
The work student gives to teachers is always their best work	1	2	3	4	5
Student thinks they are as good a friend as other kids their age	1	2	3	4	5
Student doesn't cry in class	1	2	3	4	5
If the class doesn't get the full time for recess, student is not angry	1	2	3	4	5
Student doesn't get upset when they lose	1	2	3	4	5
Student feels kids in their class are fair when they play games	1	2	3	4	5
Student doesn't yell at people	1	2	3	4	5
Student likes to help kids when they are sad	1	2	3	4	5
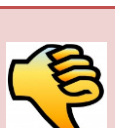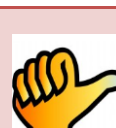
	Do not include	Little Importance	Somewhat Important	Important	Absolutely Include
Student likes to share their toys	1	2	3	4	5
Student wants to make school better a better place	1	2	3	4	5
If an activity gets cancelled student doesn't complain	1	2	3	4	5
Student waits their turn in line	1	2	3	4	5
Student feels other kids like me even if we sometimes argue	1	2	3	4	5
Student doesn't cry at recess	1	2	3	4	5
Other kids ask the student to play with them	1	2	3	4	5
Student uses their words to tell someone if I'm upset	1	2	3	4	5
Student likes to learn	1	2	3	4	5
Student likes to learn new games even if they seem hard at first	1	2	3	4	5
Student doesn't poke other kids	1	2	3	4	5
Student agrees they can be a friend with a kid that others say they don't like	1	2	3	4	5
Student feels there are many people they can talk to if they have a problem	1	2	3	4	5
The student feels important at school	1	2	3	4	5
Student invites kids to play with them	1	2	3	4	5
Student thinks kids want to be my friend	1	2	3	4	5
Student listens carefully to the teacher	1	2	3	4	5
Student thinks they are a good reader	1	2	3	4	5
Student can join in games other kids are playing	1	2	3	4	5
Student has lots of fun at recess	1	2	3	4	5
Student likes it when the teacher gives them jobs	1	2	3	4	5
When student asks kids to play with them, they say yes	1	2	3	4	5
Student feels they are special	1	2	3	4	5
Student agrees that teachers are always saying good job to me	1	2	3	4	5
Student lets other kids pick the games played during recess	1	2	3	4	5
Student thinks their teacher likes them	1	2	3	4	5
Student gets their work done when they are supposed to	1	2	3	4	5
Student likes to come up with new games to play at recess	1	2	3	4	5
Student gets an adult if kids are fighting	1	2	3	4	5
Student thinks they are a good listener to other kids	1	2	3	4	5

	Do not include	Little Importance	Somewhat Important	Important	Absolutely Include
If other kids are talking when they aren't supposed student can still get their work done	1	2	3	4	5
Student thinks they have enough friends	1	2	3	4	5
Student likes to help other kids if they are angry	1	2	3	4	5
Student likes the kids they sit with at lunch	1	2	3	4	5
Student agrees they do their best when they work	1	2	3	4	5
Student raises hand when they have a question	1	2	3	4	5
Student likes to help other kids at school	1	2	3	4	5
The student cleans up after lunch	1	2	3	4	5
Student can sit and listen to a whole story without getting up	1	2	3	4	5
The student likes coming to school	1	2	3	4	5
Student knows they are not great at every game I try	1	2	3	4	5
Student smiles a lot	1	2	3	4	5
Student agrees that when other kids are playing a game with the student, the student wants them to do their best	1	2	3	4	5
Student doesn't break toys when angry	1	2	3	4	5
Student feels the teacher notices when the student does their best work	1	2	3	4	5
Student agrees they are smart	1	2	3	4	5
Student feels teachers are helpful	1	2	3	4	5
Student can do a lot of things without help from adults	1	2	3	4	5
Student likes to help their teacher	1	2	3	4	5
Student likes how they look	1	2	3	4	5
Student feels their family cares about them	1	2	3	4	5
Student stays in seat when supposed to	1	2	3	4	5
Student tells people that they are happy	1	2	3	4	5
Student feels teacher cares about them even when the student makes a mistake	1	2	3	4	5
Student can tell someone they are upset without yelling	1	2	3	4	5
Student feels they are good at math	1	2	3	4	5
Student laughs a lot	1	2	3	4	5
Student asks kids to play with the student who look different than the student	1	2	3	4	5
Student thinks other kids don't try to hurt my feelings	1	2	3	4	5
Student likes to think about how to make school better for everyone, not just themselves	1	2	3	4	5

STUDENT SURVEY

1	👍	👎	👍	I like myself
2	👍	👎	👍	I raise my hand when I have a question
3	👍	👎	👍	I yell at people
4	👍	👎	👍	I invite kids to play with me
5	👍	👎	👍	I can do a lot of things without help from adults
6	👍	👎	👍	People at school care about me
7	👍	👎	👍	I like to help kids when they are sad
8	👍	👎	👍	My teacher notices when I do my best work

9	👍	👎	👊	I poke other kids
10	👍	👎	👊	When I ask kids to play with me they say yes

1	👍	👎	👊	I listen carefully to the teacher
2	👍	👎	👊	I use my words to tell someone if I'm angry
3	👍	👎	👊	I can join in games other kids are playing
4	👍	👎	👊	My teacher cares about me even when I make a mistake
5	👍	👎	👊	I like coming to school
6	👍	👎	👊	I like to help my teacher

7	👍	👎	✊	I do my best when I work
8	👍	👎	✊	I wait my turn in line
9	👍	👎	✊	I cry when it's time to come to school
10	👍	👎	✊	I like playing games even when I lose

1	👍	👎	✊	Other kids like me even if we sometimes argue
2	👍	👎	✊	There are many people I can talk to if I have a problem
3	👍	👎	✊	I like to learn
4	👍	👎	✊	I get upset when I lose

5	👍	👎	👍	I get my work done when I'm supposed to
6	👍	👎	👍	I can tell people how I am feeling
7	👍	👎	👍	I take turns
8	👍	👎	👍	I feel included by my friends during recess
9	👍	👎	👍	Kids at school like me
10	👍	👎	👍	I like to help other kids at school

APPENDIX F

ITEM CATEGORY RESPONSE FUNCTION GRAPHS

Item 13

Item 14

Item 16

Item 17

Item 18

Item 19

Item 20

Item 21

Item 22

Item 23
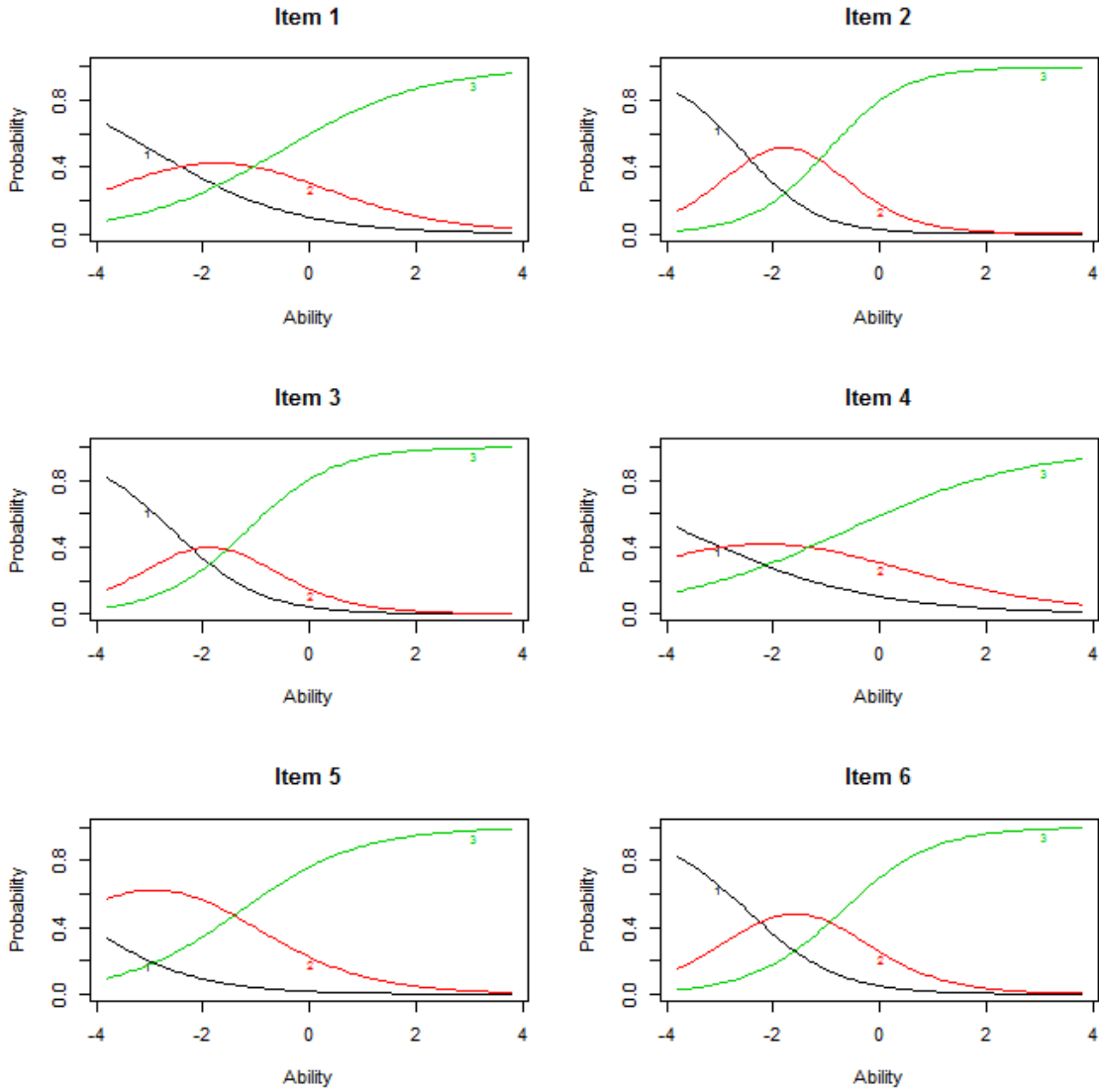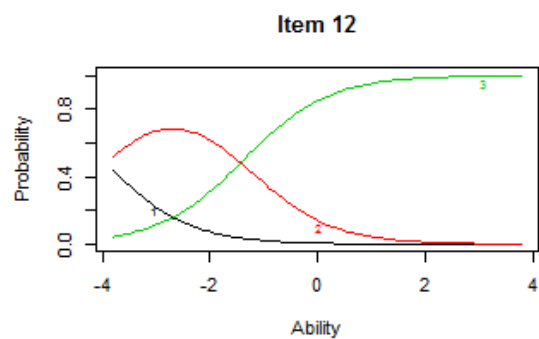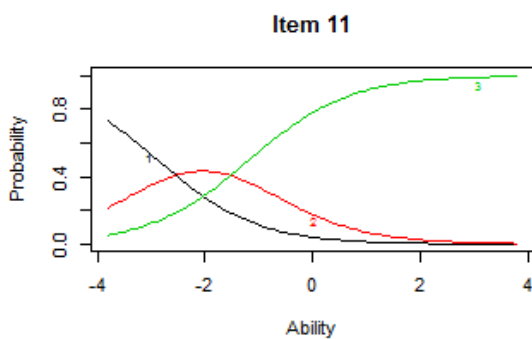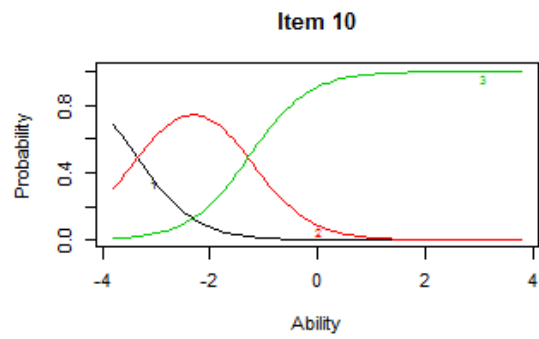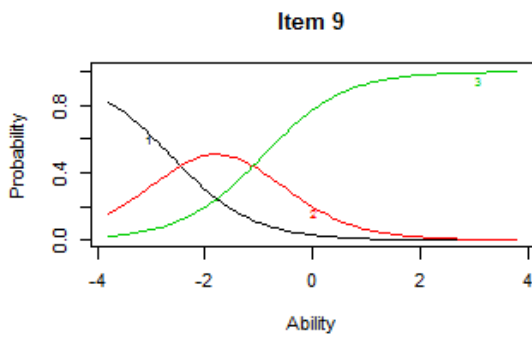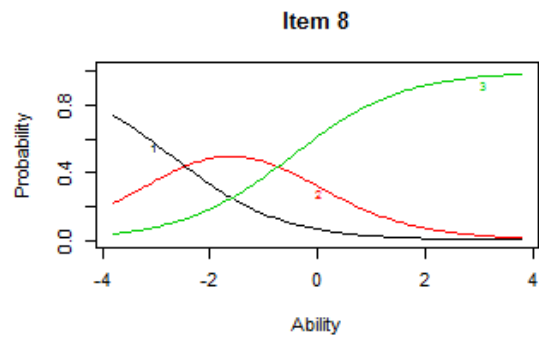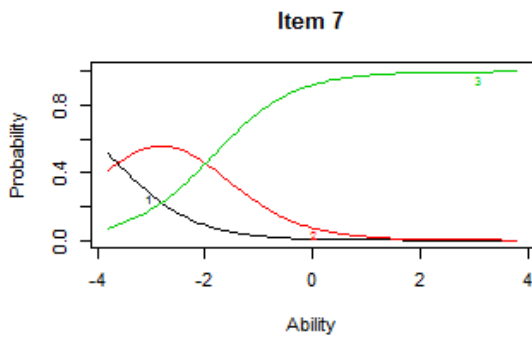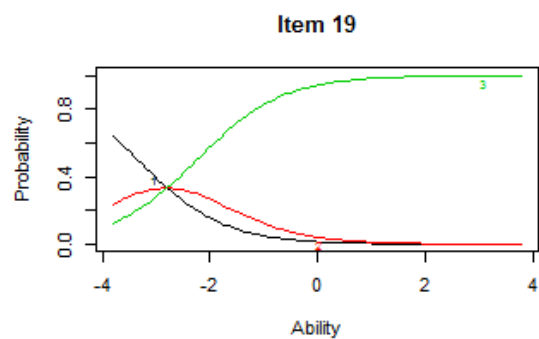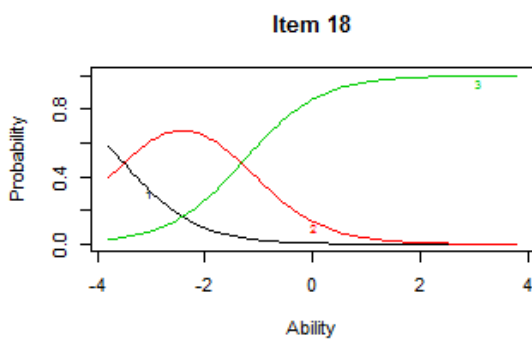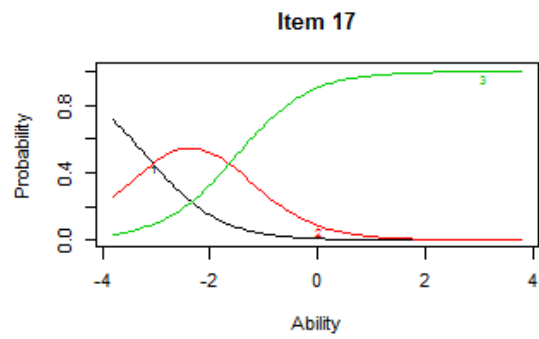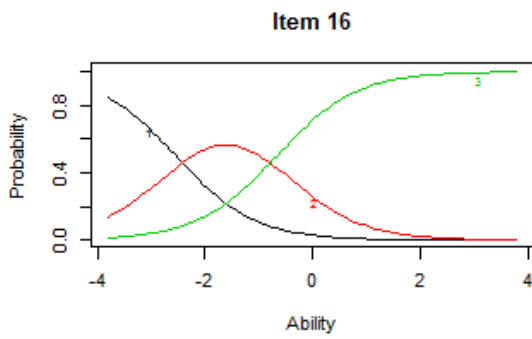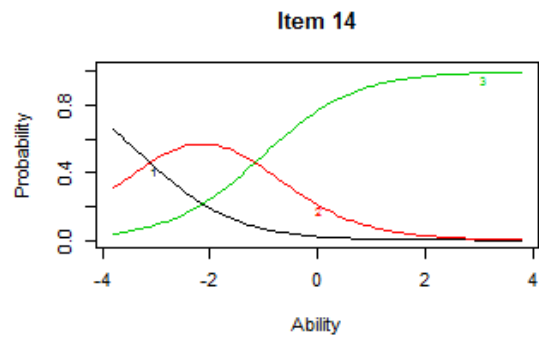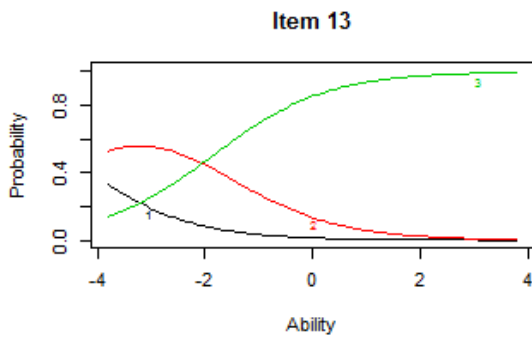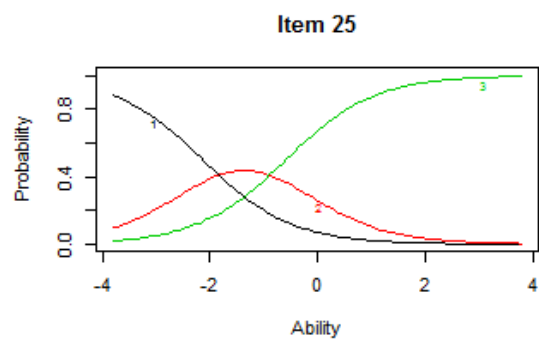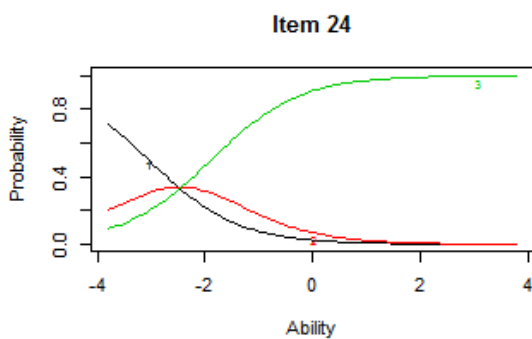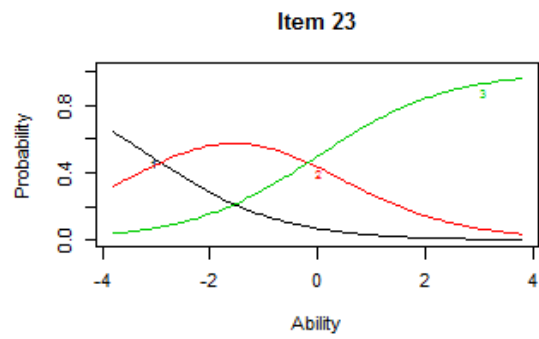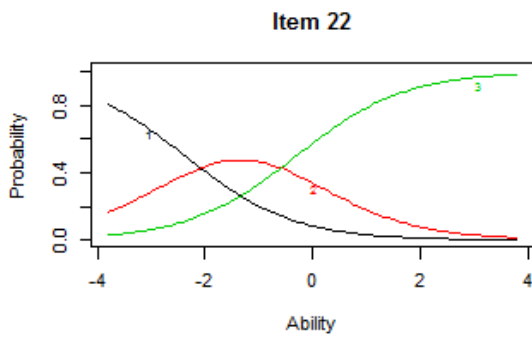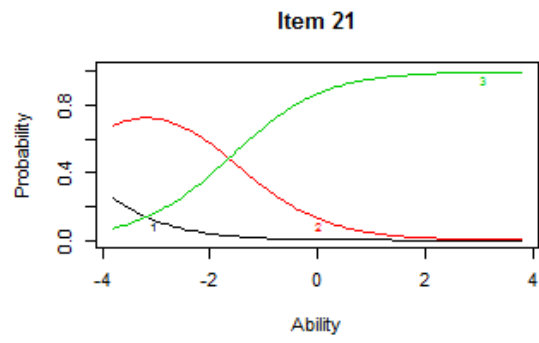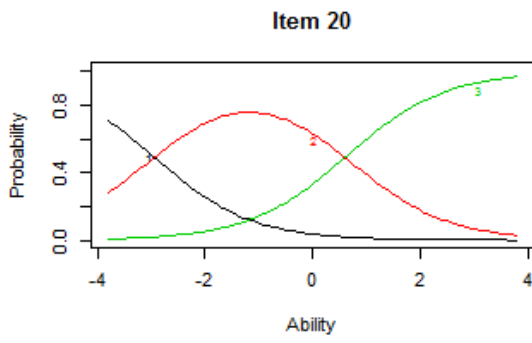
Item 24
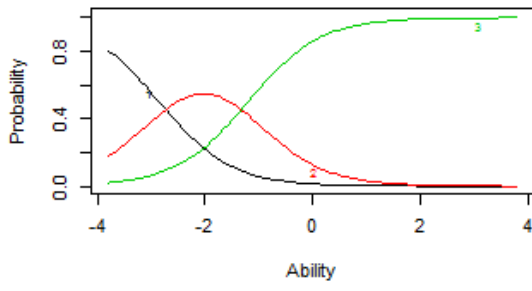
Item 25
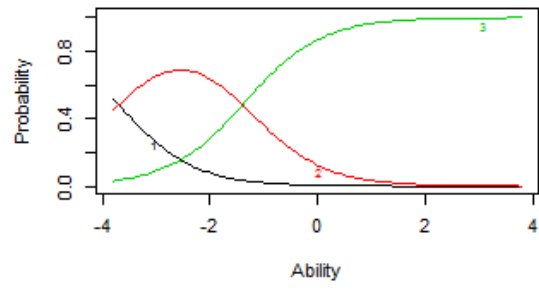
Item 26
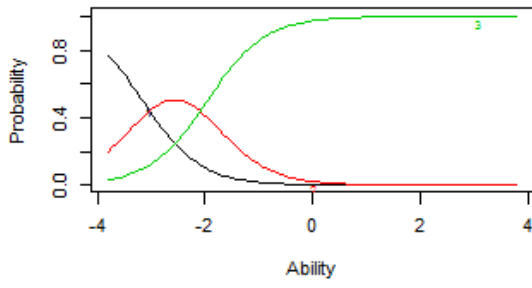
Item 27
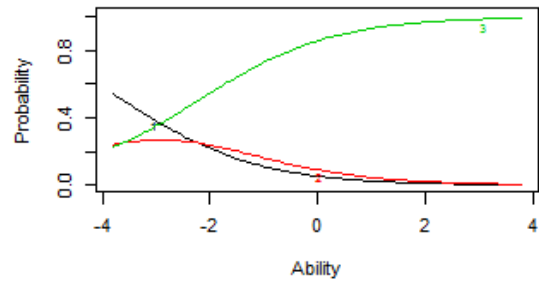
Item 28
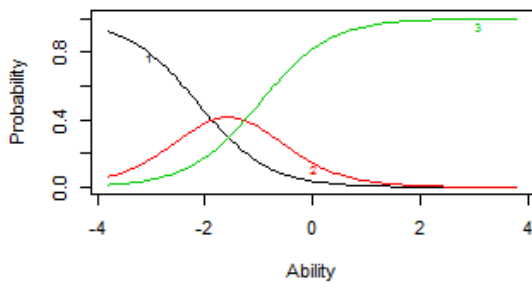
Item 29

Item 30

APPENDIX G

ITEM INFORMATION CURVES GRAPHS

107
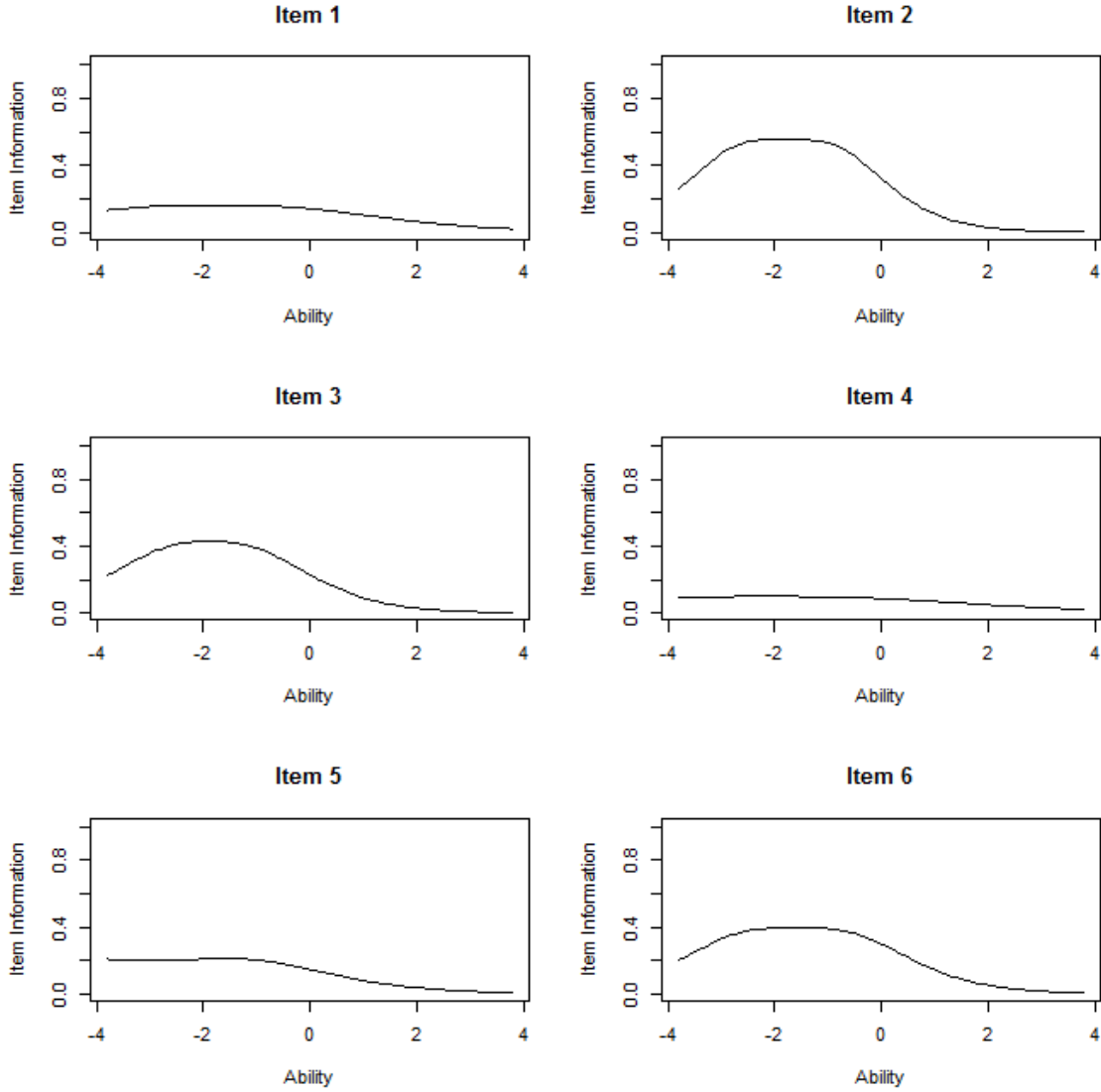
Item 7
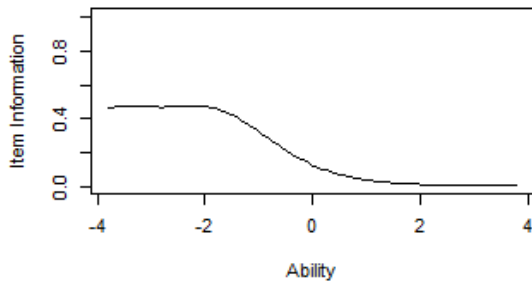
Item 8
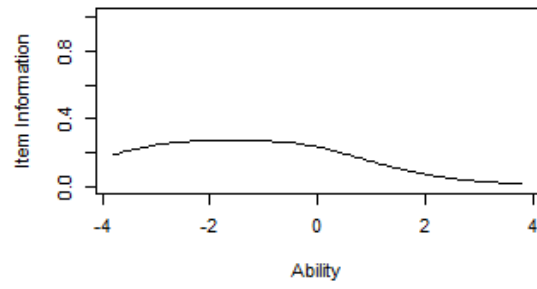
Item 9
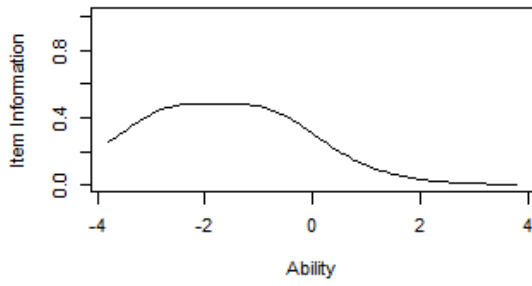
Item 10
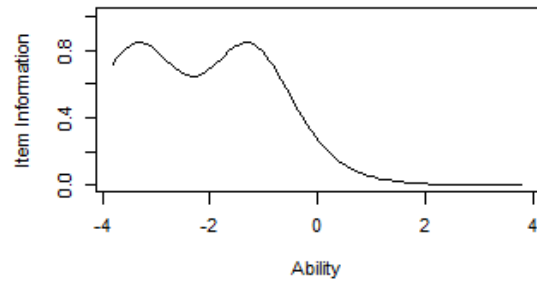
Item 11
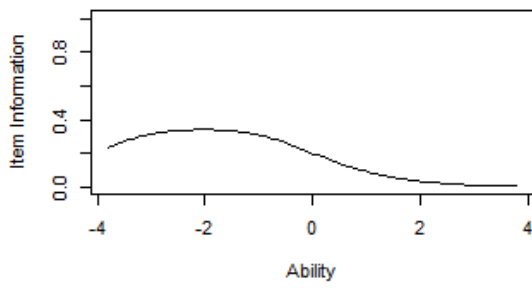
Item 12
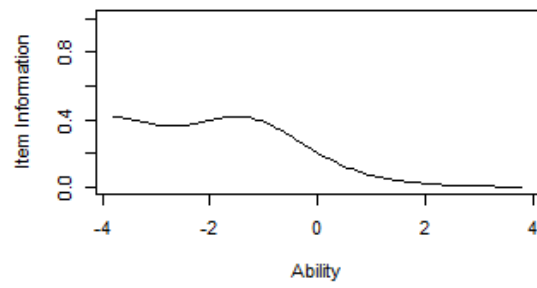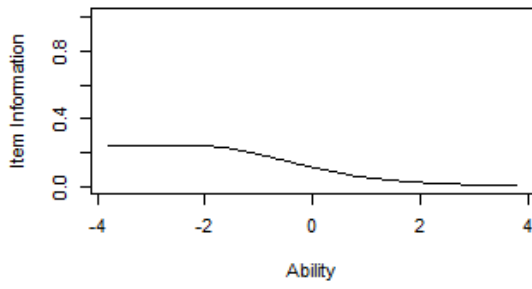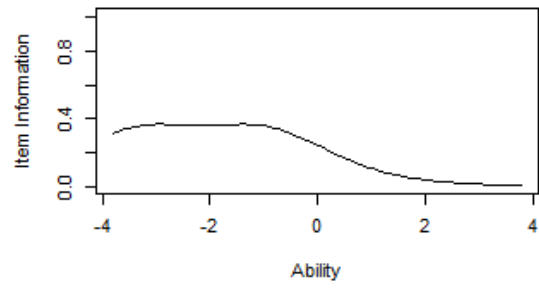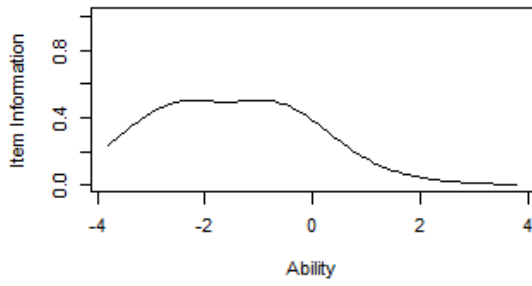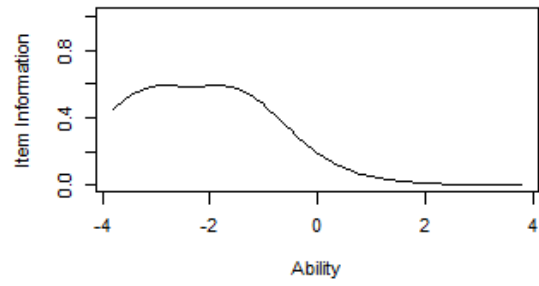
108

Item 13

Item 14

Item 16
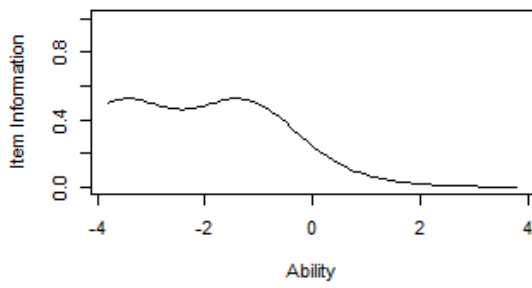
Item 17

Item 18

Item 19
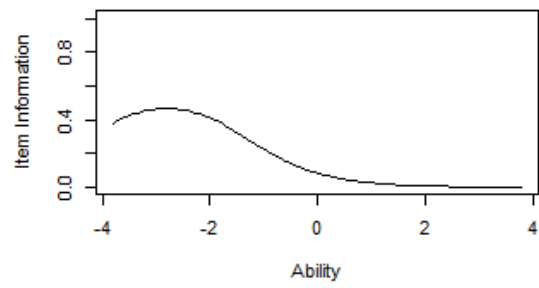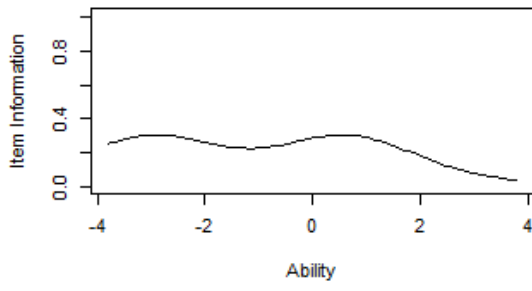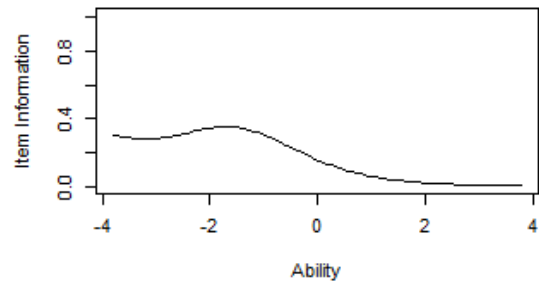
Item 20
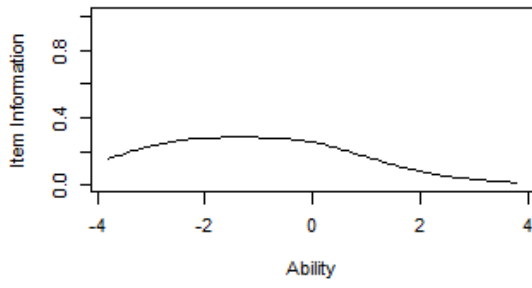
Item 21
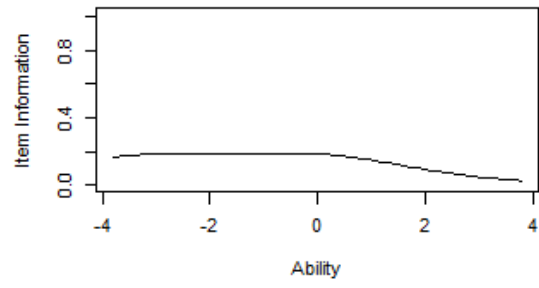
Item 22
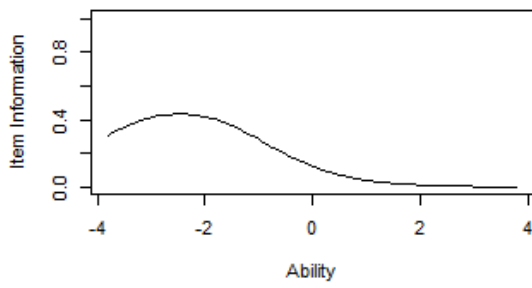
Item 23
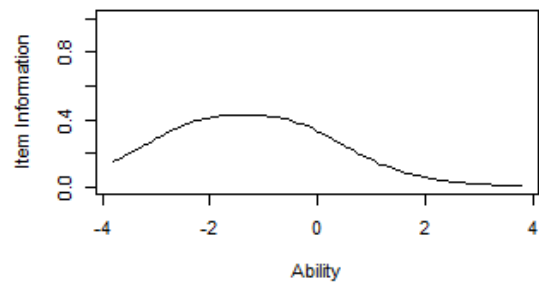
Item 24
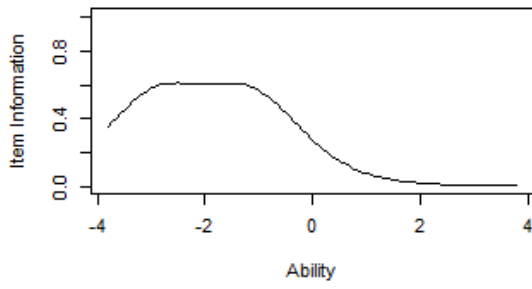
Item 25
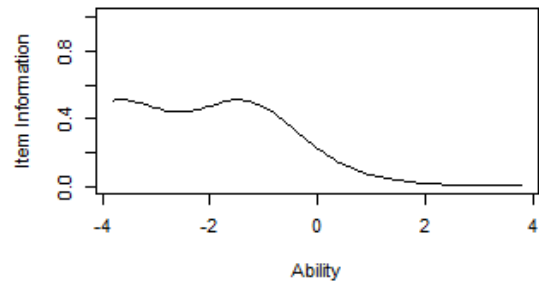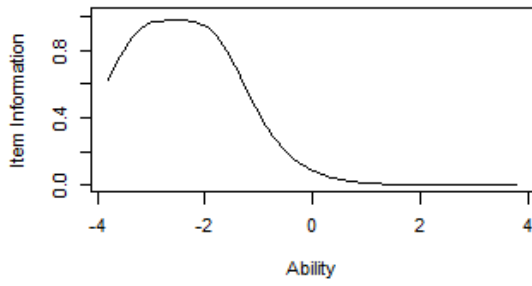
Item 26

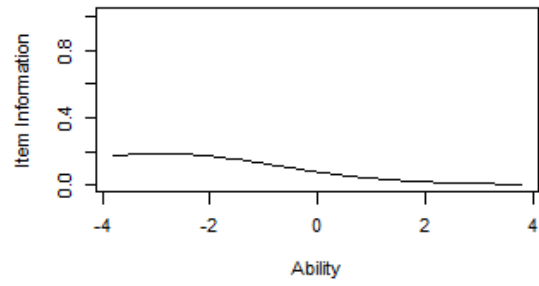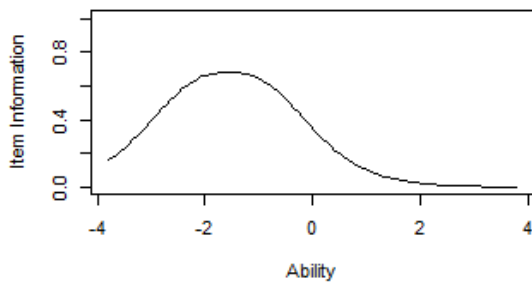Item 27

Item 28

Item 29

Item 30

REFERENCES

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85–106). New York: Psychology Press.

American Psychological Association, Task Force on Evidence-Based Practice for Children and Adolescents. (2008). *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care.* Retrieved from https://www.apa.org/practice/resources/evidence/children-report.pdf

Angold, A., Costello, E.J., Messer, S.C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, 5, 237–249.

Argyle, M. (2008). *Social encounters: Contributions to social interaction*. Piscataway, N.J.: Aldine Transaction.

Ashdown, D. M., & Bernard, M. E. (2012). Can explicit instruction in social and emotional learning skills benefit the social-emotional development, well-being, and academic achievement of young children?. *Early Childhood Education Journal*, *39*(6), 397-405.

Ayala, R. J. D. (2009). *The theory and practice of item response theory*. New York: Guilford Press.

Baker, F. B. (2001). *The Basics of Item Response Theory. Second Edition*. Retrieved from ERIC database. (ED458219)

Barber, B. K., & Schluterman, J. M. (2008). Connectedness in the lives of children and adolescents: A call for greater conceptual clarity. *Journal of Adolescent Health*, *43*(3), 209-216.

Bardi, A., Lee, J. A., Hofmann-Towfigh, N., & Soutar, G. (2009). The structure of intraindividual value change. *Journal of Personality and Social Psychology*, *97*(5), 913-929.

Baumeister, R. F. (1999). Self-concept, self-esteem, and identity. In V. J. Derlega, B. A. Winstead, W. H. Jones, V. J. Derlega, B. A. Winstead, & W. H. Jones (Eds.), *Personality: Contemporary theory and research* (2nd ed., pp. 339-375). Chicago, IL, US: Nelson-Hall Publishers.

Becker, D. R., Miao, A., Duncan, R., & McClelland, M. M. (2014). Behavioral self-regulation and executive function both predict visuomotor skills and early academic achievement. *Early Childhood Research Quarterly*, *29*(4), 411-424.

Bevel, R. K., & Mitchell, R. M. (2012). The effects of academic optimism on elementary reading achievement. *Journal of Educational Administration, 50*(6), 773-787.

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology, 78*(2), 156-168.

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

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*, 899 –911.

Blair, C., Ursache, A., Greenberg, M., & Vernon-Feagans, L. (2015). Multiple aspects of self-regulation uniquely predict mathematics but not letter–word knowledge in the early elementary grades. *Developmental Psychology*, *51*(4), 459-472.

Boman, P., Furlong, M. J., Shochet, I., Lilles, E., & Jones, C. (2009). Optimism and the school context. In R. Gilman, E. S. Huebner, M. J. Furlong, R. Gilman, E. S. Huebner, & M. J. Furlong (Eds.), *Handbook of positive psychology in schools* (pp. 51-64). New York, NY, US: Routledge/Taylor & Francis Group.

Brown, A., Bransford, J., Ferrara, R., & Campione, J. (1983). Learning, remembering, and understanding. In P. H. Mussen (Ed.), *Handbook of Child Psychology* (4th ed., Vol. 3, pp. 77-166). New York: John Wiley & Sons.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology*, *102*(4), 964-981.

Buhs, E., Ladd, G., & Herald, S. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology, 98*, 1–13.

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: Sources of variation in emergent mathematic achievement. *Developmental Science*, *14*(4), 679-692.

Caemmerer, J. M., & Keith, T. Z. (2015). Longitudinal, reciprocal effects of social skills and achievement from kindergarten to eighth grade. *Journal of School Psychology*, *53*(4), 265-281.

Caldarella, P., Christensen, L., Kramer, T. J., & Kronmiller, K. (2009). The effects of Strong Start on second grades students' emotional and social competence. *Early Childhood Education Journal, 37*(1), 51-56.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, *11*, 302-306.

Collaborative for Academic, Social, and Emotional Learning. (2017a). Federal policy. Retrieved from https://www.casel.org/federal-policy-and-legislation/

Collaborative for Academic, Social, and Emotional Learning. (2017b). History. Retrieved from https://www.casel.org/history/

Centers for Disease Control and Prevention. (2013). *Mental health surveillance among children- United States, 2005-2011*. Retrieved from https://www.cdc.gov/mmwr/preview/mmwrhtml/su6202a1.htm

Climie, E., & Henley, L. (2016). A renewed focus on strengths-based assessment in schools. *British Journal of Special Education*, *43*(2), 108-121.

Coie, J. D. (1990). Towards a theory of peer rejection. In S. R. Asher & J. D. Coie (Eds.), *Peer rejection in childhood* (pp. 365–401). New York: Cambridge University Press.

Collins, W. A. (1984). *Development during middle childhood: The years from six to twelve*. Washington, DC: National Academy Press.

Connors, E. H., Prout, J. T., Kozlowski, M., & Stephan, S. H. (2016). Student-focused supports to promote mental health and well-being. *Advances in School Mental Health Promotion*, *9*(2), 71-72.

Cowen, E. L. (1973) Social and community interventions. *Annual Review of Psychology, 24*, 423-472.

Cox, K. F. (2006). Investigating the impact of strength-based assessment on youth with emotional or behavioral disorders. *Journal of Child and Family Studies*, *15*(3), 278-292.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.

Dean, S., Britt, E., Bell, E., Stanley, J., & Collings, S. (2016). Motivational interviewing to enhance adolescent mental health treatment engagement: A randomized clinical trial. *Psychological Medicine*, *46*(9), 1961-1969.

DiPerna, J. C., Bailey, C. G., & Anthony, C. (2014). Broadband screening of academic and social behavior. In R.J. Kettler, T.A. Glover, C.A. Albers, K.A. Feeney-Kettler (Eds.), *Universal screening in educational settings: Evidence-based decision making for schools* (pp. 223-248). Washington, DC: American Psychological Association.

DiPerna, J.C., Volpe, R.J., & Elliott, S.N. (2002). A model of academic enablers and elementary reading/language arts achievement. *School Psychology Review, 31*, 298–312.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405-432.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. *Handbook of child psychology*, *4*, 643-691.

Eisenberg, N., Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., & Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal study. *Journal of Personality and Social Psychology*, *82*(6), 993-1006.

Eklund, K., Renshaw, T. L., Dowdy, E., Jimerson, S. R., Hart, S. R., Jones, C. N., & Earhart, J. (2009). Early identification of behavioral and emotional problems in youth: Universal screening versus teacher-referral identification. *California School Psychologist*, *14*89-95.

Elliott, S. N., & Gresham, F. M. (1993). Social skills interventions for children. *Behavior Modification*, *17*(3), 287-313.

Endrulat, N. R., Tom, K., Ravitch, K., Wesley, K., & Merrell, K. W. (2010, March). *Gender differences in positive social emotional functioning.* Presented at the meeting of the National Association of School Psychologists, Chicago, IL.

Epps, E. G., & Smith, S. F. (1984). School and children: The middle childhood years. In W. A. Collins (ed.), *Development during middle childhood: The years from six to twelve.* Washington DC: National Academies Press.

Epstein, M. H. (2000). The Behavioral and Emotional Rating Scale: A strength-based approach to assessment. *Assessment for Effective Intervention*, *25*(3), 249-256.

Epstein, M. H., & Sharma, F. (1998). *Behavioral and emotional scale (BERS): A strength-based approach to assessment*. Austin, TX: PRO-ED.

Epstein, M. H., Dakan, E., Oswald, D. P. & Yoe, J. T. (2001). Using strengths-based data to evaluate children's mental health programs. In M. Hernandez & S. Hodges (Eds.), *Developing outcome strategies in children's mental health: Systems of care for children's mental health* (pp.153-166). Baltimore, MD: Paul H. Brookes Publishing.

Ettelson, R., & Laurent, J. (2002). Multiple-stage screening for the identification of depression among adolescents. Retrieved from ERIC database. (ED463480)

Flahive, M. W., Chuang, Y., & Li, C. (2015). The multimedia piers-harris children's self-concept scale 2: Its psychometric properties, equivalence with the paper-and-pencil version, and respondent preferences. *Plos ONE*, *10*(8), 1-13.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*(2), 291-314.

Frieden, T. R. (2010). A framework for public health action: The health impact pyramid. *American Journal of Public Health*, *100*(4), 590–595.

Fuchs, D., & Fuchs, L. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93-99.

Fuchs, L. S., & Fuchs, D. A. (1998). Unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice*, *13*, 204–219.

Fuhs, M. W., Farran, D. C., & Nesbitt, K. T. (2013). Preschool classroom processes as predictors of children's cognitive self-regulation skills development. *School Psychology Quarterly*, *28*(4), 347-359.

Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. New Haven, CT: Yale University Child Study Center.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117–135.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*, 645 – 662.

Gough, H. G. (1953). What determines the academic achievement of high school students?. *The Journal of Educational Research*, *46,* 321-331.

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, *58*(6-7), 466-74.

Gresham, F. M., Cook, C. R., Crews, S. D., & Kern, L. (2004). Social skills training for children and youth with emotional and behavioral disorders: Validity considerations and future directions. *Behavioral Disorders*, *30*(1), 32-46.

Gresham, F. M., Vance, M. J., Chenier, J., & Hunter, K. (2013). Assessment and treatment of deficits in social skills functioning and social anxiety in children engaging in school refusal behaviors. In Dean McKay & Eric Storch (Eds.), *Handbook of assessing variants and complications in anxiety disorders* (pp. 15-28). New York: Springer.

Griggs, M. S., Rimm-Kaufman, S. E., Merritt, E. G., & Patton, C. L. (2013). The "Responsive Classroom" approach and fifth grade students' math and science anxiety and self-efficacy. *School Psychology Quarterly*, *28*(4), 360-373.

Hair, J.F. Jr., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate Data Analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Hall, C. S., & Lindzey, G. (1957). *Theories of personality*. Hoboken, NJ, US: John Wiley & Sons Inc.

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement*, *71*(4), 617-637.

Ialongo, N. S., Edelsohn, G., & Kellam, S. G. (2001). A further look at the prognostic power of young children's reports of depressed mood and feelings. *Child Development*, *72*(3), 736-47.

Ialongo, N., Edelsohn, G., Werthamer-Larsson, L., Crockett, L., & Kellam, S. (1995). The significance of self-reported anxious symptoms in first grade children: Prediction to anxious symptoms and adaptive functioning in fifth grade. *Child Psychology & Psychiatry & Allied Disciplines*, *36*(3), 427-437.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509-527.

Jenkins, J. R. (2003, December). *Candidate measures for screening at-risk students.* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention symposium, Kansas City, MO. Retrieved December 15, 2016, from http://www.nrcld.org/symposium2003/jenkins/index.html

Jones, S. M., Brown, J. L., & Aber, J. L. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, *82*(2), 533-554.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kamphaus, R. W. (2012). Screening for behavioral and emotional risk in schools: Constructs and practicalities. *School Psychology Forum, 6*, 89-97.

Katz, C., Bolton, S., Katz, L. Y., Isaak, C., Tilston-Jones, T., & Sareen, J. (2013). A systematic review of school-based suicide prevention programs. *Depression and Anxiety*, *30*(10), 1030-1045.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives Of General Psychiatry*, *62*(6), 593-602.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617-627.

Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). New York, NY: Routledge.

Kuo, E. S., Vander Stoep, A., Herting, J. R., Grupp, K., & McCauley, E. (2013). How to identify students for school-based depression intervention: Can school record review be substituted for universal depression screening?. *Journal of Child and Adolescent Psychiatric Nursing*, *26*(1), 42-52.

Laursen, E. K. (2000). Strength-based practice with children in trouble. *Reclaiming Children and Youth*, *9*(2), 70.

Leadbeater, B. J., Thompson, K., & Sukhawathanakul, P. (2016). Enhancing social responsibility and prosocial leadership to prevent aggression, peer victimization, and emotional problems in elementary school children. *American Journal of Community Psychology*, *58*(3-4), 365-376.

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research, & Evaluation*, 12(2), 1-11.

Lewis, M. (1990). Self-knowledge and social development in early life. In L. A. Pervin & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 277-300). New York, NY, US: Guilford Press.

Lonigan, C. J., Allan, D. M., & Phillips, B. M. (2017). Examining the predictive relations between two aspects of self-regulation and growth in preschool children's early literacy skills. *Developmental Psychology, 53*(1), 63-76.

Low, S., Smolkowski, K., & Cook, C. (2016). What constitutes high-quality implementation of SEL programs? A latent class analysis of Second Step® implementation. *Prevention Science*, *17*(8), 981-991.

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

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

McCarty, C. A., Wymbs, B. T., Mason, W. A., King, K. M., McCauley, E., Baer, J., & Vander Stoep, A. (2013). Early adolescent growth in depression and conduct problem symptoms as predictors of later substance use impairment. *Journal of Abnormal Child Psychology*, *41*(7), 1041-1051.

Merrell, K. W. (2008). *Behavioral, social, and emotional assessment of children and adolescents* (3rd ed.). New York: Taylor & Francis.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741-749.

Molnar, M. (2015, December 07). Half of K-12 students to have access to 1-to-1 computing by 2015-16. Retrieved February 04, 2017, from https://marketbrief.edweek.org/marketplace-k-12/half_of_k-12_students_to_have_access_to_1-to-1_computing_by_2015-16_1/

Müller, U., & Kerns, K. (2015). The development of executive function. In R. M. Lerner (ed.), *Handbook of child psychology and developmental science*. John Wiley and Sons.

National Association of School Psychologists. (2017). *NASP practice model 10 domains*. Retrieved from https://www.nasponline.org/standards-and-certification/nasp-practice-model/nasp-practice-model-implementation-guide/section-i-nasp-practice-model-overview/nasp-practice-model-10-domains

National Center on Response to Intervention (March 2010). *Essential components of RTI – A closer look at Response to Intervention*. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

Newcomer, A. R., Roth, K. B., Kellam, S. G., Wang, W., Ialongo, N. S., Hart, S. R., & ... Wilcox, H. C. (2015). Higher childhood peer reports of social preference mediates the impact of the good behavior game on suicide attempt. *Prevention Science*, *17*(2), 145-156.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.

O'Connor, E. E., McCormick, M. P., Cappella, E., McClowry, S. G., & Society for Research on Educational Effectiveness. (2014). *Getting a good start in school: Differential effects of "INSIGHTS" on the behaviors and engagement of children with challenging temperaments*. Retrieved from ERIC database. (ED562775)

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical Child and Family Psychology Review*, *10*(4), 335-351.

Parsons, J. E., & Ruble, D. N. (1977). The development of achievement-related expectancies. *Child Development*, 1075-1079.

Penfield, R. D. (2014). An NCME instructional module on polytomous item response theory models. *Educational Measurement: Issues and Practice*, *33*(1), 36-48.

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2012). Control networks and neuromodulators of early development. *Developmental psychology*, *48*(3), 827-835.

Rashid, T., & Ostermann, R. F. (2009). Strength-based assessment in clinical practice. *Journal of Clinical Psychology*, *65*(5), 488-498.

Reise, S. P., & Revicki, D. A. (2015). *Handbook of item response theory modeling: Applications to typical performance assessment*. New York, NY: Routledge.

Reivich, K. (2010). Optimism and well-being. *Communique*, *38*(7), 10-11.

Reyes, M. R., Brackett, M. A., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on targeted student outcomes for the RULER approach to social and emotional learning. *School Psychology Review, 41*, 82-99.

Reynolds, W. M. (1992). The study of internalizing disorders in children and adolescents. In W. M. Reynolds, & W. M. Reynolds (Eds.), *Internalizing disorders in children and adolescents* (pp. 1-18). Oxford, England: John Wiley & Sons.

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly*, *15*(2), 147-166.

Rosenberg, R., & Bryant, M. (2001). *Discrimination by young children across three different survey response icons*. Retrieved from ERIC database. (ED462404)

Sawyer, A. C. P., Chittleborough, C. R., Mittinty, M. N., Miller-Lewis, L. R., Sawyer, M. G., Sullivan, T., & Lynch, J. W. (2015). Are trajectories of self-regulation abilities from ages 2–3 to 6–7 associated with academic achievement in the early school years?. *Child: Care, Health and Development*, *41*(5), 744-754.

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., & Diamond, A. (2015). Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: A randomized controlled trial. *Developmental Psychology*, *51*(1), 52-66. doi:10.1037/a0038454

Schwartz S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1-65.

Seligman, M. E. P. (1991). *Learned optimism*. New York: Knopf.

Seligman, M. P., Ernst, R. M., Gillham, J., Reivich, K., & Linkins, M. (2009). Positive education: Positive psychology and classroom interventions. *Oxford Review of Education*, *35*(3), 293-311.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45*, 193–223.

Shechtman, Z., & Abu Yaman, M. (2012). SEL as a component of a literature class to improve relationships, behavior, motivation, and content knowledge. *American Educational Research Journal*, *49*(3), 546-567.

Snape, L., & Atkinson, C. (2016). The evidence for student-focused motivational interviewing in educational settings: A review of the literature. *Advances in School Mental Health Promotion, 9*(2), 119-139

Stanovich K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-406.

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*(2), 245-259.

Suldo, S. M., & Shaffer, E. J. (2008). Looking beyond psychopathology: The dual-factor model of mental health in youth. *School Psychology Review*, *37*(1), 52.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357-385.

Thorpe, P. K. (2003). *A mediation model relating teacher ratings of student achievement to student connectedness at school*. Retrieved from ERIC database. (ED476417)

U. S. Department of Education. (2001, September). *Overview of public elementary and secondary schools and districts: School year 1999-2000.* Retrieved from https://nces.ed.gov/pubs2001/overview/table05.asp

U.S. Department of Education, National Center for Education Statistics. (2016). *Digest of education statistics, 2015* (NCES Publication No. 2016-014). Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=64

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. (2009). *Fostering school connectedness: Improving student health and academic achievement*. Retrieved from https://www.cdc.gov/healthyyouth/protective/pdf/connectedness_administrators.pdf

U.S. Office of Education. (1977). Assistance to states for education for handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register*, 42, G1082–G1085.

Vander Stoep, A., McCauley, E., Thompson, K. A., Herting, J. R., Kuo, E. S., Stewart, D. G., & Kushner, S. (2005). Universal emotional health screening at the middle school transition. *Journal of Emotional and Behavioral Disorders*, *13*(4), 213-223.

Varni, J. W., Limbers, C., & Burwinkle, T. M. (2007). Literature review: Health-related quality of life measurement in pediatric oncology: Hearing the voices of the children. *Journal of Pediatric Psychology*, *32*(9), 1151-1163.

Walker, H. M., Severson, H. H., Feil., E. (1992). *Systematic Screening for Behavior Disorders (SSBD), second edition*. Retrieved from Pacific Northwest Publishing. https://pacificnwpublish.com/products/SSBD-Online.html#about

Watkins, M. L. (2008). Promoting the early identification of internalizing problems in preliterate children: Development of the Watkins Early Self-Report of Internalizing Problems. *Dissertation Abstracts International Section A*, *68*, 4609

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullotta, T. P. (2015). Social and emotional learning: Past, present, and future. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T.P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 3–19). New York: Guilford.

Weng, L. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, *64*(6), 956-972.

Whitlock, J., Wyman, P. A., & Moore, S. R. (2014). Connectedness and suicide prevention in adolescents: Pathways and implications. *Suicide and Life-Threatening Behavior*, *44*(3), 246-272.

Wilcox, H. C., Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., & Anthony, J. C. (2008). The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempts. *Drug and Alcohol Dependence*, *95*(Suppl1), S60-S73. doi:10.1016/j.drugalcdep.2008.01.005

Williams, S. T. (2008). Mental health screening and assessment tools for children: Literature review. Northern California Training Academy. Retrieved from http://humanservices.ucdavis.edu/academy/pdf/final2mentalhealthlitreview.pdf

Wise, D., & Rosqvist, J. (2006). Explanatory style and well-being. In J. C. Thomas, D. L. Segal, M. Hersen, J. C. Thomas, D. L. Segal, & M. Hersen (Eds.), *Comprehensive Handbook of Personality and Psychopathology, Vol. 1: Personality and Everyday Functioning* (pp. 285-305). Hoboken, NJ, US: John Wiley & Sons Inc.

Wray-Lake, L., Syvertsen, A. K., & Flanagan, C. A. (2016). Developmental change in social responsibility during adolescence: An ecological perspective. *Developmental Psychology*, *52*(1), 130-142.

Ysseldyke, J., Burns, M., Dawson, P., Kelly, B., Morrison, D., Ortiz, S., Rosenfield, S., Telzrow, C. (2006). *School Psychology: A blueprint for training and practice III*. Bethesda, MD: National Association of School Psychologists.

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (2004). *Building school success through social and emotional learning*. New York: Teachers College Press.

Zito, J. M., Safer, D. J., Gardner, J. F., Boles, M., & Lynch, F. (2000). Trends in the prescribing of psychotropic medications to preschoolers. *Journal of the American Medical Association*, *283*(8), 1025-1030.