Validity and Reliability of the Behavioral and Emotional Rating Scale (2nd Edition): Youth Rating Scale

Michael H. Epstein
University of Nebraska–Lincoln

Paul Mooney
Louisiana State University

Gail Ryser
PRO-ED

Corey D. Pierce
Southern Illinois University

Objective: This article reports findings of three studies addressing convergent validity and test-retest reliability of the Youth Rating Scale of the Behavioral and Emotional Rating Scale–Second Edition (BERS-2). Method: Pearson product–moment correlations were used in all three studies, the first two addressing convergent validity and the third addressing test-retest reliability. Results: Analysis indicated that (a) the six BERS-2 subscales and overall strength index were generally highly positively correlated with the social skills composite score from the Social Skills Rating System–Student Form (Secondary Level, Grades 7 to 12), (b) the BERS-2 subscales and strength index were generally moderately negatively correlated with the Problem scales of Achenbach’s Youth Self-Report, and (c) test-retest reliability coefficients over a 1-week period were all above .80. Conclusions: Results provide evidence that the BERS-2 Youth Rating Scale has acceptable psychometric properties and may be considered for use by social work practitioners in assessment and intervention activities.

Keywords: strength-based assessment; emotional and behavioral competence

Strength-based assessment has been defined as the measurement of emotional and behavioral skills and characteristics that “create a sense of personal accomplishment; contribute to satisfying relationships with family members, peers, and adults; enhance one’s ability to deal with adversity and stress; and promote one’s personal, social, and academic development” (Epstein & Sharma, 1998, p. 3). Strength-based assessment is founded on the following principles: (a) all children have strengths; (b) focusing on a child’s strengths rather than weaknesses may result in enhanced motivation and improved performance; (c) failure to demonstrate a skill should first be viewed as an opportunity to learn the skill rather than as a problem; and (d) a focus on strengths when developing educational, mental health, and social work treatment plans may result in greater acceptability by key players (Epstein & Sharma, 1998). The focus on strengths and competencies is in direct contrast to the more familiar and traditional deficit-oriented assessment models (Epstein, 1999).

Over the years, formal assessment instruments have been used predominantly to identify deficits or problems in an individual’s or a group’s performance (Epstein, Harniss, et al., 2002). In the fields of education, mental health, and social work, a number of instruments with sound psychometric properties have provided useful information about a child’s deficits or pathologies to practitioners. Although useful, assessment systems such as the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach, 1991a) and individual instruments such as the Revised Behavior Problem...
Checklist (Quay & Peterson, 1987) tend to limit the scope of the behavioral information gathered during an assessment by restricting the focus of those providing the data (Epstein, Harniss, et al., 2002). This restricted focus may hinder the collection of information critical to the development, implementation, and monitoring of effective service plans.

Strength-based assessment, on the other hand, allows for the collection of a broader range of important information related to an individual’s capabilities and weaknesses (Epstein, Harniss, et al., 2002; McConaughy & Ritter, 2002). This holistic approach to evaluation has contributed to strength-based assessment gaining wide acceptance from practitioners in several disciplines (e.g., Dunst, Trivette, & Deal, 1994; Lourie, Katz-Leavy, & Stoul, 1996; Nelson & Pearson, 1991). Strength-based assessment has been taking place informally for years (Van Den Berg & Grealish, 1996, 1998). Social workers, for example, have routinely engaged in “strength chats” with a child and those who know the child best (e.g., parents, teachers, peers). During these informal assessments, social workers ask all types of questions related to what a child does well and what a child or family member envisions for the child’s future. Although these informal assessments often have succeeded in identifying vital information, they are generally carried out without consistency and stability. The questions themselves as well as the delivery of these questions have varied from professional to professional and from setting to setting. To address concerns with the fidelity and technical adequacy of these informal strength-based assessments, standardized, norm-referenced instruments have been developed. The Behavioral and Emotional Rating Scale (BERS; Epstein & Sharma, 1998) is one such instrument.

**Behavioral and Emotional Rating Scale (BERS)**

The BERS is a psychometrically sound, norm-referenced, standardized instrument designed to aid in the process of strength-based assessment. The BERS, which consists of 52 Likert-type items, provides an overall strength index and five factor-analytically-derived subscales that assess important areas of functioning. The subscales include (a) Interpersonal Strength (e.g., reacts to disappointment in a calm manner), which measures ability to control emotions and behaviors in social situations; (b) Family Involvement (e.g., participates in family activities), which measures participation and relations with the family; (c) Intrapersonal Strength (e.g., demonstrates a sense of humor), which assesses the child’s perception of competence and accomplishment; (d) School Functioning (e.g., pays attention in class), which addresses competence in classroom tasks; and (e) Affective Strength (e.g., acknowledges painful feelings), which focuses on the ability to give and receive affect. Each item is rated on a scale of 0 to 3 (0 = not at all like the child; 1 = not much like the child; 2 = like the child; 3 = very much like the child). Scores are calculated for each strength dimension and are then combined to provide an overall strength index. Higher scores reflect greater perceived strengths.

**Technical Adequacy**

Investigations suggest the BERS is a psychometrically sound instrument. Technical adequacy has been addressed with respect to content validity, convergent validity, discriminant validity, interrater reliability, and short- and long-term test-retest reliability.

**Content validity.** The BERS underwent a rigorous validation process (Epstein, 1999). Parents and professionals developed an initial pool of 1,200 statements characteristic of perceived emotional and behavioral strengths that was eventually reduced to 127 objective statements. The list was then tested for its ability to discriminate between children with more emotional and behavioral strengths (i.e., children without emotional and behavioral disorders [EBD]) and those with less emotional and behavioral strengths (i.e., children with EBD). Finally, these analyses resulted in a prototype BERS scale that was normed on 2,176 students without EBD and 861 students with EBD.

**Convergent validity.** The degree to which the BERS measures the same constructs as other validated measures has been tested in three studies across different age ranges. The BERS has been evaluated on kindergartners (Trout, Ryan, La Vigne, & Epstein, 2003), primary elementary school students (Epstein, Nordness, Nelson, & Hertzog, 2002), and adolescents (Harniss, Epstein, Ryser, & Pearson, 1999). Comparison instruments have included the Walker-McConnell Scale of Social Competence and School Adjustment–Adolescent Version (WMS; Walker & McConnell, 1995), Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1990), Scale for Assessing Emotional Disturbance (SAED; Epstein & Cullinan, 1998), Social Skills Rating...
Discriminant validity. The extent to which the BERS is able to differentiate between children with differing strengths was evaluated in a study involving students with disabilities (i.e., EBD, learning disabilities [LD]) and students without disabilities (Reid, Epstein, Pastor, & Ryser, 2000). Results demonstrated that the BERS discriminated between the three groups.

Interrater reliability. Interrater reliability is an indicator of the consistency with which different individuals rate the same behavior. Nine pairs of teacher or teachers and classroom aides completed BERS on students from their classrooms. All correlations were highly positive and statistically significant.

Test-retest reliability. Test-retest reliability is an index of stability over time. High correlations between scores collected with an intervening interval between assessment periods increases the confidence with which test users may generalize the findings of one administration over time. Separate studies of short-term (i.e., 10 days) (Harniss et al., 1999) and long-term (i.e., 6 months) (Epstein, Hertzog, & Reid, 2001) stability were conducted. Results from both studies indicated that the BERS demonstrates stability over time.

Rationale for the Current Study

No one source of data can provide all the information necessary when assessing children. Recommended assessment practice often calls for multiple measurement methods along with multiple sources of information gathered across multiple settings. Rating scales are an invaluable component of multiaxial assessment. The use of rating scales in assessment not only allows multiple informants (e.g., parents, youth) to contribute information but also gives evaluation teams the chance to compare perspectives across informants for similarities and differences that may be relevant to intervention. In the fields of social work, education, and mental health, two of the more commonly used systems are the Achenbach scales (i.e., ASEBA) and the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992). Both of these systems allow for the consideration of multiple perspectives. Parents and teachers, for example, assess their child’s deficits, problems, and pathologies. Children, through self-report, also judge their own deficits, problems, and pathologies. The ASEBA and the BASC have demonstrated adequate psychometric properties. However, as mentioned, both are predominantly deficit-oriented systems. To this point there has yet to be an empirically validated, strength-based, alternative system to be included in the assessment of children and adolescents.

The overall aim of this article was to report on the technical adequacy of a standardized self-report measure, the Youth Rating Scale of the Behavioral and Emotional Rating Scale—Second Edition (BERS-2; Epstein, 2004). The BERS-2 includes three scales: Teacher Rating Scale (the original 52-item BERS), Parent Rating Scale, and Youth Rating Scale. The Parent and Youth Rating scales underwent a standardization process similar to the original BERS (see Epstein, 2004). The Parent and Youth Rating scales contain the original 52 strength items (five subscales) plus 5 additional items (one subscale) that assess career and vocational strengths. A Career Strength subscale was added because of research (e.g., Bullis & Yovanoff, 2002) documenting the importance of work and other community-connecting experiences to the overall adjustment of youth with emotional and behavioral challenges. The Parent and Youth Rating scales were standardized on a nationally representative sample of 1,015 parents and 896 youth without disabilities nationwide (Epstein, 2004). Factor analyses of these data identified six factors, that is, the original five factors of the BERS and the career/vocational factor, for both forms. In addition, the strength index and all six subscales on each form have demonstrated acceptable levels of internal consistency, with alpha coefficients reported at .95 for the strength index and from .79 to .88 for all six subscales (Epstein, 2004; Epstein, Ryser, & Pearson, 2002). The purpose of this article was to report on three studies investigating the psychometric properties of the BERS-2 Youth Rating Scale: (a) Studies 1 and 2 assessed the convergent validity of the BERS-2, and (b) Study 3 assessed the test-retest reliability of the BERS-2.

STUDY 1

In Study 1, we examined the convergent validity of the BERS-2 Youth Rating Scale with the Social Skills composite score from the Student Form (Secondary Level,
grades 7 to 12) of the SSRS (Gresham & Elliott, 1990). The SSRS is an instrument designed for screening and classification of students suspected of having social behavior problems as well as assisting in the development of appropriate interventions. When completed, the SSRS offers a social skills standard score and four subscale raw scores. SSRS items are phrased positively, meaning the higher the score, the more competent the individual. We anticipated moderate-to-high positive correlations between the SSRS social skills composite and the BERS-2 Youth Rating Scale.

METHOD

Participants

Participants were 49 students enrolled in the seventh grade of an urban middle school in a small, midwestern U.S. city (pop. 12,496). The sample consisted of 25 boys and 24 girls ranging in age from 12 years, 4 months to 13 years, 7 months (M = 12 years, 11 months). All reported their race as White. Five students (10%) self-reported disability status (e.g., mental retardation, learning disability).

Measures

The convergent validity of the BERS-2 Youth Rating scale was assessed by comparing its scores with the social skills composite score from the SSRS. The BERS-2 Youth Rating scale contains 57 items that assess six dimensions of emotional and behavioral competence, has a reading level of fifth grade, and can be completed in less than 15 minutes. The SSRS includes 39 items and measures socially acceptable behaviors that enable a person to effectively interact with others (Gresham & Elliott, 1990). The Social Skills scale is broken down into four subscales, which include Cooperation (i.e., behaviors such as helping others, sharing materials, and complying with rules and directions), Assertion (i.e., initiating behaviors such as asking others for information), Empathy (i.e., behaviors that show concern and respect for others’ feelings and viewpoints), and Self-Control (i.e., behaviors that emerge in conflict situations such as responding appropriately to teasing) (Gresham & Elliott, 1990). Students rate each item from 0 to 2 (0 = never; 1 = sometimes; 2 = very often) based on its perceived importance to their relationships with others. The SSRS has demonstrated acceptable reliability and validity properties (Gresham & Elliott, 1990).

Procedure

All seventh-grade parents (N = 180) received mailings detailing the purpose of the study and consent procedures along with a consent document they could sign and return if they were interested in having their children participate. Of the 180 parents sent mailings, 49 gave their consent for their child to be included in the study, representing a return rate of 27.2%. For those children whose parents had consented, two graduate students explained the data-gathering process with those children and sought their assent. All children whose parents had consented gave their assent to participate. Data were collected by two graduate students who administered the BERS-2 Youth Rating Scale and the SSRS. Data were collected on two occasions. First, the SSRS was administered to 49 seventh graders in the school’s cafeteria during the students’ morning homeroom period, which lasted approximately 45 minutes. One week later, the BERS-2 Youth Rating Scale was administered to the same students under identical circumstances.

Analysis

Pearson product–moment correlation coefficients were calculated between standard scores of all BERS-2 measures and the SSRS composite score as well as between raw scores for subscales of the BERS-2 and the SSRS. All Pearson correlations calculated from BERS-2 and SSRS standard scores were adjusted for restricted range and attenuation. With respect to range restriction, an adjustment was made because using the scores of a selected sample instead of an entire population has the effect of artificially lowering the correlation coefficient (Guilford & Fruchter, 1978; Kirk, 1990). The following adjustment was made: \( r_{c}(S/\overline{S}) \) divided by the square root of \( 1 - r_{c}^{2} + r_{c}^{2} \times (S_{c}^{2}/\overline{S}^{2}) \), where \( r_{c} \) is the correlation within the restricted group (i.e., the initial correlation calculated), \( S \) is the standard deviation of the variable on which the restriction occurs (i.e., 3 for all BERS-2 subscales and 15 for BERS-2 strength index or SSRS composite), and \( S_{c} \) is the standard deviation on the same variable in the unrestricted group (i.e., 3 for subscales or 15 for composite) (Guilford & Fruchter, 1978). With respect to attenuation, an adjustment for attenuation was calculated in the criterion only because imperfect
predictive-validity coefficients can make the comparison test appear less valid than it really is. The correction for attenuation in the criterion formula was the validity coefficient between the two tests divided by the square root of the validity coefficient for the criterion (Guilford & Fruchter, 1978).

**RESULTS**

Table 1 reports corrected and uncorrected correlations between the BERS-2 Youth Rating Scale and the SSRS as well as the proportion of variance potentially explained by the effects (PVE). In general, correlation coefficients for standard scores were highly positive and statistically significant \( (p < .01) \) after adjusting for restricted range and attenuation. A high correlation (.71) was reported between the composite scores of both instruments. Fifty percent of the variance was potentially explained by the relationship between the composite scores of the two measures. Corrected correlations ranged from a high of .73 between BERS-2 Interpersonal Strength and SSRS Social Skills to a low of .32 for BERS-2 Career Strength and SSRS Social Skills. Correlations for the raw score comparisons were also highly positive and statistically significant. PVE values ranged from a high of 53% between the BERS-2 Interpersonal Strength subscale and SSRS Social Skills composite to a low of 10% between BERS-2 Career Strength subscale and SSRS Social Skills composite.

**SUMMARY**

For a correlation coefficient to be accepted as support for a test’s validity, correlations must be statistically significant at an alpha level of .05 (Anastasi & Urbina, 1996). The BERS-2 Youth Rating Scale easily meets this criterion, with the Strength Index and all six subscales surpassing that mark when compared with the Social Skills composite score of the SSRS Student Form. Moreover, when judging magnitude, correlation coefficients above .50 are considered large, according to Hopkins (2002). In the current analyses, nearly 86% of the standard score correlations \( (i.e., 6 \text{ of } 7) \) were considered large. As expected, correlations were highest between similar scales. For example, the correlation between the BERS-2 Interpersonal Strength and the SSRS Social Skills composite was .73. Conversely, the lowest correlation \( (i.e., .32) \) was between BERS-2 Career Strength and SSRS Social Skills composite. Furthermore, findings were comparable for the subscale comparisons, with 19 of the 24 correlations statistically significant and at least moderately large in magnitude.

**STUDY 2**

In Study 2, we examined the convergent validity of the BERS-2 Youth Rating Scale with the pathology subscales and composite scores of the YSR (Achenbach, 1991c). The YSR is designed to assess a youth’s problems and competencies in a standardized format. In the current study, only data from the problem item scales are reported. The YSR was selected because it is a widely used, well-researched, and psychometrically sound self-report instrument for measuring psychopathology in children. The behavioral problem items of the YSR are phrased to reflect problem behaviors, which means the higher the score, the more maladjusted the child. Because the BERS-2 Youth Rating Scale items are all phrased to reflect positive behaviors, we anticipated moderate-to-high negative correlations with YSR problem items, particularly those of an externalizing nature.

**METHOD**

**Participants**

Participants were 42 students enrolled in the eighth grade of a middle school in a small, midwestern U.S. city. The sample consisted of 24 boys and 18 girls ranging in age from 12 years, 9 months to 14 years, 10 months \( (M = 13 \text{ years, 10 months}) \). All reported their race as White. Seven students \( (17\%) \) self-reported disability status \( (e.g., \text{LD}) \).

**Measures**

The convergent validity of the BERS-2 Youth Rating Scale was assessed by comparing its scores with scores from the YSR. The YSR contains 112 problem and 8 competence items on which youth, ages 11 to 18 years, rate themselves for how true each item is now or was within the past 6 months on a scale from 0 to 2 \( (0 = \text{not true at all}; 1 = \text{somewhat true}; 2 = \text{very true}) \). The YSR provides a series of problem behavior scores including
one total problem score, two dimension scores (Internalizing and Externalizing), and eight specific syndrome scores (Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Delinquent Behavior, and Aggressive Behavior) (Achenbach, 1991c). The YSR has demonstrated highly acceptable reliability and validity (Achenbach, 1991c).

Procedure

Parental consent and child assent procedures were identical to those in Study 1. Of the 160 parents of eighth graders who were sent information, 42 provided written consent for their children to participate, an overall return rate of 26.3%. All children whose parents had provided consent agreed to participate. Data were collected by two graduate students who administered the BERS-2 Youth Rating Scale and the YSR. Data were collected on two occasions. First, the YSR was administered to 42 eighth graders in the school’s cafeteria during the students’ morning homeroom period, which lasted approximately 45 minutes. One week later, the BERS-2 Youth Rating Scale was administered to the same students under identical circumstances.

Analyses

Pearson correlation coefficients were calculated on all standard score comparisons from the BERS-2 and the YSR. All correlations were adjusted for restricted range and attenuation as outlined in Study 1. The only difference was that the YSR standard deviation of 10 was used in the range restriction calculation instead of the SSRS standard deviation of 15.

RESULTS

Correlation coefficients between the BERS-2 Youth Rating Scale and the YSR as well as PVE are reported in Table 2. All BERS-2 correlations were negatively correlated across YSR problem scales, and most were statistically significant. Of the problem correlations, 84% (i.e., 65 of 77) were moderately to highly negative according to Hopkins’ (2002) criteria. Problem correlations ranged from −.05 to −.86. All correlations were adjusted for restricted range and attenuation. Sixteen percent of the variance between the composite scores of the two measures was potentially explained by the relationship. For the composite scores, PVE values ranged from a high of 17% between the BERS-2 Interpersonal Strength subscale and YSR Total Problems to a low of 8% between BERS-2 Affective Strength subscale and YSR Total Problems.

SUMMARY

Eight-three percent of the correlations between the BERS-2 subscales and YSR (i.e., 64 of 77) were statistically significant at an alpha level of .05 and therefore are evidence of the instrument’s validity, according to Anastasi and Urbina (1996). In terms of magnitude, more than one third of the 77 correlations (n = 26) surpassed .50 and are considered large (Hopkins, 2002). Another 52% (n = 40) were moderately negative. Six of the seven correlations between the BERS-2 Youth Rating Scale and the YSR Total Problems score were moderately negative. Equally important to the establishment of convergent validity, the BERS-2 subscales that assess behavior directed outward (i.e., Interpersonal Strength, Family...
Involvement, and School Functioning) demonstrated a stronger relationship with the YSR Externalizing scale than with the Internalizing scale. Conversely, the BERS-2 subscales that measure behaviors of an internal state (i.e., Intrapersonal Strength and Affective Strength) showed a stronger relationship with the YSR Internalizing scale than with the Externalizing scale. Correlations between the Career Strength subscale and the YSR Problem scales were generally moderately to highly negative and ranged from a high of −.70 with the Social Problems syndrome score to a low of −.14 to the Aggressive Behavior syndrome score.

**STUDY 3**

In Study 3, we examined the test-retest reliability of the BERS-2 Youth Rating Scale. As stated, test-retest reliability is an indicator of a measure’s stability over time and can be conducted over the short or long term. For measures such as the BERS-2, reliability coefficients must approximate or exceed .80 in magnitude for the instrument to be considered minimally reliable (i.e., for screening); coefficients must be .90 or greater to be considered most desirable (i.e., for making high stakes decisions such as identification or intervention) (Salvia & Ysseldyke, 1998).

**Participants**

Participants were 42 students enrolled in the sixth grade of a middle school in a small, midwestern U.S. city. The sample consisted of 23 boys and 19 girls ranging in age from 11 years, 2 months to 12 years, 9 months (M = 11 years, 11 months). All reported their race as White. Three students (7%) self-reported disability status (e.g., LD).

**Procedure**

Parental consent and child assent procedures were identical to those detailed earlier. Of the 160 parents of sixth graders who were sent information, 42 provided their written consent for their children to participate, a return rate of 26.3%. All children whose parents provided consent agreed to participate. Data were collected in the fall of 2001 by two graduate students who administered...
the BERS-2 Youth Rating Scale to 42 sixth graders on two occasions. Students completed the BERS-2 Youth Rating Scale in the school’s cafeteria during the students’ morning homeroom period, which lasted approximately 45 minutes. One week later, the same 42 students completed the BERS-2 Youth Rating Scale under identical circumstances.

RESULTS

Means and standard deviations for each testing for the BERS-2 Youth Rating Scale appear in Table 3. Pearson correlations across the samples for the six BERS-2 subscales and the overall strength index were high and all were statistically significant ($p < .0001$). Eighty-three percent of the variance between the two composite scores was potentially explained by the relationship.

SUMMARY

Test-retest reliability is an indication of a measure’s stability over time. High correlations between scores collected with an intervening interval between assessment periods increases the confidence with which test users may generalize the findings of one administration over time (Salvia & Ysseldyke, 1998). The BERS-2 Youth Rating Scale appears to be a stable measure. Reliability coefficients ranged from a low of .84 to a high of .91. Assessments such as the BERS-2 Youth Rating Scale, which may be used for screening, planning, and evaluation purposes and for which data will be reported individually, should have reliability coefficients of at least .80 (Salvia & Ysseldyke, 1998).

DISCUSSION AND APPLICATION TO SOCIAL WORK PRACTICE

The purpose of the three studies was to evaluate the convergent validity and test-retest reliability of the Youth Rating Scale of the BERS-2 (Epstein, 2004). Results of this research suggest that the BERS-2 Youth Rating scale is moderately to highly correlated with the SSRS and the YSR and possesses stability over time. With respect to convergent validity, 84% of all correlations (i.e., 91 of 108) support the BERS-2’s validity based on their statistical significance. More important, there was evidence that correlations were generally (a) at least moderately large in magnitude and (b) higher between subscales that purport to measure similar behavioral constructs. In terms of magnitude, 83% of the validity coefficients were at least moderately large in magnitude according to Hopkins (2002), who suggested that coefficients between 0.3 and 0.05 are moderate, coefficients between 0.5 and 0.7 are large, and coefficients between 0.7 and 0.9 are very large. In terms of similar constructs, for example, the BERS-2 subscale of Interpersonal Strength, which focuses on self-management and social interaction, correlated highly with the Social Skills scale of the SSRS. Furthermore, findings were generally favorable for the BERS-2 Career subscale. A high positive correlation was reported between the BERS-2 Career subscale and the SSRS Social Skills composite.

With respect to test-retest reliability, all correlations were above .80 and two were above .90. The test-retest reliability of the newest subscale (i.e., Career Strength) was .89. Reliability findings enhance the confidence with which practitioners can use the measure for screening decisions in which the data will be reported for individual children. Taken together, validity and reliability findings strengthen researchers’ and practitioners’ confidence in using the BERS-2 Youth Rating Scale as a measure of emotional and behavioral strength for adolescents in comprehensive (i.e., multimodal) evaluation processes, particularly in conjunction with acceptable levels of internal consistency (see Epstein, 2004).

Evidence of adequate technical properties for the BERS-2 Youth Rating Scale seems to add to the options available to school and other social work practitioners involved in the assessment of children. Assessing a youth’s emotional and behavioral strengths is an important part of evaluating the whole child. Current evaluation procedures, particularly in the area of youth with EBD, tend to focus on the assessment of deficits, problems, and pathologies and result in information geared toward fixing, changing, or remedying a situation (Epstein et al., 1999). Strength-based assessment, on the other hand, has the potential to focus decisions more positively on ways practitioners can support and foster youth strengths while simultaneously working to bolster areas where youth are not as strong.

Strength-based assessment has demonstrated broad appeal to stakeholders in social work (Epstein, Harniss, Robbins, et al., 2002) and been taking place informally for years (Van Den Berg & Grealish, 1998). Only recently with the advent of standardized procedures such as the BERS have social work practitioners been able to incorporate formal strength-based instruments into their assessment and evaluation processes. With the BERS-2 Youth Rating Scale demonstrating adequate
Means, Standard Deviations, Test-Retest Correlation, and Proportion of Variance Potentially Explained for the BERS-2 Youth Rating Scale

<table>
<thead>
<tr>
<th>Subscales</th>
<th>First Testing</th>
<th>Mean (M)</th>
<th>SD</th>
<th>Second Testing</th>
<th>Mean (M)</th>
<th>SD</th>
<th>r*</th>
<th>PVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpersonal strength</td>
<td></td>
<td>9.98</td>
<td>1.92</td>
<td>10.45</td>
<td>2.83</td>
<td>.89</td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td>Family involvement</td>
<td></td>
<td>11.12</td>
<td>2.47</td>
<td>11.17</td>
<td>2.77</td>
<td>.85</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td>Intrapersonal strength</td>
<td></td>
<td>10.95</td>
<td>2.00</td>
<td>11.10</td>
<td>2.55</td>
<td>.91</td>
<td>.83</td>
<td></td>
</tr>
<tr>
<td>School functioning</td>
<td></td>
<td>11.52</td>
<td>2.13</td>
<td>11.48</td>
<td>2.53</td>
<td>.89</td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td>Affective strength</td>
<td></td>
<td>10.02</td>
<td>2.30</td>
<td>10.43</td>
<td>2.61</td>
<td>.84</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>Career strength</td>
<td></td>
<td>10.69</td>
<td>1.98</td>
<td>10.61</td>
<td>2.14</td>
<td>.87</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>Strength index</td>
<td></td>
<td>104.69</td>
<td>11.09</td>
<td>106.05</td>
<td>14.96</td>
<td>.91</td>
<td>.83</td>
<td></td>
</tr>
</tbody>
</table>


Means are reported as standard scores. *p < .0001.

psychometric properties, social workers now have the means to examine strength-based data formally not only from youth but also across informants (e.g., teachers, parents, and youth), procedures that previously could only take place under the more traditional deficit-oriented assessment models. Self-report measures allow youth to have their own voice in the assessment process, particularly with respect to how they view their own competencies and problems. Cross-informant analysis of youth emotional and behavioral strengths is particularly important when one considers that no one single data source can substitute for all others in assessing areas of functioning (McConaughy, 1993).

School social workers might benefit the most from inclusion of the BERS-2 Youth Rating Scale and use of cross-informant analysis of emotional and behavioral strengths in their assessment and treatment processes. Generally, children are referred to social workers, counselors, or psychologists because of perceived and/or real problems. Use of instruments that measure strengths allows social workers and others the chance to facilitate and frame conversations with teachers, parents, and peers in terms of strengths rather than weaknesses. Use of strength-based instruments over time might also allow school and other social workers the chance to document progress or lack thereof in a broader range of areas (e.g., vocations) than could be accomplished through use of other available instruments. Such a practice would particularly be appropriate for secondary students with disabilities and their families as the transition team develops and evaluates postsecondary plans. In these instances, use of strength-based instruments would allow team members to document efforts to promote student self-determination, a critical component of effective transition practices (Greene & Kochhar-Bryant, 2003).

Several limitations to this research should be noted. First, in all three studies relatively small numbers of youth were represented, and these youth failed to represent national samples of adolescents. Findings may differ with a larger and more diverse sample. Moreover, given the number of comparisons in the validity studies, the small number of participants becomes problematic. To this end, replication of these studies with larger sample sizes is warranted. Second, all three studies were conducted primarily with participants without verified disabilities. Findings may differ with larger samples of students with verified disabilities (i.e., emotional disturbances, learning disabilities). Third, as the data were collected from youth in one state (Nebraska), the findings may not generalize to other regions. Finally, validity studies did not account for the order in which the instruments were administered to the youth. Specifically, in both validity studies, the BERS-2 was administered on the second test day. In future research, the order of the tests needs to be counterbalanced to minimize the order effects of testing.

Future research should address these limitations by investigating validity and reliability using larger samples of children with and without disabilities who are selected from all regions of the country. Second, other types of validity and reliability studies should be conducted, including studies of discriminant validity (i.e., EBD vs. learning disabilities vs. no disability) and longer term (i.e., 6 month) test-retest reliability. Finally, additional validation is needed for the Career Strength subscale, particularly in light of research (e.g., Bullis & Yovanoff, 2002) demonstrating the importance of engaging certain youth in work- or school-related activities as they move toward adulthood. Moreover, use of the BERS-2 Youth Rating Scale in the transition planning process merits further empirical review.

Despite the need for additional research and the limitations to current research, the studies reported here provide initial evidence in support of the reliability and validity of the BERS-2 Youth Rating Scale. Given the reported acceptable psychometric properties, the BERS-2 Youth Rating Scale is recommended for the following uses: to document children's emotional and behavioral strengths; to identify children with limited strengths; to target goals for an individual educational, treatment, or vocational plan; and to document progress in a strength area resulting from implementation of an individualized treatment plan.