University of Nebraska at Omaha
DigitalCommons@UNO

Special Education and Communication Disorders Faculty Publications

Department of Special Education and Communication Disorders

2009

Convergent Validity and Test-Retest Reliability of the Preschool Behavioral and Emotional Behavior Rating Scale: Parents as Respondents

Philip D. Nordness
University of Nebraska at Omaha, pnordness@unomaha.edu

Michael H. Epstein
University of Nebraska-Lincoln, mepstein1@unl.edu

Lori Synhorst
University of Nebraska-Lincoln, lsynhorst2@unl.edu

Follow this and additional works at: https://digitalcommons.unomaha.edu/spedfacpub

Part of the Educational Assessment, Evaluation, and Research Commons

Running Head: PreBERS Parents as Respondents

Convergent Validity and Test-Retest Reliability of the Preschool Behavioral and

Emotional Behavior Rating Scale: Parents as Respondents

Philip D. Nordness

University of Nebraska-Omaha

Michael H. Epstein and Lori Synhorst

University of Nebraska-Lincoln

Abstract

A number of professional organizations have called on the need for valid and reliable assessments that measure young children's strengths and competencies for the purpose of making decisions about teaching and learning, identifying areas of lesser strength, and for designing and evaluating interventions. The *Preschool Behavioral and Emotional Rating Scale* (PreBERS; Epstein & Synhorst, in press) is a standardized test designed to assess the emotional and behavioral strengths and competencies of children 3 to 5 years of age. Two studies investigated the PreBERS with parents as the primary respondents. The first study investigated the convergent validity of the PreBERS by comparing it to the *Caregiver-Teacher Report Form* (C-TRF; Achenbach & Rescorla, 2000). Because the PreBERS is a measure of emotional strengths and the C-TRF assesses problem behaviors, moderate to very large negative correlations (-.370 to -.775) were reported between the two measures. The second study investigated the test-retest reliability of the PreBERS over a one-month period. All of the correlations were over .787 indicating that the PreBERS is a stable measure across ratings. The results suggest that when parents are the primary respondents, the PreBERS is a valid and reliable measure for assessing emotional and behavioral strengths in preschool children.

Convergent Validity and Test-Retest Reliability of the Preschool Behavioral and

Emotional Behavioral Rating Scale: Parents as Respondents

The 2004 amendments to the Individuals with Disabilities Education Act (IDEA)

states that one way of enhancing the education of children with disabilities is by

strengthening the role and participation of parents and families in the assessment process

(Turnbull, Huerta, & Stowe, 2006). In fact, the current philosophy on best practices in

assessment indicates that questions about individual child functioning are best answered

through a process that allows for family input (Hall, 2005). Furthermore, IDEA requires

that parents and families be a part of any evaluation process that concerns their child and

that they are included as a member of the decision making team. Given that parents spend

the more time with their children than professionals, it stands to reason that they are more

knowledgeable about their child's functioning and behavior. Thus, parents are an

excellent resource for providing useful information about the child's development,

preferences, and social and emotional functioning (Lynch, 2005).

One approach to including parents in the assessment process is through the use of

rating scales. Rating scales serve a number of functions in the assessment process

including the gathering of baseline data and as a progress monitoring device for

measuring student outcomes during and after treatment. In addition, rating scales allow

multiple informants (e.g., parents, teachers) the opportunity to contribute information and

provide evaluation teams the chance to compare ratings across informants to identify

similarities and differences that may be relevant to identification and intervention

planning (Mooney, Epstein, Ryser, & Pierce, 2005). In response to the clear and

compelling recognized value of parent input in the assessment process, parent rating

scales have become increasingly recognized as an important resource in assessing the behavioral, social, and emotional strengths of young children (Merrill, 1999).

Historically, early childhood assessment has been conducted to identify deficits in young children and then sort them into different categories of disability or pathology (Meisels & Atkins-Burnett, 2000). While the assessment of deficits or challenges is often critical in the identification of children who may require special assistance or services, there is a potential problem that an over emphasis on deficits may ignore specific strengths, competencies, skills, and resources that a young child may possess. In fact, the particular set of skills and behaviors that a child possesses may be more important to recognize than the amount of deficits or absence of ability (Meisels, 1994). Recognizing and identifying the strengths and competencies within an individual is essential in designing, implementing, and evaluating interventions for use with young children.

The increased recognition of the value of strength-based assessment has led to a number of key recommendations from professional organizations. Specifically, in 2003 the Joint Position Statement of the National Association for the Education of Young Children and the National Association of Early Childhood Specialists in the State Departments of Education offered two key recommendations. First, all early childhood programs should include parents in the assessment process. Second, valid and reliable assessments that measure young children's strengths and competencies should be developed for the purpose of making sound decisions about teaching and learning, identifying areas of concern, and for designing educational and developmental interventions (National Association for the Education of Young Children, 2003). In addition, the Zero to Three Working Group on Developmental Assessment recommended

that the assessment process for infants and toddlers identify the child's current

competencies and strengths as well as the competencies that will contribute to continued

developmental growth (Greenspan & Meisels, 1996). The increased emphasis on

strength-based assessment has led to the need for effective methods of identifying the

assets of children who are at-risk for learning and behavior problems in early childhood.

While many early childhood prevention and intervention programs were originally

designed to build on child and family strengths, they have frequently lacked

psychometrically sound instruments to identify and assess strengths in young children

(Raver & Zigler, 1997).

The *Preschool Behavioral and Emotional Rating Scale* (PreBERS; Epstein &

Synhorst, in press) is a standardized test designed to assess the emotional and behavioral

strengths and competencies of children 3 to 5 years of age. The test has 42-items and is to

be completed by preschool teachers, parents, and other adults who are familiar with the

child. The subscales of the PreBERS were designed to assess four domains of emotional

and behavioral strengths in preschool children. The Emotional Regulation subscale

consists of 13 items that assess a child's ability to interact with others in social situations

(e.g., "shares with others", "accepts responsibility for own behavior"). The School

Readiness subscale consists of 13 items that assess a child's behaviors that are associated

with being ready to learn in school (e.g., "pays attention to task", "persists with tasks

until completed"). The Social Confidence subscale includes 9 items that assess a child's

ability to understand his or her own functioning and relationships with others (e.g., "asks

for help", "makes friends"). The Family Involvement subscale consists of 7 items that

measure a child's relationship with their family (e.g., "interacts positively with siblings",

"maintains positive relations with family"). In these areas, adults who are familiar with the child judge a series of statements on a Likert type scale (0=not at all like the child; 1=not much like the child; 2=like the child; 3=very much like the child). A total raw score is determined for each subscale by adding the individual item scores. For each subscale, raw scores are converted to standard scores with a mean of 10 and a standard deviation of 3. An overall strength score is calculated by summing the standard scores of the four subscales and converting that sum into a total strength standard score using the table provided in the manual.

The content validity of the PreBERS was determined through a detailed development process (Epstein & Synhorst, in press). First, a list of the original 52 items from the original *Behavioral and Emotional Rating Scale* (Epstein & Sharma, 1998; Epstein, 2004), a strength based measure for students 6 to 18 years of age, was reviewed by approximately 150 preschool professionals who were asked to judge the appropriateness of each item for use with children three to five years of age. Based on the feedback from these individuals, 21 items were eliminated for inclusion in the PreBERS. Second, a comprehensive literature review of the research on social-emotional development, childhood psychopathology, and risk and protective factors of children was conducted. Several excellent sources provided relevant and detailed content (e.g., DelCarmen-Wiggins & Carter, 2004; National Research Council and Institute of Medicine, 2000; Zero to Three, 2005). In addition, tests that were previously developed to assess the emotional and behavioral problems and competencies of preschool children were reviewed. For example, the items of the *Behavioral Assessment System for Children* (Reynolds & Kamphus, 2004) and the *Caregiver-Teacher Report Form* (C-TRF,

Achenbach & Rescorla, 2000) were reviewed for content, format, and wording. This review led to the addition of 39 items to the PreBERS prototype. Then, a study was conducted to determine if each individual item differentiated between children with and without disabilities. Eight items did not discriminate the two groups and those items were deleted from the prototype. Next, data on 239 preschool children who were developing typically were collected from teachers who judged each student on the 62-item PreBERS. The data were factor analyzed and 5 items were removed because they were either redundant to other items or failed to meaningfully add to a factor (Epstein & Synhorst, in press). Finally, data were collected on a nationally representative group of preschool children (N=1,308) and used to standardize and norm the PreBERS (Epstein, Synhorst, Cress, & Allen, in press). An exploratory factor analysis with a Promax solution identified four factors. This analysis led to an additional 15 items being removed based on redundancy, overlap with other items or failure to meaningfully contribute to a factor. Thus, the comprehensive content validation process resulted in a 42 item, 4-factor test.

Reliability and Validity

A psychometrically sound instrument should be able to demonstrate scores to support their validity across instruments and reliability over time in a consistent manner. Validity refers to the appropriateness, meaningfulness and usefulness of what can be said based on the scores obtained from a test (Mertens, 1998). Content validity measures the extent to which a test's items represent the construct to be measured (Salvia & Ysseldyke, 2007). The content validity of the PreBERS was established during the multi-step item development and norming processes described above. Discriminant validity examines the extent to which an instrument can differentiate between constructs. The

discriminant validity of the PreBERS was also established during the norming process when the PreBERS scores of children with and without disabilities were compared (Epstein & Synhorst, in press). As noted earlier, those findings identified 8 items that did not differentiate the two samples and they were deleted from the original prototype of the PreBERS. Another type of validity is convergent validity. Convergent validity refers to the relationship between measures of the same construct using different assessment measures (Crocker & Algina, 1986; Salvia & Ysseldyke, 2007). The greater the relationship between the two tests, the stronger the convergent validity. While the convergent validity of the PreBERS and the *Caregiver-Teacher Report Form* (C-TRF, Achenbach & Rescorla, 2000) has been assessed with teachers (Epstein & Synhorst, in press), it has not been investigated with both parents serving as the primary respondents. Therefore, one purpose of the present study was to assess the convergent validity of the PreBERS and the C-TRF with parents as the primary responders on both tools.

The concept of reliability refers to "the extent to which individual differences in test scores are attributable to 'true' differences in the characteristics under consideration and the extent to which they are attributable to chance error" (Anastasi and Urbina, 1977, p. 84). The reliability of the PreBERS was established in several earlier studies (Epstein & Synhorst, in press). Specifically, the internal consistency of the PreBERS was assessed using the norming data where the Cronbach coefficient alphas for each of the subscales and total strength scores were calculated separately by age (3, 4, and 5 year olds) and gender (male and female). The coefficient alphas were found to be highly acceptable (range .838 - .982). In addition to achieving acceptable internal reliability, an instrument must demonstrate other forms of reliability. Inter-rater and short-term test-retest

reliability (i.e., less than 1 month) were demonstrated in other studies (Epstein &

Synhorst, in press). In those studies, teachers or paraprofessionals completed the

PreBERS on preschool children. The correlations were over .80 demonstrating the test is

stable over a short time and is consistent between raters. However, the reliability of the

PreBERS using parents as responders has not been researched. Thus, the second purpose

of the present study was to assess the reliability of the PreBERS using parents as the

primary raters. Prior to the research, we received the approval of the Institutional Review

Board at the University of Nebraska-Lincoln. IRB approved informed consent and

confidentiality procedures were followed in these studies.

<center>Methods</center>

Participants

Ninety-six preschool children and their parents participated in this study. The

participants were recruited from preschool classrooms located in a medium-sized city in

the Midwest. The children were recruited from preschool classes in eleven public schools

and two non-profit organizations serving working parents. Of the 96 preschoolers, 61

(63.5%) were male and 35 (36.5%) were female, and ranged in age between 3 and 6 years

of age (M=4.71, SD=.599). The ethnicity of the population was split between Caucasian

(85.4%), Hispanic/Latin American (1%), African American (3.1%), Asian (1%), Native

American (1%), multiracial (6.2%), and other (2.3%). All of the participants spoke

English as their primary language.

Measure

The convergent validity of the PreBERS was assessed by correlating the

PreBERS with the *Caregiver-Teacher Report Form* (C-TRF, Achenbach & Rescorla,

2000). The C-TRF is a 100-item measure, and is completed by teachers or caregivers familiar with the child. The C-TRF is used to assess problem behavior and yields a total problem behavior score and internalizing and externalizing broadband behavior scores. For this study, the Internalizing, Externalizing, and Total Problems scores of the C-TRF were used for analysis. The C-TRF was chosen because of its extensive use as a measure of behavior in preschool children and its sound psychometrics. The C-TRF has well-established levels of validity and reliability (Achenbach & Rescorla, 2000). Because the PreBERS is a measurement of emotional strengths and the C-TRF assesses problem behaviors, we would expect a negative correlation between the subscales of the instruments. Therefore, we would expect a child who scores high on the PreBERS for Emotional Regulation would likely score low on the C-TRF for Externalizing behavior.

As stated earlier, the PreBERS is a 42-item rating scale that assesses four areas of emotional and behavioral strengths in preschool children. A parent or caregiver familiar with the child rates each item on a Likert type scale (0=not much like the child to 3=very much like the child). Standard scores are calculated for each subscale (M=10, SD=3) and total strength score (M=100, SD=15).

Procedure

Approximately 50 preschool teachers attending a workshop on assessment of children in early childhood special education programs were asked to participate. The authors of the PreBERS conducted the workshop and solicited the teachers. From this group, 13 preschool teachers agreed to assist and signed IRB approved consent forms. Then, information packets were sent home to parents (N=132) of all of the preschool children in the classes. The packet included an informational letter describing the purpose

of the study, a consent form and a stamped, self-addressed envelope to be returned to the

researchers. Ninety-six parents (73% response rate) signed and returned consent forms

for their child to participate. Next, the parents were asked to complete the PreBERS and

the *Caregiver-Teacher Report Form* (Achenbach & Rescorla, 2000), according to the

instructions that appear in each test. First, they were mailed a packet containing

instructions on how to complete the ratings and a copy of each instrument. Each parent

was given two weeks to complete the ratings. Parents who did not return the tests within

that time period were given a reminder telephone call. Parent packets were returned on all

96 children. Second, two weeks after completing the first PreBERS form, parents were

mailed a second packet containing instructions and a PreBERS form to be completed.

Again, parents were given two weeks to complete the form and those who did not

received a reminder phone call. Packets were returned on 96 children with a mean

elapsed time between ratings of 37 days.

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

Convergent Validity

Pearson correlation coefficients for each of the four subscales and Total Strength

score of the PreBERS and the Internalizing, Externaling and Total Problem Scores of the

C-TRF are presented in Table 1. Correlations across each of the scores were statistically

significant (p < .0001). The correlations ranged between -.370 and -.775. According to

Hopkins (2002), correlation coefficients between .30 and .50 are considered moderate,

between .50 and .70 are large and those between .70 and .90 are very large. All of the

correlation coefficients in this study ranged between ranged between moderate and very

large with a median correlation of -.616.

For a correlation coefficient to be cited as evidence of validity it should meet or exceed .35 in strength (Hammill, Brown, & Bryant,1989). All of the correlations across the subscales exceeded this standard. As mentioned earlier, because the PreBERS is a measurement of emotional strengths and the C-TRF assesses problem behaviors, all of the correlations demonstrated negative relationships. It should be noted that the correlations fell within a moderate to large range of scores (-.370 to -.775). This finding was not surprising as externalizing and internalizing symptoms in young children are often less distinct from each other, and may often tend to occur together (Merrell, 1999). In fact, Achenbach and Rescorla (2000) noted that in preschool samples, children who scored high in externalizing behavior problems were likely to score high on internalizing problems as well.

Test Retest Reliability

The means and standard deviations for the reliability coefficients are reported in Table 2. Pearson correlations are presented for each of the four subscales of the PreBERS as well as the Total Strength Score. Correlations across each of the scores were statistically significant (p < .0001) and ranged from a low of .787 for the Social Confidence to .920 for the Total Strength Score. Eighty-five percent of the variance between the two composite scores was potentially explained by the relationship.

According to Hopkins (2002), reliability coefficients between .70 and .90 are considered very large. In the present study on the reliability of the PreBERS all of the correlations were over .787. Very large test-retest reliability correlations across intervals of about one month, such as in the present study, increase the confidence with which practitioners may generalize the findings of one administration to another. The results

from the present study suggest that the PreBERS is a relatively stable measure in the

short term for assessing the strengths of young children when parents are the primary

respondents. Assessments like the PreBERS, which may be used for screening,

identification and planning purposes, must be able to demonstrate stability over time if

those results are to assist in making meaningful decisions on how to enhance the

strengths of young children (Salvia & Ysseldyke, 2007).

Discussion

The findings from these two studies into the convergent validity and short-term

reliability of the *Preschool Behavioral and Emotional Rating Scale* (Epstein & Synhorst,

in press) with parents serving as the primary respondents provides further evidence of the

psychometric status of the instrument for assessing the strengths of young children.

Given the increasing emphasis on including parents in the assessment process, the results

suggest that when parents are the primary respondents, the PreBERS is a valid and

reliable measure for assessing the strengths of young children. In regards to convergent

validity, all of the correlations with the C-TRF were statistically significant and exceeded

the .35 standard to support its validity in assessing what it purports to measure. In regards

to reliability, previous research on the PreBERS has supported its internal consistency,

inter-rater reliability and short-term test-retest reliability with teachers serving as the

primary raters. In this study, the short-term reliability of the PreBERS was assessed with

parents serving as the primary raters. All of the correlations were very large and

demonstrated significant stability over a period of one month. This suggests that parents

can accurately identify the emotional and behavioral strengths of preschool children and

that these strengths are consistent from one administration to another.

An increasing number of professional organizations and public policy initiatives have acknowledged the importance of strength-based assessment by calling for tests that assess the competencies and strengths in young children. Strength-based assessment is an integral part of the assessment process because it identifies the strengths and competencies that can serve as the foundation for future growth and development and lessen the influence of negative life experiences for the child (Rutter, 1985). Therefore, the results from a strength-based instrument such as the PreBERS can be useful for designing, implementing, and evaluating positive behavioral interventions and supports for young children. In addition, the information derived from a strength-based assessment can be used to assist in identifying goals and objectives for individualized child and family services plan (IFSP) and/or an individual educational program (IEP) for the child.

In lieu of these applications, strength-based assessment is an ideal approach for assessing young children who might be considered at-risk for emotional and behavioral deficits. Factors that can put children at-risk for emotional and behavioral deficits can include biological variables (e.g., attention deficit/hyperactivity disorder (ADHD), developmental delay) and environmental variables (e.g., poverty, abuse) that influence child development. Therefore, strength-based assessment and the PreBERS would be useful for programs that serve children who may evince one or more of these risk factors such as early childhood special education programs and Head Start programs.

Limitations

There are several limitations to the findings that should be noted. First, characteristics of the children sampled were not representative of the larger U.S. preschool population. Specifically, young children from racial and ethnic minorities were

underrepresented. Second, these studies were conducted in a mid-sized, mid-western

community. Therefore, it is possible that the findings will not generalize to parents and

children in other geographical regions. Third, the convergent validity of the PreBERS

was assessed with only a single instrument, namely the C-TRF. The degree of validity

may vary with other social adjustment measures for young children. Fourth, the study

used a convenience sample of parents who agreed to participate in this study and

therefore the nature of the sample may have influenced the results. For this reason, the

data do not represent the input of the parents who choose not to participate. Finally, while

previous research has established the ability of the PreBERS to discriminate between

children with and without disabilities (Epstein & Synhorst, in press), the sample in this

study included a broad range of children that did not include specific cases of children

with disabilities.

Future Research

Although a number of recent studies (e.g., Epstein & Synhorst, in press) have

examined the psychometric characteristics of the PreBERS, additional research is needed.

First, researchers should continue to examine the convergent validity of the PreBERS

with other behavioral assessments appropriate for primary age children and with parents

as the primary raters. For example, the convergent validity of the PreBERS chould be

examined with measures such as the *Preschool and Kindergarten Behavior Scale*

(Merrell, 2002) and the *Minnesota Child Development Inventory* (Ireton, 1992). In

addition, individual subscales of the PreBERS could be assessed with other specific

measures. For example, the School Readiness subscale could be contrasted with the

Dynamic Indicators of Basic Early Literacy Skills (Good & Kaminski, 2000), a test that

measures emergent literacy skills. Second, while a number of studies have examined the convergent validity of the PreBERS, research is needed on the predictive validity of the measure. For instance, a child's score on the School Readiness subscale could be compared to their actual school performance at the conclusion of one or more years of schooling to examine the predictive validity of the PreBERS. Third, additional research is needed on the long-term reliability of the PreBERS with parents as the primary respondent to examine how consistent the measure is over a long period of time. Finally, researchers should continue to examine the psychometric characteristics with larger and more diverse samples, including children from different cultural, racial, and economic backgrounds as well as children with disabilities.

Clinical Implications

Measuring the emotional and behavioral strengths of young children is an important part of the assessment process of gathering information about an individual in order to make important decisions. <u>With the support of additional research, the PreBERS could potentially serve a number of uses for teachers, parents, and practitioners.</u> First, the PreBERS <u>could potentially</u> be used to identify children with limited strengths who might be at-risk for school failure. As mentioned earlier, this would include children with a variety of biological and environmental risk factors, but also children who display a persistent pattern of externalizing or internalizing behaviors that impede their ability to maintain developmentally appropriate social relationships, or their ability to participate in organized activities in a socially appropriate manner. Second, the PreBERS can be completed independently by parents and teachers to compare and contrast a child's emotional and behavioral strengths across home and school settings. This could help

facilitate increased collaboration in the assessment process between parents and teachers

and lead to interventions that enhance social and emotional skills across settings. Third,

the PreBERS could be an excellent tool for developing strength-based goals and

interventions. For example, a preschooler who scores high on the Family Involvement

subscale, but low on the School Readiness subscale may benefit from an intervention that

incorporates the family as part of the treatment plan for improving school performance.

Such an intervention may include improving communication between the school and

parents with a "talk-book" that goes home with the child every night and details the

events of their day, or suggesting books or other school readiness materials that may

assist the parent and child in their pre-school and kindergarten preparation. Ideally, by

maximizing the family involvement strength, the staff may be able to enhance the child's

school readiness. Finally, the PreBERS could be used as a pre and post test measure to

determine the effectiveness of an intervention on an individual child or group of students.

For example, to determine the effectiveness of a new schoolwide positive behavioral

support program, a principal could have parents and teachers complete the PreBERS on

all of the students at the beginning of the school year and at the completion of the school

year to determine the effectiveness of the program as a whole and on an individual basis.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont Department of Psychiatry.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle, NJ: Prentice Hall.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.

DelCarmen-Wiggins, R., & Carter, A. (Eds.). (2004). *Handbook of infant, toddler, and preschool mental health assessment*. Oxford: Oxford University Press.

Epstein, M. H., & Sharma, J. (1998). *Behavioral and Emotional Rating Scale: A Strength-Based Approach to Assessment*. Austin, TX: PRO-ED.

Epstein, M. H., & Synhorst, L. (in press). *Preschool Behavioral and Emotional Rating Scale*. Austin, TX: PRO-ED.

Epstein, M. H., Synhorst, L., Cress, C. J, & Allen, E. (in press) Development and standardization of a test to measure the emotional and behavioral strengths of preschool children. *Journal of Emotional and Behavioral Disorders*.

Good, R., & Kaminski, R. (2000). *Dynamic Indicators of Basic Early Literacy Skills*. Loveland, CO: Sopris West.

Greenspan, S. I., & Meisels, S. J. (1996). Toward a new vision for the developmental assessment of infants and young children. In S. J. Meisels & E. Fenichel (Eds.), *New visions for the developmental assessment of infants and young children* (pp. 231-66). Washington, DC: ZERO TO THREE.

Hall, L. J. (2005). Early childhood assessment. McLoughlin, J. A. & Lewis, R. B., In

Assessing students with special needs (6th ed.) (pp. 548-578). Upper Saddle River,

    NJ: Pearson Education, Inc.

Hammill, D. D., Brown, L., & Bryant, R. R. (1989). *A consumer's guide to tests in print*.

    Austin, TX: PRO-ED.

Hopkins, W. G. (2002). A scale of magnitudes for effect statistics. In *A new view of*

    statistics. Retrieved June 13, 2002, from

    http://www.sportsci.org/resource/stats/effectmag.html

Ireton, H. (1992). *The Child Development Inventory Manual*. Minneapolis, MN:

    Behavior Science Systems.

Lynch, E. W. (2005). Parent and family involvement. McLoughlin, J. A. & Lewis, R. B.,

    In *Assessing students with special needs (6th ed.)* (pp. 530-547). Upper Saddle

    River, NJ: Pearson Education, Inc.

Meisels, S. J. (1994). Designing meaningful measurements for early childhood. In B. L.

    Mallory & R. S. New (Eds.), *Diversity in early childhood education: A call for*

    more inclusive theory, practice, and policy (pp. 205-25). New York: Teachers

    College Press.

Meisels, S. J., & Atkins-Burnett, S. (2000). The elements of early childhood assessment.

    In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood*

    intervention (pp. 231-257). New York: Cambridge University Press.

Merrell, K. W. (2002). *Preschool and Kindergarten Behavior Scales* (2nd edition).

    Austin, TX: PRO-ED.

Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children and*

    adolescents. Mahwah, NJ: Lawrence Erlbaum Associates.

Mertens, D. M. (1998). *Research methods in education and psychology: Integrating*

   diversity with quantitative and qualitative approaches. Thousand Oaks, CA: Sage.

Mooney, P., Epstein, M. H., Ryser, G., & Pierce, C. D. (2005). Reliability and validity of

   the Behavioral and Emotional Rating Scale-Second Edition: Parent Rating Scale.

   Children & Schools, 27, 147-155.

National Association for the Education of Young Children (2003, November). *Early*

   Childhood Curriculum, Assessment, and Program Evaluation: Building an

   effective, accountable system in programs for children birth through age 8.

   Retrieved October 24, 2007, from http://www.naeyc.org/about/positions/cape.asp

Raver, C. C., & Zigler, E. F. (1997). Social competence: An untapped dimension in

   evaluating Head Start's success. *Early Childhood Research Quarterly, 12*, 363-

   385.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal*

   of Orthopsychiatry, 57, 316-331.

Salvia, J., & Ysseldyke, J. (2007). *Assessment (10th ed.).* Boston: Houghton Mifflin.

Turnbull, R., Huerta, N., & Stowe, M., (2006). *The Individuals with Disabilities*

   Education Act as amended in 2004. Upper Saddle Creek: Pearson Education

Zero To Three (2005). *Diagnostic classification of mental health and developmental*

   disorders in infancy and early childhood (rev. ed.). Washington, DC: Zero to

   Three Press.

Table 1

Correlation Coefficients Between Subscales for the PreBERS and the Caregiver – TRF with Parents as Responders

Subscale	Internalizing	Externalizing	Total Problem
Emotional Regulation	-.617	-.775	-.738
School Readiness	-.536	-.574	-.616
Social Confidence	-.444	-.370	-.432
Family Involvement	-.548	-.668	-.672
Total Strength Score	-.615	-.689	-.705

Note. All were statistically significant at the $p < .0001$ level.

Table 2

Test-Retest Reliability of the PreBERS with Parents as Responders

	First Testing		Second Testing			Adjusted R Square (% Variance)
	M	SD	M	SD	R	
Emotional Regulation	9.00	2.356	9.09	2.280	.877	.769
School Readiness	9.13	2.558	9.21	2.731	.883	.779
Social Confidence	10.22	2.258	10.28	2.434	.787	.619
Family Involvement	9.87	2.273	9.95	2.292	.793	.629
Total Strength Score	97.18	12.769	97.38	13.236	.920	.846

Note. All of the correlations were statistically significant at the p < .0001 level.