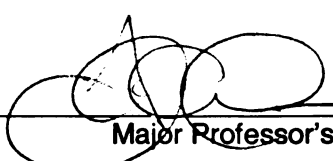This is to certify that the
dissertation entitled

THE VALIDITY OF THE DEVEREUX EARLY CHILDHOOD
ASSESSMENT WITHIN A HEAD START SAMPLE

presented by

MY THI LIEN

has been accepted towards fulfillment
of the requirements for the

_____Ph.D._____ degree in _____School Psychology_____

_____
Major Professor's Signature

_____8-21-06_____
Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

THE VALIDITY OF THE DEVEREUX EARLY CHILDHOOD ASSESSMENT
WITHIN A HEAD START SAMPLE

By

My Thi Lien

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology, & Special Education

2006

ABSTRACT

THE VALIDITY OF THE DEVEREUX EARLY CHILDHOOD ASSESSMENT WITHIN A HEAD START SAMPLE

By

My Thi Lien

Evidence in the current literature suggests the need to engage in early social-emotional screening to identify children who display risk factors for developing mental health disorders. The purpose of this study was to examine the validity of the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999) within a Head Start sample and to compare the DECA with the Michigan Risk Factors Student Identification Criteria (MRFSIC), the screening tool currently being use within the sample. In regard to the DECA, support was found for all of the types of validity-related evidences that were examined within this study: *content, substantive, structural,* and *generalizeability.* The internal consistencies and mean T-scores on the DECA for the Head Start sample closely resembled those from the DECA Standardization Sample for all three protective factors, Total Protective Factors, and Behavior Concerns. However, the correlation from the Head Start sample on the Total Protective Factors Scales and the Behavior Concerns Scale differed quite remarkably in comparison to what was reported in the DECA technical manual. Although both samples indicated an inverse relationship between Total Protective Factors and Behavior Concerns, the strength of the relationship differed (Head Start Sample: r = -.39; DECA Standardization Sample: r = -.65). Results also indicated that the same three factors exist on the DECA for the Head Start sample ratings, but with differences in item loadings onto the three factors. Multiple regression analysis indicated that as a whole, MRFSIC scores did not significantly predict higher scores on the DECA

Behavior Concerns Scale or the Total Protective Factors Scale. However, four individual risk factors on the MRFSIC significantly predicted higher scores on the DECA Behavior Concerns Scale, and eight individual risk factors on the MRFSIC significantly predicted lower scores on the Total Protective Factors Scale. There was not a clear relationship between the DECA and the MRFSIC warranting more research to determine whether the DECA can replace the MRFSIC. The use of validated screening tools is discussed as an important method of preventing future maladaptive behaviors by accurately identifying those children who may benefit from services provided within Head Start programming.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

Introduction

In recent reports on children's mental health, it was estimated that 20% of children have diagnosable psychiatric disorders with up to 11% of children exhibiting significant impairment in family, school, and social relationships [United States Department of Health and Human Services (USDHHS), 1999]. Evidence in the current literature suggests the need for a greater understanding of child psychopathology, with particular attention given to the early identification of children who experience risk factors for developing social, emotional, and behavioral disorders (e.g., Campbell, 1995; Fantuzzo, McWayne, & Bulotsky, 2003; National Advisory Mental Health Council's Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment, 2001).

The nature and extent of developmental changes occurring during toddlerhood and the preschool years underscore the potential for children to set out on either a positive, adaptive developmental course or to develop adjustment problems that may be longstanding (Campbell, 1995). A growing body of evidence indicates that behavior problems identified in the preschool years often persist (e.g., Campbell & Ewing, 1990; Webster-Stratton & Hammond, 1998), that a significant proportion of children do not grow out of their childhood difficulties (e.g., Loeber & Dishion, 1983; Moffitt, 1990), and for a majority of children, mental health problems go unidentified [National Institutes of Health & Mental Health (NIH/MH), 1998]. Many children who are unidentified consequently become part of the juvenile delinquency system [Children's Defense Fund (CDF), 2004].

Effective methods are needed for identifying and treating children who are vulnerable and currently underserved by traditional mental health service delivery systems (Fantuzzo et al., 2003). Specifically, greater attention to the development of early detection and prevention approaches for emotional or behavioral difficulties is warranted. Early screening and prevention activities within the school setting may help curb maladjustment and address critical issues of wellness across the lifespan (Cowen, 1997; DuPaul, 2003; Fantuzzo et al.; Masten, 2003; Ringeisen, Henderson, & Hoagwood, 2003; Roeser & Eccles, 2000; Squires, 2000). Most school-based mental health treatment systems unfortunately remain more reactive in their referral process rather than proactive. Misidentification or significant delays in identification are common (Steinberg & Knitzer, 1992). Thus, emotional or behavioral problems become much more severe before they are identified, and the occurrence of secondary disorders becomes more likely. Current practices in schools do not reflect a developmental framework of psychopathology in which early signs or symptoms of psychopathology are assumed detectable and possibly more effectively treated in early childhood rather than waiting until full-blown psychiatric disorders appear in later years (Coie et al., 1993).

From a developmental perspective, Head Start is an ideal setting within which early screening and intervention strategies should be targeted (Baydar, Reid, & Webster-Stratton, 2003; Webster-Stratton & Hammond, 1998). Emotional and behavioral difficulties are more widespread and pronounced in children from low-income, highly stressed families like those served in the Head Start Program (Campbell, 1997). Studies have shown that children with these characteristics are most at-risk for developing

maladaptive pathways that may lead to serious emotional and behavioral difficulties (Brooks-Gunn, Duncan, & Aber, 1997; Campbell).

Head Start plays a crucial role as a comprehensive, collaborative, community-based service delivery program for low-income children and their families. Since Head Start is often the first intensive contact with the system of care for low-income families in their communities, these programs are in an ideal position to proactively screen for, and accurately detect some of the earliest manifestations of behavioral and mental health problems. From a developmental perspective, the earlier risk status or actual behavioral or mental health problems can be detected, the more likely prevention and intervention efforts will be effective.

Research has identified a need for the inclusion of data-based decision-making approaches in the early identification of mental health problems in the Head Start population (NIH/MH, 1998). More recent emphasis has been placed on the need to assess social, emotional, and behavior development within this population. For instance, Head Start Performance Standards (USDHHS, 2001) require that all children receive social, emotional, and behavioral screening within 45 days of the child's entry into the program. A screening captures a child's overall social, emotional, and behavioral development for the purpose of determining whether further evaluation in a given area is needed. In contrast to a screening, an evaluation is a more in-depth process to determine a child's progress and service planning. In addition, Head Start has documented standards for child mental health which include Head Start working collaboratively with parents to promote positive mental health development (USDHHS, 1996). See Table 1 for Head Start Mental Health-Related Performance Standards.

Table 1

Devereux Early Childhood Assessment Program as it Relates to National Head Start Performance Standards (LeBuffe, & Naglieri, 1999)

National Head Start Performance Standards	Devereux Early Childhood Assessment Program
1304.20(b) Child Health and Developmental Services (1) Screening for developmental, sensory, and behavioral concerns, in collaboration with each child's parent, and within 45 calendar days of the child's entry into the program, grantee and delegate agencies must perform or obtain linguistically and age appropriate screening procedures to identify concerns regarding a child's developmental, sensory (visual and auditory), behavioral, motor, language, social, cognitive, perceptual, and emotional skills (see 45 CFR 1308.6(b)(3) for additional information). To the greatest extent possible, these screening procedures must be sensitive to the child's cultural background.	The DECA is a standardized, valid and reliable instrument for children ages 2-5 that measures child protective factors as well as screens for behavioral concerns. The DECA, to be completed by both staff and parents, provides a tool to better understand a child's behavioral and social strengths and needs as well as determine when further assessment is necessary. Having been normed on a representative sample of children in the United States, the instrument is sensitive to children's cultural backgrounds. The instrument is available in English and Spanish. Requests for versions in other languages should be forwarded to Devereux, see contact information at the end of this document.
1304.20(b) (2) Grantee and delegate agencies must obtain direct guidance from a mental health or child development professional on how to use the findings to address identified needs.	The scoring mechanism of the DECA is one source of information that may be relayed to a mental health professional or child development specialist, when appropriate, and with parental consent. Available to program staff is electronic access to an early childhood/mental health specialist. This service provides staff with immediate feedback and helps facilitate the referral process. In addition, ongoing technical support is offered through Devereux for on-site or distance learning.
1304.20(b) (3) Grantee and delegate agencies must utilize multiple sources of information on all aspects of each child's development and behavior, including input from family members, teachers, and other relevant staff who are familiar with the child's typical behavior.	The Devereux model emphasizes that both staff and parents administer the DECA and then jointly develop consistent plans for the home and school environments that promote children's resilience, based on multiple sources of information. Step one in the DECA Program helps teachers understand why and how to collect information on children.
1304.21(a) (3) (i) (A) Grantee and delegate agencies must support social and emotional development by: Encouraging development which enhances each child's strengths by building trust; (B) Fostering independence; (C) Encouraging self-control by setting clear, consistent limits, and having realistic expectations;	The primary focus of the DECA Program is to promote protective factors in preschoolers. These protective factors will help the child grow into a resilient adult. One of the three protective factors that the DECA Program will assess in children and work to build is Attachment, or, the mutual, strong, and long-lasting relationship between a child and significant adults such as parents, family members, and teachers. The second protective factor the DECA Program will assess and build in children is: Initiative, or, a child's ability to use independent thoughts or actions to meet his or her needs. The third protective factor the DECA Program will assess and build in children is: Self-control, or, the ability to experience a range of feelings and express them using the words and actions that society considers appropriate.

1304.24 (a) Mental Health Services (1) Grantee and delegate agencies must work collaboratively with parents for issues related to parents education by: (i) Soliciting parental information, observations, and concerns about their child's mental health.	Because the DECA is completed by both parents and staff, communication about the child can be accomplished through completing and discussing its results. This communication can occur through both formal and informal methods.
1304.24 (a) (1) (ii) Sharing staff observations of their child and discussing and anticipating with parents their child's behavior and development, including separation and attachment issues.	Observation and collecting information about children and their environment is a key step in the DECA Program. Staff learns the importance of appropriate observation, how to share this information with families, and how to use these observations to plan with families to promote children's attachment, self-control and initiative. The parent guide includes a chapter on how to foster attachment in preschoolers.
1304.24 (a) (1) (iii) Discussing and identifying with parents appropriate responses to their child's behaviors	The parent guide includes positive techniques of guidance for parents and staff to increase a child's attachment, self-control and initiative. Rather than focus on negative behaviors, the model identifies protective factors that parents and teachers can support and enhance together.
1304.24 (a) (1) (iv) Discussing how to strengthen nurturing, supportive environments and relationships in the home and at the program	DECA Program resources are designed to assist staff and parents in helping children develop attachment, self-control and initiative by focusing on the areas of environment, daily routine, supportive interactions, partnerships with families, and activities and experiences.
1304.24 (a) (1) (v) Helping parents to better understand mental health issues	The parent guide includes activities that assist parents in learning about and promoting children's resilience. With teacher's involvement in the DECA Program, staff will be better prepared to discuss these sensitive mental health issues with parents.
1304.24 (a) (1) (vi) Supporting parents participation in any needed mental health interventions	Through a culturally competent approach, the Devereux model emphasizes to staff the importance of communicating with parents about mental health issues and providing parents with information about partnering with staff and other mental health resources to promote children's social and emotional development. Regularly scheduled family activities are strongly encouraged in the Devereux model.
1304.24 (a) (2) Grantee and delegate agencies must secure the services of mental health professionals on a schedule of sufficient frequency to enable the timely and effective identification of and intervention in family and staff concerns about child's mental health	The DECA Program may be used as one source of information to recommend the referral of a child to a mental health specialist. The model also provides electronic access to an early childhood/mental health specialist. This access will provide support to the staff until an on-site mental health professional is available. In addition, the electronic technical assistance will guide staff and parents in their efforts to secure the assistance of a mental health professional in their community.
1304.24 (a) (3) (i) (3) Mental health program services must include a regular schedule of on-site mental health consultation involving the mental health professional, program staff, and parents on how	The Devereux model supports and enhances regularly scheduled on-site mental health consultation. The DECA produces two types of profiles, an individual profile and a classroom profile. Based on these profiles generated by the assessment, teachers implement

to: (i) Design and implement program practices responsive to the identified behavioral and mental health concerns of an individual child or group of children	strategies that are specifically geared to address the needs of the individual child as well as the entire class.
1304.24 (a) (3) (ii) Promote children's mental wellness by providing group and individual staff and parent education on mental health issues	The Devereux model provides staff and parent education on mental health through the Strategies and electronic technical assistance. While this does not take the place of an on-site mental health specialist, it provides feedback to staff and parents regarding their immediate concerns.
1304.24 (a) (3) (iii) Assist in providing special help for children with atypical behavior or development	The DECA includes a behavioral concerns screener which can serve as information for teachers and parents about a child's atypical behavior. This screen alone should not be used to make a decision about a child's behavior, but should serve as a guide regarding the need for further assessment. Devereux is nearing completion of a full assessment of behavioral concerns and protective factors. This assessment, the DECA-Clinical or DECA-C is due to be published in spring 2002. In addition, the Observation Journal, a reproducible DECA resource, offers rich planning forms, including several devoted to helping children change challenging behaviors.
1308.6 (b) Screening, the first step in the assessment process, consists of standardized health screening and developmental screening which includes speech, hearing, and vision. It is a brief process, which can be repeated, and is never used to determine that a child has a disability. It only indicates that a child may need further evaluation to determine whether the child has a disability. Rescreening must be provided as needed. **(1)** Grantees must provide for developmental, hearing, and vision screening of all Early Head Start and Head Start children within 45 days of the child's entry into the program. This does not preclude starting screening in the spring, before program services begin in the fall. [Note: The 1308 standards were written prior to the specific requirements for behavioral, emotional, and social skills screening cited in 1304.20(b)(1)] **(2)** Grantees must make concerted efforts to reach and include the most in need and hardest to reach in the screening effort, providing assistance but urging parents to complete screening before the start of program year. **(3)** Developmental screening is a brief check to identify children who need further evaluation to determine whether they may have disabilities. It provides information in three major developmental areas: visual/ motor, language and cognition, and gross motor/body awareness for use along with	The DECA Program is to be completed by parents and teachers for ALL children in the classroom. Each child's individual strengths and needs in the areas of attachment, initiative, self-control, and behavior concerns will be discovered. Children with low protective factors who might have previously slipped though the cracks will be identified and incorporated in the planning process, even when they lack behavioral concerns. This aspect of the DECA makes the assessment unique in comparison with other instruments that assess social and emotional health. It is a tool to help identify children early in order to prevent future social and emotional problems.

observation data, parent reports, and home visit information. When appropriate standardized developmental screening instruments exist, they must be used. The disabilities coordinator must coordinate with the health coordinator and staff who have the responsibility for implementing health screen-ing and education staff who have the responsibility for implementing developmental screening.	

Despite the increased attention to children's mental health concerns, many children's needs are not being met. The President's New Freedom Commission on Mental Health (PNFCMH, 2003) reported that about 5 to 9 percent of children have serious emotional disturbance, and many of them do not get the help they need because of inconsistent services, including lack of treatment options. The Commission's report specifically pointed to the problems of youth in the child welfare and juvenile justice systems and those of younger children (CDF, 2004). Now is the time to identify effective tools that can assist Head Start in meeting their federal standards for early mental health screening. The role of early screening and identification procedures may serve to set at-risk children on a positive developmental trajectory.

The present study aims to address several key gaps within the current literature on early childhood mental health. Particularly, evidence for the validity of the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999) is examined, and the DECA is discussed as a potential screening tool for the early identification of risk and protective factors associated with psychopathology in this population. This study, therefore, helps to identify an additional screening tool that may be reliable and valid for use within low-income, preschool populations. The literature review section is presented as follows. First, a discussion of developmental psychopathology theories and research on risk and resilience are presented. Second, the need for early screening of mental

health is explored. Third, the rationale for including Head Start programs in mental health screening is provided with a description of Head Start's current practices regarding mental health screening. Fourth, a review of rating scales commonly used for screening in the preschool population is provided with a discussion of important issues in assessment to consider. Finally, research questions and hypotheses for the current study are presented. Following the literature review section, methods for the current study are described, results are displayed, and a discussion of the results including implications and directions for future research is presented.

Developmental Psychopathology

Developmental psychopathology provides a strong theoretical basis for helping to understand the impact of risk and protective factors that influence the expression of psychopathology across time (Cicchetti, 1993; Masten, 2003; Sameroff, 2000). Developmental psychopathology refers to the study of maladaptive behavioral and emotional deviance in the developmental tasks, sequences, and processes that characterize human growth (McConaughy & Achenbach, 1990). In developmental psychopathology, theorists emphasize the importance of developmental processes, context, and the influence of multiple interacting events and processes in shaping adaptive and maladaptive development (Mash & Dozois, 2003; Sameroff).

Shirk, Talmi, and Olds (2000) identify four main principles of developmental psychopathology:

1. Most mental disorders result from the interplay of multiple psychological, social, and biological processes.
2. Most, if not all, mental disorders can be reached from different developmental pathways or through different pathogenic processes.
3. Exposure to early developmental hazards increases the probability of developing mental disorders, but continuity depends on subsequent events or conditions that maintain maladjustment.
4. Mental disorders are embedded in multiple contexts that are mutually influential. (p.838)

The developmental psychopathology framework is used as a way of thinking about problems that arise from many different causes, take different forms across development, and result in different outcomes. In other words, individuals will experience the same events differently depending on their level of functioning across all domains of psychological and biological development. Accordingly, various problems will have

different meanings for an individual depending on both their nature and the timing of the experience (Rutter, 1989; Sameroff, 2000).

Risk & Protective Factors

Examinations of risk and resilience are essential in order to understand the role that various biological, psychological, and environmental factors play in fostering or inhibiting positive adaptation, as well as for informing theories of development (Cicchetti & Toth, 1997). This section a) defines risk and resilience b) presents the current research on risk, and c) presents the current research on resilience.

The term *"at-risk"* refers to a set of presumed cause-and-effect dynamics that place the child or adolescent in danger of negative future events. At-risk refers to a situation that is not necessarily current but that can be anticipated in the absence of intervention (McWhirter, McWhirter, McWhirter, & McWhirter, 1998). Because the area of developmental psychopathology is concerned with the detection of developmental deviation before an actual disorder crystallizes, as well as with the course of disorders once exhibited, knowledge derived from research in this area possesses considerable relevance for application to the prevention and treatment of risk and psychopathological conditions (Masten, 2003).

The study of successful adaptation is integral in understanding the etiology, prevention, and treatment of problems in development (Luthar & Cicchetti, 2000; Masten & Curtis, 2000; Rutter & Sroufe, 2000). This central tenet of developmental psychopathology emerged from pioneering investigations of children at-risk for maladaptation due to perinatal hazards, parental psychopathology, psychosocial disadvantage, and loss (Masten, Best, & Garmezy, 1991). The study of resilience thus

arose from the study of risk as researchers realized that there were children flourishing in the midst of adversity (Garmezy, 1974; Rutter, 1979; Werner & Smith, 1982). These early researchers recognized that such children could teach us better ways to reduce risk, promote competence, and shift the course of development in more positive directions. Therefore, it is important to focus especially on the protective processes that bring about changes in life trajectories from risk to adaptation (Rutter, 1987). Studies of children at-risk for psychopathology played a key role in bringing both the phenomenon of resilience and the integrative framework for developmental psychopathology to the forefront of theoretical and research attention (Garmezy, 1971; Garmezy, 1974; Masten & Garmezy, 1986; Masten, 1989; Rutter, 1979).

Professionals have recognized that protective factors in early childhood have a crucial role in determining subsequent adjustment or maladjustment to life stresses (Masten, 2003; Webster-Stratton & Hammond, 1998). During the past 35 years, a number of researchers have focused on the study of resilience in children. Researchers have described the term *resilience* in three ways: good developmental outcomes despite high-risk status, sustained competence under stress, and recovery from trauma (Masten & Reed, 2002; Rutter, 1990; Werner, 1993). Under each of these conditions, researchers have focused their attention on protective factors, or mechanisms that buffer a person's reaction to a stressful situation or chronic adversity so that his or her adaptation is more successful than would be the case if the protective factors were not present. Resilience, therefore, is conceptualized in terms of protective factors, or positive factors that can cushion the extent to which risk factors negatively impact children's development. The

discussion of the role of protective factors in the literature of child psychopathology is relatively recent (Rutter, 1979; Werner & Smith, 1982, 1992).

Risk Research. Three general categories encompass the areas of research on risk factors (Yoshikawa, 1994): child-centered, family-centered, and contextual factors. Child-centered factors include genetic vulnerability, sex, perinatal risk, temperament, cognitive abilities, and school achievement. Family-centered factors include parenting, attachment, child maltreatment, and marital conflict. Finally, contextual factors include family and community SES and community crime and violence.

Researchers have examined infant precursors of behavior problems in toddler and preschoolers studied prospectively. The focus of this work has been primarily on perinatal problems, maternal perceptions of infant temperament, and co-occurring risk factors. In general, studies indicate that in low risk samples (e.g., Bates, 1987), maternal ratings of infant difficulties across the first year of life do predict later ratings of behavior problems, especially externalizing problems in boys (e.g., Bates & Bayles, 1988). Similarly, Goldberg, Corter, Lojkasek, and Minde (1990) reported that in a sample of low birth weight, premature children, maternal ratings of infant temperament at 1 year were the best predictors of maternal and teacher ratings of behavior problems at age 4. It is widely accepted that early temperamental difficulties may be a precursor of later behavior problems, although it is usually argued that infant characteristics interact with parenting quality to produce good or poor outcomes (Campbell, 1995; Kim-Cohen, Moffitt, Caspi & Taylor, 2004; Patterson, Kupersmidt, & Vaden, 1990). It is important to note that past research has attempted to understand childhood difficulties through assessing parental perceptions. Including parental perceptions has been widely used because parents are

viewed as having the best understanding of their children. There are however, some concerns that parental perceptions of their children's problems may be colored by factors such as the parent's own dysfunction, family distress, and parental disciplinary styles (e.g., Campbell, Pierce, March, & Ewing, 1991; Emery, 1982).

Shaw, Keenan, and Vondra (1994) followed a sample of high risk, low income mothers and their infants prospectively. Mother-infant interaction was observed in several laboratory contexts at 12, 18, and 24 months to derive measures of infant demandingness in a frustrating situation, aggression and noncompliance, as well as maternal responsiveness. For boys, but not girls, observed noncompliance and lower levels of maternal responsiveness at earlier assessments predicted global ratings of observed aggression at 24 months. Earlier aggression and the interaction between aggression and maternal responsiveness predicted maternal ratings of externalizing symptoms at 36 months. In addition, parental disagreement over childrearing was an especially strong predictor of externalizing problems in boys, whereas poor marital adjustment predicted both externalizing and internalizing problems in girls at 36 months (Shaw, Vondra, Dowdell-Hommerding, Keenan, & Dunn, 1994). Moreover, higher levels of family adversity assessed in infancy and toddlerhood predicted higher ratings of internalizing and externalizing symptoms.

Prospective studies from infancy also have examined biological risk factors such as anoxia, prematurity, low birth weight, and other birth complications as predictors for later problems. A majority of the studies have focused on cognitive outcomes rather than behavior problems, but several studies have linked premature birth and low birth weight with slightly higher rates of attentional and behavior problems at preschool age (e.g.,

13

Rose, Feldman, Rose, Wallace, & McCarton, 1992). In general, these prospective studies from infancy to preschool age suggest that difficult early behavior and perinatal problems are associated with the development of problems in young children, but only in concert with environmental adversity.

Resiliency Research. Most of the studies on individual resilience and protective factors in children have been short-term, focusing on middle childhood and adolescence. An exception is the Kauai Longitudinal Study, which followed a cohort of participants over several decades (Werner, 1993). Results from this study indicated that approximately 30% of the survivors in this study were considered high-risk children because they were born in chronic poverty, had experienced perinatal stress, and lived in family environments characterized by chronic discord, divorce, or parental psychopathology. Two-thirds of the children who had experienced four or more such risk factors by age 2 developed serious learning or behavior problems by age 10 or had delinquency records, mental heath problems, or pregnancies by age 18. The more compelling results from this study were that one third of the children who had experienced four or more of these risk factors developed instead into competent, confident, and caring adults.

Radke-Yarrow and Brown (1993) have also attempted to study resilience using a longitudinal design. Using measures systematically obtained in a 10-year period, case studies were developed on 18 resilient children with healthy adaptation throughout development and on 26 troubled children with serious persistent problems. All children in this study had family risks of affective illness in both parents and a highly chaotic and disturbed family life. Well-adapted children of well-functioning parents and families

14

were a comparison group. This study found that resilient and control children were very similar on most measures. Troubled children as a group had lower scores of intelligence, were more often shy, had poor academic achievement, and had a history of poor peer relationships. Resilient children elicited more positive reactions from teachers, were more likely to be the favored child in the family, and had more positive self-perceptions. In this study, child assertiveness, high intelligence, and high achievement motivation provided a means for children to develop positive relationships with parents, teachers, and peers.

In another longitudinal study, Masten and colleagues (1999) followed 205 children from an urban community over 10 years. Multiple methods and informants were used to assess three major domains of competence from childhood through adolescence (academic achievement, conduct, and peer social competence), multiple aspects of adversity, and major psychosocial resources. Results of this study concluded that (a) the development of competence is related to psychosocial resources, (b) good resources are less common among children growing up in the context of adversity, (c) if reasonably good resources are present, competence outcomes are generally good, even in the context of chronic, severe stressors, and (d) maladaptive adolescents tend to be stress-reactive and have a history of adversity, low resources, and broad-based competency problems. These findings add to the growing longitudinal evidence suggesting that effective parenting (e.g., well-functioning parent-child relationships) and cognitive skills are general protective factors for development that may be particularly important for overcoming serious chronic adversity.

Kim-Cohen and colleagues (2004) studied genetic (i.e., twin-study) and environmental contributions to young children's resilience and vulnerability to SES deprivation. In an epidemiological cohort of 1,116 five-year-old twin pairs, the researchers found that some children exposed to socioeconomic deprivation (SES) are resilient and function better than expected, given the level of deprivation they have experienced. In this study, children's resilience was assessed by the difference between their actual score and the score predicted by their level of SES deprivation. Results of this study indicated that maternal warmth, stimulating activities, and children's outgoing temperament promoted positive adjustment in children exposed to SES deprivation. These findings suggest that protective processes operate through both genetic and environmental effects. Since this study examined children who were experiencing socioeconomic deprivation, the results may be helpful in understanding maternal warmth, stimulating activities, and child temperament in a population like Head Start that is characterized by low SES.

The Need for Early Screening

The President's New Freedom Commission on Mental Health (2003) emphasized in their report on mental health that early childhood is a critical period for the onset of emotional and behavioral impairments. The impetus for improving screening and assessment is partly attributed to the alarming number of very young children being expelled from preschools and child care facilities for disruptive behaviors (CDF, 2004). These concerns highlight the need for a strong research base for understanding mental health disorders in children younger than 5 years of age (Campbell, 1995; Fantuzzo et al., 2003; Hooks, Mayes, & Volkmar, 1988).

Evidence suggests that preschool children experience emotional and behavioral problems at prevalence rates comparable or higher than that of older children. For example, Campbell (1995) estimated that 10 to 15% of preschoolers have mild to moderate behavior problems that interfere with their school progress. Other empirical studies indicate that approximately 20% of preschool children exhibit moderate to clinically significant emotional and behavioral difficulties (Campbell; Lavigne et al, 1996; PNFCMH, 2003). Finally, Hooks et al. (1988) found even higher prevalence rates of psychiatric disorders among children. Participants in this study were children under the age of 5 who presented at mental health facilities for evaluation and treatment. Results indicated that 32% of children were diagnosed with emotional disorders (e.g., anxiety disorders, depression, adjustment disorders, attachment disorders), 4.1% with disruptive behavioral disorders (e.g., attention deficit disorder, oppositional defiant disorder, conduct disorder), and 10.9% with parent-child difficulties. Thus, it appears that impaired functioning in social-emotional and behavioral arenas do impact developmental outcomes. Providing early mental health services promotes a positive developmental trajectory for at-risk children.

Prevention and intervention programs are most effective when they are provided early before problems become chronic and severe (Bear, Webster-Stratton, Furlong, & Rhee, 2000). If children are misidentified, or not identified as in need of help, they will either not get any help, or they will not get appropriate help. Since early onset problematic behavior is predictive of long term and often high cost problems, this has fiscal implications as well as implications for child and family functioning (Coie, Miller-Johnson, & Bagwell, 2000). For example, Mrazek and Haggerty (1994) reported the

estimated costs for treating mental health disorders in 1990 was $147 billion, in addition

to the $164 billion spent on alcohol and drug abuse. Aside from financial expenditures of

mental health treatment, those interested in the prevention of disorders from occurring

argue that preventing human suffering and loss of personal productivity caused by mental

disorder is itself a social value that outweighs other cost-benefit considerations (Coie et

al., 2000). Thus, it is financially and socially important to pay attention to signs of

emotional and behavioral problems in children.

Rationale for Including the Head Start Program

Theoretically, the most effective detection programs should begin in preschools

where children can be screened annually for learning or behavioral problems (Forness et

al., 2000). Most school programs, however, do not use school-wide screening or

detection but rather wait until kindergarten or much later when children are first referred.

Recent studies indicate that most children are initially served only in their middle-to-late

elementary school years, with very few children identified or systematically served in

early childhood (NIH/MH, 1998).

Head Start is an ideal setting within which early screening and intervention

strategies should be targeted (Baydar et al., 2003; Webster-Stratton & Hammond, 1998).

Head Start has been suggested as the most effective venue for the initiation of a system of

early identification and prevention because risk factors such as broken families,

underemployment, and limited access to health or social services are most severe (Lopez,

Tarullo, Forness, & Boyce, 2000; Mrazek & Haggerty, 1994). In addition, new federal

regulations for health screening now focus more explicitly on the detection of mental

health disorders that include Head Start staff and parents in the process (CDF, 2004; PNFCMH, 2003).

A growing body of evidence indicates that preschool children experience emotional and behavioral difficulties at similar prevalence rates as those of older children. Two independent studies reported that 20% of preschool children exhibit moderate to significant emotional and behavioral problems (Lavigne et al, 1996; Pianta & Caldwell, 1990). Furthermore, it has been shown that these problems are relatively persistent and predictive of future developmental maladjustment (Campbell, 1997). Empirical studies indicate that young children living in high-risk environments are most likely to manifest emotional and behavioral maladjustment. Major risk factors associated with emotional and behavioral problems include: poverty, living in a single-female headed household, and exposure to multiple stressors associated with densely populated urban settings (Campbell; Lavigne et al.; Patterson et al., 1990).

Research suggests that poverty increases the likelihood of impaired functioning for children (Costello & Angold, 1996; McLeod & Shanahan, 1996). Compared with community samples, preschool children from low-income families show higher rates of psychiatric disorders, a pattern similar to school-age children and adolescents from low-income families (Keenan, Shaw, Walsh, Delliquardri, & Giovannelli, 1997). Webster-Stratton (1998) reported that 23% of her Head Start sample had behavior problems in the clinical range. Moreover, many Head Start families are coping with both adverse social and economic conditions and exposure to community violence (Takanishi & DeLeon, 1994). Mental health services for diverse, low-income families need to be strengthened

to respond to the increasing challenges of this population (Piotrowski, Collins, Knitzer, & Robinson, 1994).

Head Start currently enrolls more than 900,000 children in the US annually and has served over 22 million children since it began in 1965 (USDHHS, 2004). Head Start is a child development program that serves children from birth to age five and their families. The Program's goals include promoting social competence and school readiness, targeting families that are at or below the federal poverty level. Many of the students enrolled in this program are children who already experience risk for later behavioral and health problems. In light of recent research that indicates maternal warmth, the presence of stimulating activities, and children's outgoing temperament as promoting positive adjustment in children exposed to low socio-economic status (Kim-Cohen et al., 2004), collaboration with parents and families may be one method for increasing these important protective factors in Head Start children.

Anderson (1983) studied the prevalence rates of behavioral and emotional disturbance among 462 Head Start children. Using the criteria outlined in the Preschool Behavior Questionnaire (PBQ), 31.6% of children were identified as exhibiting behavioral or emotional concerns that warranted further action. Results from this study indicated that a high proportion of children from economically disadvantaged backgrounds, such as those in this Head Start population, experience significant behavioral and emotional difficulties. Anderson concluded that many children in this sample will begin their initial elementary school years at considerable risk for school maladjustment due to behavioral or emotional interference with the learning process.

Current Head Start Practices. Yoshikawa and Knitzer (1997) have documented the typical mental health intervention currently available in Head Start. In their description, a mental health consultant conducts observations in Head Start classroom, typically at least once or more per year. The observation serves to assess the mental health environment of each classroom. Mental health consultants also receive requests from teachers for consultation; however, requests are usually made after mid-year. Thus, it may be spring or the following school year until a child receives appropriate attention and treatment. This description is typical of the majority of Head Start programs nationally (Yoshikawa & Knitzer). There is general agreement among the Head Start community that mental health services are a weak link in Head Start's comprehensive services strategy (Piotrkowski et al., 1994; Yoshikawa & Knitzer). This is due, in part, to difficulties in assessing and diagnosing preschool children (e.g., because of delayed speech and language, delayed motor development) and to concerns about labeling, insufficient staff training, noncompetitive salaries, a shortage of community resources, a lack of clarity within Head Start about how best to deliver mental health services, and the absence of a consistent vision, leadership, and support at the national level (Piotrkowski et al.).

Piotrkowski et al. (1994) conducted a telephone survey with a random sample of 101 Head Start programs. The interviews were designed to gather information about mental health needs, strategies for delivering mental health services, staffing, needed services, funding, referrals, and other related matters. Results of this survey indicated that classroom behavioral concerns were among the top in mental health concerns for Head Start children and families (37%). Most respondents (88%) indicated that there

were insufficient mental health resources in their communities, including lack of funding and affordable mental health professionals and barriers to service, such as waiting lists, poor transportation, and lack of culturally competent providers. In addition to needing more consultant time, more than half of the respondents (54%) expressed the need for a full-time mental health professional on staff to provide mandated services. A majority of respondents also stated the need for more resources for mandated mental health services and training for staff (71%). Thus, current Head Start practices lack a standardized, data-based, and systematic approach to early screening for mental health difficulties.

Local Head Start Practices. Although Head Start has always included the mental health needs of children within its framework of goals and services, additional attention and resources are needed to address the full range of mental health issues adequately (Knitzer, 2000; Yoshikawa & Knitzer, 1997). Current screening procedures in the local Head Start sample include a criterion-based tool referred to as the Michigan Risk Factors Student Identification Criteria (MRFSIC). The MRFSIC is 25-item checklist of risk factors. Sample items on this checklist include: low birth weight, physical/sexual abuse or neglect, and family history of low school achievement or dropout (see Appendix D for a full listing of the MRFSIC Checklist). As each child enrolls in the Head Start Program, the parents (typically the biological mother) respond to whether the factors exist or not. In order to be determined socially or emotionally at-risk, a parent must indicate the presence of two or more risk factors. According to the MRFSIC, risk is defined as a child having two or more of the risk factors present. The MRFSIC does not allow for ratings of the degree to which a risk factor is present; it only indicates whether a risk factor is present.

Early identification of significant behavior problems has taken on increasing importance with the recognition that untreated problem behaviors are precursors to severe mental health problems later in life. Head Start programs can directly provide efficient and cost-effective early prevention and intervention for high-risk children and their families. However, existing research methods have not been developed or empirically tested for populations of children and families that Head Start serves (Fantuzzo, McDermott, Manz, Hampton, & Alverez Burdick, 1996). Based on dependable early screening, more targeted interventions can then be made available to those children and families.

Issues in Assessment

Limited attention is given to multicultural issues in assessment and research (e.g., Feil, Walker, Severson, & Ball, 2000; Garcia Coll et al., 1996; Sue, 1999), which has direct implications on the use of these tools within Head Start. For example, Sue argued for the need for all research to address external validity issues, while Feil and colleagues stated a need for cross-cultural research in screening tools for young multicultural children. Cicchetti and Toth (1997) highlight other critical issues of cultural diversity and potential bias in both identification and treatment. Available data suggest that children from African-American backgrounds tend to be over-identified in school or mental health systems while children from Hispanic or Asian-American families tend to be under-identified (Cicchetti & Toth). Both school and mental health professionals often make a number of diagnostic mistakes with children from culturally diverse backgrounds (e.g., Serna, Nielsen, Mattern, & Forness, 2002).

Problems in bias become even more critical when dealing with very young children whose symptoms are in the formative stage. In a population such as Head Start, consideration needs to be given to issues of multiculturalism and external validity of tools used due to the diverse characteristics (i.e., low-income, diverse racial make-up) of this population. With increased interest in early intervention, researchers and clinicians have a critical need for reliable and valid measures of children's behavior for use with diverse populations. Establishing the construct validity and reliability of a measure for a given population is important in establishing the utility of a measure (Edwards, Whiteside-Mansell, Conners, & Deere, 2003). A recommended research strategy for obtaining more information about ethnic minority children's behaviors is to develop and validate culturally relevant assessment instruments (Guerra & Jagers, 1998; Mendez, McDermott, & Fantuzzo, 2002). An efficient approach to obtaining information about large number of at-risk children is to utilize behavior rating scales.

The Use of Rating Scales. Behavior rating scales have been called a "best practice" in assessing child behavioral and emotional problems (McConaughy & Ritter, 2002). They are one of the most efficient ways to identify behavioral strengths and weaknesses (Knoff, 2002). Standardized behavior rating scales and checklists are the most frequently used instruments by psychologists in assessing emotional and behavioral difficulties in children (Barkley, 1988; Stinnett, Havey, & Oehler-Stinnett, 1994).

In the recent decade, psychometric properties of ratings scales have generally increased, thus, they are an ideal tool to be used as part of the screening and identification process for children referred for possible special education services (Busse & Gresham, 1993). Although there are a number of well-established behavioral rating scales for

school-aged children, until recently there have been few scales with adequate technical properties specifically developed for preschool-aged children (Funderburk, Eyberg, Rich & Behar, 2003).

Several reasons exist for using rating scales as an initial screener for psychopathology. Ratings scales can be used unobtrusively to evaluate behaviors within home and school settings. In addition, rating scales are generally time efficient, easily administered and scored, and can facilitate links between assessment and interventions. Rating scales also allow for the involvement of multiple stakeholders in assessment activities. For example, Greenfield, Iruka, and Munis (2004) suggest that the use of rating scales can promote communication between the school and family by offering parents and teachers a common language to discuss a child's strengths and weaknesses. Such a collaborative approach offers a comprehensive understanding of the child across multiple contexts.

Issues of Validity. Traditional discussions of validity in measurement have described three distinct types of validity: content, criterion-related, and construct. DeVellis (2003) describes content-related validity as the degree to which the domain measured by a test is represented by the test items. Criterion-related validity is described as the degree to which an item or scale is associated with some criterion or "gold standard." Finally, construct-related validity measures the extent to which a test measures what it purports to measure.

More recent conceptualizations of validity do not distinguish different types of validity, but instead define validity as a unified concept. Specifically, validity is the overall evaluative judgment of the degree to which empirical evidence and theoretical

rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment (Messick, 1989, 1995). Messick's conception of validity posits that validity is not a property of the test or assessment but instead, a meaning of the test scores.

> *"These scores are a function of the test items or stimulus conditions, but also of the persons responding as well as the context of the assessment. In particular, what needs to be valid is the meaning or interpretation of the score, as well as any implications for action that this meaning entails. The extent to which score meaning and action implications hold across persons or population groups and across settings or contexts is a persistent and perennial empirical question. This is the main reason that validity is an evolving property and validation is a continuing process."* (Messick, 1995, p. 741)

Thus, it can never be stated that a test is valid, but only that there is evidence for the validity of its use for the purposed purpose. The American Educational Research Association (1999) has adopted this unitary definition of validity in its most recent publication of the Standards for Educational and Psychological Testing.

Messick (1989, 1995) describes six types of evidence that can exist to support the overall construct validity of a test. These six aspects function as general validity criteria or standards for all educational and psychological measurement.

1. Content: Includes evidence of content relevance, representativeness, and technical quality
2. Substantive: Refers to theoretical rationales for the observed consistencies in test responses, including process models of task performance, along with empirical evidence that the theoretical processes are actually engaged by respondents in the assessment tasks
3. Structural: Appraises the fidelity of the scoring structure to the structure of the construct domain at issue
4. Generalizeability: Examines the extent to which score properties and interpretations generalize to and across population groups, settings, and tasks, including validity generalization of test criterion relationships
5. External: Includes convergent and discriminant evidence from multitrait-multimethod comparisons as well as evidence of criterion relevance and applied utility

6. Consequential: Appraises the value implications of score interpretation as a basis for action as well as the actual and potential consequences of test use, especially in regard to sources of invalidity related to issues of bias, fairness, and distributive justice (Messick, 1989, p. 745)

Although all six types of validity described by Messick are important in assessing the overall validity of an instrument, given the limited data available, the current study does not address *external-* or *consequential- related* evidence. Messick (1989) provided some possibilities for assessing the six types of validity evidence, including using judgmental and logical analyses (e.g. based on theory), correlational or covariance analyses (e.g., correlations, factor analysis, path analysis, structural equation modeling, ANOVA), analysis of processes underlying item or task performance (e.g., protocol analysis, chronometric analysis, cognitive correlates), analysis of group differences over time (e.g., test-retest, changes in score level and variability), and by altering test scores in theoretically predicted ways (e.g., experimental treatment and manipulation of conditions). Ultimately, however, he noted that there are a wide variety of data and analyses that are pertinent, but that "data and analyses are to be advanced that are relevant to the score-based interpretations and decisions being made. The validation task is to accumulate a preponderance of evidence for or against the proposed interpretation or use" (p.50).

Validity of Early Childhood Behavior Assessment Tools. Bracken, Keith, and Walker (1998) evaluated thirteen different preschool measures of social-emotional functioning and found that most of the rating scales were limited in critical and primary areas of technical adequacy. The following aspects of technical adequacy reported in the examiner's manual for each instrument were examined: standardization sample, subtest

and scale internal consistency and stability, inter-rater reliability, average subtest and scale ceiling/floor, item gradient, and evidence of validity.

Three tables were created to provide an overview of the instrument reviews. Table 2 presents an overall picture of the standardization samples of the 13 tools evaluated by Bracken et al. (1998). Table 3 presents psychometric properties of the 13 tools, including reliability coefficients, test-retest reliabilities, inter-rater reliabilities, and evidence for validity. Table 4 provides an overall picture of the extent to which the 13 tools met the criteria set by Bracken et al. Information available from the DECA Technical Manual (LeBuffe & Naglieri, 1999) in regard to the 13 criteria is also presented (in bold) in this table to provide for comparison. The criteria used in the evaluation were:

a) a median subtest internal consistency of .80 or great
b) scale internal consistency of .90 or greater
c) a scale stability coefficient of .90 or greater
d) an average subtest floor or ceiling at least 2 standard deviations beyond the normative mean subtest score
e) a scale floor or ceiling at least 2 standard deviations beyond the normative mean score
f) an item gradient no steeper than three items per standard scored standard deviation
g) validity data available/presented in the test manual (without evaluation)
h) the instrument's standardization samples were evaluated according to size, representativeness, and whether the respective instruments were nationally or locally/regionally normed. (p. 154)

Table 2

Standardization Sample Characteristics of Behavior Rating Scales Reviewed by Bracken et al. (1998) and Printz et al. (2003)

Scale	Age level	Sample size	Sample geographics
ASQ:SE*	Pre-K	3014	NR
BASC			
Parent	4-8 years	3483	26 states/4 regions
Teacher	4-8 years	2401	26 states/4 regions
BITSEA*	12-36 months (b)	1237	NR
Burks'	3-4 years	127	1 country
	Kindergarten	337/31 (a)	1 country
CBCL			
Competence	6-18 years	2116	4 regions
Problem	4-18 years	2368	4 regions
CRS			
Parent	3-17 years	578	NR
Teacher	3-17 years	578	NR
Parent	6-14 years	NR	NR
Teacher	4-12 years	9583	Canada
DECA	**2-5 years**	**2000**	**4 regions**
ESP*	Pre-K – K	2853	8 states
Eyberg			
ECBI	2-12 years	512	1 state
SESBI	Preschool	55	1 state
Louisville	4-6 years	287	1 country
PQB	3-6 years (b)	496	2 states
PIC	3-5 years	192	NR
	6-16 years	2390	1 state
PKBS	3-6 years	2955	16 states/ 4 regions
PKBS*	3-6 years	3313	17 states/4 regions
SSRS			
Parent	Pre-K- 12th gr.	1335/259 (b)	18 states/4 regions
Teacher	Pre-K- 12th gr.	992	18 states/4 regions
SSRS*	3-8 years (b)	6456	4 regions
TABC			
Parent	3-7 years	1381	4 states/3 regions
Teacher	3-7 years	577	2 regions
TABS*	11-71 months (b)	1000	33 states
Vineland	birth – 18 years	3000	1 state
Walker	1st - 3rd grades	852	1 state

Adapted from Bracken et al., 1998. Note: * = based on review by Printz et al., 2003; NA= no total test score; NR= not reported. (a)= Students/Teachers; (b)= Manual does not report ages for which the test was normed, only that it is appropriate for this age group. ASQ:SE= Ages and Stages Questionnaire: Social Emotional; BASC= Behavior Assessment System for Children; BITSEA= Brief Infant Toddler Social and Emotional Assessment; Burks'= Burks' Behavior Rating Scales; CBCL= Child Behavior Checklist; CRS= Conners' Rating Scale; ESP= Early Screening Project; Eyberg= Eyberg Child Behavior Checklist; Louisville= Louisville Behavior Checklist; PBQ= Preschool Behavior Questionnaire; PIC= Personality Inventory for Children; PKBS= Preschool and Kindergarten Behavior Scales; SSRS= Social Skills Rating System; TABC= Temperament Assessment Battery for Children; TABS= TABS Screener for the Temperament and Atypical Behavior Scale; Vineland= Vineland Adaptive Behavior Scales; Walker= Walker Problem Behavior Identification Checklist

Table 3

Psychometric Properties of Behavior Rating Scales Reviewed by Bracken et al. (1998) and Printz et al. (2003)

Scale	Age level	Total test r	Test- Retest	Inter-rater	Validity evidence
ASQ:SE*	NR	NR	.94	NR	NR
BASC		NR		.70-.76/.38-.53/.39-.58/.61-.69 (a) (i)	CR; D; F
Parent					
	Girls 4-5	.89/.84/.87/.91 (a)	NA	NA	
	Boys 4-5	.89/.87/.87/.92 (a)	NA	NA	
	Girls 6-7	.88/.88/.91/.92 (a)	NA	NA	
	Boys 6-7	.90/.85/.93/.92 (a)	NA	NA	
	4-5 years	NA	.70/.38/.39/.61 (a)	NA	
	6-7 years	NA	.76/.53/.58/.69 (a)	NA	
Teacher				.47-.79/.62-.69/.89/.76-.80/.53-.84 (b) (k)	
	Girls 4-5	.93/.90/.90/.95 (a)	NA	NA	
	Boys 4-5	.95/.91/.90/.95 (a)	NA	NA	
	Girls 6-7	.91/.90/.92/.96/.96 (b)	NA	NA	
	Boys 6-7	.93/.89/.92/.95/.97 (b)	NA	NA	
	4-5 years	NA	.95/.90/.90/.95 (a)	NA	
	6-7 years	NA	.91/.81/.94/.96/.93 (b)	NA	
BITSEA*					CR; CC; P
Competence	NR	NR	.85	.61 (i); 59 (j)	
Problem	NR	NR	.87	.68 (i); 28 (j)	
Burks'	NR	NA	NA	.57-.92 (j)	CN; CR; F; RR; CG
CBCL			.90-.92/.82-.94 (f)	.78 (i)	CN; CS; CR; D
Competence			NA	NA	
	Girls 4-11	.62	NA	NA	
	Boys 4-11	.57	NA	NA	
Problem			NA	.76 (i)	
	Girls 4-11	.96	NA	NA	
	Boys 4-11	.96	NA	NA	
CRS	NR	NA	NA	.85 (g) (i)	CS; CC; P; D

					CN, CS, CR, CG, P
DECA					
Parent	Preschool				
Total Protective Factors		.91	.74	.21 (i); .29 (j); .69 (k)	
Behavior Concerns		.71	.55	NA	
Teacher	Preschool				
Total Protective Factors		.94	.94	NA	
Behavior Concerns		.80	.68	NA	
ESP*	NR	NR	.72	.87	CR;CC
Eyberg					CS;CV;D
ECBI	2-12	NA	.80-.86/.85-.88 (e)	.79-.86 (i) ; .85-.87 (k)	
SESBI	Pre-K – K	NA	.90/.89 (e)	NA	
Louisville	4-6	NA	NA	NA	CN; CS; CR
PQB	NR	NR	NR	NR	CR
PIC		NA	.90-.97	.57-.69 (g) (i); .99 (g) (j); .66 (g) (l)	CS; P; CR
Profile Scal.	NR	NA	NA	NA	
Factor Scal.	NR	NA	NA	NA	
PKBS	NR	.96/.96 (c)	.58-.69/.86-.78 (c)	.48/.59 (c) (k); .38/.16 (c) (j)	CN; CS; CR
PKBS*	NR	.84-.97	.62-.87	.36-.63	CR;CC
SSRS				.22-.36	CN; CS; CR; S
Parent	Preschool	.90/.73 (c)	NA	NA	
	Elementary	.87/.87 (c)	.87/.65 (c)	NA	
Teacher	Preschool	.94/.82 (c)	NA	NA	
	Elementary	.94/.88/.95 (c)	.85/.84/.93 (d)	NA	
SSRS*					
Parent	NR	.87-.90/.73-.87 (c)	.87/.64 (c)	NR	CN; CS; CR; S
Teacher	NR	.93-.94/.82-.86 (c)	.85/.84 (c)		
TABC					CR
Parent	72-99 months	NA	NA	.34 (g) (k)- .45 (i)	
Teacher	30-80 months	NA	NA	NA	
TABS+	NR	NR	.81-.94	NA	NR
Vineland			.89-.93	.74 (i)	CN; CS; CR
	2-0 – 2-11	.97	NA	NA	
	3-0 – 3-11	.96	NA	NA	
	4-0 – 4-11	.96	NA	NA	
	5-0 – 5-11	.94	NA	NA	
	6-0 – 6-11	.95	NA	NA	
Walker	NR	.98	.66-.80 (g)	.81 (g) (k)	CN; CS; CR; F; I

Adapted from Bracken et al., 1998. Note. * = based on review by Printz et al., 2003; NA = no total test score; NR = not reported. (a) = Externalizing Problem/Internalizing Problem/Adaptive Skills/Behavior Symptoms; (b) = Externalizing Problems/Internalizing Problems/School Problems/Adaptive Skills/Behavior Symptoms; (c) = Social Skills/Problem Behavior Scales; (d) = Social Skills/Problem Behaviors/Academic Competence; (e) = Intensity/Problem Scales; (f) = male/female; (g) = Reported in manual as results of independent studies; (h) = Mean reliability of 4 subtests, manual cites additional studies for various factors; (i) = Parent-parent; (j) = Parent-teacher; (k) = teacher-teacher; (l) = Parent-student. ASQ:SE= Ages and Stages Questionnaire: Social Emotional; BASC= Behavior Assessment System for Children; BITSEA= Brief Infant Toddler Social and Emotional Assessment; Burks'= Burks' Behavior Rating Scales; CBCL=Child Behavior Checklist; CRS= Conners' Rating Scale; ESP= Early Screening Project; Eyberg= Eyberg Child Behavior Checklist; Louisville= Louisville Behavior Checklist; PBQ= Preschool Behavior Questionnaire; PIC= Personality Inventory for Children; PKBS= Preschool and Kindergarten Behavior Scales; SSRS= Social Skills Rating System; TABC= Temperament Assessment Battery for Children; TABS= TABS Screener for the Temperament and Atypical Behavior Scale; Vineland= Vineland Adaptive Behavior Scales; Walker= Walker Problem Behavior Identification Checklist; CN= Content Validity; CS= Construct Validity; CR= Criterion-Related Validity; CC= Concurrent Validity; CV= Convergent Validity; CG= Contrasted Groups; P=Predictive Validity; F= Factorial Validity; S= Social Validity; I= Item Validity; RR= Rater-Rater Agreement.

Table 4

Summary of Psychometric Criteria of Behavior Rating Scales Reviewed by Bracken et al,. 1998 and Printz et al., 2003

Instrument	Criteria								
	1	2	3	4	5	6	7	8	9
ASQ:SE*	Y	NR	NR	NR	NR	NR	NR	NR	Y
BASC	Y	P	P	P	Y	Y	Y	N	Y
BITSEA*	Y	NR	NR	NR	NR	NR	NR	N	Y
Burks'	N	N	NA	NA	NA	NA	NA	P	Y
CBCL	Y	N	P	P	Y	Y	Y	N	Y
CRS	N	Y	NA	NA	Y	P	NA	N	Y
DECA	Y	Y	N	NR	NR	NR	NR	N	Y
ESP*	Y	NR	NR	NR	NR	NR	NR	N	Y
Eyberg	N	Y	NA	P	NA	NA	NA	N	Y
Louisville	N	Y	NA	NA	Y	Y	NA	N	Y
PBQ	N	N	N	N	NA	NA	NA	N	Y
PIC	N	P	NA	Y	Y	N	NA	P	Y
PKBS	Y	Y	Y	N	NA	Y	Y	N	Y
PKBS*	Y	NR	P	NR	NR	NR	NR	N	Y
SSRS	Y	P	P	P	NA	Y	Y	N	Y
SSRS*	Y	NR	P	NR	NR	NR	NR	NR	Y
TABC	N	P	NA	NA	Y	Y	NA	N	Y
TABS*	Y	NR	NR	NR	NR	NR	NR	P	Y
Vineland	Y	N	Y	P	Y	Y	Y	N	Y
Walker	N	N	Y	N	Y	Y	Y	N	Y

Adapted from Bracken et al., 1998; * = reviewed by Printz et al., 2003; N= Criterion not met; NA= Criterion not applicable; NR= Criterion not reported; P= Criterion partially met; Y= Criterion met at most or all age levels; Criteria: 1. Large, national standardization sample; 2. Median subtest reliability ≥ .80; 3. Global scale reliability ≥ .90; 4. Global scale stability ≥ .90; 5. Average subtest standard score associated with deviant extreme scale ≥ +2SD or ≤ -2SD; 6. Item gradient ≥ 1; 3 SD per one raw score point; 7. Global scale standard score associated with deviant extreme of scale ≥ +2SD or ≤ -2SD; 8. Inter-rater reliability ≥ .90; 9. Evidence of validity in test manual ASQ:SE= Ages and Stages Questionnaire: Social Emotional; BASC= Behavior Assessment System for Children; BITSEA= Brief Infant Toddler Social and Emotional Assessment; Burks'= Burks' Behavior Rating Scales; CBCL=Child Behavior Checklist; CRS= Conners' Rating Scale; ESP= Early Screening Project; Eyberg= Eyberg Child Behavior Checklist; Louisville= Louisville Behavior Checklist; PBQ= Preschool Behavior Questionnaire; PIC= Personality Inventory for Children; PKBS= Preschool and Kindergarten Behavior Scales; SSRS= Social Skills Rating System; TABC= Temperament Assessment Battery for Children; TABS= TABS Screener for the Temperament and Atypical Behavior Scale: Early Childhood Indicators of Developmental Dysfunction; Vineland= Vineland Adaptive Behavior Scales; Walker= Walker Problem Behavior Identification Checklist

In relation to the Head Start population, Printz, Borg, and Demaree (2003) recommended six tools for assessing social, emotional, and behavioral difficulties that meet the requirements of the Head Start Performance Standards. These six tools are: Ages and Stages Questionnaire: Social Emotional (ASQ:SE; Squires, Bricker & Twombly, 2002); Brief Infant Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan, & Carter, 2000); Early Screening Project (ESP; Walker, Severson, & Feil, 1995); Preschool and Kindergarten Behavior Scales (PBKS; Merrell, 1994); Social Skills Rating System (SSRS; Gresham & Elliott, 1990); TABS Screener for the Temperament and Atypical Behavior Scale: Early Childhood Indicators of Developmental Dysfunction (TABS; Neisworth, Bagnato, Salvio, & Hunt, 1999). Tables 2- 4 above provide an overview of the technical adequacy of these six tools presented with the 13 tools reviewed by Bracken and colleagues (1998). The tools reviewed specifically by Printz et al. are marked with an asterisk.

It should be noted that in their respective reviews, Bracken et al. (1998) and Printz et al. (2003) used the traditional definition of validity in their evaluations, not the definition recommended in the *Standards for Educational and Psychological Testing* (AERA, 1999). The discussions on validity in these review articles do not adequately address the evidence available for each of these tools to support the intended interpretation of test scores for the proposed purpose. Thus, evidence of the use of these tools in specific populations and for specific purposes (i.e., screening in Head Start) is missing. Consequently, Bracken et al. for example, is only able to provide an overall picture of the psychometric soundness of each tool without respect to validity for the use of these tools within Head Start.

One of the major disadvantages of many rating scales, including those recommended by Printz et al. (2003) and Bracken et al. (1998), is that they are not practical for use within school settings for several reasons including: time for administration, lack of personnel to administer, lack of resources in obtaining appropriate tools. For example, Cicchetti and Toth (1997) highlight the promise of the Early Screening Project (ESP) for school-wide screening in preschool settings. Although the ESP appears to be effective in general preschool settings, such a tool would not be practical in a setting like Head Start in which many students are at-risk or are already experiencing difficulties, highlighting the importance of cultural validity in selecting assessment tools for Head Start (Sue, 1999; Feil et al., 2000). In addition, the time (e.g., approximately one hour to administer) and resources needed to implement and maintain the ESP may present a challenge for Head Start programs.

The Devereux Early Childhood Assessment

The Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999) is a relatively new tool that shows promise for assessing childhood difficulties. The DECA is a standardized, norm-referenced behavior rating scale that possibly addresses the need for new measures and methodologies for assessing developmental psychopathology. The DECA evaluates within-child protective factors (age 2-5) with a 37 item rating scale. The rater, generally someone who has daily contact with the child, is asked to respond on a rating scale (Never, Rarely, Occasionally, Frequently, Very Frequently) how often a behavior has been observed in the past four weeks. On average, the DECA takes approximately 15 to 20 minutes to complete. Items on the DECA are listed in Appendix C. The DECA includes items that produce two different scores:

Behavioral Concerns and Total Protective Factors (comprised of Initiative, Self-Control, and Attachment subscales). DECA results can be displayed graphically by individual or by groups.

Researchers recommend that both early assessment and intervention should focus on protective factors as well as risk factors (Werner & Smith, 1982, 1992). The DECA provides an overview of strengths and weaknesses of a child's protective factors. It helps to identify children who may be exhibiting emotional and/or behavioral problems, assists early childhood programs to develop strength-based programs to foster healthy social and emotional growth, and assists Head Start programs in meeting *Program Performance Standards*, particularly in regard to behavioral screenings, support of social-emotional development, and child mental health services (LeBuffe & Naglieri, 1999).

One clear advantage of the DECA is that it is linked to an early intervention program called the DECA Program. Within this program, the DECA assessment results are used to plan and set goals for individual children (or the classroom as a whole, if appropriate) in order to promote children's social and emotional well-being. Based on the individual (or classroom) strengths and weaknesses on the DECA (i.e., scores on the Total Protective Factors, Initiative Scale, Self-Control Scale, Attachment Scale, and Behavior Concerns Scale) the DECA Program lays out strategies for classroom teachers and parents to use that foster resilience. The DECA and DECA Program clearly aim to promote the development of positive behaviors despite risk factors that may exist in a child's life; therefore, the DECA is based on developmental psychopathology and resilience theories in that the instrument is directly linked to strategies that can

significantly alter a child's developmental trajectory by building on the child's present level of protective factors and while decreasing risk factors.

Although the DECA was not one of the six tools evaluated by Printz et al. (2003), it should be noted that it was listed as a commonly used social, emotional, and behavioral assessment tool. Printz et al. do not provide the criteria they used for selecting which tools to review (and which tools not to review); however, one reason that the DECA was excluded may be because there has been no independent research conducted on the validity and utility of this instrument for assessing risk and protective factors in children. Unlike many other tools that assess only deficits (e.g., what a child *cannot* do), the DECA is a strength-based instrument that helps teachers and parents understand the nature of a child's social, emotional, and behavioral functioning (e.g., what a child *can* do). Theoretically, the DECA is set up to meet the National Head Start Performance Standards in assessing children's functioning in these three domains (LeBuffe & Naglieri, 1999), and may serve as a tool to help professionals meet national mental health standards. Thus, validation studies of the DECA are warranted. Appendix A outlines the National Head Start Performance Standards alongside the description of how the DECA addresses the specific standard. A review of the psychometric properties of the DECA is reported in Tables 2-4 alongside the instruments reviewed by Bracken et al. (1998) and Printz et al. The information is presented in the same tables to allow for comparisons among all instruments. DECA information is provided in bold text.

Problem Statement and Hypotheses

Key gaps exist in the mental health research related to children from birth to age five, especially low-income children such as those served by Head Start (Lopez et al., 2000). The purpose of this study was to examine the validity of the DECA for use within a low-income Head Start sample in a mid-Michigan city. Questions regarding the validity of the DECA exist, as no known independent studies have been conducted on the DECA within the Head Start population. Furthermore, the DECA is compared with the current screening tool (the Michigan Risk Factors Student Identification Criteria; MRFSIC) being used within this Head Start sample.

Research Questions & Hypotheses

1. Is there adequate validity evidence for the DECA to be used within the Head Start sample to screen for children who are at-risk for developing social, emotional, or behavioral difficulties? It was hypothesized that validity evidence (Messick, 1989, 1995) is adequate for the DECA to be used in the Head Start sample.

2. Is there evidence to support the use of the DECA when compared to the Michigan Risk Factors Student Identification Criteria, the current tool being used in Head Start? It was hypothesized that evidence will exist to support the use of the DECA.

Method

Participants

An existing data set with information collected from 1208 children and their

parents' was used for this study. Parental responses on DECA forms and data from the

MRFSIC for the same children were used. The sample consisted of parents of children

in a Head Start program from multiple sites (i.e., three counties) in a diverse, Mid-

Michigan area. The data was collected at the time of enrollment into the 2004-2005 Head

Start Program year. The children in this sample range in age from 2 years to 5 years with

an average age of 3.64 years. The proportion of male and female children in the sample

was approximately equal with 49.3% males and 50.7% female. Information regarding

children's ethnic backgrounds was unavailable for this sample of children. However,

Head Start enrollment characteristics for the previous year (i.e., 2003-2004 year) were as

follows: American Indian (<1%), Asian (2.3%), African American (28.8%), Caucasian

(43.0%), Hispanic (10%), and Multi-racial (15.4%). These rates differ from the National

Head Start enrollment demographics for the same year (USDHHS, 2004): American

Indian (3.2%), Asian, (2.9%), African American (31.5%), Caucasian (27.6%), Hispanic

(30.6%), and Multi-racial (4.2%), indicating an underrepresentation of American Indian,

African American, and Hispanic children and an overrepresentation of Caucasian and

multi-racial children in the current Head Start sample when compared to national level

enrollment demographics.

Measures

The Devereux Early Childhood Assessment. LeBuffe and Naglieri (1999) report

the Devereux Early Childhood Assessment (DECA) being highly reliable for assessing

preschool children's protective factors. The DECA Protective Factors (Initiative, Self-Control, Attachment) was standardized on 2000 preschool aged children (1017 teacher ratings; 983 parent ratings). The Behavioral Concerns Scale was standardized on a sample of 1108 preschool children ages two to five years (567 teacher ratings; 541 parent ratings). LeBuffe and Naglieri report that the standardization sample for both the Protective Factors and the Behavioral Concerns Scale closely resembled the two- to five-year old population in the United States in terms of gender, geographic region of residence, race, ethnicity, and socioeconomic status. However, it should be noted that national Head Start population characteristics differ greatly from the national population of two- to five-year old children. Thus, more information is needed to understand the validity of the DECA within the Head Start population since the DECA standardization sample does not closely resemble Head Start population characteristics.

The items on Protective Factors Scales were identified using factor analysis. Factor loadings above .34, those which accounted for 10% or more of the variance, were used, resulting in the three Protective Factors Scales labeled Initiative, Self-Control, and Attachment. Thus, the analyses show that each of the Protective Factors Scale is comprised only of items with substantial loadings on the scale in which they were placed. The ten items on the Behavior Concerns Scale were selected from the 77 problem behaviors used in the standardization process, selected based on their psychometric properties and their wide representation of behavior problems. Further details can be found in the DECA Technical Manual (LeBuffe & Naglieri, 1999).

Internal consistency of the protective factor scales (i.e., Initiative, Self-Control, Attachment) ranged from .76 to .86 for parents and .85 to .90 for teachers. The overall

Total Protective Factor internal reliability values of .91 (parents) and .94 (teachers) exceeds Bracken's (1987) "desirable standard" for internal consistency. In addition, test-retest reliabilities and inter-rater reliability values were reported for the protective factors scales. Test-retest reliabilities over a 24-hour period ranged from .55 to .80 for parents and from .87 to .91 for teachers on the three protective factors scales, with overall Total Protective Factors at .74 (parents) and .94 (teachers). Interrater reliability coefficients from ratings provided by parents and teachers for the three protective factors scales ranged from a low of .19 (Attachment; parent-teacher) to .77 (Self-Control; teacher-teacher). Total Protective Factors interrater reliability reported were: .21 (parent-parent), .69 (teacher-teacher), and .29 (parent-teacher).

Internal consistency values were also reported for the Behavior Concerns Scale (LeBuffe & Naglieri, 1999). Internal consistencies of the Behavior Concerns Scale were .71 for parents and .80 for teachers. For the Behavior Concerns Scale, test-retest reliabilities over a 24-hour period were .55 for parents and .68 for teachers, and interrater reliability coefficients were .44 (parent-parent), .62 (teacher-parent), and .23 (parent-teacher). All values on the test-retest and inter-rater reliabilities were significant at the .01 level. A review of the psychometric properties of the DECA is reported in Tables 2-4 alongside the instruments reviewed by Bracken et al. (1998) and Printz et al. (2003). The information is presented together to allow for comparisons among all instruments.

Validity measures are also reported by the LeBuffe and Naglieri (1999) in the DECA Technical Manual to ensure that the DECA measures the intended protective factors in preschool children. Content-related validity assesses the degree to which the domain measured by a test is represented by the test items. Content-related validity is

measured by comparing the items on the test to that of other well established measures or research findings in the same area. Since the DECA is the first published rating scale of within-child protective factors, such comparisons could not be made. The authors instead based the content of the DECA on the resilience literature and results from focus groups of parents and teachers. Construct-related validity measures the degree to which an instrument accurately measures the intended construct. The construct validity of the DECA was tested by comparing the Protective Factor Scales and the Behavioral Concerns Scale. Results indicated that protective factors and problem behaviors were inversely related (r= -.65). Finally, criterion-related validity assesses the degree to which the scores on an assessment instrument predict an individual's performance on an outcome or criterion measure or the status or group membership of an individual (LeBuffe & Naglieri). Results of criterion-related validity studies indicated that the overall Total Protective Factors score was able to correctly classify 69% of the children as either being part of a clinical or a non-clinical sample, while the Behavioral Concerns Scale correctly predicted group membership for 71% of the children. According to the authors, these percentages are comparable to the classification accuracy of well-established scales of symptomatic behavior such as the Devereux Scales of Mental Disorders (Naglieri, LeBuffe, & Pfeiffer, 1995). In examining criterion-related validity the authors report that children received similar ratings on the DECA, regardless of race; thus, the DECA is appropriate for use with minority children. Table 2 above provides a listing of the types of validity reported for the DECA alongside the validity evidence for the instruments reviewed by Bracken et al. (1998) and Printz et al. (2003). It is noted that

despite widely varying behaviors within the preschool population (e.g., ages 2-5), only group norms are provided, rather than age norms.

Two independent reviews of the DECA were reported in the 15th Edition of the Mental Measurements Yearbook. Buhs (2003) reported the following about the DECA:

> *"The developers have published a useful behavioral assessment targeted at identifying the strengths or adaptive characteristics of young children. The DECA is quick and reliable for early childhood practitioners to use and requires minimal training to administer and score. Interpretation guidelines are thorough, contain appropriate cautionary language, and seem likely to provide effective means for giving constructive feedback to parents and practitioners..."* (p. 81)

Chittooran (2003), a second reviewer reported:

> *"The DECA is the first published measure of its kind and is a useful addition to the growing arsenal of tools available to assess social-behavioral functioning in young children. It is theory-based, psychometrically sound, demands minimal training and time for administration and interpretation, and links assessment to intervention..."* (p. 83).

Michigan Risk Factors Student Identification Criteria. The Michigan Risk Factors Student Identification Criteria (MRFSIC) is 25-item checklist of risk factors. Sample items on this checklist include: low birth weight, physical/sexual abuse or neglect, and family history of low school achievement or dropout (see Appendix B for a full listing of the Michigan Risk Factors Student Identification Checklist). The MRFSIC does not allow for ratings of the degree to which a risk or protective factor is present; it only indicates whether a risk factor is present. On average, the MRFSIC takes approximately 15 to 20 minutes to complete. No reports of reliability or validity are available for this instrument at this time.

Data Collection Procedures

The DECA and MRFSIC forms were collected for each child at the time of enrollment into the 2004-2005 Head Start Program year as part of enrollment procedures (i.e., if a child enrolled in Head Start at the beginning of the year, the information was collected at that time; if a child enrolled in the second half of the year, information was collected in January). Head Start staff members were responsible for administering the instruments. Data was provided to researchers at Michigan State University for data analysis purposes. The present author was not involved in the data collection stage.

Analysis Strategy

Statistical analyses were conducted on the existing data set to test for evidence of validity for the DECA for its use within a diverse, mid-Michigan Head Start sample. In particular, analyses were conducted to assess whether the statistical properties of the DECA data derived from the current study are comparable to that of the standardization sample reported in the DECA technical manual (LeBuffe & Naglieri, 1999). In addition, analyses were conducted to assess the psychometric properties of the MRFSIC to allow for a comparison of this tool with the DECA. Although Messick's (1989, 1995) conceptualization of the six types of validity evidence is presented in detail within the literature review, the current study evaluated four of the six types of validity discussed including: *content, substantive, structural,* and *generalizeability. External- and consequential- related* validities were not measured as part of this study.

Results

Results for Research Question 1

Internal consistency (Cronbach's alpha) values were calculated for the Head Start sample parental ratings on the DECA to provide *generalizeability* (by examining the extent to which scores and interpretations from the HS Sample generalize to the DECA standardization sample), *content* (by providing technical quality and content relevance), and *substantive evidence* (by examining theoretical rationales for the observed consistencies) for validity (Messick, 1989, 1995). Internal consistency is the extent to which the items on a scale measure the same underlying construct. In addition, standard error of measurement values were also used as a quality measure of the reliability scores. The standard error of measurement is an estimate of the amount of error in the observed score, expressed in standard score units. Internal reliability and standard errors of measurement from the Head Start sample were compared to those calculated from parental ratings in the standardization sample of the DECA. Table 5 below summarizes Cronbach's alpha and standard error of measurement values for the DECA Standardization Sample and the present Head Start Sample (parental ratings) on the Protective Factors and Behavior Concerns scales of the DECA. As presented in Table 5, the internal consistencies on the DECA for the Head Start Sample closely resemble those from the DECA Standardization Sample for Protective Factors and Behavior Concerns Scales.

Table 5

Internal Consistency for DECA Standardization and Head Start Samples

	DECA Standardization Sample		Head Start Sample (N=1208)	
	Cronbach's Alpha	Standard Error of Measurement	Cronbach's Alpha	Standard Error of Measurement
Protective Factors	.91 (N=983)	2.97	.91	2.98
Initiative	.84	4.03	.83	4.07
Self-Control	.86	3.74	.85	3.87
Attachment	.76	4.91	.77	4.84
Behavior Concerns	.71 (N=541)	5.40	.71	5.43

Pearson's product-moment correlation was used to examine *content-*, *substantive-*, and *structural- related* validity evidence for the DECA by providing support for technical quality and representativeness, theoretical rationales for observed consistencies, and fidelity of the scoring structure to the construct domain at issue, respectively (Messick 1995). This procedure indicated that Total Protective Scale scores on the DECA for the Head Start Sample were inversely related to DECA Behavior Concerns scale, $r= -.39$, $p \leq .00$ (N= 1208). This value has moderate magnitude according to Cohen's (1988) guidelines. Theoretically, children with higher Total Protective Scale scores on the DECA should have lower Behavior Concerns Scale scores. Inversely, children with high DECA Behavior Concerns scales should have low Total Protective Scale scores. However, the correlation between DECA Total Protective Factors Scale score and DECA Behavior Concerns Scale score is less strong for the Head Start sample when compared to the DECA standardization sample ($r= -.65$). Z-score conversions were conducted to determine the significance of the difference between these two correlations,

which resulted in insignificant findings (p > .05). Thus, although appearing large, the difference between the correlations for the Head Start Sample and the DECA Standardization Sample is not significant.

Exploratory factor analysis was conducted to determine the extent of *content-* and *structural- related evidence* for validity described by Messick (1989, 1995). Factor analysis provides support for *content-related* evidence by examining the content relevance of the test items and *structural-related* evidence by examining the fidelity of the scoring structure. Exploratory factor analysis is a procedure that examines for a pattern of relationships (e.g., clusters or factors) within an instrument. This analysis evaluated the same three-factor model for protective factors- Initiative, Self-control, and Attachment- reported in the standardization sample. As presented in Table 6, results indicated that the same three factors exist in the Head Start sample ratings; however, there were differences in which items loaded onto each of the three factors. Specifically, items 22, 32, and 37 loaded onto different factors. All DECA items with a factor loading of .34 or greater are including in Table 6.

Table 6

DECA Item Factor Loadings for DECA Standardization and Head Start Samples

Factors	DECA Standardization Sample	Head Start Sample
Initiative	2, 3, 7, 12, 16, 19, 20, 24, 28, **32**, 36	2, 3, 7, 12, 16, 19, 20, **22**, 24, 28, 32, 36, **37**
Self Control	4, 5, 13, 21, 25, 30, 33, 34	4, 5, 13, 21, 25, 30, 33, 34
Attachment	1, 6, 10, 17, **22**, 29, 31, **37**	1, 6, 10, 17, 22, 29, 31, **32**, 37

Note: *Item 22*: Ask adults to play with or read to him/her; *Item 32*: Ask other children to play with him/her; *Item 37*: Show an interest in what children/adults are doing

Generalizeability evidence for validity (Messick, 1989, 1995) was assessed by applying procedures to determine the scale score differences among the Head Start sample in comparison to the standardization and community samples presented in the DECA Manual (LeBuffe & Naglieri, 1999). This procedure provided mean t-scores and standard deviations for the data to provide evidence for whether the Head Start Sample scores differed from the DECA Identified or Community Samples (see Table 7). The DECA Identified Sample was a group of 95 children who had formal DSM-IV diagnosis of social, emotional, or behavioral disorder. The DECA Community sample consisted of children from the general population that had no current diagnosis.

The mean and standard deviation values for the Head Start sample more closely resemble the standardization sample mean of 50 and standard deviation of 10 reported by in the DECA Technical Manual (LeBuffe & Naglieri), with the exception of the Behavior Concerns score which was higher than both the DECA standardization sample and Community Sample means. T-tests were conducted to examine the significance of mean scores on the DECA for the Head Start Sample compared with the DECA Identified Sample and DECA Community Sample, respectively. The Head Start Sample T-scores on the DECA differed significantly on all scales when compared with the DECA Identified Sample [Initiative: $t(1301) = 7.81$, p< .0001; Self-Control: t $(1301) = 10.53$, p< .0001; Attachment: $t(1301) = 4.41$, p< .0001; Total Protective Factors: $t(1301) = 8.82$, p< .0001; Behavior Concerns: $t(1301) = 6.87$, p< .0001]. However, the Head Start Sample only differed significantly from the DECA Community Sample on the Behavior Concerns Scale [$t(1292) = 2.71$, p= .007]. Significant differences are marked with an asterisk in Table 7 below.

Table 7

Mean T-scores and Standard Deviations for DECA and Head Start Samples

Mean T-Scores, SD, & Difference Statistics	DECA Identified Sample (N=95)	DECA Community Sample (N=86)	Head Start Sample (N=1208)
Initiative			
Mean	41.2*	48.6	49.28
Stand. Dev.	9.8	9.2	9.70
Self-Control			
Mean	38.9*	49.1	50.13
Stand. Dev.	10.2	10.0	9.99
Attachment			
Mean	41.9*	47.0	47.14
Stand. Dev.	10.5	11.3	11.21
Total Protective Factors			
Mean	38.5*	47.3	47.91
Stand. Dev.	9.9	10.0	10.02
Behavior Concerns			
Mean	65.4*	55.7*	58.54
Stand. Dev.	8.8	9.3	9.41

Note: An asterisk (*) on DECA columns indicates a significance level of $p < .01$ when compared to the Head Start Sample means (third column)

Results for Research Question 2

Pearson's correlation indicated a significant relationship between Total Behavior

Concerns and Total MRFSIC scores, r= .14, p≤ .00 (N= 1208). According to guidelines

set forth by Cohen (1988), this correlation would be considered of small magnitude. This

was expected, given that both scales measure behavioral risk. Thus, further analyses

were warranted. Multiple regression was applied to analyze the MRFSIC, the tool

currently being used by the local Head Start Program. This analysis was conducted to

provide *content- related evidence* (i.e., content relevance and technical quality).

Multiple regression is a statistical procedure that is used to make predictions based on a

regression model. This statistical procedure was used to determine whether as a whole,

the MRFSIC is related to higher scores on the DECA Behavior Concerns Scale (i.e., does

a high score on the MRFSIC predict a high score on the DECA Behavior Concerns

Scale?). Typically, multiple regression would be used to predict an existing criterion

(i.e., MRFSIC scores would be predicted from DECA scores). However, because of no

reliability or validity information is available for the MRFSIC, it would not be

meaningful to predict these scores from the DECA scores. That is, even if DECA scores

significantly predicted MRFSIC scores, because the reliability and the validity of the

MRFSIC is unknown, we would question the relevance of the results. Thus, MRFSIC

scores were used to predict DECA scores because some data is available regarding the

reliability and validity of the DECA.

Total MRFSIC score was entered into the regression analysis with DECA

Behavior Concerns Scale score being the outcome variable. Results indicated that Total

score on the MRFSIC did not significantly predict higher scores on the DECA Behavior

Concerns Scale. In addition, multiple regression was used to determine whether any individual items on the MRFSIC predicted higher scores on the DECA Behavior Concerns Scale. Thus, individual MRFSIC items were each entered into the regression model with DECA Behavior Concerns score being the outcome variable. Table 8 reports the Unstandardized and Standardized Coefficients and the significance levels for the multiple regression analyses performed for Behavior Concerns Total Scale scores.

Table 8

Significant MRFSIC Predictor Items of Behavior Concerns Total Scores

	Unstandardized Coefficients		Standardized Coefficients	Significance Level
	B	Std. Error	Beta	alpha
Developmentally immature	1.41	.68	.07	p≤ .05
Destructive or violent temperament	2.99	.53	.20	p≤ .00
Language deficiency or immaturity	1.12	.51	.08	p≤ .05
Family history of low school achievement or dropout	.96	.40	.09	p≤ .05

As presented in Table 8, regression analysis indicated that four of the 25 risk factors on the MRFSIC significantly predicted higher scores on the DECA Behavior Concerns Scale. These four items were: a) Developmentally immature, b) Destructive or violent

51

temperament, c) Language deficiency or immaturity, and d) Family history of low school

achievement or dropout.

Multiple regression analysis was also employed to examine whether any

individual items on the MRFSIC predicted lower scores on the DECA Total Protective

Factors Scale. Each MRFSIC item was entered into the regression model with the DECA

Total Protective Factors Scale score as the outcome variable. As presented in Table 9,

results indicated that eight individual items from the MRFSIC significantly predicted

lower scores on the DECA Total Protective Factors Scale: a) Developmentally immature,

b) Nutritionally deficient, c) Diagnosed handicapping condition, d) Lack of stable support

system of residence, e) Destructive or violent temperament, f) Language deficiency or

immaturity, g) Non-English or limited English speaking household, and h) Family history

of low school achievement or dropout. Table 9 reports the Unstandardised and

Standarised Coefficients and the significance levels for the multiple regression analyses

performed for Total Protective Factors scores.

Table 9

Significant MRFSIC Predictor Items of Total Protective Factors Scores

	Unstandardised Coefficients		Standardised Coefficients	Significance Level
	B	Std. Error	Beta	alpha
Developmentally immature	-5.27	1.77	-.10	p≤ .00
Nutritionally deficient	-2.80	.95	-.12	p≤ .00
Diagnosed handicapping condition	-5.77	2.17	-.09	p≤ .00
Lack of stable support System of residence	-3.26	1.65	-.07	p≤ .05
Destructive or violent temperament	-5.79	1.38	-.15	p≤ .00
Language deficiency or immaturity	-4.71	1.34	-.13	p≤ .00
Non-English or Limited English speaking household	-4.31	1.37	-.11	p≤ .00
Family history of low school achievement or dropout	-2.73	1.05	-.10	p≤ .00

Discussion

To date, key gaps exist in the mental health research related to children from birth to age five, especially low-income children such as those served by Head Start (Campbell, 1995; Fantuzzo et al, 2003; Lopez et al., 2000). The purpose of this study was to examine the validity of the DECA for use within a diverse, mid-Michigan Head Start sample. The DECA has previously been understudied, especially within the Head Start population. This study adds to the current literature in that it is the first known independent study to examine validity of the DECA. In addition, this study is unique because it attempted to understand child psychopathology and developmental outcomes from a risk *and* resiliency framework, whereas previous research has a limited focus on resiliency. This study also included a large sample size (N=1208) of Head Start children, and thus, has implications for the use of the DECA within the Head Start population for screening children for social-emotional and behavioral difficulties.

Validity evidence was examined according to Messick's (1989, 1995) conceptualization of validity. This is unlike previous research (i.e., Bracken et al, 2003; Printz et al., 1998) on early childhood screening tools that have taken a more traditional definition of validity. Because validity was conceptualized differently in this study it is difficult to present a point-by-point comparison of the results of this study to the other measures reviewed by Bracken et al. and Printz et al. However, comparisons will be made where appropriate and the relevance discussed throughout the following section. It should be noted that if validity were examined in the traditional approach (see description under *Issues of Validity* in the literature review), the results of this study would indicate the presence of both construct and content validities.

Messick described six types of evidence that can exist to support the overall construct validity of a test: *content, substantive, structural, generalizeability, external,* and *consequential.* These six aspects function as general validity criteria or standards for all educational and psychological measurement. Although Messick's conceptualization of the six types of validity evidence is presented in detail within the literature review, the current study evaluated four of the six types of validity discussed including: *content, substantive, structural,* and *generalizeability.* Varying degrees of support for each of the related validities was found. Given the limitations of the available dataset, *external-* and *consequential- related* validities were not measured as part of this study.

Four validities that were examined as part of this study were supported: *content, substantive, structural,* and *generalizeability.* The presence of these validity-related evidences supports the use of the DECA within the Head Start sample. Table 10 summarizes the analyses conducted to provide support for each type of validity.

Table 10

Summary of Support for Validities

Type of Validity	Supporting Analysis
Content	Internal Consistency; Correlation; Multiple Regression
Substantive	Internal Consistency; Correlation; Factor Analysis
Structural	Correlation; Factor Analysis
Generalizeability	Internal Consistency; Scale Score Differences; t-tests

The internal consistencies found on the DECA for the Head Start sample (Initiative: .83; Self-Control: .85; Attachment: .77; Total Protective Factors .91; Behavior Concerns .71) were comparable to those found for the DECA standardization sample (Initiative: .84; Self-Control: .86; Attachment: .76; Total Protective Factors .91; Behavior Concerns .71) on the three protective factors scales, Total Protective Factors, and Behavior Concerns. Because the Behavior Concerns Scale reliability score for the Head Start sample (.71) is low according to the technical standards and those set forth by Bracken et al. (2003), compared to the other instruments reviewed by Bracken et al. and Printz et al. (see Table 3), the internal consistency values fall within a similar range. For instance, the internal consistencies of the Protective Factor Subscales found for the Head Start sample were higher than a similar scale (i.e., Social Competence) reported for the standardization sample of the CBCL [.62 (girls ages 4-11) and .57 (boys ages 4-11)], one-well known and widely-used tool. The Head Start sample internal consistencies fell in the same range for competence scores (i.e., protective factors) when compared to another widely recognized and used tool, the SRSS (competence: .87). Thus, it should be reiterated that the Behavior Concerns Scale of the DECA should only be used as a screener to identify children who are in need of additional data to inform referrals. Using the Total Protective Factors scores may be a more appropriate way to identify children in this Head Start sample who do not necessarily need additional services, given that the internal consistencies for this scale was quite high (.91).

Based on the data from the Head Start sample, the DECA would have also passed six of the eight psychometric criteria set forth by Bracken et al., including: median subtest internal consistency of .80 or greater (Initiative, Self-Control, & Attachment),

scale internal consistency of .90 or greater (Total Protective Factors), average subtest floor or ceiling at least 2 standard deviations beyond the normative mean subtest score, scale floor or ceiling at least 2 standard deviations beyond the normative mean, validity data available, and instrument's standardization sample evaluated.

Furthermore, as expected, the Head Start sample (Total Protective Factors: 47.91, SD: 10.02; Behavior Concerns: 58.54, SD: 9.41) more closely resembled the DECA Community Sample (Total Protective Factors: 47.3, SD: 10.0; Behavior Concerns: 55.7, SD: 9.3) on mean T-scores and standard deviations, rather than the DECA Identified Sample (Total Protective Factors: 38.5, SD: 9.9; Behavior Concerns: 65.4, SD: 8.8) of children who have been formally diagnosed with a DSM-IV disorder. The internal consistency and mean scale scores provided *content-, substantive-,* and *generalizeability-evidence* for overall validity of the DECA. These results indicate that, while the Head Start sample is a sample of children already considered to be at-risk, the children in this sample may not have developed the degree of behavioral difficulties that are exhibited in children with formally diagnosed behavioral disorders. Thus, the presence of protective factors could explain why the children in this Head Start sample exhibit behavioral risk and concern at levels which are comparable to the average preschool child.

Based on the results from this study, it appears that there is an inverse relationship between protective factors (i.e., resiliency) and behavioral concern (i.e., risk) as is consistent with theories associated with developmental psychology (Luthar & Cicchetti, 2000; Masten, 2003; Masten & Curtis, 2000; Rutter & Sroufe, 2000; Webster-Stratton & Hammond, 1998). As the score on the Total Protective Factors Scale increased, the score on the Total Behavior Concern Scale decreased. However, the correlations from the

Head Start sample on the Total Protective Factors Scales and the Behavior Concerns

Scale differed quite remarkably in comparison to what is reported in the DECA technical

manual. Although both samples indicated an inverse relationship between Total

Protective Factors and Behavior Concerns, the strength of the relationship differed (Head

Start Sample: r = -.39; DECA Standardization Sample: r = -.65). A greater difference

exhibited between risk and protective factors would result in a stronger inverse

relationship between protective factors and risk, which was not as substantial within this

Head Start sample.

The rationale for this finding is curious and raises a number of interesting

questions. Because children in Head Start are already considered an at-risk population, it

might be that the children in this Head Start sample have higher behavioral difficulties

(i.e., risk) but are also perceived by their parents as having more protective factors

compared to what would be expected from the stronger inverse relationship reported in

the technical manual. This has strong implications for educators and mental health

professionals to foster and build upon the protective factors (resiliency) in children in a

unique at-risk population such those within the Head Start Program. Another possible

explanation for the weaker correlation found for the Head Start sample is that there may

have been interpretation issues that may influence how an individual from a non-English

speaking home rated items on the DECA. This possibility in itself raises a number of

other issues regarding ecological and cultural differences that may need to be further

considered in future research. A final possibility for the difference in strength of the

relationship may be due to restricted range. The standard deviations were higher for the

Head Start group when compared to the DECA groups that may have been influenced by

outliers, which could result in a restricted range that is causing a weaker correlation for the Head Start group. This explanation would fit expectations—the Head Start group might be expected to have more behavior problems generally-- and perhaps a ceiling effect caused restriction of range on the Behavior Concerns Scale. Thus a weaker correlation could be more theoretically linked to the unique characteristics of this sample (i.e., low income, non-English speaking) which raises additional questions regarding the representation of cultural factors within the resiliency research.

Factor analysis from both samples revealed the same 3-factors on the DECA, although there were differences in where three items loaded. It should be noted that the three items in question were all protective factors (*Item 22*: Ask adults to play with or read to him/her; *Item 32*: Ask other children to play with him/her; *Item 37*: Show an interest in what children/adults are doing), thus where the items load does not impact the ratings of risk on the Behavior Concerns scale. It is interesting to find virtually identical factor loadings given the differences in the two samples (i.e., Head Start sample characterized by high poverty, limited language, etc.), but theoretically the factor loadings make sense. These results indicate that the DECA has adequate *content-* and *structural- related* validity for measuring three distinct protective factors in this Head Start sample: Initiative, Self-Control, and Attachment.

The second hypothesis stated that there would be evidence to support the use of the DECA when compared to the MRFSIC. Although *content-related* evidence was found to support the validity of the DECA for use with this sample, there was not a clear relationship between the DECA and the MRFSIC for identifying children who may be in need of additional support services within Head Start. Thus, the use of the DECA over

the MRFSIC is not fully supported by the results of this study, yet closer attention to what defines risk on the MRFSIC is necessary as current Head Start criteria indicated that almost the entire population is at-risk and does not provide information to users about who may be at greater risk for presenting struggles within the classroom.

Results of multiple regression analyses indicated that Total Score on the MRFSIC did not significantly predict higher scores on the DECA Behavior Concerns Scale or lower scores on the DECA Total Protective Factors Scale. It appears that while some individual items from the MRFSIC significantly predict scores on the DECA, as a whole, it is unable to significantly predict DECA Behavior Concerns scores or DECA Total Protective Factors scores. Thus, these findings reveal some initial evidence that risk on the MRFSIC and the DECA are poorly related. These analyses provided *content-related* evidence for the overall validity of the DECA. The poor link found between risk on these two instruments could also be explained by the kinds if risk items included on each tool. Close examination of the items on the tools reveal risk items that are child-related (i.e., internal) and ecologically-related (i.e., environmental). For instance, while both tools contain items that would be considered ecological in nature, the MRFSIC contains several items that are more child-related (i.e., long-term chronic illness, developmental immaturity, destructive temperament) when compared to the items on the DECA. Clinically speaking, examining the presence of risk factors alone may not be adequate for determining which children are in need of additional services. In other words, both risk and protective factors need to be examined, and the degree to which these are present may be equally important. The combined results of multiple regression, however, indicate that more internal risk factors (i.e., nature) significantly predicted DECA

Behavior Concerns and Total Protective Factors scores; thus, funding may wish to be directed toward services to diminish individual, internal risk (e.g., parent education on biological risk, biological interventions).

Furthermore, it should be noted that according to the MRFSIC criteria, all 46 classrooms who participated in this study would be considered at-risk and in need for further services. Using the DECA criteria of a Behavior Concerns scale T-Score of 60 or higher, eight of the 46 classrooms would be considered in need for further support services based on classroom mean data on the DECA. It is unlikely that all 46 classrooms in this sample are in need of additional services, and it is also unlikely that there are enough resources available to employ additional services for all classrooms. Thus, it appears that the DECA may provide a more useful means of identifying children in need of additional supports when compared to the MRFSIC. It is important to reiterate that the Behavior Concerns Scale on the DECA is only a screener for risk. That is, it should not be used for identifying children who should be referred for additional services, but instead be used to indicate those who warrant additional attention, such as more data gathered from teacher reports and classroom observations.

Understanding risk and the impact of risk on children's developmental of psycho-pathology is important. In some instances, risk factors can be reduced or eliminated. In other cases, risk cannot be adequately changed to a degree that alters a child's develop-mental outcomes. Issues of resiliency or protective factors might be more easily addressed. Developmental psychopathology exerts that protective factors may diminish the degree to which risk manifests itself in a child's life. A developmental perspective provides a theoretical framework for Head Start researchers to identify and develop

competence. The developmental psychopathology perspective emphasizes the importance of human development in the context in which the development occurs and seeks to understand human development in terms of changes in the multifaceted nature of functions over time. Development is multifaceted in that it is understood by looking at the central tasks that children are expected to perform involving their physiological, cognitive, emotional, and social capacities. Understanding the role that context plays in development is an essential component in this model. As such, it is important to have contextually relevant measures of individual competencies and measures of positive influences in the environment that foster competent functioning. The results of this study suggest that resiliency theory does not adequately address the influence of protective factors on groups who may differ culturally from the norm. Cultural issues experienced by the children in this Head Start sample (e.g., ethnicity, socio-economic status, and language differences) deserve more attention within resiliency work. This study contributed only a small piece of what is needed to understand the whole picture of resiliency in culturally-diverse children.

The DECA is the only known rating scale available that comprehensively assesses protective factors in young children. By focusing on building these important protective factors in children, we may be able to alter the trajectory of development in a positive fashion. This study is only a beginning point for evaluating tools that may be valid and practical for use within Head Start programming. Given that the DECA has been linked to the National Head Start Performance Standards it has the potential to assist Head Start staff in meeting these national standards. The fact that it compares favorably to other rating scales used with preschoolers in terms of validity evidence, the DECA should be

further examined for inclusion in the mental health services provided by Head Start and other preschool programs.

Limitations & Future Research

There are a number of limitations to this study that impact our findings. Two of Messick's (1989, 1995) validity evidences were not examined in this study. Having linked data for this current dataset of children who were actually referred for additional services or who received additional support would provide further validity evidence for the predictive power of the DECA for identifying children in need of services in Head Start. Specifically, a linked dataset would have allowed for a closer examination of Messick's (1989, 1995) *external-* and *consequential- related* validity. Regardless, it should be noted that at this point, *external-* and *consequential- related* evidence is still needed for the DECA and *all* tools recommended and reviewed by Bracken et al. (2003) and Printz et al. (1998). It should also be noted that this study employed several different statistical analysis procedures to determine validity support. Three of the validity-related evidences had multiple means of support through multiple statistical procedures (i.e., *content, substantive, structural*), while *generalizeability-related* validity evidence had more limited support. See Table 10 for a summary. The limited support available was in part due to the limited information available in the dataset (which restricted the types of analyses that could be conducted, rather than due to poor results itself). Future research should be conducted to provide further support for the six validity-related evidences, especially those in this study that had more limited support.

In addition, ethnic background information for children in the Head Start Sample would have allowed for additional evaluation of patterns of risk and protective factors

among minority and underserved groups. Information regarding children's ethnic backgrounds was unavailable for this sample of children. However, Head Start enrollment characteristics for the previous year (i.e., 2003-2004 year) differ from the National Head Start enrollment demographics for the same year (USDHHS, 2004) and suggests that American Indian, African American, and Hispanic children are underrepresented in current Head Start sample and that Caucasian and Multi-racial children are overrepresented when compared to national- level enrollment demographic characteristics. Thus, caution should be taken when generalizing these results to the entire population of Head Start children.

There are disadvantages inherent to the use of any rating scale that should be noted. In specific, this study examined parental ratings of their children on the DECA and the MRFSIC. It is not known if parents were influenced to answer in a particular way (i.e., Did parents overrepresent their children's problems to gain more access to services? Did parents underrepresent their children's risk factors to reduce social stigma?). It may have also been useful if teacher ratings of children were available in this study to provide for additional reliability evidence (i.e., test-retest, interrater reliability). However, as indicated by LeBuffe and Naglieri (1999), more research in general is needed on the predictive validity of parental ratings on instruments such as the DECA. Limitations also existed on the MRFSIC as there is no data available on the reliability or validity of this tool. Thus, as found in the literature review of childhood behavior rating tools, best practice is not always followed in selecting tools and establishing cut-off criteria for determining those who are at-risk for social-emotional and behavioral difficulties. The importance and need for multiple raters and multiple methods should

also be stressed when gathering information on children for the basis of decision-making regarding services and programming.

This study focused on screening in Head Start because of the identified gap in the research and because the literature indicated that children in Head Start are developmentally amenable to screening due to language and motor development. Review of the literature on precursors of psychopathology suggest that Head Start is an appropriate starting point to systemically screen for children who may be in need of additional support services (Baydar et al., 2003; Webster-Stratton & Hammond, 1998). However, beginning screening activities even within Early Head Start may be useful in beginning the early prevention and intervention process for children who are at-risk.

Identifying early risk factors is only the first step in addressing childhood social and behavior problems. Once the significant risk factors are determined, prevention efforts can be implemented to students who are identified as experiencing these risk factors. The DECA is part of a comprehensive assessment system referred to as the DECA Program. Alongside the DECA rating scale, the Devereux Foundation has developed a supporting curriculum that involves systematic classroom observations and interventions to help children increase their protective factors (i.e., Initiative, Self-Control, Attachment). This current study examined only the DECA rating instrument, the first step of the comprehensive DECA Program. Future research should also be conducted on the *external* validity of the entire DECA Program for serving the social, emotional, and behavioral needs of Head Start children. For instance, research may shed further information regarding whether classroom staff find the DECA to be easily administered and interpreted.

Devereux Early Childhood Assessment (DECA) [LeBuffe & Naglieri, 1999]

1. Act in a way that made adults smile or show interest in him/her

2. Do things for himself/herself

3. Choose to do a task that was challenging for her/him

4. Listen to or respect others

5. Control his/her anger

6. Respond positively to adult comforting when upset

7. Participate actively in make-believe play with others (dress-up, etc.)

8. Fail to show joy or gladness at a happy occasion

9. Touch children/adults inappropriately

10. Show affection for familiar adults

11. Have temper tantrums

12. Keep trying when unsuccessful (act persistent)

13. Handle frustration well

14. Have no reaction to children/adults

15. Use obscene gestures or offensive language

16. Try different ways to solve a problem

17. Act happy or excited when parents/guardian returned

18. Destroy or damage property

19. Try or ask to try new things or activities

20. Start or organize play with other children

21. Show patience

22. Ask adults to play with or read to him/her

23. Have a short attention span (difficulty concentrating)

24. Focus his/her attention or concentrate on a task or activity

25. Share with other children

26. Fight with other children

27. Become upset or cry easily

28. Say positive things about the future (act optimistic)

29. Trust familiar adults and believe what they say

30. Accept another choice when his/her first choice was unavailable

31. Seek help from children/adults when necessary

32. Ask other children to play with him/her

33. Cooperate with others

34. Calm herself/himself down when upset

35. Get easily distracted

36. Make decisions for himself/herself

37. Show an interest in what children/adults are doing

Appendix B

Michigan Risk Factors Student Identification Criteria (MRFSIC)

1. Low birth weight

2. Developmentally immature

3. Physical and/or sexual abuse and neglect

4. Nutritionally deficient

5. Long-term chronic illness

6. Diagnosed handicapping condition

7. Lack of stable support system of residence

8. Destructive or violent temperament

9. Substance abuse or addiction

10. Language deficiency or immaturity

11. Non-English or limited English speaking household

12. Family history of low school achievement or dropout

13. Family history of delinquency

14. Family history of diagnosed family problems

15. Low parental/sibling educational attainment or illiteracy

16. Single parent

17. Unemployed parent/parents

18. Low family income

19. Family density

20. Parental loss by divorce or death

21. Teenage parent

22. Chronically ill parent (physical, mental, emotional)

23. Incarcerated parent

24. Housing in rural or segregated area

25. Family relocation two or more times last 12 months

References

American Educational Research Association. (1999). *The standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Anderson, D. R. (1983). Prevalence of behavioral and emotional disturbance and specific problem types in a sample of disadvantaged preschool-aged children. *Journal of Clinical Child Psychology, 12,* 130-136.

Barkley, R. A. (1988). Child behavior rating scales and checklists. In M. Rutter, A.H. Tuma, & I. S. Lann (Eds.) *Assessment and diagnosis in child psychopathology* (pp. 113-155). New York, NY: The Guilford Press.

Bates, J. E. (1987). Temperament in infancy. In J. Osofsky (Ed.), *Handbook of infant development, 2nd Ed.* (pp. 1101-1149). New York, NY: John Wiley.

Bates, J. E., & Bayles, K. (1988). Attachment and the development of behavior problems. In J. Belsky, & T. Nezworski (Eds.), *Clinical implications of attachment.* (pp. 253-299). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baydar, N., Reid, M. J., & Webster-Stratton, C. (2003). The role of mental health factors and program engagement in the effectiveness of a preventive parenting program for Head Start mothers. *Child Development, 74,* 1433-1453.

Bear, G. G., Webster-Stratton, C., Furlong, M., & Rhee, S. (2000). In K. M. Minke & G. G. Bear (Eds.), *Preventing school problems-promoting school success* (pp. 1-69). Bethesda, Maryland: National Association of School Psychologists.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment, 5,* 313-326.

Bracken, B. A., Keith, K., & Walker, K. C. (1998). Assessment of preschool behavior and social-emotional functioning: A review of thirteen third-party instruments. *Journal of Psychoeducational Assessment, 16,* 153-169.

Briggs-Gowan, M., & Carter, A. (2000). *Brief infant toddler social and emotional assessment.* New Haven, CT: Yale University.

Brooks-Gunn, J., Duncan, G. J., & Aber, L. (1997). The effects of poverty on children. *The Future of Children: Children and Poverty, 7,* 55-71.

Buhs, E. S. (2003). Review of the Devereux Early Childhood Assessment. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *Mental Measurements Yearbook, 15th ed.* (pp. 312-314). Lincoln, NE: Buros Institute of Mental Measurements.

Busse, R. T., & Gresham, F. M (1993). Behavior rating scales: Issues of use and development. *School Psychology Review, 22,* 313-321.

Campbell, S. B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology & Psychiatry, 36,* 113-149.

Campbell, S. B. (1997). Behavior problems in preschool children: Developmental and family issues. *Advances in Clinical Child Psychology, 19,* 1-26.

Campbell, S. B., & Ewing, L. J. (1990). Hard-to-manage preschoolers: Adjustment at age nine and predictors of continuing symptoms. *Journal of Child Psychology & Psychiatry,* 31, 871-889.

Campbell, S. B., Pierce, E., March C., & Ewing, L. J. (1991). Noncompliant behavior, overactivity, and family stress as predictors of negative maternal control in preschool children. *Development and Psychopathology, 3,* 175-190.

Children's Defense Fund (2004). *The state of America's children 2004.* Washington DC: Children's Defense Fund.

Chittooran, M. M. (2003). Review of the Devereux Early Childhood Assessment. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *Mental Measurements Yearbook,* 15th ed. (pp. 314-316). Lincoln, NE: Buros Institute of Mental Measurements.

Cicchetti, D. (1993). Developmental psychopathology: Reactions, reflections, projections. *Developmental Review,* 13, 471-502.

Cicchetti, D., & Toth, S. L. (1997). Transactional ecological systems in developmental psychopathology. In S. S. Luthar, J. A. Burack, et al. (Eds.), *Developmental psychopathology: Perspectives on adjustment, risk, and disorder* (pp .317-349). New York, NY: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd Ed.* Hillsdale, NJ: Lawrence Earlbaum Associates.

Coie, J. D., Miller-Johnson, S., & Bagwell, C. (2000). Prevention science. In A. Sameroff, M. Lewis, & S. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 93-112). New York, NY: Kluwer Academic/Plenum Publishers.

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, J. J., et al., (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist, 4,* 1013-1022.

Costello, E. J., & Angold, A. (1996). Developmental psychopathology. In R. B. Cairns, & G. H. Elder (Eds.), *Developmental Science* (pp. 168-189). New York, NY: USsity Press.

Cowen, E. L. (1997). Schools and the enhancement of children's wellness: Some opportunities and some limiting factors. In R. P. Weissberg, T. P. Gullotta, R. L. Hampton, B. A. Ryan, & G. R. Adams (Eds.), *Healthy children 2010: Establishing preventive services* (pp. 97-123). Thousand Oaks, CA: Sage Publications, Inc.

DeVellis, R. F. (2003). *Scale development: Theory and applications, 2nd Ed.* Thousand Oaks, CA: Sage Publications.

DuPaul, G. J. (2003). Commentary: Bridging the gap between research and practice. *School Psychology Review, 32,* 178-180.

Edwards, M., Whiteside-Mansell, L., Conners, N. A., & Deere, D. (2003). The unidimensionality and reliability of the Preschool and Kindergarten Behavior Scales. *Journal of Psychoeducational Assessment, 21,* 16-31.

Emery, R. E. (1982). Interparental conflict and the children of discord and divorce. *Psychological Bulletin, 92,* 310-330.

Fantuzzo, J. W., McDermott, P. A., Manz, P. H., Hampton, V. R., & Alverez Burdick, N. (1996). The Pictorial Scale of Perceived Competence and Social Acceptance: Does it work with low-income urban children? *Child Development, 67,* 1071-1084.

Fantuzzo, J., McWayne, C., & Bulotsky, R. (2003). Forging strategic partnerships to advance mental health science and practice for vulnerable children. *School Psychology Review, 32,* 17-37.

Feil, E. G., Walker, H., Severson, H., & Ball, A. (2000). Proactive screening for emotional/behavioral concerns in Head Start preschools: Promising practices and challenges in applied research. *Behavioral Disorders, 26,* 13-25.

Forness, S. R., Serna, L. A., Nielsen, E., Lambros, K., Hale, M. J., & Kavale, K. A. (2000). A model for early detection and primary prevention of emotional or behavioral disorders. *Education and Treatment of Children, 23,* 325-345.

Funderburk, B. W., Eyberg, S. M., Rich, B. A., & Behar, L. (2003). Further psychometric evaluation of the Eyberg and Behar rating scales for parents and teachers of preschoolers. *Early Education & Development, 14,* 67-81.

Garcia Coll, C., Lamberty, G., Jenkins, R., McAdoo, H. P., Crnic, K., Wasik, B. H., et al. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development, 67,* 1891-1914.

Garmezy, N. (1971). Vulnerability research and the issue of primary prevention. *American Journal of Orthopsychiatry, 41,* 101-116.

Garmezy, N. (1974). The study of competence in children at-risk for severe psychopathology. In E. J. Anthony & C. Koupernik (Eds.), *The child in his family (Vol. 3): Children at psychiatric risk* (pp. 77-97). New York, NY: Wiley.

Goldberg, S., Corter, C., Lojkasek, M., & Minde, K. (1990). Prediction of behavior problems in 4 year olds born prematurely. *Development and Psychopathology, 2,* 15-30.

Greenfield, D. B., Iruka, I. U., & Munis, P. (2004). Assessment of social competence in high-risk preschoolers: Evaluation of the Adaptive Social Behavior Inventory (ASNI) across home and school settings. *Journal of Psychoeducational Assessment, 22,* 220-232.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System.* Circle Pines, MN: AGS Publishing.

Guerra, N. G., & Jagers, R. (1998). The importance of culture in the assessment of children and youth. In V. C. McLoyd, & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 167-181). Mahwah, NJ: USum Associates.

Hooks, M. Y., Mayes, L. C., & Volkmar, F. R. (1988). Psychiatric disorders among preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 623-627.

Keenan, K., Shaw, D. S., Walsh, B., Delliquadri, E., & Giovannelli, J. (1997). DSM-III-R disorders in preschool children from low-income families. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 620-627.

Kim-Cohen, J., Moffitt, T. E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development, 75,* 651-668.

Knitzer, J. (2000). Early childhood mental health services: A policy and systems development perspective. In J. P. Shonkoff, & S. J. Meisels (Eds.), *Handbook of early childhood intervention, 2nd Ed.* (pp. 416-438). New York, NY: Cambridge University Press.

Knoff, H. M. (2002). Best practices in personality assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp. 1281-1302). Bethesda, MD: National Association of School Psychologists.

Lavigne, J. V., Gibbons, R. D., Christoffel, K. K., Arend, R., Rosenbaum, D., Binns, H., et al. (1996). Prevalence rates and correlates of psychiatric disorders among preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 204-214.

LeBuffe, P. A., & Naglieri, J. A. (1999). *Devereux Early Childhood Assessment* (DECA). Villanova, PA: The Devereux Foundation.

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin,* 94, 68-99.

Lopez, M. L., Tarullo, L. B., Forness, S. R., & Boyce, C. A. (2000). Early identification and intervention: Head Start's response to mental health challenges. *Early Education & Development, 11,* 265-282.

Luthar, S. S. & Cicchetti, D. (2000). The construct of resilience: Implications for interventions and social policies. Development & Psychopathology, 12, 857-885.

Mash, E. J., & Dozois, D. J. A. (2003). Child psychopathology: A developmental systems perspective. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 3-71). New York, NY: Guildford Press.

Masten, A. S. (1989). Resilience in development: Implications for the study of successful adaptation for developmental psychopathology. In D. Cicchetti (Ed.), *The emergence of a discipline: Rochester Symposium on Developmental Psychopathology Vol. 1* (pp. 261-294). Hillsdale, NJ: Erlbaum.

Masten, A. S. (2003). Commentary: Developmental psychopathology as a unifying context for mental health and education models, research, and practice in schools. *School Psychology Review, 32,* 169-173.

Masten, A. S., Best, K. M, & Garmezy, N. (1991). Resilience and development: Contributions from the study of children who overcame adversity. *Development & Psychopathology, 2,* 425-444.

Masten, A. S., & Curtis, W. J. (2000). Integrating competence and psychopathology: Pathways toward a comprehensive science of adaption in development. *Development & Psychopathology, 12,* 529-550.

Masten, A. S., & Garmezy, N. (1986). Risk, vulnerability, and protective factors in developmental psychopathology. In B. B. Lahey, & A. E. Kazdin (Eds.), *Advances in clinical child psychology, vol. 8.* (pp. 74-88). New York, NY: Plenium Press.

Masten, A. S., Hubbard, J. J., Gest, S. D., Tellegen, A., Garmezy, N., & Ramirez, M. (1999). Competence in the context of adversity: Pathways to resilience and maladaptation from childhood to late adolescence. *Development and Psychopathology, 11,* 143-169.

Masten, A. S., & Reed, M. J. (2002). Resilience in development. In C. R. Snyder, & S. J. Lopez (Eds.), *Handbook of positive psychology.* Oxford, England: Oxford University Press.

McConaughy, S. H., & Achenbach, T. M. (1990). Contributions of developmental psychopathology to school services. In T. B. Gutkin, & C. R. Reynolds (Eds.). *The handbook of school psychology,* 2nd Ed. (pp. 244-268). Oxford, England: John Wiley & Sons.

McConaughy, S. H., & Ritter, D. R. (2002). Best practices in multidimensional assessment of emotional or behavioral disorders. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp.1303-1320). Bethesda, MD: National Association of School Psychologists.

McLeod, J. D., & Shanahan, M. J. (1996). Trajectories of poverty and children's mental health. *Journal of Health & Social Behavior, 37,* 207-220.

McWhirter, J. J., McWhirter, B. T., McWhirter, A. M., & McWhirter, E. H. (1998). *At-risk youth: A comprehensive response.* Pacific Grove, CA: Brooks/Cole Publishing Company.

Mendez, J. L., McDermott, P., & Fantuzzo, J. (2002). Identifying and promoting social competence with African American preschool children: Developmental and contextual considerations. *Psychology in the Schools, 39,* 111-123.

Merrell, K. W. (1994). *Preschool and Kindergarten Behavior Scales.* Austin, TX: PRO-ED.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement, 3rd Ed.* (pp. 13-103). New York, NY: Macmillan.

Messick, S. (1995). Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50,* 741-749.
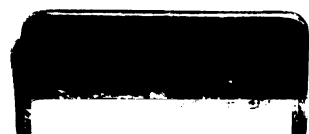
Moffitt, T. E. (1990). Juvenile delinquency and attention deficit disorder: Boys' developmental trajectories from age 3 to age 15. *Child Development, 61,* 893-910.

Mrazek, P. G., & Haggerty, R. J. (Eds.) (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research.* Washington DC: National Academy Press.

Naglieri, J. A., LeBuffe, P. A., & Pfeiffer, S. I. (1995). *The Devereux Scales of Mental Disorders.* San Antonio, TX: Psychological Corporation.

National Advisory Mental Health Council's Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment (2001). *Blueprint for change: Research on child and adolescent mental health.* Rockville, MD: National Institute of Mental Health.

National Institutes of Health, & Mental Health (1998). *Bridging science and service: A report by the National Advisory Mental Health Council's Clinical Treatment and Services Research Workgroup.* Washington DC: National Institute of Mental Health.

Niesworth, J. T., Bagnato, S. J., Salvio, J., & Hunt, F. M. (1999). *TABS screener for the Temperament and Atypical Behavior Scale: Early childhood indicators of developmental dysfunction.* Baltimore, MD: Brookes Publishing Co.

Patterson, C. J., Kupersmidt, J. B., & Vaden, N. A. (1990). Income level, gender, ethnicity, and household income composition as predictors of children's school-based competence. *Child Development, 61,* 485-494.

Pianta, R. C., & Caldwell, C. B. (1990). Stability of externalizing symptoms from kindergarten to first grade and factors related to instability. *Development and Psychopathology, 2,* 247-258.

Piotrkowski, C. S., Collins, R. C., Knitzer, J., & Robinson, R. (1994). Strengthening mental health services in Head Start. *American Psychologist, 49,* 133-139.

President's New Freedom Commission on Mental Health. (2003). *Achieving the promise: Transforming mental health care in America.* Washington, DC: US Department of Health and Human Services.

Printz, P., Borg, A., & Demaree, M. A. (2003). *A look at social, emotional, and behavioral screening tools for Head Start and Early Head Start.* Washington, DC: Center for Children and Families, Education Development Center, Inc.

Radke-Yarrow, M., & Brown, E. (1993). Resilience and vulnerability in children of multiple- risk families. *Development and Psychopathology, 5,* 581-592.

Ringeisen, H., Henderson, K., & Hoagwood, K. (2003). Context matters: Schools and the "research to practice gap" in children's mental health. *School Psychology Review, 32,* 153-168.

Roeser, R. W., & Eccles, J. S. (2000). Schooling and mental health. In A. Sameroff, M. Lewis, & S. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 135-156). New York, NY: Kluwer Academic/Plenum Publishers.

Rose, S. A., Feldman, J. F., Rose, S. L., Wallace, I. F., & McCarton, C. (1992). Behavior problems at 3 and 6 years: Prevalence and continuity in full-terms and preterms. *Development and Psychopathology, 4,* 361-374.

Rutter, M. (1979). Protective factors in children's responses to stress and disadvantage. In M. W. Kent, & J. E. Rolf (Eds.), *Primary prevention of psychopathology, Vol. 3: Social competence in children* (pp. 49-74). Hanover, NH: University Press of New England.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry, 57,* 316-331.

Rutter, M. (1989). Pathways from childhood to adult life. *Journal of Child Psychology & Psychiatry, 30,* 23-51.

Rutter, M. (1990). Isle of Wight revisited: Twenty-five years of child psychiatric epidemiology. *Annual Progress in Child Psychiatry & Child Development,* 131-179.

Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: Concepts and challenges. *Development & Psychopathology, 12,* 265-296.

Sameroff, A. J. (2000). Developmental systems and psychopathology. *Development and Psychopathology, 12,* 297-312.

Serna, L., Nielsen, E., Mattern, N., & Forness, S. R. (2002). Use of different measures to identify preschoolers at-risk for emotional or behavioral disorders: Impact on gender and ethnicity. *Education and Treatment of Children, 25,* 415-437.

Shaw, D. S, Keenan, K., & Vondra, J. I. (1994). Developmental precursors of externalizing behavior: Ages 1 to 3. *Developmental Psychology, 30,* 1-10.

Shaw, D. S., Vondra, J. I., Dowdell-Hommerding, K., Keenan, K., & Dunn, M. (1994). Chronic family adversity and early child behavior problems: A longitudinal study of low income families. *Journal of Child Psychology & Psychiatry, 35,* 1109-1122.

Shirk, S., Talmi, A., & Olds, D. (2000). A developmental psychopathology perspective on child and adolescent treatment policy. *Development and Psychopathology, 12,* 835-855.

Steinberg, Z., & Knitzer, J. (1992). Classrooms for emotionally and behaviorally disturbed students: Facing the challenge. *Behavioral Disorders, 17,* 145-156.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12,* 331-350.

Squires, J. K. (2000). Identifying social/emotional and behavioral problems in infants and toddlers. *The Transdisciplinary Journal, 10,* 107-119.

Squires, J. K., Bricker, D., & Twombly, E. (2002). *Ages and Stages Questionnaire: Social Emotional.* Baltimore, MD: Brookes Publishing Co.

Sue, S. (1999). Science, ethnicity, and bias: Where have we gone wrong? *American Psychologist, 54,* 1070-1077.

Takanishi, R., & DeLeon, P. H. (1994). A Head Start for the 21st century. *American Psychologist, 49,* 120-122.

United States Department of Health and Human Services (1996). *Head Start mental health-related performance standards.* Washington, DC: Administration of Children, Youth, and Families/Head Start Bureau.

United States Department of Health and Human Services (1999). *Mental health: A report of the Surgeon General.* Washington, DC: U.S. Government Printing Office.

United States Department of Health and Human Services (2001). *Head Start performance standards and other regulations.* Washington, DC: Administration of Children, Youth, and Families/Head Start Bureau.

United States Department of Health and Human Services (2004). *Head Start program fact sheet.* Washington, DC: Administration of Children, Youth, and Families.

Walker, H. M., Severson., H. H., & Feil, E. G. (1995). *Early Screening Project.* Longmont, CO: Sopris West Educational Services.

Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology, 66,* 715-730.

Webster-Stratton, C., & Hammond, M. (1998). Conduct problems and level of social competence in Head Start children: Prevalence, pervasiveness, and associated risk factors. *Clinical Child & Family Psychology Review, 1,* 101-124.

Werner, E. E. (1993). Risk, resilience, and recovery: Perspectives from the Kauai Longitudinal Study. *Development and Psychopathology, 5,* 503-515.

Werner, E. E., & Smith, R. S. (1982). *Vulnerable but invincible: A longitudinal study of resilient children and youth.* New York, NY: McGraw-Hill Book Company.

Werner, E. E., & Smith, R. S. (1992). *Overcoming the odds: High risk children from birth to adulthood.* Ithaca, NY: Cornell University Press.

Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family support and education on chronic delinquency and its risks. *Psychological Bulletin, 115,* 28-54.

Yoshikawa, H., & Knitzer, J. (1997). *Lessons from the field: Head Start mental health strategies to meet changing needs.* New York, NY: National Center for Children in Poverty, Columbia School of Public Health.