Pergamon

Journal of School Psychology, Vol. 32, No. 4, pp. 419–425, 1994

Printed in the USA
0022-4405/94 $6.00 + .00

Behavior Assessment System for Children

Jonathan Sandoval and Adriana Echandia
University of California, Davis

The Behavior Assessment System for Children (BASC) is a well-designed and useful set of measures for the assessment and identification of school-age children with emotional disturbances and behavioral disorders. The BASC consists of five measures designed to gather information about a child or adolescent from a variety of sources (teachers, parents, direct observations, students, and historical records). In combining this information into an integrated system, the BASC uniquely attempts to provide a multidimensional understanding of a child. This review summarizes the technical qualities of the BASC and critiques its usefulness for practicing school psychologists. Although there are shortcomings, the BASC is made up of some of the best measures of their kind and represents an approach of choice for identifying children with emotional and behavioral disorders in schools.

Keywords: Personality assessment; Rating scales; Internalizing problems; Externalizing problems.

The Behavior Assessment System for Children (BASC) is an integrated set consisting of a self-report, a teacher rating scale, a parent rating scale, a developmental history, and an observation protocol all designed to assess children for the differential diagnosis and educational treatment of emotional and behavior disorders. The focus is on both adaptive and maladaptive behavior, and different sources and methods are available to the clinician to converge on an estimate of functioning. The various components of the BASC (the acronym is pronounced "Basque") are the Self-Report of Personality (SRP), the Teacher Rating Scale (TRS), the Parent Rating Scale (PRS), the Structured Developmental History (SDH), and the Student Observation System (SOS). The materials consist of a manual and forms along with two optional microcomputer programs. A copy of the manual with sample forms is available for $50; the manual without forms is $45. The forms can be ordered with hand scoring keys (25 forms, $20) or without (25 forms, $12.50) for use with the computer scoring programs. Two programs will be marketed to score and produce reports for the system, the *BASC Enhanced Assist* with unlimited usage on DOS machines for $225 and the *BASC Plus Assist*, which includes an enhanced narrative report but which costs $50 for the first 50 uses and $40 for each succeeding 50 uses. The later program was not available for review. We believe one of the computer programs is a worthwhile investment for the serious user.

The self-report and the two rating scales yield scores in *T*-score units and percentiles based on a national norm group, by gender in the norm group, and in comparison with a group of seriously emotionally disturbed children. All three scales feature an F index indicating the tendency to be unusually negative in the judgments made. The

Address correspondence and reprint requests to Jonathan Sandoval, Division of Education, University of California at Davis, Davis, CA 95616-8479.

SRP additionally contains an L index (fake good) and a V index (nonsensical items) to determine the usefulness of the scores. The ratings produce various scales and composites that can be used psychometrically, in that composites can be compared to determine if they are different, and the scales can be compared to identify ipsative strengths and weaknesses.

The Structured Developmental History yields no score but consists of answers given by a primary parental caretaker to a structured interview or questionnaire covering demographic data for the parents, pregnancy and birth history, developmental history, medical history, family medical history, social history, and educational history. The interview-questionnaire was adapted from other such histories.

The Student Observation System consists of three components: a behavior checklist, a time sampling record, and a place for anecdotal observations. It is intended for observations of 15-minute periods. The behavior checklist presents 65 or more specific behaviors in 13 categories ranging from Peer Interaction to Inappropriate Sexual Behavior. Frequency and disruptiveness of the behavior during the observation session is rated. The time sampling record provides a space for recording which of four adaptive behaviors and which of nine problem behaviors is occurring during a 3-second interval every 30 seconds over the 15-minute observation session. In addition, the SOS provides space for free observations and anecdotes. This is one of the first such widely published observation systems and will enable many school psychologists to document what is often done as an informal activity. Psychologists who use this observation scale probably should have training and supervision in observation techniques and practice with this category system, since the SOS is somewhat more complicated to use than the other components of the BASC.

Both the SDH and the SOS have no norms and little specific guidance for how they are to be interpreted, previous training of the users being required to make them useful. These two parts of the BASC system require no further comment except for praise of their inclusion and warning of the general problem of recording and interpreting the very low frequency behaviors that are to be found on both measures.

TECHNICAL QUALITIES

Test Development and Theory

It is safe to assert that the BASC has been carefully developed and represents a synthesis of what is known about developmental psychopathology and personality development. The items for all of the components have been derived from a review of the relevant literature and collected clinical experience. In addition, on the basis of empirical data during standardization, scales were refined and shortened. Sophisticated data reduction methods have been used throughout test development.

Each of the clinical scales on the TRS and PRS has an established rationale and items with high content validity. The clinical scales consist of Aggression, Anxiety, Attention Problems, Atypicality (psychoticism), Conduct Problems, Depression, Hyperactivity, Learning Problems, Somatization, and Withdrawal. The Hyperactivity, Aggression, and Conduct Problems scales are combined in the Externalizing Problems composite, and the Anxiety, Depression, and Somatization scales are combined into the Internalizing Problems composites. Attention Problems and Learning Problems make up a third composite, School Problems. These three dimensions of difficulty are found in a number of studies of the affective development. Finally, the Hyperactivity, Aggression, Anxiety, Depression, Attention Problems, and Atypicality scales are combined into a single Behavioral Symptoms Index, which the authors, unfortunately, compare to the *g* on intelligence tests. (Although a number of personality measures do

have an overall adjustment scale, most personality researchers would be uncomfortable with this analogy.)

Both the TRS and PRS also contain adaptive scales, which focus on more positive aspects of behavior. The importance of these factors has also been supported by a substantial literature. The scales are Adaptability, Leadership, Social Skills, and (on the TRS alone) Study Skills. The scales on all measures are combined in an Adaptive Skills composite. It is important to consider the positive aspects of behavior, and having this additional measure of adaptive behavior is a welcome feature of BASC, particularly if it proves useful in the identification of mental retardation.

The self-report measure, the SRP, also yields clinical and adaptive scales. The clinical scales are Anxiety, Attitude to School, Attitude to Teachers, Atypicality, Depression, Locus of Control, Sensation Seeking, Sense of Inadequacy, Social Stress, and Somatization. Somatization and Sensation Seeking are found only on the adolescent version. Four of these 10 are labeled the same as scales on the PRS and TRS. The adaptive scales are Interpersonal Relations, Relations With Parents, Self-Esteem, and Self-Reliance. Five composite scores are identified, four of which are scaled: the School Maladjustment Composite (Attitude to School + Attitude to Teacher + Sensation Seeking), the Clinical Maladjustment Composite (Atypicality + Locus of Control + Somatization + Social Stress + Anxiety), the Personal Adjustment Composite (Relations With Parents + Interpersonal Relations + Self-Reliance + Self-Esteem), and the Emotional Symptoms Index (ESI: Social Stress + Anxiety + Depression + Sense of Inadequacy + the inverse of Interpersonal Relations + the inverse of Self-Esteem). The fifth composite follows from the authors' suggestion to monitor the Social Stress, Anxiety, and Depression scales. If all scales fall above 1.5 SD, this triad may be a special index of suicidal risk. The SRP protocol also yields a small number of items (16 out of 186 on the adolescent SRP), which, if selected, indicate rare and unusual ideation, which can be followed up in a clinical interview.

All of the contents of the scales were selected from research findings, other measures, and clinical experience. The items were constructed with the help of professionals (including teachers) and students, and were carefully evaluated for readability, acceptability, and comprehensibility. All items were required to contribute to only one scale, and to serve across ages and across measures on the rating scales. Prior to standardization, pilot items were subjected to careful item analysis designed to assure that they contributed to measurement and discrimination. After standardization, covariance structure analysis was used to evaluate the placement and functioning of each item. Studies were also done to ensure that the items had strong correlations with hypothetical latent variables. The correlations, published in Tables 10.4 and 10.5 of the final version of the manual, illustrate that the items perform very well. Each of the scales and composites were similarly established by contemporary latent trait analysis techniques and factor analyses. The finding of a principal factor on the scales led to the creation of the Behavioral Symptoms Index; however, the Emotional Symptoms Index was constructed by selecting from the scales loading on composites, rather than on the first factor. The items were also subjected to bias analyses to detect gender and ethnic discrepancies in performance so that the final version of the scale contained no biased items.

We were impressed with the care used to construct and verify the content and items that appear on the BASC. The authors have set the standard in test construction for this kind of scale to be used with the childhood population. Undoubtedly new constructs will be developed in this field, and some of the constructs selected for the BASC may turn out to lack utility, but the authors have produced a modern instrument for use by child psychologists.

Norming—Standardization Procedures

The BASC TRS, PRS, and SRP were normed on a sample of children from across the United States at 116 testing sites. At each site, two classrooms per grade and four randomly selected children per classroom (two boys, two girls) were targeted for teacher ratings (and two for parent ratings), while the entire class in Grades 3–12 completed the SRP. Thus 9,861 SRPs, 3,483 PRSs and 2,401 TRSs were collected. Approximately half the subjects were male, half female. A sufficient number of subjects were available at each age level to provide useful norms. It should be noted that minority children were overrepresented at the younger ages, but that scores were weighted to create norms characteristic of the U.S. population. The developers seem to believe there is little difference in the percentage distributions of the U.S. population across age groups for different race/ethnicity groups, a proposition that is probably not true for minority groups such as Latinos. Weighting was also done to make the sample match the population with respect to maternal education, geographical region, and special educational placement. The standardization sample was reasonably close to the target, so the weighting process probably is appropriate and did not result in great distortions.

A clinical norm sample was also included that consists of children with identified emotional and behavioral problems, the most prevalent of which were behavior disorder and attention-deficit hyperactivity disorder. The clinical norms were collected for the TRS on 693 children, for the PRS on 401 children, and for the SRP on 411 children from 39 U.S. and Canadian sites. This sample was not adjusted, and males and whites were overrepresented. The sample is described by gender, age, race/ethnicity, and geographic region.

The standardization of the instruments was quite good for measures of this sort. It should be noted, however, that a large number of the participating schools were private Roman Catholic institutions, or were associated with universities. To the extent that these institutions are selective, there may be a bias in the norms. In this norming, as in many others, children with a Spanish-speaking family background are lumped together as Hispanic. Both this and the category of Asian may be too heterogeneous to be of use in the future as these groups grows in size in the United States.

The scale norms were constructed by using linear *T*-score scaling. This procedure is advisable because many of the constructs being measured are probably not normally distributed, and this metric preserves underlying distribution. Users are cautioned not to interpret these *T* scores, particularly those around the mean, in the same way they would a *T* score on a measure of intellectual functioning. The percentile estimates are also adjusted to the skewness of the distribution.

The user has a choice of the general, the clinical, and (within these) gender-based norms. The manual's rationale for general rather than gender norms seems strained, and we anticipate gender norms will be more useful to clinicians in the long run. We also found curious the injunction to add 3 *T*-score points to the ratings by fathers of social skills for preschoolers and older children, although the manual provides an explanation.

Reliability

The manual reports internal consistency reliabilities, test–retest reliabilities, and interrater reliabilities for the rating scales, and the first two of these for the self-report. Internal consistency reliability is quite good, averaging .80 but increasing with the age of the child to almost .90 for adolescents. Reliabilities are similar across gender but are highest for externalizing and adaptive ratings. The internal consistency reliabilities for the clinical sample are similar to those for the general sample. The composites are

more reliable than the scales, almost all alpha coefficients falling in the low to middle .90s. The picture is similar across measures; almost all scales and all composites have high enough reliability to use in making decisions about individuals. The least reliable scales tended to be Somatization and Anxiety for the younger ages on the TRS and Atypicality and Conduct Problems on the PRS.

One-month test-retest reliabilities are also quite high, in some cases exceeding the alpha estimates, except on the SRP, where they were slightly lower. These test-retest estimates, too, are typically in the middle .80s to the middle .90s over a 1-month period. The pattern of correlations matched the pattern in the internal consistency estimates. The lowest reliabilities (low .70s) were for the adolescent level of the PRS. One study of ratings of a clinical sample of young children over 7 months yielded median correlations of .69, suggesting stability over a longer time period.

The interrater reliabilities for the TRS were reasonably high for the child version but low on some scales of the preschool version. The composite values on the child TRS ranged from .69 (Internalizing Problems) to .89 (School Problems). As would be expected from other studies, comparing ratings from parents on the PRS yielded only moderate correlations, with median values of .46, .57, and .67 for the preschool, child, and adolescent forms, respectively. The lowest values were for Internalizing Problems. Parents vary in how they see children and how they view them in settings with differing demands, so they often do disagree. The preschool reliabilities may not be high enough for the scores to be accurately profiled. The user should exercise caution in the use of the BASC with the very young.

Validity

The validity data reported on the scales consist of factorial validity, concurrent validity with other behavior measures and known clinical group scoring on the measure. As mentioned previously, the authors have used sophisticated modern techniques to construct their scales and study the structure of the relationships between scales. They started out with a model of relationships and modified it by using the data provided by the standardization sample. The confirmatory factor analysis techniques have validated the authors' conceptualization of what is being measured, but future studies on different samples will be needed to further validate the model of Internalizing, Externalizing, and Adaptive Skills.

The BASC manual reports correlational studies between the TRS and five other well known teacher report forms: Achenbach's Teacher's Report Form, Quay and Peterson's Revised Behavior Problem Checklist, the Conners Teacher Rating Scales, Burks' Behavior Rating Scales, and the Teacher Rating Scale of the Behavior Rating Profile. The correlations between the TRS and the Achenbach scales are quite high, in the .80s and .90s. The correlations are lower with the other scales, but still relatively high in parallel areas, particularly the measures of externalizing and school behavior.

The PBS is correlated fairly highly with the Child Behavior Checklist, particularly on the externalizing scales (.71–.84), is moderately correlated with the Personality Inventory for Children-Revised and with the Conners Parent Rating Scales (PBS Hyperactivity and Externalizing both correlate .56 with the Conners Hyperactivity Index). These are the three most commonly used parent ratings of children's problems.

The SRP has been correlated with the MMPI, Achenbach's Youth Self-Report, the Student Rating Scales of the Behavior Rating Profile, and the Children's Personality Questionnaire. The correlations between the SRP composites and selected MMPI scales are reasonably high for an adolescent population, with average correlations across the scale ranging from .52 to.60. The SRP composites correlate reasonably well with the Youth Self-Report, but other correlations suggest the two measures are

measuring somewhat different constructs. The same is true of the Student Rating Scales but not of the Children's Personality Questionnaire. The correlations between the SRP and the Children's Personality Questionnaire are only moderately high, suggesting that these two measures are different.

The validity of the BASC is also reported in terms of the performance of carefully diagnosed groups of children with Conduct Disorder, Behavior Disorder, Depression, Emotional Disturbance, Attention-Deficit Hyperactivity Disorder, Learning Disability, Mild Mental Retardation, and Autism. On balance, the profiles of the groups on the various BASC measures, using the general norms, make sense and support the construct validity of the scale. Interestingly, the SRP measures do not differentiate Attention Deficit, Learning-Disabled or retarded children very well, although the PRS and TRS do.

Finally, additional convergent and discriminant validity are addressed by presenting the correlation matrix for the TRS and PRS and for the SRP with both the TRS and PRS. There are many reasons why the various sources of information in the system may not agree and other methods beside correlation are needed to investigate agreement. Nevertheless the data in the manual support the use of the test. There are low to moderate correlations between the student self-reports and ratings by parents and teachers and fairly good agreement between teacher and parent ratings except at the preschool age.

UTILITY FOR SCHOOL PSYCHOLOGISTS

In general, we found the materials easy to use in the school situation. The manual is user-friendly in the way it is organized, and it employs graphics well to convey technical information.

The parent rating scales can be readily understood and completed by parents with minimal assistance from the professional administering the scale. The protocols are efficiently arranged and attractive. The rating scales are easy to score, thanks to the tear-apart protocols, although we would opt for the computer program, to save time and increase accuracy of scoring. Having the summary table and profile chart on the same page was a useful feature. The scores derived from the various components of the system are familiar and seem consistent with theory and other measures.

We were concerned about the level of reading ability and vocabulary demanded for 8- to 11-year-olds on the Self-Report of Personality–Child (SRP-C). Younger children may need some terms defined or words read to them by the examiner.

We had no difficulty using the software package, BASC Enhanced Assist. Keying in the data from the protocol was a bit awkward in that the screen did not match the form. Instead, the items are grouped in sets of five in order that the person doing data entry not get lost. Rather than copy the form, the data entry person must mentally transpose data marked as circled letters to numbers, which gets tedious and introduces an extra source of error. We were also disconcerted that it is not easy to go back and check and correct data; instead data must be reentered from the beginning.

One of the few places where the BASC does not live up to its promise is in providing a way to compare all of the data collected systematically. There is no easy way to integrate the information obtained from the rating scales with the observational protocol or with the developmental history. This shortcoming greatly limits the advantage of having an all-encompassing assessment system, since the components cannot be readily compared. As well done as the developmental history and observation scale are, in a sense they stand alone. On the other hand, no system should be established that excuses psychologists from thinking and making careful judgments about the individuals they are assessing, so in a way, perhaps, it is best to force the interpreter to use judgement and theory in understanding a child.

It is the case that data obtained from the parent and teacher rating scales can be easily compared if the user has also purchased the computer program. These differences cannot be examined statistically by the hand-scorer. In any case, the self-report data cannot be compared statistically with the parent and teacher ratings.

In summary, the BASC is one of the most useful and sophisticated of all the new measures available to those wishing to assess school-age children. Of the different components of the BASC, the parent and teacher reports appear to be the most valuable for practitioners, but the self-report measure should not be overlooked. The measures are not radically different from familiar rating and self-report instruments and they break no new ground; instead they represent a refinement of current technology. The one caution is that the BASC should be used carefully with preschoolers. The measures for this age range are not as accurate and reliable as they are for older children and adolescents. The BASC measures are state of the art for identifying children and adolescents with emotional disturbance and behavioral disorders and they deserve a place in every school and child psychologist's test library.