Behavior Evaluation Scale, Third Edition
(BES-3)

Publisher/Date:
• Hawthorne Educational Services, Inc., 800 Gray Oak Drive, Columbia, MO 65201. Published, 2005.

Purpose:
• Behavior rating scales designed to identify serious emotional-behavioral disorders and contribute to service delivery, for ages 4-19.

Provides:
• Four formats (Home Version-Long Form, Home Version-Short Form, School Version-Long Form, and School Version-Short Form) completed by the appropriate raters assess a student’s social-emotional and behavioral functioning in a variety of dimensions which are aligned with elements included in State- and Federal definitions of Emotional-Behavioral Disability. Subscales of the measure include Learning Problems, Interpersonal Difficulties, Inappropriate Behavior, Unhappiness/Depression, and Physical Symptoms/Fears. A “Total Score” (composite) is obtained, which provides a global behavioral estimate.

Standardization Issues:
• The Home Version was normed on 4,643 children and the School Version was normed on 5,124 children. Both samples were matched to 2000 US Census data. Individual cell sizes for age and grade groupings all included about 100 students.

Reliability and Validity Issues:
• Internal-consistency reliabilities of .80s-.90s are reported by subtest and total scores on both versions. Interrater reliability was acceptable (.78). Reliability data was reported for total samples/scales and not further determined at each age/grade-level. Test-retest data is questionable, based on the low number studied (n=24). Content validity was determined by expert judgment, and factor-analysis supported the test’s construct validity. Concurrent validity data compared the BES-3 forms to other behavioral scales (Devereaux, BRP-2) and resulted in moderate-to-high correlations (.78-.94 for Total scores, and .62-.95 for individual scales). Factor-analyses conducted of the individual scales did not convincingly support the interpretation of the subscales as “distinct constructs.”

Additional Points:
• The test publisher offers a companion intervention manual and parent’s guide to suggest solutions for problems at school and home. While the
resources may offer suggested places to start and to consider, in terms of interventions, the interventions listed are not clearly indicated as evidence-based.

- There was a slight (but acceptable) overrepresentation of males in the norm sampling for the School Version.
- While the standardization sample included children with emotional-behavioral difficulties, the exact number and nature was not reported in the manual. If this subgroup was in fact underrepresented, there is a possibility of overestimating behavioral difficulties.
- Norm tables include gender-specific data.
- Use of both Home- and School-Versions (as well as multiple school-informants) provides a more ecological view of the child’s functioning. The authors clearly indicate that the scales cannot be used as sole data sources for diagnostic purposes, but should be part of a more comprehensive evaluation.
- Mean subtest scaled scores of 4-5 were typical for the children with behavior disorders, while the overall standardization group yielded scaled scores in the range of 10. However, this evidence of discriminative-validity was based on groups not matched to important demographic variables.
- Test-retest reliability was found to be lowest on the Home Version’s Unhappiness/Depression Scale (.65); below Bracken’s minimal-acceptability level of .70.
- The sensitivity of the forms is questionable for use in progress monitoring.
- The test’s simplicity and generally-adequate psychometric properties make it an appropriate screener or part of a multi-method, multisource, and ecological assessment.
- The scales only measure the presence or absence of pathology, and do not assess the presence of adaptive and pro-social behaviors, which can be important variables in designing appropriate treatment plans.