EXTENDING RELIABILITY EVIDENCE FOR ADJUSTMENT SCALES
FOR PRESCHOOL INTERVENTION (ASPI) SCORES
AMONG AT-RISK RURAL PRESCHOOLERS

A Dissertation in
School Psychology
by
Caitlin Cells

© 2017 Caitlin Cells

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Doctor of Philosophy

May 2017
The dissertation of Caitlin Cells was reviewed and approved* by the following:

Barbara A. Schaefer  
Professor-in-Charge, School Psychology Program  
Associate Professor of Education  
Dissertation Advisor  
Chair of Committee

Shirley A. Woika  
Associate Professor of Education

Paul L. Morgan  
Professor of Education

Scott D. Gest  
Professor of Human Development and Family Studies

*Signatures are on file in the Graduate School
ABSTRACT

Behavior problems are a primary concern in educational settings, and early identification and intervention are key to addressing them. The purpose of this study was to provide evidence of psychometric integrity of the Adjustment Scales for Preschool Intervention (ASPI; Lutz, Fantuzzo, & McDermott, 2002) scores using an at-risk rural preschool sample. This study extended previous research that has established acceptable inter-rater reliability evidence of the ASPI for use with an urban sample (Bulotsky-Shearer & Fantuzzo, 2004). Teachers and assistant teachers of 57 preschoolers from a rural Head Start Center in Pennsylvania completed ASPI behavior rating scales on their students. Resulting scores were analyzed to establish inter-rater and internal consistency reliability of scores. Across multiple reliability estimates, all scales were statistically significant ($p < .05$); however, all five problem behavior domains did not consistently indicate acceptable internal consistency evidence of at least .70 and acceptable inter-rater reliability evidence in the form of intraclass correlations of .75 or greater and interclass agreement of .80 or greater. Ratings of Overactive behaviors, especially aggression and oppositionality, demonstrated stronger reliability evidence than Underactive behaviors.
# TABLE OF CONTENTS

LIST OF TABLES.........................................................................................................................v

ACKNOWLEDGEMENTS............................................................................................................vi

Chapter 1. INTRODUCTION AND LITERATURE REVIEW.........................................................1
  Outcomes of Early Behavior Problems .................................................................................. 2
  Risk Factors for Behavior Problems .................................................................................... 10
  Evaluation of Current Behavioral Assessment Tools .......................................................... 17
  Rural and Urban Characteristics ........................................................................................... 24
  Adjustment Scales for Preschool Intervention ...................................................................... 31
  Rationale and Purpose ......................................................................................................... 33

Chapter 2. METHOD ................................................................................................................ 35
  Participants ............................................................................................................................ 35
  Measures .............................................................................................................................. 36
  Procedure ............................................................................................................................ 40
  Data Analyses ..................................................................................................................... 41

Chapter 3. RESULTS..................................................................................................................44
  Analytic Strategy .................................................................................................................. 44
  Preliminary Analyses for Reliability Estimates ................................................................. 44
  Intraclass Correlation of Teacher and Assistant Teacher Behavior Ratings .................... 45
  Interrater Agreement between Teacher and Assistant Teacher ......................................... 45
  Internal Consistency of ASPI Scores ................................................................................. 46

Chapter 4. DISCUSSION .......................................................................................................... 51
  Limitations ........................................................................................................................... 54
  Implications for Research and Practice .............................................................................. 56
  Directions for Future Research ......................................................................................... 57
  Conclusions ........................................................................................................................ 59

REFERENCES ......................................................................................................................... 60

Appendix A ............................................................................................................................... 80
  Sample ASPI Questions ....................................................................................................... 80

Appendix B ............................................................................................................................... 81
  Teacher Demographic Questionnaire ................................................................................. 81

Appendix C ............................................................................................................................... 82
  Child Demographic Questionnaire ..................................................................................... 82

Appendix D ............................................................................................................................... 83
  Tables ................................................................................................................................... 83
LIST OF TABLES

Table 1: Descriptive Statistics for ASPI domains using t-tests for Equality of Means…………..83
Table 2: Intraclass Correlation Coefficients for ASPI Domains .......................................84
Table 3: Cronbach’s Alpha for ASPI domains by Rater and Gender .................................85
Table 4: Cronbach’s Alpha of Individual Items if Removed and Corrected Item-Total Correlations.................................................................86
ACKNOWLEDGEMENTS

I would first like to thank my dissertation committee. To Dr. Barbara Schaefer, my advisor and chair, thank you for your positive outlook and support throughout this journey. Your encouragement has been very motivational. Thank you also to my other committee members, Dr. Paul Morgan, Dr. Scott Gest, and Dr. Shirley Woika, for your guidance and insight throughout this process.

I would like to thank Dr. Paul McDermott for granting me permission to use the ASPI for this research. Thank you to all of the Head Start teachers, assistant teachers, and families for participating in the study. I would also like to thank my graduate assistant, Kelsey Spanjers, for your assistance in the data collection process.

Just as importantly, I would like to thank my family. To my parents, Vincent and Michele Cells, words cannot express the impact you have had in shaping who I’ve become. You’ve always modeled and taught me the importance of education and working hard, even when it would be easier to give up. You saw my potential and ensured that I reached it. To my sister, Meaghan Cells, you have always strived to be the best version of yourself in everything you do, which has encouraged me to be a better version of myself.

To my cohort, I would not have made it through this journey without you. You have made me laughed, helped me learn, and grown with me. I will never forget and always appreciate this journey that we took together. A special thank you to Dr. Nelli El-Ghazal and Ashley Morse for your friendship over these years. You have been two of my loudest cheerleaders. Thank you for dropping everything whenever I needed you and believing in me during the times when I didn’t believe in myself.
The completion of this dissertation would not have been possible without the support and encouragement of my colleagues at Loudoun County Public Schools. Thank you to Dr. Heather Applegate. The completion of this dissertation is a direct result of your encouragement, persistence, and follow-through. I would also like to my internship supervisor, Dr. Kristi Hockensmith, for checking in on my progress over the years and being invested in my success as if it were your own. Thank you to my fellow interns, Dr. Todd Robinson and Erin Balanc, as well as my colleagues, Ellen McGraw and Dr. Susan Book, for always believing in me and being willing to listen to my joys, fears, and frustrations throughout this process.

To all of my friends, family, colleagues, professors, and teachers not mentioned by name, who have helped, guided, and encouraged me along the way, you have shaped who I am and the path I have taken. I am forever grateful.
Chapter 1: Introduction and Literature Review

Behavior problems are a growing concern in school settings, and school personnel are working to identify these behaviors in children at younger and younger ages so as to proactively address these problems. Behavior problems in young children negatively impact academics (Arnold et al., 2006), social interactions (Fantuzzo, Bulotsky, McDermott, Mosca, & Lutz, 2003), and attitudes toward school (Coolahan, Fantuzzo, Mendez, & McDermott, 2000). Additionally, these problems can be lasting through school and adulthood (Stevenson & Goodman, 2001). Teachers can also be negatively impacted by their students’ behavior (Webster-Stratton, Reid, & Stoolmiller, 2008). Behavior problems that preschoolers exhibit are generally categorized as internalizing or externalizing behavior problems. For young children in a preschool setting, the source of these classroom behavior problems has been divided further into problems in structured learning, problems in peer interaction, and problems in teacher interaction (Bulotsky-Shearer, Fantuzzo, & McDermott, 2008).

Not only can behavior problems impact how children function in school, but many risk factors make some students more likely to display problem behaviors than others. Child, parent, and family variables like socioeconomic status, poverty, ethnicity, parental psychopathology, and quality of childcare all influence the likelihood that children will act out starting at a young age. These children are considered at-risk. Other elements such as whether a child lives in a rural or urban setting may also impact his or her behavior in school.

Given these factors, high quality measurement tools that are appropriate for these children are needed in order to validly assess these behaviors. This study examined the appropriate use of one such tool, the Adjustment Scales for Preschool Intervention (ASPI; Lutz,
Fantuzzo, & McDermott, 2002), with the aim to expand its utility with not only urban, but also rural at-risk populations.

Estimating the prevalence of behavior problems in the general preschool population is difficult, and rates vary between researchers (Huaqing & Kaiser, 2003). Many studies on the rates of mental health and behavior problems have failed to include children under the age of five (Lopez, Tarullo, Forness, & Boyce, 2000). However, best estimates of problem behaviors among preschoolers of low socioeconomic status (SES) are around 30%, in comparison to only 3 to 6% for the general preschool population (Huaquing & Kaiser, 2003). Other research has indicated that 10% of both low- and high-income preschoolers demonstrate aggression problems (Kupersmidt, Bryant, & Willoughby, 2000) or more extreme problem behaviors (Lavigne et al., 1996). For DSM III Axis I disorders, the prevalence rate rises to almost 22% (Lavigne et al., 1996). Given estimates that almost one in four Head Start students exhibits externalizing behaviors (Harden et al., 2000), a clear need for early identification and treatment of these problems exists.

**Outcomes of Early Behavior Problems**

**Academic effects.** Links have consistently been found between behavior problems and negative academic outcomes. Specifically, problems with aggression have been found to predict disruptive peer play, while hyperactive/inattentive problems have predicted lower cognitive skills (Fantuzzo et al., 2003). Additionally, children who are chronically sad or disruptive when they are young receive less instruction, learn less from their classmates, and are less engaged as learners -- and these children with the most problems are the least likely to receive the help they need (Arnold et al., 2006). Preschool behavior problems in structured learning, peer interaction, and teacher interaction also lead to negative academic and social outcomes. Specifically,
problems in structured learning situations for at-risk preschoolers have been found to predict a lack of connection with classmates and ongoing struggles with learning (Bulotsky-Shearer et al., 2008). Behavior problems in preschoolers have also been linked to poor work-related skills such as following directions, staying on task, and organizing materials (McClelland, Morrison, & Holmes, 2000). Problems with self-regulation have also been associated with reading problems in preschoolers (Blair & Razza, 2007). Conversely, executive functioning skills in preschool students have been found to predict school readiness (Graziano, Garb, Ros, Hart, & Garcia, 2015). In one study of preschoolers from middle- and low-income families, researchers found that behavior problems in both groups of preschoolers, regardless of SES, were associated with worse emergent literacy outcomes. Specifically, inattention was associated with poorer print knowledge and phonological sensitivity among low-income students; similarly, inattention was associated with poorer print knowledge, oral language skills, and phonological memory for middle-income students (Lonigan et al., 1999). Further, low-SES children with the greatest behavioral needs in play and peer interactions demonstrate lower skills in science, math, and literacy when compared to their peers with high peer interaction; differences which are maintained across the school year (Bell, Greenfield, Bulotsky-Shearer, & Tracey, 2016).

In contrast, prosocial skills are associated with positive academic outcomes. Kindergarteners with better social behavior have higher literacy outcomes (Miles & Stipek, 2006). Research has also found that both prosocial skills and internalizing behaviors are important in predicting later academic success in the early elementary school years (Bain & Agostin, 1997). Findings of one study with urban Head Start preschoolers indicated that behavior problems in the classroom alone did not distinguish those students at risk for academic problems from those not at risk; however, when combined with academic and motor skill indicators,
behavior/social skills did predict early academic problems (McWayne, Fantuzzo, & McDermott, 2004)

Behavioral regulation, which includes behaviors such as working memory, attention, following directions, and inhibitory control, can also impact educational outcomes. Higher rates of behavioral regulation in preschool and kindergarten predict better reading, mathematics, and vocabulary in first grade (Sektnan, McClelland, Acock, & Morrison, 2010). Even when controlling for age, gender, and native language, higher levels of behavioral regulation in preschoolers were found to be associated with higher levels of literacy, mathematics, and vocabulary (McClelland et al., 2007). Additionally, those students who showed greater gains in behavioral regulation in their pre-kindergarten year showed concurrent gains in their literacy, math, and vocabulary skills (McClelland et al., 2007).

Conversely, behavioral regulation also can serve as a mediator between poor emotional regulation in preschool and academic problems in later school years (Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003). Learning-related behaviors, which include behavioral regulation, have also been found to have a long-term impact on academics. These learning behaviors in kindergarten are unique predictors of reading and math achievement through sixth grade (McClelland, Acock, & Morrison, 2006). Additionally, those students with poor regulation skills in kindergarten have worse reading and math achievement in sixth grade, and the gap widens as they age (McClelland et al., 2006). Behavioral regulation in preschoolers has even been found to predict early academic success when controlling for IQ and maternal education (Howse et al., 2003). These findings indicate that appropriate school behaviors are important to positive academic outcomes.
**Social effects.** Problems with behavior in preschool also have a negative effect on these students’ social interactions. In one study, students with overactive behaviors, (e.g., aggression, opposition) at the beginning of the year were found to be socially disruptive at the end of the year. Conversely, those students with underactive behavior problems, (e.g., social reticence, withdrawal) were disengaged from their peers at the end of the year (Fantuzzo et al., 2003). Similarly, other research has found that preschoolers who evidence problems with teachers or classmates have social problems with their peers at the end of their preschool year (Bulotsky-Shearer et al., 2008).

Additionally, problems with social problem solving have been found to be associated with behavior problems in preschool. Specifically, relational aggression has been shown to lead to psycho-social maladjustment in both preschool boys and girls (Harden et al., 2000). Similarly, the temperament of preschoolers has been found to be associated with early behavior problems (Lavigne et al., 1996). In another study, both boys and girls with high levels of relational aggression were found to be less likely to share or help others than their non-aggressive peers (Crick, Casas, & Mosher, 1997). Conversely, preschoolers who demonstrated positive peer play had higher levels of attention, motivation, and persistence (Coolahan et al., 2000).

**Long-term outcomes.** The impact of behavior problems in preschool can impact students in later school years, although the findings of these long-term effects are mixed. A longitudinal study of children from 10 sites across the country found that children’s externalizing behaviors reported by mothers and teachers decreased between two and nine years old, with the largest drop occurring between the ages of two and seven (Miner & Clarke-Stewart, 2008); however, other research has demonstrated the relative stability of behavior and behavior problems over time. Bub, McCartney, and Willett (2007) found that externalizing behavior problems at two
years old correlated strongly with externalizing behavior problems in kindergarten. In another study of infants, researchers found that children who were rated by their parents as having clinically significant behavior problems at 12 to 36 months were more likely to have behavior problems in elementary school, thus providing additional support for early assessment and identification of behavior problems (Briggs-Gowan & Carter, 2008). Kerr, Lunkenheimer, and Olson (2007) also found that in a sample of children from the Midwest, both externalizing and internalizing problems in preschoolers remained stable from three to five years old. Furthermore, in a national study Morgan, Farkas, and Wu (2009) found that children who demonstrated high levels of externalizing problem behaviors in kindergarten were four times more likely than their peers to also display these behaviors in third and fifth grade. The researchers also found that children already exhibiting high levels of internalizing behaviors were three times more likely to display internalizing problem behaviors in third and fifth grade as well.

Behavior problems in preschoolers have also been found to be predictive of later academic difficulties. Preschoolers with behavior problems in learning contexts have been found to have worse literacy and language outcomes in kindergarten and first grade even when controlling for cognitive ability (Bulotsky-Shearer & Fantuzzo, 2011). Those preschoolers with peer and teacher interaction problems also had lower phoneme segmentation in first grade (Bulotsky-Shearer & Fantuzzo, 2011). In a sample of Canadian students, researchers found that aggressive and prosocial behavior in preschool predicted cognitive self-control, which includes the ability to plan and evaluate, in first grade (Normandeau & Guay, 1998). Additionally, anxious-withdrawn behavior (Normandeau & Guay, 1998) and inhibition have been found to directly influence school achievement (Blair & Razza, 2007). Behaviors of children in kindergarten, such as hyperactivity and classroom behavior, have been found to influence their
reading success in fourth grade (Spira, Bracken, & Fischel 2005). In their study of former Head Start students, Spira et al. (2005) found that students who, in kindergarten, struggled with reading but had good classroom behavior, demonstrated greater gains in reading by fourth grade than their peers with poor classroom behavior in kindergarten. Additionally, results of the study indicated that classroom behavior in kindergarten, along with other language skills such as oral language and print knowledge, were predictive of later reading ability (Spira et al., 2005).

Preschool behaviors are associated with later math outcomes in addition to reading and literacy outcomes. Effortful control and inhibition in preschoolers from low-income homes have been found to be associated with math ability and letter knowledge in kindergarten (Blair & Razza, 2007). Although most findings have supported the relationship between early behavior problems and later academic achievement, Duncan et al. (2007) found that early math and reading skills were stronger predictors of achievement in early and mid-elementary grades than behavior problems. Though attention skills were a less powerful predictor of later achievement, evidence was not found that internalizing and externalizing behaviors predicted achievement. However, these researchers only examined academic success, and behavioral problems may contribute to other indicators of overall school success (Duncan et al., 2007). Additionally, Grimm, Steele, Mashburn, Burchinal, and Pianta (2010) noted problems in the Duncan et al. study, such as informant variation in behavior rating scales and improperly categorizing behavior problems as linear, which may have produced these contrary results. While Grimm et al. (2010) did not demonstrate that behavior was a strong predictor of academic achievement, their results indicated that higher rates of internalizing behavior had small to moderate relationships with lower rates of math and reading achievement.
Many types of behavior problems can also lead to victimization by peers in later years. Young children with problem behaviors including aggression, impulsivity, and anxiety have been found to be victimized into early and middle childhood (Schwartz, McFadyen-Ketchum, Dodge, Pettit, & Bates, 1999). A similar predictive relationship has been found between behavior problems with peers and disruptive play at the end of the preschool year and into kindergarten. Students with these behavior problems were often unable to initiate or maintain satisfactory peer play relationships (Bulotsky-Shearer, Dominguez, Bell, Rouse, & Fantuzzo, 2010).

Not only can behavior problems lead to negative outcomes for children in school, but they can have enduring effects as adults. Behavior problems in preschool can affect students’ success later on in life. In one study, early disruptive behaviors in Canadian students were found to predict later school dropout, even when controlling for other student characteristics (Vitaro, Larocque, Janosz, & Tremblay, 2001). Additionally, behavior problems in preschool have been found to predict crime as adults. Externalizing behavior problems in preschoolers, like tantrums, have been associated with adult criminal convictions, mostly for violent offenses, although the variance explained by these behaviors was not indicated (Stevenson & Goodman, 2001). Furthermore, in a longitudinal study of New Zealanders, researchers found that anti-social behavior in children age 5 to 11 was associated with higher rates of sexual activity before age 16 and high-risk sexual activity by the age of 21. These findings were mediated by having delinquent friends (Ramrakha et al., 2007).

Behaviors specifically related to interactions with teachers also impact later achievement. Negative teacher-student relationships in kindergarten predict academic and behavior problems later on into middle school, and have been found to be an even stronger predictor of academic outcomes than the poor behavior of the child alone (Hamre & Pianta, 2001). Kindergarteners
who had a poor relationship with their teacher and had high levels of behavior problems were more likely to also have poor work-habits and more disciplinary referrals (Hamre & Pianta, 2001). Given that teacher perceptions of their relationship with students have not been found to vary significantly between preschool and kindergarten (Pianta & Stuhlman, 2004), these effects often are true with preschoolers as well. Small effects have been found between preschool teacher-child relationship quality and academic achievement in first grade (Pianta & Stuhlman, 2004). This trend is important because it emphasizes the importance of early behaviors as predictors of later school success or failure. Since a growing body of research is aimed at addressing behavior problems, the measures used to assess these problems in such young children need to be examined and refined to more accurately assess the levels of behavior problems existing in these students.

Although behavior problems can be lasting into later years and adulthood, there is variability in the developmental stability of these behaviors over time. Certain behaviors, such as hyperactivity, are generally consistent from early childhood (Spira et al., 2005); however, other behavior problems, such as externalizing behaviors and physical aggression, frequently decline as children grow and develop, in the absence of other risk factors (Côté, Vaillancourt, LeBlanc, Nagin, & Tremblay, 2006; Miner & Clarke-Stewart, 2008).

**Attitude toward school.** Behavior problems exhibited by preschoolers can not only negatively impact the students, but can have a negative impact on teachers as well. Teachers have reported feeling that they have difficulty teaching academic skills in their classrooms when so many students have conduct problems (Webster-Stratton et al., 2008). A national survey of kindergarten teachers reported that one half of kindergarteners had minor to serious adjustment
problems. Teachers also reported that half of their students had a deficit of academic skills and problems following directions (Rimm-Kaufman, Pianta, & Cox, 2000).

Behavior problems can also impact students’ attitude toward school. Preschoolers who demonstrated high levels of positive peer interactions and attention to tasks also had low levels of aggression, received feedback well, and had better math outcomes. Conversely, those who demonstrated disengaged behavior had difficulties connecting to academic activities (Fantuzzo, Bulotsky-Shearer, McDermott, McWayne, & Frye, 2007). Additionally, preschoolers displaying problems interacting with peers appropriately also showed an unwillingness to accept help from teachers and a lack of effort for school tasks (Coolahan et al., 2000).

**Risk Factors for Behavior Problems**

As articulated by Webster-Stratton (1996), three common categories define risk for child behavior problems: child, parenting, and family variables. Among the child risk variables are disruptive behavior, inattention, and aggression. Parenting risk variables include poor parenting and negative attitudes, while family risk variables consist of SES, marital problems, and parental mental illness. While some of these risk factors are distal to the development of behavior problems and represent general, underlying risk, others are more directly causal. Even though any student could develop behavior problems, some children are more at-risk than others. Low-income preschoolers have been found to be at a higher risk for emotional and behavioral disorders than children from high-SES backgrounds (Brooks-Gunn & Duncan, 1997; Deater-Deckard, Dodge, Bates, & Pettit, 1996; Feil et al., 2005; Huaqing & Kaiser, 2003; Lavigne et al., 1996). Specifically, children in Head Start were found to have higher levels of both internalizing and externalizing problem behavior in one study. Children in that Head Start sample, however, did have more clinical ratings for internalizing problems than externalizing problems (Kaiser,
Hancock, Cai, Foster, & Hester, 2000). In another study which consisted of a sample of African American children, family income was found to uniquely predict antisocial behavior (Luster & McAdoo, 1994).

Additionally, family income has been found to be a strong predictor of behavior problems in children at age five; income was found to be a stronger predictor than even maternal education and ethnicity (Duncan, Brooks-Gunn, & Klebanov, 1994). Furthermore, kindergarteners from low-SES have been found to be more at risk for behavior problems in fifth grade (Morgan et al., 2009). Results of a national longitudinal study found that young children with a mother who did not go to college, from single parent homes, or from families receiving federal assistance were less likely to get along with their classmates and make friends (West, Denton, & Germino-Hausken, 2000). Family income not only directly impacts behavior problems in young children, but other factors mediate that relationship. The relationship between high family income and fewer behavioral problems in a sample of preschoolers who had been pre-term babies has been found to be mediated by a cognitively stimulating home environment, positive parenting, and lower rates of maternal stress (Linver, Brooks-Gunn, & Kohen, 2002).

Research on whether the length of time a child lives in poverty influences emotional and behavioral outcomes suggests that longer exposure to poverty distally increases the risk of behavior problems. In general, negative links have been made between the length of time a child is exposed to social risk factors and his or her behavioral outcomes. Children exposed to risk for greater periods of time have demonstrated less prosocial behavior and more behavior problems through fourth grade (Burchinal, Roberts, Zeisel, Hennon, & Hooper, 2006). More specifically, children who have lived in poverty for their entire lives are at greater risk for both externalizing and internalizing behaviors than children who have never been poor (Brooks-Gunn & Duncan,
1997). Additionally, children who lived in poverty for one to four years are also at risk for externalizing and internalizing behavior problems, but with smaller effects for children who lived in short-term poverty (Brooks-Gunn & Duncan, 1997). In a longitudinal study examining the length of time a child spent living in economic hardship, those children who lived in persistent poverty started out behind their peers and never caught up on various school indicators, including behavior problems. Additionally, findings from that study indicated that boys living in persistent poverty were more likely to exhibit behavior problems than girls (Bolger, Patterson, Thompson, & Kupersmidt, 1995). Other research has indicated that the effects of poverty on behavior problems are up to 80% greater for children who have lived in constant poverty than those who have experienced temporary poverty (Duncan et al., 1994). Not only is poverty a risk-factor for behavioral problems, but children who grow up poor also have worse physical health outcomes, have increased rates of learning disabilities and developmental delays, and lower school achievement (Brooks-Gunn & Duncan, 1997). These children are also at a higher risk for harsh parental discipline and witnessing violence (Huaquing & Kaiser, 2003). Additionally, low-income Head Start preschoolers have demonstrated higher overall levels of aggression than those in non-Head Start schools, but no differences in the percentages of aggression between the two groups have been found (Kupersmidt et al., 2000). This pattern is counter to most findings that indicate that young children from low-SES families are more at risk for behavior problems. Even though a significant difference was not found between the percentage of Head Start and non-Head Start children engaging in aggression, low income preschoolers individually demonstrated a greater frequency of physical aggression (Kupersmidt et al., 2000). A strong body of research supports the influence of income and poverty as a risk factor for later behavior problems in children, but other factors put children at-risk as well.
Findings about ethnicity as a risk factor for behavior problems are mixed. Some research has found links between poor behavioral regulation with ethnic minority status and SES (Sektnan et al., 2010). Other research indicates, however, that teachers overrate aggression among White preschoolers, while parents underrate social skills among African Americans (Feil et al., 2005). Other research has found that teachers rate Black children higher than White children on exhibiting problem and externalizing behaviors, however. When parents rate problem behavior in their own children, differences between Black and White children have not been found (Miner & Clarke-Stewart, 2008; West et al., 2000). In fact, of parent ratings, the biggest differences in rates of problem behavior are between Asian children and their Black, White, and Hispanic peers (West et al., 2000). Even the relationship between externalizing behavior problems and discipline is different between ethnic groups. For Caucasian children, more harsh discipline results in more externalizing problems, but for African American children, no association has been found between harsh discipline and problem behaviors (Deater-Deckard et al., 1996). Race is also intertwined with socio-economic status, and the effects of race on behavior problems are often linked with the effects of SES. Some research on behavior problems in young children found that three-fifths of Blacks lived in neighborhoods with at least one-fifth of the population who was poor (Duncan et al., 1994). Researchers also have noted that Blacks were more likely to be poor for longer periods of time than Whites (Duncan et al., 1994; Linver, et al., 2002). Other research on African-American children found that children who were exposed to more risk factors, such as negative maternal characteristics, family structure, and home environment, had higher rates of antisocial behavior problems (Luster & McAdoo, 1994).

Parental characteristics are often related to risk for behavior problems in preschoolers. Psychopathology in parents has been found to be related to behavior problems in their young
children (Harden et al., 2000; New, Razzino, Lewin, Schlumpf, & Joseph, 2002). Maternal personality issues also influence behavior problems (Shaw, Owens, Vondra, Keenan, & Winslow, 1996). Specifically, maternal depression has been found to indirectly relate to disruptive behavior in preschoolers (Spieker et al., 1999), including low-level depression symptoms that do not meet the threshold for disorder (Conners-Burrow, Swindle, Mc Kelvey, & Bokony, 2014). It also relates to poor behavioral regulation in preschool children and subsequently in kindergarten and first grade (Sektnan et al., 2010). Similarly, maternal self-esteem has been found to predict behavior problems in African American children (Luster & McAdoo, 1994). Additionally, low maternal education, maternal insensitivity, and harsh discipline have all been linked to higher levels of externalizing behaviors in preschoolers and young children (Miner & Clarke-Stewart, 2008). Authoritarian, negative parenting by mothers, including yelling and spanking, is not only associated with disruptive behavior problems in preschoolers (Webster-Stratton, 1996), but is also associated with a maintenance of those behaviors over time (Spieker et al., 1999). Further, parental moral disengagement has also been associated with higher levels of externalizing behaviors in young children. Specifically, in one Italian study, parents who tended to blame others frequently were more likely to have aggressive and oppositional preschoolers (Camodeca & Taraschi, 2015). Parenting has been found to be a mediator between risk factors in African American children and behavior problems later on in school (Burchinal et al., 2006). Additionally, links have been found between poor behavioral regulation in first graders and low maternal education (Sektnan et al., 2010). However, children with high behavioral regulation have been found to display higher academic achievement regardless of whether or not they are at risk for behavior problems (Sektnan et al., 2010).
Home and family environment are risk factors that also influence the development of behavior problems in preschoolers and young children. Children who come from a single-parent home have been found to have higher levels of behavior problems (Deater-Deckard et al., 1996; Lavigne et al., 1996; Morgan et al., 2009). Other research has supported these findings by indicating that both transitioning from a two-parent to one-parent home as well as being raised from birth by a single mother both influence the development of behavior problems in young children (Duncan et al., 1994). In addition to the number of parents in the home, the number of siblings in the home has been found to predict behavior problems in African American children, such that those who came from smaller families had fewer behavior problems than those from large families (Luster & McAdoo, 1994). Preschool students in Head Start who have lived with someone who has been arrested are also at a higher risk for both internalizing and externalizing behavior problems (Ziv, Alva, & Zill, 2010). Similarly, children exposed to violence are at a greater risk for problem behaviors, especially clinical problems (Eiden, 1999). The rate of exposure is also an important factor in the rate of behavior problems of these children. The frequency of exposure to violence has been found to be a predictor of behavior problems in children (Eiden, 1999). Violence in the home, in the form of parental aggression, conflict, or child-rearing disagreements, has also been found to be a risk factor for behavior problems in presholers (New et al., 2002; Shaw et al., 1996). Conversely, in a study of Head Start students, researchers found that children with home environments with parental affection and emotional responsiveness were less likely to have behavior problems and were more likely to receive needed mental health services (New et al., 2002).

In addition to home and family characteristics, the quality of childcare is important in behavioral outcomes later on in school. Effective child care has been found to be a protective
factor against risk for behavior problems between kindergarten and third grade (Burchinal et al., 2006). Students enrolled in Head Start, on average, have greater reductions in externalizing behaviors than other comparable preschool centers, though results vary individually between Head Start centers (Bloom & Weiland, 2015). Researchers have found that children with higher quality preschool experiences had lower rates of behavior problems up to five years later (Peisner-Feinberg et al., 2001). While these effects were found for a diverse range of children, the effect of childcare on problem behaviors was more pronounced for preschoolers with mothers who had less education (Peisner-Feinberg et al., 2001). Similarly, a lack of interaction with other classmates has been found to influence later behavior problems. In a study of kindergarteners with no, low quality, average quality, or high quality friendships as reported by their parents, researchers found that having no friendships was associated with higher rates of externalizing behaviors for those students. Conversely, those students with high-quality friendships were found to demonstrate more prosocial behaviors than children in any of the other groups (Engle, McElwain, & Lasky, 2011). Classroom characteristics also play a role as a risk factor for behavior problems. Large student-teacher ratios as well as longer school days were associated with higher teacher ratings of problem behaviors (Mashburn, Hamre, Downer, & Pianta, 2006).

Along with parent and family risk factors, child characteristics can pose a risk for behavior problems. One important child-level risk factor for behavior problems in preschoolers is gender. Boys have been found to have overall higher rates of behavior problems than girls (Bulotsky-Shearer et al., 2008; Morgan et al., 2009; Spieker, Larson, Lewis, Keller, & Gilchrist, 1999). Specifically, both teachers and mothers have rated boys from two to nine years old as having more externalizing behaviors than girls (Miner & Clarke-Stewart, 2008). Another child
characteristic that can put them at risk for behavior problems is developmental delays. When compared with preschoolers without them, those children who had developmental delays were up to five times more likely to have worse social skills and higher rates of problem behavior (Merrell & Holland, 1997). Other problems that affect children from a young age, such as disorganized attachment as an infant, can also lead to problem behaviors in preschool (Shaw et al., 1996). Conversely, high child intelligence, emotionality, sociability, and likeability by others is related to lower behavior problems in young boys who are otherwise at risk (Owens, Shaw, Giovannelli, Garcia, & Yaggi, 1999).

**Evaluation of Current Behavioral Assessment Tools**

The early identification of behavior problems is important so difficulties can be addressed before these students fall too far behind, and quality behavior rating scales designed to assess young children are an important first step in that process. One of the reasons why early assessment of and intervention for behavior difficulties are needed is that parents of many young children are not taking advantage of the services available to them, potentially due to access and cost of services. In a study involving a nationally representative sample of children age three to 17, researchers found that only 21% of children who need mental health services actually use them, and the underutilization of services is greatest amongst Latinos (Kataoka, Zhang, & Wells, 2002). Additionally, other research has indicated that universal screenings for young children are effective for early identification of young children in need of interventions for behavior problems (Essex et al., 2009). Proper assessment tools are an important part of that process.

**Limitations.** Quality measurement tools with scores that have strong reliability and validity evidence are important when assessing behavior problems among preschoolers. Some commonly used tools for this age group include the following: Behavior Assessment System for
Children – Second Edition (BASC-2; Reynolds & Kamphaus, 2004), Preschool and Kindergarten Behavior Scales- Second Edition (PKBS-2; Merrell, 2002), the Preschool Behavioral and Emotional Rating Scale (PreBERS; Epstein & Synhorst, 2009), and Social Skills Rating System (SSRS; Gresham & Elliot, 1990). If behavior problems are not assessed and consequently not diagnosed properly, it may result in ineffective interventions and inappropriate labels. This issue is especially relevant to low-income and minority populations. Having empirical or theoretical support that an intervention will be effective with the minority population is important prior to testing an intervention designed for the majority culture on that ethnic minority population (Fisher et al., 2002). This support does not always exist, however.

One of the barriers to using high-quality measurement tools to assess behavior problems in preschoolers is that access to the most updated measures and assessment procedures in urban or low-income populations is scarce (Cooper et al., 2008). Consequently, inappropriate measures may be used with these children.

Additionally, when selecting assessment tools for use with minority and low-income students, the way race, ethnicity and culture are conceptualized must be considered (Fisher et al., 2002). Researchers need to make every effort to use tools that have been normed on as many of these children as possible, and when this is not possible, researchers, and even practitioners, must be sure they are operationally and conceptually assessing the mental health needs they are intending to study (Fisher et al., 2002). One rating scale that reveals these limitations when used with minorities is the Child Behavior Checklist (CBCL; Achenbach, 1991). The items on the CBCL have not been found to properly cover the range of problem behaviors exhibited by African American children. Researchers found that while the CBCL identified about half of the problem behaviors exhibited by their sample of African American children, it did not measure
approximately 20 behaviors exhibited by these students, such as not cooperating, disrespecting others, and refusing to complete schoolwork (Lambert, Rowan, Lyubansky, & Russ, 2002). This may limit the cultural relevance of this tool when making classification and intervention decisions for these students. Another measure designed to assess problem behaviors and ADHD in children, the ADD-H: Comprehensive Teacher’s Rating Scale (ACTeRS; Ullmann, Sleator, & Sprague, 2000), did not report the gender, SES, or ethnicity of the norm sample, thus potentially limiting its generalizability. Additionally, no validity evidence was reported (Demaray, Elting, & Schaefer, 2003).

One of the problems with some of the measurement tools currently in used in research to assess young children is that many researchers do not consider developmental issues when using them in their studies (Fantuzzo & Mohr, 1999). A developmental approach considers changes in children as they grow and how experiences shape their development (Fantuzzo & Mohr, 1999). Some of the main developmental issues that need to be considered for preschoolers include an understanding of the child in multiple domains, an understanding of various developmentally appropriate behaviors, and the importance of longitudinal data (Fantuzzo & Mohr, 1999).

Additionally, children with severe behavior problems have a variety of presentations of their problems that are not always easy to categorize (Drotar, 2002). The variance in how behaviors present themselves, coupled with rapid development, limits the effectiveness of behavior assessment tools for preschoolers. Another problem with the measures used to assess behavioral problems concerns the use of psychological checklists. When psychological checklists are used, especially with low-income, ethnic minority populations, they may be biased and not representative of those children (Fantuzzo & Mohr, 1999). These limitations parallel those
limitations found when using the behavior rating scales mentioned above (Demaray et al., 2003; Lambert et al., 2002).

Since behavior problems in the early years have not been found to decrease over time unless interventions are implemented (Briggs-Gowan & Carter, 2008; Kerr, Lunkenheimer, & Olson, 2007; Morgan, Farkas, & Wu, 2009), tools used to assess these problems should predict later behavior problems as well. Research on the predictive utility of behavior rating scales has been mixed. One study by Martin-Storey, Serbin, Stack, and Schwartzman (2009) examined a sub-sample of Canadian children whose parents were involved in a longitudinal study. Participants were assessed for behavior problems in preschool using the Behavioural Style Observational System (BSOS; Karp, Serbin, Stack, & Schwartzman, 2004) and again in early elementary school using the CBCL. The researchers found that the BSOS was effective in predicting later behavior problems in elementary school, but did not predict internalizing behavior problems in elementary school (Martin-Storey et al., 2009). Another study involving Canadian twins examined the validity of preschool parent rating scales in predicting teacher-rated hyperactive and impulsive behaviors in elementary school. Researchers found that parent ratings in preschool moderately predicted teacher ratings in the early elementary years accounting for 14.3 percent of the variance (LeBlanc et al., 2008).

**Strengths.** Behavior rating scales in general have some advantages over other behavior assessment tools. First, they are increasingly one of the most common assessment instruments used. Results of one study found that behavioral rating scales are used in 60 to 90 of assessments conducted by school psychologists, with 75% of psychologists using them in at least 8 of their last 10 assessments (Shapiro & Heick, 2004). Second, rating scales allow psychologists and researchers to collect data on a child’s behavior from people in different settings at different
times, including parents, teachers, and self-reports (Shelton & Barkley, 1994). They also allow clinicians to examine how severe a child’s symptoms are (Shelton & Barkley, 1994). Third, these raters can provide information about a child over a long period of time, thus documenting less frequent behaviors that might not occur during an observation (Shelton & Barkley, 1994). However, results of rating scales are based on the subjective views of the rater and not on purely objective observations, per se (Mandal, Olmi, & Wilczynski, 1999), in that teachers may rate a student they dislike more harshly than his or her behaviors actually are and conversely may rate other students more positively to avoid stigmatizing them with a label.

**Teacher rating scales.** Teacher rating scales, specifically, are a commonly used method for assessing behavior problems in preschoolers. Advantages of teacher rating scales as measurement tools are that teachers spend a great amount of time with the students during the day so they know the children better than an outside observer, and teachers are able to compare the target child with other students in the class. They are also generally more knowledgeable about children of that age given their experience. One study found that teacher ratings of severe behavior problems in young children were the best predictor of behavior problems two years later (Webster-Stratton, 1996). Additionally, teacher rating scales are cheap, quick to administer, and require little training (Volpe, Briesch, & Gadow, 2011). However, a critique of the use of certain rating scales to assess behavior problems in preschoolers is that scores of these measures have little reliability and validity evidence for young age groups even though they are being used to assess infants and preschoolers (Drotar, 2002). Although rating scales used to assess preschoolers have been found to have stronger technical properties than other preschool assessment tools like speech and language tests, they fare worse when compared with preschool cognitive tests (Bracken, Keith, & Walker, 1998). One study found that newer rating scales’
scores such as the BASC and the SSRS had stronger reliability and validity evidence than older ones, thus suggesting that clinicians should avoid assessing young children with these older rating scales (Bracken et al., 1998). For example, scores for competence on the CBCL demonstrated internal consistency of only .57 for boys and .62 for girls (Bracken et al., 1998). Additionally, the Attention Deficit/Hyperactivity Disorder Test (ADHDT; Gilliam, 1995) scores, designed to assess ADHD and behavior problems in children and young adults, was found to have missing inter-rater reliability estimates and questionable test-retest reliability (Demaray et al., 2003).

This questionable reliability and validity evidence for such scores used in research is especially evident for low-income, culturally diverse, and minority populations (Lopez et al., 2000). In a matched study of elementary teachers from a low-income inner city school, findings indicated that both Black and White teachers rate Black students as having more adjustment problems, fewer abilities, and worse academic outcomes (Pigott & Cowan, 2000). The findings also indicated that Black teachers rated all students more positively than White teachers. This potentially suggests that they view the importance of education and teacher-student connections for low-income children as more crucial to success, and thus had a higher tolerance for behavior problems than their White counterparts (Pigott & Cowan, 2000). In a similar study comparing low income preschoolers to middle income preschoolers, the findings indicated that teacher ratings of behavior problems were much higher for children from low income than from middle income, but ratings by trained observers were consistently lower in both the low and middle income groups (Phillips & Lonigan, 2010). The ethnicity of teachers may also impact how different groups of students are rated. Non-minority teachers have been found to report higher levels of student problems in classrooms with mostly minority students than minority teachers.
reported in similar classrooms (Rimm-Kaufman et al., 2000). Research on ethnic bias is mixed, however, with some studies finding different behavior ratings based on teacher and student ethnicity and others not finding a difference (Mason, Gurses, & Ney, 2014).

Teacher rating scales have also been criticized because research has found that some of the rating scales currently in use specifically for referrals and identification in preschools underreport the levels of behavior problems in these students (Fantuzzo et al., 2003). This limits the utility of teacher rating scales because in order to avoid labeling children, many of them may not being receiving proper services, and false negatives are created. Additionally, many teacher rating scales are designed to assess the number of behavioral symptoms, but not the intensity or frequency of these issues. This can contribute to problems with attributing behavior problems to specific contexts (Fantuzzo, et al., 2003). Teacher characteristics can also influence their rating of behaviors in their preschoolers. Research has found differences amongst teachers in their ratings of children (Mashburn et al., 2006). More specifically, experienced teachers are more likely to rate children as demonstrating more problem behavior, while teachers with greater self-efficacy are more likely to report pro-social skills (Mashburn et al., 2006).

Another important limitation of rating scales is that there is often disagreement amongst raters about the behaviors of the same child in both the same and different contexts. Inter-rater reliability estimates for behavior rating scales very often fall below the recommended .80 reliability criteria. Factors such as halo effects and situational influences often contribute to the variability in reliability estimates (Mandal et al., 1999). Teacher rating scales report a unique view of child behavior problems that is not often the same as the ratings of parents. While teacher ratings are better estimates of externalizing problems, parents are a better source for perceptions of internalizing problems (Konold & Pianta, 2007). Parents and teachers also often
report different rates of behavior problems (Ruffalo & Elliott, 1997). Miner and Clarke-Stewart (2008) found that mothers rated their children as having more externalizing problems than teachers did and that mothers and teachers ratings were often not related. These uncorrelated ratings may be due in part, however, to the difference experiences that the raters have with the children; yet, some research has found consistency across mother, father, and teacher ratings of preschool behavior (Kerr, Lunkenheimer, & Olson, 2007). Additionally, research on the consistency of parent rating scales is mixed. Results of one study indicated that parent ratings are generally consistent, but only weak relationships exist between mothers’ ratings and fathers’ ratings with teachers’ ratings of student behaviors (Ruffalo & Elliott, 1997). Results of another study found, with the exception of the CBCL, large differences between parent ratings across four different parent rating scales of preschool behavior (Walker & Bracken, 1996). Additionally, ratings were found to be more consistent between parents for more extreme behaviors exhibited by preschoolers and for girls (Walker & Bracken, 1996).

Rural and Urban Characteristics

Scores on scales used with at-risk rural students should be validated by including them in the norm sample because differences exist between rural and urban populations. Most research, however, has focused on disadvantage from an urban perspective (Brown, Copeland, Costello, Erkanli, & Worthman, 2009). Even though rural areas comprise a large geographic area, they account for a small relative and actual proportion of the U.S. population. Thus, a dearth of research exists on children from these areas (Maher, Frestedt, & Grace, 2008). In fact, as recently as 2012, only about 20% of students nationally attended schools in rural districts (Strange, Johnson, Showalter, & Klein, 2012). Even in national public-use files, information about rural children is often not supplied because so few children living in rural areas participate in these
studies, and specific demographics about participants is kept confidential (Capizzano & Fiorillo, 2004).

There is not one agreed upon way to define rurality (Coladarci, 2007). Rural classification systems use differing variables to determine rurality. Some researchers classify by population size, distance from an urban area, economic activity, or other factors (Coladarci, 2007). The context of a rural area varies from place to place even when measuring rurality with these standard variables, making rurality more heterogeneous than is acknowledged (Coladarci, 2007). Even though a consistent definition of rurality across the literature does not exist, the U.S. Census Bureau defines rural areas as “those areas with a population of less than 25,000 located outside of central metropolitan statistical areas and metropolitan statistical areas” (Grace et al., 2006, p. 3). This has been considered one of the more precise definitions of rurality, and thus was used in the current study (Capizzano & Fiorillo, 2004). Furthermore, the U.S. Department of Agriculture and the WWAMI Rural Health Research Center have used the U.S. Census Bureau’s definition of rurality to create 33 rural and urban census categories called Rural-Urban Commuting Areas (RUCAs; Virginia Department of Health [VDH] Office of Minority Health and Public Health Policy & Virginia Rural Health Association, 2008). Other definitions of rurality include the Isserman definition of rurality, which defines a rural county as one with less than 500 people per square mile and no urban city with 10,000 or more people (VDH et al., 2008).

One difference between rural and urban residents is that children from rural areas often have a greater attachment to the place they have grown up in than non-rural children. Findings from a study of nationally representative children living in rural areas indicated that fewer rural students wanted to go to a four-year college than non-rural students, even when accounting for
race, parent income, and parent education (Howley, 2006). Additionally, results of the study found that rural parents were more likely to turn down jobs out of town because they would have to move away from their community (Howley, 2006). Rural parents also perceive their neighborhoods to be safer than those who live in non-rural areas (Grace et al., 2006). Another family difference between rural and urban residents is that rural children have generally younger parents (Rogers, 2005). The United States Department of Agriculture’s (USDA) Economic Research Service, which collected national statistics on rural children and families, reported that 19.3% of rural children had a parent younger than 29 years old, compared to 15.7% of non-rural children (Rogers, 2005). This often leads to these children having lower parental income and education levels than those from urban areas (Rogers, 2005; Stanley, Comello, Edwards, & Marquart, 2008). Specifically, rural parents are less likely to have a bachelor’s degree or have an annual income of over $75,000 (Grace et al., 2006).

In one study of rural and urban middle- and high-school students from across the US, researchers found that students from urban areas participated in fewer extracurricular activities than rural students, both in school and outside of school (Stanley et al., 2008). In fact, parents from rural communities were also found to participate in more activities with their children at home than non-rural parents (Grace et al., 2006). With regard to parenting and parental characteristics, parents from rural areas were less likely to report becoming frustrated with their children. This is especially evident with parents with low incomes. However, rural parents who were White or with a higher income level were more likely to report that they spank their children weekly (Grace et al., 2006). Additionally, rural parents were more likely to report higher levels of externalizing problems in their young children than their urban and suburban counterparts (Sheridan, Koziol, Clarke, Rispoli, & Coutts, 2014). They were also less likely to
provide their preschoolers with emotional support (Sheridan et al., 2014). Further, maternal education was found to predict vocabulary knowledge and early literacy in rural areas, but not urban areas (Bojczyk, Rogers-Haverback, Pae, Davis & Mason, 2015).

Another difference between rural and urban children involves health disparities. Since rural families generally tend to be poorer than urban families and face logistical challenges to accessing services, they have historically been less able to provide health care for their children. In 2001, only 12% of urban children lacked health care coverage, whereas 22% of rural children were not insured (Rogers, 2005). With the recent implementation of the Affordable Care Act, however, this has begun to change. According to the U.S. Department of Health & Human Services, health care coverage in rural counties increased 8% between 2013 and 2015, and the number of rural residents who were not able to find health care coverage dropped by 6% (Avery, Finegold, & Xiao, 2016). Black children from rural areas were also found to be less likely than non-rural Black children to receive a medical check-up in pre-kindergarten (Grace et al., 2009). In addition to maintaining proper health for their children, rural mothers, especially White mothers, were more likely to report being depressed (Grace et al., 2009).

Physical activity is another domain of health where differences exist by setting. In a study of 3,416 elementary school students from Iowa, children from rural areas were found to be more active than urban students, especially during lunch time, but they were found to be more overweight than their urban peers (Joens-Matre et al., 2008). In a national study of children in kindergarten through high school, Lutfiyya, Lipsky, Wisdom-Behounek, & Inpanbutr-Martinkus (2007) found that children who live in rural areas are 25% more likely to be overweight or obese than children from urban areas, even when controlling for other factors like using technology and not being physically active. Transportation difficulties in rural areas may not only affect access
to physical activity centers, but other places such as daycare centers, grocery stores, and hospitals. This might lead to rural children being more overweight than urban children. In addition to rural children being more at-risk for overweight and obesity, these overweight rural children also are more likely to live in poverty, have no health insurance, and receive little medical care (Lutfiyya et al., 2007).

Specific school-based differences have also been found between rural and urban students. Daycare centers in rural areas have a smaller child to student ratio than those in urban areas (Maher et al., 2008). The sizes of schools, in general, are also smaller in rural areas, but the percentage of free and reduced lunches is greater (Stanley et al., 2008). Teachers were also found to be more likely to report behavior and adjustment problems in their kindergarteners in rural schools than urban schools (Rimm-Kaufman et al., 2000). These rural kindergartners are also enrolled in special education at higher rates than their urban counterparts (Grace et al., 2006); however, rural students with disabilities tend to spend more time in the general education setting than their urban peers (Brock & Schaefer, 2015). Additionally, one researcher found differences in educational aspirations between rural and urban students. In a study of students from metropolitan and non-metropolitan areas in the U.S., Howley (2006) found that while rural students aspired to four-year degrees at slightly higher rates than non-rural students, rural students also aspired to graduate from high school only at higher rates than non-rural students. Additionally, rural students in the study were found to aspire to less graduate education than their non-rural counterparts, indicating that overall rural students have lower educational attainment goals (Howley, 2006). In contrast, one study found that for African American children, achievement was negatively associated with living in an urban area (Luster & McAdoo, 1994). Additional educational benefits have been found for kindergarteners living in rural areas when
compared to those living in urban areas. They are more likely to be enrolled in a pre-
kindergarten Head Start class, have greater social competence, be in an organized classroom, and
receive more developmental evaluations (Grace et al., 2006). Specifically, over twice as many
Black children from rural areas attend Head Start as their non-rural counterparts (Grace et al.,
2006). When looking further at ethnicity and locality, White kindergartners from rural areas have
more orderly classrooms and more full-day kindergarten than their counterparts in urban areas.
Interestingly, Black kindergartners from rural areas have fewer internalizing problems and are
more likely to have a teacher who has taken a class in early childhood education than their urban
counterparts (Grace et al., 2006). Black kindergartners from rural areas were found to only be
able to identify a third of the beginning sounds in words that non-rural Black kindergartners
could identify, however (Grace et al., 2006). A racial dynamic also exists that contributes to
differences within rural communities and not just between rural and non-rural areas. For
instance, rural African American and Hispanic high schoolers perceive more barriers to complete
their education than their rural white peers (Irvin, Byun, Meece, Farmer, & Hutchins, 2012).

Findings on behavior and mental health problems in rural and urban areas have been
mixed. One Australian study found that people living in rural areas were more likely to commit
suicide than those living in urban areas (Sankaranarayanan, Carter, & Lewin, 2010). In another
study of rural youth in Tennessee, researchers found that rural adolescents were more often
diagnosed with emotional disturbance than urban adolescents. They were also more likely than
urban adolescents to be in multiple risk groups for behavior problems (Heflinger & Hoffman,
2009). In fact, rural youth accounted for 43% of all of those with an emotional disturbance
(Heflinger & Hoffman, 2009). Further, another study found that rural children were slightly more
likely to have mental health problems and behavioral difficulties than their urban counterparts
(Lenardson, Ziller, Lambert, Race, & Yousefian, 2010). They were also more likely to be frequently affected by these conditions and have less access to comprehensive mental health treatment than urban children (Lenardson et al., 2010).

Conversely, other researchers have not found that differences existed overall between rural and non-rural children and their rates of behavior problems, though it should be noted that rural children in that study who were from the West or were Native American were found to demonstrate greater rates of internalizing behaviors than their non-rural counterparts (Grace et al., 2006). Additionally, no differences were found between rural and urban parents in a study on African American’s perceptions on mental health. Both groups explained their lack of utilization of mental health services as due to the fear of their children being labeled with mental health problems (Mukolo & Heflinger, 2011). Furthermore, research has indicated that when compared to suburban students, urban children were more likely to react with anger when corrected; however, those location effects were not easily distinguished from other child variables like SES (Schaefer, 2004). Thus, due to these mixed findings, rating scales used on rural populations must be examined to confirm that they are appropriate for use with that subpopulation.

Although a body of research exists on differences between rural and urban settings, differences also exist across rural communities. Rural areas are not homogenous, and differences exist in rural patterns across the U.S. For example, rural poverty has been found to vary by region, with the South (especially the Southeast and Mid-South Delta) having the highest rates of poverty and poverty gaps with its metro counterparts (Johnson & Strange, 2007; United States Department of Agriculture Economic Research Service [USDA ERS], 2016). Additionally, the states where rural education is more important to the overall educational performance of the state are mostly in the Plains (Johnson & Strange, 2007). Further, rural students in more urban states
on the east and west coasts tend to underperform their rural counterparts from more traditionally rural states, as their needs are not the focus (Johnson & Strange, 2007). In the Southeast and Southwest, there is more rural student diversity, which is associated with lower SES and worse educational outcomes for these students (Johnson & Strange, 2007).

Even though some differences exist between rural and urban areas, there are other areas were no differences have been found. School adjustment, care from teachers, and care from parents have not been found to differ between urban and rural students (Stanley et al., 2008). Additionally, rural and urban children did not show differences in having family routines or discussing religion (Grace et al., 2006). Further, both rural and urban Head Start students have similar levels of vocabulary knowledge and early literacy skills, even though both groups have shown weaknesses in these areas when compared to the normative group (Bojczyk et al., 2015). Rural and urban students also aspire to the same amount of post-secondary education including two- and four-year degrees (Howley, 2006). Additionally, parental warmth and appropriate discipline have not been found to differ based on whether children live in rural or urban areas (Pinderhughes, Nix, Foster, & Jones, 2001).

**Adjustment Scales for Preschool Intervention**

Presently, a number of teacher rating scales are available to assess behavior problems in preschoolers. These rating scales allow for student behavior to be compared with appropriate peer behavior, allow behavior in different settings to be distinguished, and are easy to administer. Until recently, however, many measures have not measured the intensity of behaviors observed, have inadequate psychometric properties (e.g., CBCL, ADHDT), and have restricted norms for use with at-risk preschoolers (e.g., CBCL, ACTeRS). The Adjustment Scales for Preschool
Intervention (ASPI; Lutz et al., 2002) is a measure that was designed to address these limitations found in other rating scales and allows children to be rated in different classroom contexts.

Lutz et al. (2002) sought to create a measure of emotional and behavioral problems with high reliability and validity evidence that could be used in urban preschool settings, and to determine whether different aspects of problem behavior varied by the child’s gender and age. Gender and age specifically were analyzed because research has found that boys and girls exhibit different types of behavior problems and that younger preschoolers have more behavior problems than older preschoolers (Coolahan et al., 2000). They developed the ASPI as a behavioral assessment tool to address many of the drawbacks found in other measures of preschool behavior (Lutz et al., 2002). The ASPI specifically targets low-income students and routine classroom situations, as opposed to the more commonly assessed out-of-context, illness-based symptoms, while also using teacher-friendly language. It also refines the distinction of underactive preschool problem behaviors not measured in other behavior scales (Lutz et al. 2002). These scale characteristics have not yet been available to teachers rating student behavior in rural settings. Given that young rural students have unique behavioral needs, which include being provided less emotional support, having higher reported externalizing problems, and being enrolled in higher rates of special education than their non-rural peers, it is important to have a rating scale specifically for use with them (Grace et al., 2006; Sheridan et al., 2014). Appropriate early identification is important for these students to have better outcomes since outcomes of early behavior problems can be lasting into adulthood, especially if not identified and treated (Stevenson & Goodman, 2001).

Five behavioral dimensions are assessed via the ASPI: Aggressive, Oppositional, Inattentive/Hyperactive, Withdrawn/Low Energy, and Socially Reticent (Lutz et al., 2002). Since
this scale was developed for use with at-risk urban preschoolers, it is unknown if teacher ratings demonstrate acceptable reliability evidence for other children, notably rural preschoolers. Given that rural students have some behavior differences from their non-rural peers, such as higher rates of teacher-reported behavior problems (Rimm-Kaufman et al., 2000), greater social competence (Grace et al., 2006), and higher rates of externalizing problems (Sheridan et al., 2014), reliability differences may exist in the corresponding domains on the ASPI (Aggressive, Oppositional, Inattentive/Hyperactive, and Socially Reticent) when used on rural preschoolers as well.

**Rationale and Purpose**

Given the many similarities and differences in rural and urban communities, and the need for instruments’ scores psychometric characteristics to be explored for a sample of like individuals prior to use with those children, it must be determined whether the use of the ASPI is appropriate with both populations. Initial scale development of the ASPI focused solely on preschool children in urban areas (Lutz et al., 2002). The purpose of the current study is to provide reliability evidence for the ASPI scores using an at-risk rural preschool sample. Establishing both internal consistency and inter-rater reliability of scores is important. Internal consistency reliability determines whether the scores on the scale designed to measure the same construct are similar. Inter-rater reliability evidence is also important because relevant differences exist between teachers and assistant teachers that may cause their ratings to vary. Both interclass and intraclass approaches to inter-rater reliability are helpful in determining the consistency of scores between raters: an interclass approach calculates the strength and directionality of agreement while an interclass approach calculates the absolute agreement amongst pairs of raters. In prior Head Start research, ethnic differences between the teachers and
assistant teachers have emerged, with the assistant teachers being more ethnically similar to the students. In one study of preschoolers in Head Start, 30% of the teachers were identified as Black compared to 100% of assistant teachers who identified as Black (Fantuzzo, Coolahan, Mendez, McDermott, & Sutton-Smith, 1998). The assistant teachers are also commonly former Head Start parents (Bulotsky-Shearer & Fantuzzo, 2004; Fantuzzo et al., 1998). Since research is mixed on rural and urban differences, determining whether the ASPI reliably measures the same behaviors in both groups is important if it is being used to assess both populations. If behaviors are not properly identified and measured for a given group, inappropriate diagnoses and interventions may occur for those students. Thus, establishing inter-rater and internal consistency reliability evidence of scores on the measure with a rural sample is the next step.

Given the need for further research and psychometric support for the ASPI, the research questions are: (a) What is the inter-rater reliability for the ASPI scores when used to assess behavior problems in preschoolers in rural classrooms? (b) What is the internal consistency reliability for the ASPI scores when used to assess behavior problems in preschoolers in rural classrooms?

The research hypotheses are as follows:

Hypothesis 1: The scores on the ASPI will evidence acceptable inter-rater reliability evidence in the form of intraclass correlations of .75 or greater and interclass agreement of .80 or greater between teachers and assistant teachers.

Hypothesis 2: The ASPI score will evidence acceptable internal consistency of at least .70 for the sample as a whole and by gender.
Chapter 2: Method

Participants

The sample consisted of participants from five classrooms at a Head Start center in rural Pennsylvania, with a total of 57 students participating. Students were recruited from six classrooms with approximately 18 students in each class, with a return rate of informed consent from 55% of eligible students at the site. Student participants were from low-income homes; children are eligible for Head Start if they meet age criteria and are from families with incomes below the poverty line, are homeless, are in foster care, or are eligible for public assistance (Head Start Act, as amended, 2007). Student demographic data was not provided for one student. After missing demographic data was removed, the sample was comprised of 50% males and 50% females. The students ranged in age from three to five years old ($M = 3.34$, $SD = .64$). Of the sample, 87.5% of student participants were White, 1.8% were Black, 1.8% were Hispanic, and 8.9% were Multi-racial. This sample distribution by ethnicity differs from the nationwide distribution, but more closely reflects the national rural and small town ethnicity distribution (Housing Assistance Council, 2012). According to the National Head Start Association (NHSA; 2010), 40% of enrolled students in 2010 were White, 29% were Black, 8% were biracial or multi-racial, 6% were American Indian or Pacific Islander, and 17% were Other.

The analysis of student behavior was conducted by Head Start teachers and assistant teachers. Six teachers and six assistant teachers initially consented to participate. None of the parents in one of the six classes gave consent for their children to participate; therefore, five teachers and five assistant teachers participated in the study and completed rating scales on their students. All of the participating teachers and assistant teachers were White and female, which is generally representative of the population. Prior educational research consisting of teacher
samples has demonstrated that preschool and elementary teachers are overwhelmingly female, ranging from 90 to 96 percent of the total sample (Lutz et al., 2002; Pianta & Stuhlman, 2004). Additionally, according to the U.S. Department of Education (2016), national data indicates that elementary and secondary teachers in the United States are predominantly White, comprising 82 percent of all teachers. Teachers and assistant teachers ranged in age from 22 to 55 years old ($M = 36.80$, $SD = 12.24$). Assistant teachers were older, on average, ($M = 40.20$, $SD = 16.30$) than teachers ($M = 33.40$, $SD = 6.50$). Their years of teaching experience ranged from two to 22 years ($M = 7.80$, $SD = 5.65$). Assistant teachers averaged more years of teaching experience ($M = 9.00$, $SD = 7.71$) than teachers ($M = 6.60$, $SD = 2.97$). Of the sample of teachers and assistant teachers, 20% were high school graduates, 20% had some college, 30% had two-year college degrees, and 30% had four-year college degrees. All teachers had completed a two- or four-year degree, which is roughly commensurate with the national Head Start teacher average of 93% with degrees (Head Start Program Facts, 2012). All assistant teachers had completed high school, some college, or a two-year degree.

Measures

**Assessment Scales for Preschool Intervention (ASPI; Lutz et al., 2002).** The ASPI is a teacher rating scale designed to provide a multi-dimensional assessment of emotional and behavioral problems in preschoolers. Both positive and negative behaviors are assessed via teacher observations and ratings. Teachers must have worked with a student for at least two months to be able to complete the scale. The rating scale includes 24 contextual situations explaining 144 behavioral descriptors. Teachers completing the scale mark “yes” or “no” to indicate whether they observed each behavior from the student in the past two months.
The original norm sample consisted of 829 children from urban Head Start preschools in the Northeast, of which three quarters were African American. The participating teachers were mostly female, with a range of education and experience. The majority of the teachers were White. The Head Start teachers were asked to rate each preschooler using the Penn Interactive Peer Play Scale (PIPPS), the ASPI, and a demographic questionnaire, as well as a Conners Teacher Rating Scale (CTRS-28) for one girl and one boy in each class (Lutz et al., 2002).

Five behavioral dimensions are assessed via the ASPI: Aggressive, Oppositional, Inattentive/Hyperactive, Withdrawn/Low Energy, and Socially Reticent. The authors identified dimensions via exploratory principal components analysis. The factors were rotated via varimax and equamax rotations. The following criteria were used to decide how many factors to retain: scree plot, eigenvalue of at least 1, five items with loadings of at least .40, internal consistency of at least .70, minimized intercorrelation of unit-weight factors, and psychologically meaningful factors. The final solution went through a series of promax rotations. The final oblique solution was compared to the final orthogonal solution to select the one with the fewest interfactor correlations and complex loadings. To verify the factor structure arrived at in the principal components analysis, two follow-up procedures were conducted: comparison of the specificity and error variance, and principal components cluster analysis. The five factors were subsumed by two higher order factors: Overactive (Aggressive, Oppositional, Inattentive/Hyperactive) and Underactive (Withdrawn/Low Energy and Socially Reticent). The Aggressive factor consisted of 22 items describing rough behaviors. The Oppositional factor was composed of 11 items categorizing moody or controlling behaviors. The Inattentive/Hyperactive factor categorized inattentive, hyperactive, and impulsive behaviors via 10 items. The Withdrawn/Low Energy factor was composed of 18 items describing sluggishness and a lack of connection to class
activities. Additionally, the 12 items on the Socially Reticent factor described shy behaviors (Lutz et al., 2002).

Differences in the prevalence of both problem behaviors and learning behaviors between demographic groups were found among school-age youth (McDermott & Schaefer, 1996; Schaefer, 2004), and this was true for the ASPI, as well. Younger children were found to be more socially reticent, withdrawn/low energy, and inattentive/hyperactive, while boys were found to demonstrate more aggressive and inattentive/hyperactive behaviors. Each of the five dimensions displayed adequate internal consistency reliability ranging from .78 to .92. Scores on the ASPI also demonstrated adequate convergent validity evidence with both the PIPPS and the CTRS-28.

*Scale development.* The ASPI was developed by the researchers by adapting the Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993) for preschoolers. In order to adapt the ASCA for use with preschoolers, the rating scale was reviewed by 10 Head Start professionals, including teachers and administrators with training in early childhood education. Each reviewer was instructed to read each item and remove anything that was irrelevant for preschoolers and add any preschool related topic they thought was important (Lutz et al., 2002). Based on that review, five items were removed and 14 items were added to the contextual situations. Wording changes were also made to reflect preschool-age students. Upon final review, the Head Start reviewing staff noted that the ASPI was teacher-friendly, reflected observable classroom behaviors, and was non-offensive (Lutz et al., 2002). Raw score totals for each problem behavior domain were then aggregated and converted into corresponding normalized $T$ scores (Lutz et al., 2002). $T$ scores of 70 or greater, which are two standard deviations above the mean, were recommended as diagnostic cutoffs for each of the
problem behavior domains. They were noted to be appropriate for use with low-income preschool populations in identifying significant emotional and behavioral difficulties. (Bulotsky-Shearer & Fantuzzo, 2004).

**Reliability and validity evidence of scores.** Each of the five dimensions displayed adequate internal consistency reliability. The Cronbach’s alpha was .92 for Aggressive, .85 for Withdrawn/Low Energy, .78 for Oppositional, .79 for Inattentive/Hyperactive, and .79 for Socially Reticent. Strong convergent and divergent validity evidence with the PIPPS and CTRS-28 was established (Lutz et al., 2002). Additionally, in establishing content validity, 22 items that were displayed in higher frequencies of preschoolers were removed from the scale because the purpose of the scale was to identify problem behaviors. Then the principle components analysis (PCA) was run, which identified the five factors and established content validity of the scores (Lutz et al., 2002).

With regard to age and gender differences, boys displayed more aggressiveness and inattention/hyperactivity. Also, younger children were more socially reticent, withdrawn, and inattentive/hyperactive (Lutz et al., 2002). The ASPI also provided an estimate of the frequency of problem behaviors in different school settings for each student. Results of the study found that more behavior problems were noted during child-led and group activities, and fewer behavior problems were noted in one-on-one situations (Lutz et al., 2002). One of the limitations of this measure, however, is that it was only validated with at-risk urban samples.

A more recent study sought to extend the reliability and validity evidence for the ASPI. Bulotsky-Shearer and Fantuzzo (2004) examined the use of the ASPI as a measure of behavior problems in preschoolers in urban Head Start classes. Teachers and assistant teachers were each provided with an ASPI to complete for each child in their class and were experimentally blind to
the purpose of the experiment. High levels of agreement were found for all five behavioral dimensions, with Aggressive, Oppositional, and Inattention/Hyperactivity having the strongest agreement. Inter-rater reliability estimates ranged from .49 to .76 (Bulotsky-Shearer & Fantuzzo, 2004). As with the initial study, reliability estimates were calculated for the ASPI in urban Head Start classrooms. The validity of scores was also confirmed by establishing convergent validity with the Revised Edition of the School Observation Coding System (REDSOCS), a direct observational measure of behavior (Bulotsky-Shearer & Fantuzzo, 2004).

**Demographic questionnaire.** Demographic information was collected for each student, teacher, and assistant teacher. Each questionnaire included questions about gender, age, and ethnicity. Additionally, for teachers and assistant teachers, educational background questions about years of teaching experience and highest degree completed were included.

**Procedure**

Participants were recruited for the study via a letter sent from the principal researcher to the director of Head Start Centers that met study recruitment criteria. If no response was given, the principal researcher followed up with a phone call within the next few weeks. To encourage a high rate of return, all centers were offered a summary of the current research on preschool behavior problems as an incentive for participation in the study. Additionally, all Head Start Centers were offered an evening workshop or in-service presentation to their staff on behavioral interventions. Obstacles to a high return rate included prior partnerships of Head Start Programs with other universities, a lack of response from Head Start contacts even with multiple attempts to reach out, and staff time constraints with other professional commitments and responsibilities.

As with the initial scale development, teachers must have worked with a student for at least two months for them to participate. A letter of informed consent was distributed to all
teachers and assistant teachers prior to their participation in the study. Both the teacher and assistant teacher in a classroom had to agree to participate for the classroom to be included in the study. A letter of informed consent was also sent home to the parents of the students in participating classrooms. This informed consent form had to be completed in order for their child to participate and be rated by his/her teacher. A demographic questionnaire was also included with the informed consent for parents to complete on their participating child.

The researcher’s graduate assistant distributed an assessment packet to each teacher and assistant teacher to complete with instructions on how to complete them. The packet consisted of an ASPI for each participating student in that class. Additionally, each teacher and assistant teacher was also given a demographic questionnaire to complete. Both teachers and assistant teachers were instructed to complete the surveys independently by not collaborating with each other while doing so. After five weeks, the graduate assistant collected the packets from each teacher. Data was distributed in April 2013 and collected in June 2013.

**Data Analyses**

This study investigated the strength of inter-rater and internal consistency reliability evidence of the ASPI scores for use with an at-risk rural preschool population. First, preliminary analyses were run on the data to ensure that the assumptions of independence of observation and normality were met. To address normality, the criteria for skewness and kurtosis were set at less than the absolute value of $\geq 2.0$ for skewness and $\geq 7.0$ for kurtosis (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Missing data were removed via listwise deletion if they were missing at random (MAR). Next, inter-rater reliability was determined via interclass (which calculates the strength and directionality of agreement) and intraclass (which calculates the absolute agreement amongst pairs of raters) approaches. A two-way random effects model of
intraclass correlation (ICC) was utilized to determine the absolute agreement between teacher
and assistant teacher ratings of preschool problem behaviors (Shrout & Fleiss, 1979). ICC was
selected because it measures both the type and magnitude of variance (Suen, 1988). Portney and
Watkins (2000) established criteria for acceptable intraclass correlation coefficients as follows:
A correlation of .75 or greater is Good, between .5 and .74 is Moderate, and .49 or lower is Poor.
Pearson’s $r$ was used as the interclass approach to determine the proportion of variance and
establish positive linear relationships between raters in each domain. Sattler’s (2001) criteria for
using assessment tools as part of the clinical and psychoeducational decision-making process is
set at .80; thus, the desired acceptable inter-rater reliability was set at .80 or greater. $T$ tests were
also run to determine if mean differences existed between teachers and assistant teachers across
each problem behavior domain, since it is necessary to confirm that mean ratings across
observers are not statistically significantly different in establishing interclass agreement for
interval data (McDermott, 1988). Further, Cronbach’s alpha was utilized to establish acceptable
internal consistency of scores of at least .70 (George & Mallery, 2003). George and Mallery
(2003) recommend the following guidelines for internal consistency as follows: Cronbach’s
alpha of .90 or greater is Excellent, .80 to .89 is Good, .70 to .79 is Acceptable, .60 to .69 is
Questionable, .50 to .59 is Poor, and less than .50 is Unacceptable. Internal consistency
reliability coefficients were calculated for the whole sample, as well as by gender. Gender
differences have already been established for the ASPI with the urban sample (Lutz et al., 2002);
thus, data were also examined for gender effects with rural students. The primary hypothesis of
this study was that the scores on the ASPI will evidence acceptable inter-rater reliability
evidence in the form of intraclass correlations of .75 or greater and interclass agreement of .80 or
greater between teachers and assistant teachers. The second hypothesis was that the ASPI score
will evidence acceptable internal consistency of at least .70 for the sample as a whole and by gender. Data were analyzed using the statistical software program SPSS version 24.
Chapter 3: Results

Analytic Strategy

Data were analyzed using intraclass correlations (ICC), Pearson’s $r$, and Cronbach’s alpha to determine intra-rater agreement, inter-rater agreement, and internal consistency, respectively. Descriptive statistics (means and standard deviations) were calculated for scores where appropriate. Independent sample $t$-tests were computed to determine if differences in means between raters in each domain were statistically significant. The a priori level for statistical significance was set at .05.

Preliminary Analyses for Reliability Estimates

The data were screened for missing values. No data were missing or removed. No outliers were removed. The assumptions of independence of observation and normality were met for all ASPI variables (Oppositional, Aggressive, Inattentive/Hyperactive, Withdrawn/Low Energy, and Socially Reticent scales for both teacher and assistant teacher raters). The criteria for skewness and kurtosis were set at less than the absolute value of 2.00 for skewness and less than the absolute value of 7.00 for kurtosis (Fabrigar, Wegener, MacCallum, & Strahan, 1999). No variables were skewed at more than the absolute value of 1.20 or kurtotic at greater than the absolute value of 1.25. The prevalence rate of significant problem behavior as indicated by $T$-scores of 70 or greater ranged from 0 to 4 percent for teachers and 0 to 5 percent for teaching assistants, reflecting national prevalence rates (Huaquing & Kaiser, 2003). Table 1 provides means and standard deviations for each rater (teacher and assistant teacher) across each of the five ASPI problem behavior domains. As reported in Table 1, none of the means across any of the five domains evidence a statistically significant difference between raters; therefore, significant differences in how teachers and assistant teachers on average rated their students’
behavior problems were not identified for any domain. These means were generally commensurate with normative sample scale means (Lutz et al., 2002), with the exception of oppositional and inattentive/hyperactive behaviors, which were slightly higher in the current investigation for assistant teacher ratings.

**Intraclass Correlation of Teacher and Assistant Teacher Behavior Ratings**

A two-way random model of intraclass correlation was run for each of the five ASPI problem behavior domains (Aggressive, Oppositional, Inattentive/Hyperactive, Socially Reticent, and Withdrawn/Low Energy; Shrout & Fleiss, 1979) to establish reliability between teacher and assistant teacher ratings of student behavior. ICC estimates were calculated separately for both the expected reliability of the mean rating, which is the inter-rater reliability of all ratings on average, and of a single rater, which is used to determine how accurate a single rater of one student’s behavior would be (Shrout & Fleiss, 1979). All five domains were positive and statistically significant at \( p < .05 \) for both the mean correlation and single observer. Table 2 provides these intraclass correlation coefficients. Per Portney and Watkins’ (2000) criteria, the Inattentive/Hyperactive domain evidenced Poor agreement, and all other domains (Aggressive, Oppositional, Socially Reticent, and Withdrawn/Low Energy) evidenced Moderate agreement when estimating for single raters. Additionally, the Aggressive and Oppositional domains demonstrated Good agreement, and Inattentive/Hyperactive, Withdrawn/Low Energy, and Socially Reticent domains evidenced Moderate agreement, when estimating for the scale on average.

**Interrater Agreement between Teacher and Assistant Teacher**

Bivariate correlations were calculated between teacher and assistant teacher ratings to determine the strength and directionality of inter-rater agreement. Pearson \( r \)’s were positive and
statistically significant across all domains \( r = .69, .66, .46, .56, .54 \) for Aggressive, Oppositional, Inattentive/Hyperactive, Withdrawn/Low Energy, and Socially Reticent, respectively). Widaman (1993) reported that measured variable reliabilities of .60 to .85 are commonly acceptable for research. Only the Aggressive and Oppositional domains met acceptable criteria.

**Internal Consistency of ASPI Scores**

**Teacher Ratings.** All internal consistency calculations for each rater are provided in Table 3. Calculations are provided for the total sample as well as males and females separately. Per George and Mallery’s (2003) criteria, the Overactive factor consisting of Aggressive, Oppositional, and Inattentive/Hyperactive behaviors evidenced Acceptable internal consistency of .70 or greater \( (\alpha = .88, .79, \text{ and } .73, \text{ respectively}) \). The Underactive factor, which consists of Withdrawn/Low Energy and Socially Reticent behaviors evidenced Questionable internal consistency of scores \( (\alpha = .65 \text{ and } .69, \text{ respectively}) \).

The internal consistency of scores was also examined for male and female students separately. For female student behavior, the following domains evidenced Acceptable to Good internal consistency: Aggressive \( (\alpha = .88) \), Withdrawn/Low Energy \( (\alpha = .80) \), Oppositional \( (\alpha = .78) \), and Inattentive/Hyperactive \( (\alpha = .81) \). The Socially Reticent domain \( (\alpha = .63) \) evidenced Questionable internal consistency, however. For males, the Aggressive \( (\alpha = .87) \) and Oppositional \( (\alpha = .80) \) domains evidenced Good internal consistency, but the Withdrawn/Low Energy \( (\alpha = .61) \), Socially Reticent \( (\alpha = .65) \), and Inattentive/Hyperactive \( (\alpha = .64) \) domains demonstrated Questionable internal consistency.

The scale was examined further to determine if the internal consistency of scores would improve from the removal of individual items. Per Leech, Barrett, and Morgan (2008), items
with corrected item-total correlations of less than .30 should be removed. In the Aggressive domain the removal of items 98 “Disrupts games by fooling around,” 123 “Unkind to smaller or weaker children,” and 142 “Without warning or apparent reason he/she may throw an object across the room, sweep things onto the floor, etc.” would result in minimal improvements in internal consistency. In the Oppositional domain item 99 “Refuses to take turns” was not endorsed by any teacher (thus not accounting for any variability) and was not included in the calculation of internal consistency for this domain. All other items contributed to establishing acceptable internal consistency for the Oppositional scale. In the Inattentive/Hyperactive domain, items 14, 81, and 105 were poorly correlated with the total of the other items, but only the removal of items 81 “Doesn’t complete projects” and 105 “Rather loud but not disruptive in free play” would improve the internal consistency of scores in this scale. In the Withdrawn/Low Energy domain, items 19 “Too lethargic to ask for help,” 84 “Sits meekly, afraid to budge,” and 106 “Is too timid to join in during free play” were not endorsed by any rater and were not included in the internal consistency estimates. Additionally, items 47, 71, 74, 100, 115, and 129 were poorly correlated with the sum of the other items in this domain. Only the removal of items 71 “Listless, seems unmotivated,” 74 “Only gets involved in activities with an adult’s help,” and 115 “Ignores all other students” would increase the internal consistency in this scale if removed, however. In the Socially Reticent domain, items 17, 31, 40, 67, and 110 did not evidence moderate correlation with the sum of the other items; however, the internal consistency of scores would only be improved by the removal of items 17 “Freezes up and doesn’t answer questions” and 110 “Needs teacher assistance to get involved in free play”.

**Assistant Teacher Ratings.** As with teacher ratings, for assistant teachers, internal consistency reliability was adequate for the Overactive factor; the Aggressive domain (α = .93)
demonstrated Excellent reliability, and the Oppositional ($\alpha = .77$) and Inattentive/Hyperactive ($\alpha = .70$) domains demonstrated Acceptable reliability. The Underactive factor demonstrated Questionable internal consistency reliability for both the Withdrawn/Low Energy ($\alpha = .65$) and Socially Reticent ($\alpha = .64$) domains.

Internal consistency was once again evaluated for male and female students separately, for assistant teacher ratings. For male students, the Overactive factor demonstrated Adequate to Excellent internal consistency of scores for Aggressive ($\alpha = .93$), Oppositional ($\alpha = .80$), and Inattentive/Hyperactive ($\alpha = .76$) domains. The Underactive factor evidenced Questionable internal consistency in the Withdrawn/Low Energy ($\alpha = .66$) and Socially Reticent ($\alpha = .65$) domains. For female students, Aggressive ($\alpha = .93$) and Oppositional ($\alpha = .76$) behaviors evidenced Excellent and Acceptable internal consistency, respectively, the Inattentive/Hyperactive domain ($\alpha = .63$) demonstrated Questionable internal consistency, and for the Socially Reticent domain ($\alpha = .50$), it was Poor. For female students in the Withdrawn/Low Energy domain, internal consistency was also Poor ($\alpha = .11$). Only five of the 17 items that constitute the Withdrawn/Low Energy scale were endorsed by assistant teachers. The other 12 items were therefore not included in the internal consistency calculations for assistant teacher ratings of females in this domain.

Individual items were also examined to determine their contribution to establishing adequate internal consistency. In the Aggressive domain, items 101 and 123 were poorly correlated with the total of the other items; however, the removal of these items would result in minimal increases in the internal consistency of scores. In the Oppositional domain, item 125 “Tells on others to gain teacher’s favor” was poorly correlated with the sum of the other items in this domain and would improve the internal consistency of scores if removed. In the
In the Inattentive/Hyperactive domain, items 26, 105, and 112 were not moderately correlated with the total of the other items in this domain; the removal of items 26 “Much too talkative with teacher” and 105 “Rather loud but not disruptive during free play” would improve the internal consistency. In the Withdrawn/Low Energy domain, items 48 “Too lacking in energy to be troublesome,” 77 “Lacks physical energy when working with hands,” 84 “Sits meekly, afraid to budge,” 96 “Sluggish, apathetic in games,” and 115 “Ignores all other students” were not endorsed by any rater and were not included in the internal consistency estimates. Items 19, 47, 61, 69, 71, 83, and 129 were not moderately correlated with the sum of the other items in this domain. Items 47 “Never any trouble because he/she is so timid,” 69 “Cannot work up the energy to face anything new,” 71 “Listless, seems unmotivated,” 83 “Sits lifelessly during teacher directed time,” and 129 “Does not stand up for self during conflicts” would increase the internal consistency in this scale if removed. In the Socially Reticent domain, item 28 “Aloof, seldom says anything” was not included in the Cronbach’s alpha calculation since it was not endorsed by any assistant teacher. Additionally, items 8, 67, 73, 110, and 116 evidenced poor correlation with the sum of the other items for this domain. The internal consistency of scores would only be improved by the removal of items 73 “At times does not participate in activities” and 116 “Sometimes wanders off by him/herself,” however.

Certain items were repeatedly identified as ones that, once removed, would result in the internal consistency of scores; this was the case for teacher ratings, assistant teacher ratings, and when the sample was divided by gender. Five items were consistently identified for removal across raters. In the Inattentive/Hyperactive domain, the removal of item 105 would improve the internal consistency if removed, and in the Aggressive domain, the removal of item 123 would improve internal consistency. In the Withdrawn/Low Energy domain, items 71, 84, and 115 were
either not endorsed or would increase the internal consistency if removed. Table 4 lists the
corrected item-total correlation and Cronbach’s alpha if the item was deleted for all items that
did not evidence adequate internal consistency evidence.
Chapter 4: Discussion

Initial scale development of the ASPI focused solely on preschool children in urban areas; therefore, the aim of this study was to provide reliability evidence for the ASPI scores using an at-risk rural preschool sample to determine whether the psychometric properties were also robust in that sample due to inherent differences between rural and urban communities. Prior research has revealed differences between rural and urban communities including lower levels of education and income in rural areas (Grace et al., 2006; Rogers, 2005; Stanley et al., 2008), less healthcare for rural children (Lutfiyya et al., 2007), higher rates of special education enrollment in rural schools (Grace et al., 2006), and greater reporting of behavior and adjustment problems by teachers of rural students (Rimm-Kaufman et al., 2000). This suggests possible differences in definition, measurement, and use of the ASPI between these areas.

The focus of the primary hypothesis was that the scores on the ASPI would demonstrate acceptable reliability evidence in the form of intraclass correlations of .75 or greater and interclass agreement of .80 or greater between teachers and assistant teachers. Statistically significant inter-rater reliability between teacher and assistant teachers on all five ASPI problem behavior domains was established via intraclass correlations. The strength of that relationship was Moderate for all domains with the exception of inattentive/hyperactive behaviors for a single rater (which was Poor), meaning that for any given rater of a student’s behavior, the result should be moderately reliable, with the exception of inattentive/hyperactive behaviors. For the scale mean, the strength of the relationship was Moderate to Good for all domains. Since reliability estimates for the mean rating demonstrated stronger reliability evidence across all domains than reliability for a single rater, this suggests that averaging the ratings of a student’s teacher and
assistant teacher for clinical use may deliver more consistent estimates of a student’s classroom behavior than for a single rater; though, this is currently uncommon in assessment practice.

Positive interrater agreement was also established for ratings of all five problem behaviors through an interclass approach. The strongest relationship was identified for oppositional and aggressive behaviors as compared to Underactive or inattentive/hyperactive behaviors. Acceptable interclass reliability estimates were not established with this small sample for use in clinical decision making, but suggest possible use for research purposes.

The second hypothesis was that the ASPI scores would evidence acceptable internal consistency of at least .70 by gender and for the sample as a whole. The scores on the Overactive factor were internally consistent for both teacher and assistant teacher ratings, but demonstrated Poor to Questionable internal consistency for the Underactive factor. Similarly, for ratings of male and female students separately, Overactive behaviors were generally more internally consistent than Underactive behaviors. Specifically, for both male and female students, oppositional and aggressive behaviors demonstrated Acceptable internal consistency and socially reticent behaviors did not. Differences were noted between genders across raters regarding inattentive/hyperactive and withdrawn/low energy behaviors, however. For female students, teacher ratings were internally consistent for both inattentive/hyperactive and withdrawn/low energy behaviors; while for male students, teacher assistant ratings were only consistent for inattentive/hyperactive behaviors. It is possible that these gender differences across the inattentive/hyperactive and withdrawn/low energy domains indicate that the ASPI measures behaviors differently for male and female students or that male and female students express these conditions differently; however, given the inconsistency in pattern between genders across raters, these differences are more likely accounted for by limitations due to a small sample size.
Some of the inconsistency in internal consistency between genders may also be due to the inclusion of certain items that were recommended for removal from the scale per criteria of Leech et al. (2008). One reason for the poor internal consistency may be because many of these items used phrasing that may have been interpreted differently by different raters or made them more hesitant to endorse the item such as “listless,” “afraid to budge,” and “weaker children.” More likely, however, is that since many of these items were not endorsed by any rater, there likely was not an opportunity to observe some of these behaviors in a small sample. These items may reflect less common behaviors, but they may still be important in distinguishing certain behavior problem domains from other behaviors. Therefore, while the internal consistency would improve if these items were removed from the current sample, they may have more importance if the scale was administered to a larger sample comparable to the normative sample.

All scales were statistically significant across multiple reliability estimates, but the strength of those relationships, and thus the practical utility of the scale, varied. Across all three measures of reliability, ratings of Overactive behaviors, especially aggression and oppositionality, were more reliable than ratings of Underactive behaviors. The other Overactive scale measuring inattentive/hyperactive behaviors did evidence some inconsistency, however, in that it evidenced acceptable internal consistency reliability but did not evidence acceptable inter-rater reliability. In general, this finding regarding Overactive behaviors is consistent with previous research on the ASPI with urban samples, which found a similar trend; oppositional, aggressive, and inattentive/hyperactive behaviors demonstrated the strongest agreement between raters (Bulotsky-Shearer & Fantuzzo, 2004). Overactive behaviors are generally more apparent to observe and can be disruptive to instruction, which may explain why the educators were more consistent in their ratings of those behaviors compared to Underactive behaviors.
Based on the reliability of scores obtained from this sample, the utility of the ASPI in practice has not yet been established for use with rural preschoolers as part of the clinical decision-making process via individual ratings of student behavior, as the scores do not meet best practice reliability criteria of .70 for lower stakes assessments and .90 for high stakes assessments as indicated by De Champlain (2010). Sattler (2001) similarly outlined .80 as criteria for using assessment tools as part of the clinical and psychoeducational decision-making process. Since consistency of scores is needed for making decisions about diagnoses, interventions, and program planning for students, acceptable reliability evidence has not been consistently established across three methods for the use and interpretation of the ASPI in practice in a rural setting. However, based on criteria outlined by Schmitt (1996), the ASPI may be acceptable for research purposes, with Cronbach’s alpha of less than .70 often being considered acceptable. Additionally, since stronger reliability estimates were established for average ratings, this suggests the possible clinical utility of the scale by averaging two observer ratings for a common setting (e.g., classroom).

Limitations

While previous research has established acceptable inter-rater reliability evidence of scores on the ASPI for use with an urban sample (Bulotsky-Shearer & Fantuzzo, 2004), the current investigation did not yield the same results. These differences may be related to not only rural and urban differences, but also a combination of demographic and sample size differences. The previous study by Bulotsky-Shearer and Fantuzzo (2004) consisted of 199, mostly African American, preschoolers from a large urban district in the northeastern United States. The current study, however, consisted of a smaller sample size of mostly White students from one rural site in the Northeast. Specifically, the small sample size limited the endorsement of certain items that
are less commonly observed, but contribute importantly to the classification and identification of extreme problem behaviors in each domain. Additionally, although the current sample was generally ethnically representative of rural and small town preschool students, according both to the Housing Assistance Council (2012) and Strange et al. (2012), it was not representative of either the initial norm sample of students or the sample used to establish initial reliability information, which were both predominantly African American samples (Lutz et al., 2002). Therefore, differences in findings from the previous reliability study may also reflect an interaction of rurality and ethnicity. The impact of ethnic differences in behavior is supported by previous research that has found differences in how teachers rate behavior problems between their White and Black students (Feil et al., 2005; Miner & Clarke-Stewart, 2008; West et al., 2000). Future research should attempt to establish reliability evidence of the ASPI with a more diverse sample of rural preschoolers. Similarly, given that all teacher raters in the current investigation were White females, the findings may reflect a White cultural perspective of problem behavior and may not be generalizable to a diverse range of preschool teachers and assistants at large, given that Black teachers tend to rate their students more positively than White teachers (Pigott & Cowan, 2000).

Another limitation of the current study is that the ASPI was initially developed for use with children ranging in age from three to six years old; however, no six-year-olds were included in the current study’s sample. As reported by Head Start (2012), less than two percent of Head Start enrollment nationally consists of six-year-olds, so this impact is likely small. Additionally, the prevalence rates of significant behavior problems in the current sample reflected the national rates, and did not reflect the higher rates of problem behavior found among at-risk Head Start students (Harden et al., 2000; Huaquing & Kaiser, 2003). This may have contributed to
differences in reliability between the rural and urban samples. Further, given the small sample size and insufficient reliability evidence, validity could not be established, which is necessary for the scale to be utilized on rural students in practice.

**Implications for Research and Practice**

When doing research with rural samples, comparing the results to similar studies with an urban sample is considered an acceptable strategy for establishing whether findings are an inherently rural phenomenon, since the urban studies serve as a basis for comparison (Coladarci, 2007). This study added to the body of literature that supports why this comparison between similar studies with rural and urban samples is important. Specifically, as previously noted, results of the current study did not establish consistent support for the use of the ASPI with rural students, even though its use was confirmed for urban students (Bulotsky-Shearer & Fantuzzo, 2004). Even though this study offered a desired comparison between rural and urban settings, the base rate of problem behaviors in the current sample reflected that of the national sample, not of an at-risk sample; therefore, the difference in findings may partially be impacted by the difference in prevalence of problem behaviors between settings. Practitioners should consider the base rate of problem behaviors in the normative sample when deciding which rating scales to use for screening purposes compared to individual student decision-making for their population of students; this selection can impact both their identification of problem behavior and selection of interventions for these students.

Another avenue for research is to develop a different scale that has utility across both rural and urban settings to diversify its use in the field of behavioral assessment; researchers and scale developers should include children from both rural and urban communities. Further, since acceptable internal consistency evidence of scores has been established for use in research, the
ASPI is a new tool that can be used in research on behavior problems of at-risk rural preschoolers. Further, as previously noted, a system could be established to allow ratings of two teachers to be averaged for observations in a given setting. This approach may limit individual rater perspective, but may provide a more consistent view of a student’s problem behaviors.

Problem behaviors need to be properly identified and measured so that there are appropriate diagnoses and interventions for students with behavioral concerns. Since strong reliability evidence was not established for the ASPI with a rural sample, like it was with an urban sample, the findings suggest that there may be a difference between how teachers conceptualize behavior problems in these two areas. The implications of these findings are limited, however, by the small sample size, the lack of potential exposure to some less common, but important, indicators of significant problem behavior, and low base rates of problem behaviors. Even so, school districts, preschools, and Head Start Centers in rural areas will want to ensure that the behavior rating scales they use for screenings and early identification have included rural students in the initial scale development before using them with their students.

**Directions for Future Research**

Future research should aim to address some of the demographic limitations of the present study. This investigation extended reliability information on a rural sample of predominantly White Head Start students and teachers from the northeastern United States. Future research on the reliability of ASPI scores with rural preschoolers should include a larger, more ethnically diverse sample of students and teachers from different rural regions of the country. Additionally, low-income students from non-Head Start centers and home-based care should be included in future studies so the instrument can be used on a wider range of at-risk preschool students who are also in need of early identification of problem behavior. This is important because those
students who lack a specific early care arrangement, which is more common for at-risk, low-SES children, generally have worse learning behaviors entering kindergarten than their peers who do attend a preschool center (Rathbun & Zhang, 2016). Additionally, since parent input plays a large part in the early childhood care and the identification of behavior problems, researchers should develop a parent form of the ASPI with similar contextual situations.

Further, even though acceptable reliability was not consistently established on the ASPI in a rural setting, future studies should aim to extend reliability information in other ways, such as establishing test-retest reliability to determine consistency of ratings over time. Once strong reliability estimates are established, future research should also establish validity information for use in rural areas.

The ASPI may have predictive utility in identifying problem behaviors that later result in a behavioral or emotional condition such as Attention Deficit Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), Conduct Disorder, Generalized Anxiety Disorder or Major Depressive Disorder; therefore, future research should investigate the effectiveness of the scale in identifying these early behavioral indicators that are later formally diagnosed. Across genders, the establishment of acceptable score reliability for inattentive/hyperactive and withdrawn/low energy behaviors was inconsistent; therefore, clinical thresholds should be identified for males and females separately to supplement the thresholds that already exist for the total sample. Future studies should continue to establish and expand psychometric support for the ASPI in the form of reliability and validity evidence, so it can serve a need in the early identification of problem behaviors across a geographically and ethnically diverse range of at-risk preschool populations.
Conclusion

The early identification of behavior problems is important since they are a primary concern in educational settings and can impact student learning. The purpose of this study was to extend the reliability evidence of the ASPI (Lutz et al., 2002) scores using an at-risk rural preschool sample since prior research has established acceptable reliability and validity evidence for its use in urban settings, but not with rural students (Bulotsky-Shearer & Fantuzzo, 2004; Lutz et al., 2002). Sufficient reliability evidence was not consistently established for the use of the ASPI in rural settings, though it may have use with this population in research or when averaging rater scores. This limited reliability evidence may be due, in part, to study limitations such as a small sample size and differences in the ethnic makeup of the sample as compared to the initial normative sample. Future research is recommended to replicate this research with a larger sample of geographically and ethnically diverse rural preschoolers.
References


doi:10.1002/icd


doi:10.1177/0734282906290594


doi:10.1002/pits


Appendix A

Sample APSI questions adapted from Lutz et al. (2002)

Directions: After each question, there are several descriptions of behaviors children may display. Fill in the circle beside any description that fits the child’s behavior over the past month. For each question, mark as many descriptions as apply to the child. If no descriptions apply, then do not fill in any circles for that question.

A.1. Sample 1: How does this child cope with new learning tasks?

- Has a happy-go-lucky attitude to every problem
- Charges in without taking time to think or follow instructions
- Approaches new tasks with caution, but tries
- Won’t even attempt it if he/she senses a difficulty
- Likes the challenge of something difficult
- Cannot work up the energy to face anything new

A.2. Sample 2: How is this child at free play/individual choice?

- Engages in appropriate activities
- Rather loud but not disruptive
- Is too timid to join in
- Disturbs others’ fun
- Wants to dominate and have his/her own way
- Starts fights and rough play
- Needs teacher assistance to get involved
- Usually plays by him/herself
- Moves quickly from one activity to another
Appendix B

Teacher Demographic Questionnaire

Instructions: Please try to answer all of the following questions. The answers you give will be kept confidential.

1. What is your gender? _______Female _______Male

2. What is your age? __________

3. Please select the ethnicity that best describes you:
   - White
   - Black
   - Hispanic
   - Asian
   - Alaska Native/Native American
   - Other:_____________________

4. What is the highest level of education you have completed?
   - High School Graduate
   - Some College
   - 2-Year College Degree
   - 4-Year College Degree
   - Master’s Degree
   - Doctoral/Professional Degree

   How many years of teaching experience do you have?______________
Appendix C

Child Demographic Questionnaire

*Instructions*: Please try to answer all of the following questions about your child. The answers you give will be kept confidential.

1. What is your child’s gender? ______Female  ______Male

2. What is your child’s age? __________

3. Please select the ethnicity that best describes your child:

   - White
   - Black
   - Hispanic
   - Asian
   - Alaska Native/Native American

Other:_________________________
Appendix D

Table 1.

*Descriptive Statistics for ASPI Domains Using T-tests for Equality of Means*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Teacher</th>
<th>Assistant teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Aggressive</td>
<td>52.07</td>
<td>7.95</td>
</tr>
<tr>
<td>Oppositional</td>
<td>51.51</td>
<td>8.44</td>
</tr>
<tr>
<td>Inattentive/hyperactive</td>
<td>51.96</td>
<td>8.16</td>
</tr>
<tr>
<td>Withdrawn/low energy</td>
<td>49.21</td>
<td>6.80</td>
</tr>
<tr>
<td>Socially reticent</td>
<td>49.70</td>
<td>7.89</td>
</tr>
</tbody>
</table>

*p < .05

*Note. ASPI = Adjustment Scales for Preschool Intervention*
Table 2.

*Intraclass Correlation Coefficients for ASPI domains*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Single</th>
<th>Average</th>
<th>$f$</th>
<th>$df_1$</th>
<th>$df_2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive</td>
<td>.68</td>
<td>.81</td>
<td>5.34</td>
<td>56</td>
<td>56</td>
<td>.001</td>
</tr>
<tr>
<td>Oppositional</td>
<td>.64</td>
<td>.78</td>
<td>4.80</td>
<td>56</td>
<td>56</td>
<td>.001</td>
</tr>
<tr>
<td>Inattentive/hyperactive</td>
<td>.43</td>
<td>.60</td>
<td>2.67</td>
<td>56</td>
<td>56</td>
<td>.001</td>
</tr>
<tr>
<td>Withdrawn/low energy</td>
<td>.56</td>
<td>.72</td>
<td>3.52</td>
<td>56</td>
<td>56</td>
<td>.001</td>
</tr>
<tr>
<td>Socially reticent</td>
<td>.54</td>
<td>.70</td>
<td>3.32</td>
<td>56</td>
<td>56</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Note.* ASPI = Adjustment Scales for Preschool Intervention
Table 3.

*Cronbach’s Alpha for ASPI domains by Rater and Gender*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Teacher</th>
<th></th>
<th></th>
<th></th>
<th>Assistant teacher</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>M</td>
<td>F</td>
<td></td>
<td>T</td>
<td>M</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Aggressive</td>
<td>.88</td>
<td>.87</td>
<td>.88</td>
<td></td>
<td>.93</td>
<td>.93</td>
<td>.93</td>
<td></td>
</tr>
<tr>
<td>Oppositional</td>
<td>.79</td>
<td>.80</td>
<td>.78</td>
<td></td>
<td>.77</td>
<td>.80</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>Inattentive/hyperactive</td>
<td>.73</td>
<td>.64</td>
<td>.81</td>
<td></td>
<td>.70</td>
<td>.76</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>Withdrawn/low energy</td>
<td>.65</td>
<td>.61</td>
<td>.80</td>
<td></td>
<td>.65</td>
<td>.66</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>Socially reticent</td>
<td>.69</td>
<td>.65</td>
<td>.63</td>
<td></td>
<td>.64</td>
<td>.65</td>
<td>.50</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* ASPI = Adjustment Scales for Preschool Intervention; T = Total ($N = 56$), M = Males ($n = 28$), F = Females ($n = 28$)
Table 4.

*Cronbach’s Alpha of Individual Items if Removed and Corrected Item-Total Correlations*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Item</th>
<th>Teacher $\alpha$</th>
<th>Teacher CITC</th>
<th>Teaching assistant $\alpha$</th>
<th>Teaching assistant CITC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive</td>
<td>123</td>
<td>.88</td>
<td>.15</td>
<td>.93</td>
<td>.28</td>
</tr>
<tr>
<td>Inattentive/hyperactive</td>
<td>105</td>
<td>.78</td>
<td>-.07</td>
<td>.74</td>
<td>-.15</td>
</tr>
<tr>
<td>Withdrawn/low energy</td>
<td>71</td>
<td>.67</td>
<td>-.07</td>
<td>.66</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>115$^a$</td>
<td>.67</td>
<td>-.07</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

*Note.* Item 84 was recommended for removal from the scale but was not included in the table above because it was not endorsed by any rater. CITC = Corrected item-total correlation.

$^a$Teaching assistant data was not included for this item because no teaching assistants endorsed that behavior for any students.
VITA

Caitlin Cells
caitlin.s.cells@gmail.com

EDUCATION
Doctoral Candidate, School Psychology, The Pennsylvania State University, Anticipated Graduation 2017
M.Ed. The Pennsylvania State University, University Park, PA 2010
B.A Villanova University, Villanova, PA 2007

CERTIFICATION
School Psychologist, Virginia Department of Education Certified 2012

PROFESSIONAL EMPLOYMENT
School Psychologist, Loudoun County Public Schools 2012-Present
Psychologist Intern, Loudoun County Public Schools 2011-2012

PRACTICUM EXPERIENCE
School Psychologist Clinician, Central Fulton School District, McConnellsburg, PA 2011
Mobile Clinic, The Pennsylvania State University, University Park, PA 2011
State College Practicum Experience, State College Area School District, State College, PA 2009-2010
Penn State CEDAR Clinic, University Park, PA 2008-2010
Book Buddies Tutoring Program, Bellefonte Elementary School, Bellefonte, PA 2007

TEACHING EXPERIENCE
Teaching Assistant, Department of Biobehavioral Health, The Pennsylvania State University, University Park, PA 2009-2011
Teaching Assistant, Human Development and Family Services, The Pennsylvania State University, University Park, PA 2008-2009
Morgan Academic Support Center for Student Athletes (MASCSA) Student-Athlete Tutor, The Pennsylvania State University, University Park, PA 2007-2008

PRESENTATIONS