CHAPTER 3

Interpretation

This chapter provides a review of interpretive strategies for the BASC that we find helpful. A "top down" approach is used that is consistent with the original BASC manual (Reynolds & Kamphaus, 1992), but individual subscale interpretations are emphasized, followed by a discussion of applying and interpreting the BASC to assess change. As we noted in the manual, the individual subscales of the TRG, PRS, and SRP are the most powerful and useful clinical components of the BASC's formal, standardized assessment. First, a preliminary review of scaling is presented in relation to its meaning for BASC interpretation, the validity of the scales, and their impact on interpretation.

UNDERSTANDING AND USING LINEAR T SCORES

Raw scores on a test, rating scale, or other psychometric device are difficult to interpret without consulting multiple tables and calculating a host of sometimes complex transformations—tedious and cumbersome tasks. The use of raw scores is complicated further when the scale items reflect phenomena that may be associated with, or affected by, age or developmental level of the examinee. In addition to the demographic, contextual, and related information needed to interpret scores on psychological and behavioral tests, knowledge of the scale and unit of measurement employed is also a necessity. Raw-score distributions on a multiscore instrument such as the BASC, for example, will change from one subscale to another. For this and other reasons (related
to ease of interpretability and applications of scores outside of research environments, raw scores are commonly transformed into a standardized or scaled score. Common transformations for IQ and achievement tests include normalized, age-corrected, and deviation-scaled scores of the Wechsler genre, where composite scores (e.g., Verbal IQ, Full Scale IQ) are set to a predetermined mean of 100 and a standard deviation of 15. Although there is a rational reason why these metrics were chosen decades ago, the metrics themselves are largely arbitrary. Raw-data scores from subsets of intelligence scales and related batteries (e.g., Test of Memory and Learning [TOMAL]; Reynolds & Bigler, 1994) are commonly transformed into scaled scores with a mean of 10 and a standard deviation of 3. These choices, again, are largely by convention.

In the assessment of personality and behavior, it has become conventional to use $T$ scores, which denote standardized or scaled scores, set to a predetermined mean of 50 and a standard deviation of 10, and expressed on an interval scale of measurement. Interval scales of measurement are ranked above nominal and ordinal scales (e.g., age equivalents and grade equivalents) for their utility because they convey more detailed information that can be manipulated by basic mathematical operations.

The BASC standard scores for all scales are expressed as linear $T$ scores that are on an equal interval scale with corresponding percentiles. Linear $T$ scores have particular characteristics that are necessary to understand in order to interpret these scores correctly across subscales. Linear transformations preserve the shape of the latent or underlying distribution of raw scores. In contrast, normalization of score distributions, through what are popularly called area transformations (fitting the area under the bell curve), is often applied when there is a strong theoretical reason to assume that the true population distribution is normal or “Gaussian” in shape. Under this assumption, any deviation from normality is viewed as error variance created by sampling error. Butcher, Dahlstrom, Graham, Tellegen, and Kaemmer (1989) have argued that it is undesirable to use scaled or standardized scores that have not been normalized. However, if the true distribution of scores, as reflected in the raw score distribution, is non-normal, then there is little justification, beyond convenience, for setting the distribution artificially so that it replicates a normal curve. Linear $T$ scores have not been transformed to fit the Gaussian or normal curve.

Research on more than 33,000 children examined with the tryout and standardized versions of the BASC, and prior research with other scales, clearly show that behavioral variables such as aggression, hyperactivity, attention, and sensation seeking have distributions that are decidedly non-normal.
THE CLINICIAN’S GUIDE TO THE BEHAVIOR ASSESSMENT SYSTEM FOR CHILDREN

(e.g., see Reynolds & Kamphaus, 1992). As we have discussed elsewhere (Reynolds, 2001a; Reynolds & Kamphaus, 1992), measures of psychopathology frequently show a grouping of individuals in the normal range with a long (skewed) tail stretching to the right side of the distribution. In order to accurately assess the targeted behavior, it is often necessary to design content so that discrimination is most acute in the tail of the distribution. In many domains (e.g., aggression), the majority of people tested produce scores that lump into a normal position on the curve; there are strong theoretical reasons to believe this is how it should be. Since non-normal distributions reflect the true state of behavior in the population, linear T scores were chosen to preserve these relationships.

Each scaling method has its strengths and limitations. The advantage to normalizing the raw score distribution is the resulting consistency in the percentile ranks associated with specific T scores across scales. If scores have been normalized, any given T score (e.g., 60, 70, 80, etc.) has the same percentile rank, regardless of the subscale. In relation to the BASC, this procedure means that all the scales (Aggression, Anxiety, Depression, Sensation Seeking, and so on) have precisely the same number of people who score 2 standard deviations (SDs) above or below the mean. This is not true of raw or transformed T scores, where each scale has a different distribution, and the percentile rank changes across scales for a common T score. For example, for boys ages 8–11 years, a T score of 63 on the Aggression scale has a percentile rank of 87, but the same T score on the Antisocial scale has a percentile rank of 91, and a percentile rank of 89 on the Leadership and Study Skills scale. Because we see these variations as representing true differences in the percentages of individuals showing these behaviors at a common distance from the mean (the linear T score reflects the distance of a raw score from the norm group’s mean raw score expressed in SD units), preserving these differences is advantageous.

EVALUATING AND INTERPRETING THE BASC VALIDITY SCALES

The first consideration when interpreting the standardized BASC scales (all levels of the TRS, PHS, and SSP) is to assess the various validity scales. The different BASC forms have different types of validity scales that correspond with the most concerns or problems of potential dissimulation by an examinee. On the hand-scored versions of the BASC, three indexes of dissimulation are available, although not all are provided on all forms: (1) the F or Infrequency,
index: the $L$, or Lie versus social desirability, index; and (3) $V$, the Validity index. On the computer-scored version, a $C$, or Consistency, index is available that examines internal discrepancies in responding to similar items. Though each index is designed for a different purpose, all are intended to detect response biases or (in the case of the $C$ index) unreliable responding. Before beginning an interpretation of the BASC composite or subscale scores, the examiner should review the $V$ index to determine whether an interpretable profile has been obtained, and/or whether follow-up with the informant or the examinee is necessary to clarify the meaning of any elevations (or their absence) in the profile of the referred child.

The $L$ index appears only on the SRP-A, reflecting adolescents' tendencies to deny or minimize emotional or behavioral problems to adults. $L$-index items, when answered in the keyed direction, reflect denial of common problems or flaws that pertain adolescence: For example, "I have some bad habits," "My parents are always right," and "I always think before I act." While adolescents may actually possess some of 14 unusually virtuous characteristics represented on the $L$ index, one who responds in the keyed direction to eight or more of the items is likely not being truthful. Raw scores of 10 or more raise serious concerns about the validity of the profile. Scores as low as 5 represent at least a moderate degree of defensive idealization although the protocol remains valid overall, and the clinician should look for the presence of more severe problems than those reported in the peak areas of the profile. We suggest the following guidelines for interpreting SRP-A raw scores on $L$.

<table>
<thead>
<tr>
<th>Raw score on $L$</th>
<th>Potential interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5–7</td>
<td>Defensive response set; Peak scores likely represent more serious problems than revealed.</td>
</tr>
<tr>
<td>8–9</td>
<td>Caution in the interpretation of profile: Problems are likely deflated, and follow-up interview is warranted.</td>
</tr>
<tr>
<td>10–14</td>
<td>Extreme caution is necessary: The profile is likely invalid; such scores may reflect a lack of cooperation, frank denial, &quot;faking good,&quot; or emotional immaturity and psychological naivety.</td>
</tr>
</tbody>
</table>

In deciding upon the most appropriate interpretation it is useful to review specific item content. Sometimes a repeat administration after such a review, and the interview with the adolescent, will clarify matters.

The $F$ index appears on all standardized BASC scales and at all levels of the TRS-P, $C$, and $A$; PRS-P, $C$, and $A$; and the SRP-C and $A$. Classically con-
structured, the BASC F indexes are composed of items that are infrequently endorsed in the normal population and have lower correlations with one another than to other items on the various scales. Thus, not only are items on the F Index infrequently endorsed, they are items that are least likely to form a coherent or unified construct or syndrome of psychopathology. In a valid profile, elevated F scores are associated with the presence of multiple severe disorders or a frank, psychotic disturbance; however, sometimes they represent a "plea for help" due to acute, psychological distress on the part of the examinee (in the case of the SRP) or distress or exaggeration with the behavior of the examinee (in the case of the TRS or PRS). As clinicians, we have encountered teachers who are so frustrated with a student that an exorbitant number of problems are rated severely to try and get the child removed from the class. Such rating styles will typically cause an elevated F score. The most common interpretation of an elevated F score is an attempt to create an excessively negative impression. However, beyond those noted above, other reasons can result in an elevated F, including reading difficulty, failure to follow instructions, random responding, or a failure to take the task seriously. We suggest the following guidelines for interpreting raw scores on F.

<table>
<thead>
<tr>
<th>Raw scores on F</th>
<th>Potential interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRP 2–3</td>
<td>Mildly exaggerated but clearly valid profile; some</td>
</tr>
<tr>
<td>TRS 2</td>
<td>distress evident; negative evaluation by self or others</td>
</tr>
<tr>
<td>PRS 1</td>
<td>apparent</td>
</tr>
<tr>
<td>SRP 4–5</td>
<td>Probably significant problems evident; distress and</td>
</tr>
<tr>
<td>TRS 3–4</td>
<td>some exaggeration present; caution necessary to avoid</td>
</tr>
<tr>
<td>PRS 2–3</td>
<td>overinterpretation of profile.</td>
</tr>
<tr>
<td>SRP 6–7</td>
<td>Severe, acute distress; invalid responding: frank</td>
</tr>
<tr>
<td>TRS 5–8</td>
<td>psychosis, malingering, or failure to follow instructions.</td>
</tr>
<tr>
<td>PRS 4–6</td>
<td></td>
</tr>
</tbody>
</table>

Once again, we must emphasize the need for a follow-up interview with the respondent to clarify the meaning of an elevated F score, since it may range from a true plea for help to psychosis to frank malingering or a purposeful attempt to portray the individual in an unduly negative manner.

The V index appears only on the SRP in its two versions. It consists of 5 and 6 nonsensical items for the SRP-C and SRP-A, respectively. The V-index items, listed in Table 3.1, are clearly implausible or nonsensical. They are occasionally misinterpreted by children, however, and we find it useful to question them about their keyed responses to these items.

High scores—defined as equal to or greater than 2 on the V index—
TABLE 3.1. SRP: Vindex Items

<table>
<thead>
<tr>
<th>Child level</th>
<th>Adolescent Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>I drink 30 glasses of milk every day.</td>
<td>I have just returned from a nine-month trip on an ocean liner.</td>
</tr>
<tr>
<td>I have never been in a car.</td>
<td>I have not seen a car in at least six months.</td>
</tr>
<tr>
<td>I have never been to sleep.</td>
<td>I take a plane trip from New York to Chicago at least twice a week.</td>
</tr>
<tr>
<td>I have no teeth.</td>
<td>Superman is a real person.</td>
</tr>
<tr>
<td>Superman is a real person.</td>
<td>Television does not really exist.</td>
</tr>
</tbody>
</table>

indicate misinterpretation of the items, bizarre mentation, lack of cooperation, inability to read, random responding, or a floridly psychotic state that should be evident in conversation.

The C index is available only via computer scoring and is listed on the first page of any PRS, TBS, or SRP printout. This index compares responses to thematically similar items, calculates the difference between each possible pair of common items, and compares those deviations to the extent to which individuals were consistent within the standardization sample. This score provides an indication of the relative reliability of an individual respondent. Cautionary statements are generated by the computer-scoring algorithm and guidance for interpretation is given on page 1 of the BASC Plus or BASC Enhanced Assist computer printout.

Once the examiner has determined that a valid or useful profile has been obtained, the next step is to decide whether to follow a top-down interpretive approach, first examining the composite scores and working down to the individual subscales (as detailed in Reynolds & Kamphaus, 1992, Ch. 9), or to move more directly to the subscales.

**INTERPRETING COMPOSITE SCORES**

Originally, we conceptualized a top-down approach to BASC interpretation that began with the various composite scores—for example, Behavioral Symptom Index, Emotional Symptom Index, Clinical Maladjustment, and the like. Ten years of applying the BASC in our own clinics, interacting with literally thousands of BASC users, and seeing how the BASC is used in the research literature has caused us to rethink this approach. We now view the composite scores as more limited, offering specific types of information but not the
The individual scales are perhaps the most powerful interpretive force available for the rating scales and the SRP. These scales are highly interpretable because of the manner in which they were developed, with a particular emphasis on the content validity of the individual scales. The highly recognizable content of most of the scales helps render their interpretation intuitive. How-
even it still useful to review the congruence between the individual scale and its composite to assess the breadth of the problems seen.

It may be necessary to evaluate the item content of a scale to understand the reasons for a high or low score, particularly a score that is inconsistent with the composite score. An example would be a child referred for attention problems who obtains a high Atypicality score on the TSK, in addition to high Attention Problems and Hyperactivity scores. The Atypicality items should be evaluated carefully to determine the reasons for the elevation. Items describing insinuative or hyperactive behaviors might be responsible, such as “Eats things that are not food” (the child may chew on erasers in class) or “Sings or hums to self” (this is not an unusual behavior for a child with motor hyperactivity). In this case, the high T score would not be indicative of psychotic thought processes but of motor hyperactivity, consistent with a diagnosis of ADHD. On the other hand, if items describing behavior such as “tries to hurt self” or “Says things that are not there” have been marked as occurring frequently, the examiner would need to evaluate other sources (such as other BASC components) to determine whether there is evidence of a psychotic disorder.

The normative analysis of the scales and the inspection of the critical items provide a strong basis for interpretation. However, normative analysis does not capture all of the useful information in a child’s profile. For example, a child whose clinical-scale T scores are all above 60 and whose adaptive-scale T scores are all below 40 is clearly having extraordinary difficulties, but some of the problems may be considerably worse than others.

Next we present methods for a more in-depth interpretation of the individual subscales, using a three-level model. Throughout this discussion on interpretation, it is important to keep in mind the need for a simultaneous consideration of the context in which the behaviors occur and the information gleaned from the SDQ and the clinical interview with the child.

Three Levels of Scale Interpretation

In general, we formulate three interpretations at the scale level: (1) clinically significant problems requiring diagnosis and/or treatment (and/or the child’s strengths), (2) subthreshold problems requiring monitoring and/or preventive intervention, and (3) problems or adaptive skills requiring no specific action. We will demonstrate how we use these interpretations via hypothetical examples of one child with different scores on the Depression scale of the KBIT.
1. "Jasper's" Depression scale T score of 70 and his symptoms of depression—which include expressions of hopelessness and suicidal ideation (without a specific plan), frequent tearfulness, decreased appetite, weight loss, inability to enjoy previously pleasurable activities, and decreased energy—are severe enough to warrant a diagnosis of major depressive episode, moderate. Symptomology of this severity requires immediate treatment. Individual psychotherapy is recommended along with a referral to a psychiatrist to determine whether psychopharmacotherapy is appropriate and medically safe for the young man.

2. Jasper is unhappy most of the time, as indicated by multiple sources of information, including a Depression scale T score of 63 on the SRP. He complains of tearfulness, decreased appetite, and some feelings of hopelessness. Although he does not experience sufficient symptomatology to warrant a diagnosis, his symptoms are beginning to impact his relations with parents and teachers, primarily in the form of increased irritability. It is recommended that he be monitored periodically by the school psychologist, his parents, and his pediatrician to ensure that these problems do not worsen. Simultaneously, preventive intervention, in the form of cognitive-behavioral therapy and supportive counseling, is recommended at school and in the community, if his parents deem it necessary as well.

3. Jasper does not display significant depressive symptomatology (Depression scale T score of 55) as noted by teachers, parents, and himself. He does admit to experiencing some occasional unhappiness with a depth and frequency that is within normal limits at this time.

Of course, most youth possess more than one problem area, and some of these scales do not lead directly to case formulation. Withdrawal, Study Skills, Social Skills, Leadership, Learning Problems, Interpersonal Relations, Self-esteem, Social Stress, Attitude toward Parents, Teachers, and School, Sense of Inadequacy, and Aggression are examples of scales of this type. Three interpretations are offered for the Aggression scale, again using "Jasper."

1. Jasper displays significant problems with aggression, including verbal cursing, threatening, and bullying (T score = 70). In addition, the school history indicates that he has been cited for pushing and punching others and has been suspended twice this year twice for fighting. He requires treatment for these problems in both school and community.

2. Jasper has some problems with verbal aggression, including cursing at peers and anger outbursts (in which he throws objects) primarily noted at
home and in the community (T score = 61, BASC PRS-L). These problems should be monitored to see if they respond to the new parenting strategies being adopted by his parents. If they do not, additional interventions should be implemented to reduce aggression. No significant aggression problems are reported in the more structured environment of school at this time (T = 51, BASC TRS-L).

5. Jasper does not display significant problems with aggression, as rated by teachers and his parents (T = 51 on both TRS and PRS).

As noted, BASC Interpretation is described in detail in Reynolds and Kamphaus (1992). However, we develop additional interpretive guidance as new research findings become available. The following proposals for additional interpretive steps are offered to supplement those found in the BASC manual.

Three Interpretive Steps

In an effort to simplify interpretation and yet ensure a thorough approach we advise the following steps as a starting point.

• Step 1. Write down (or highlight, circle, or in some way mark) all scales with T scores at or above 60 for the clinical scales and at or below 40 for the adaptive scales. Scores in this range are identified as either “at-risk” or “clinically significant” on BASC printouts. By utilizing scores in this range for screening, diagnostic, and treatment purposes, we are likely to reduce “false negatives” and errors in the direction of “false positives.” In addition, utilization of scores in this range allows for the identification of subsyndromal problems that may require treatment because they are associated with functional impairment (Cantwell, 1996).

We generally do not list the composite scores at this stage unless we are trying to answer a question that is associated with a composite score (see Reynolds & Kamphaus, 1992). Types of these questions include, for example, “Is Victor’s behavior generally worse than that of other boys his age?” (BSI), or “Does Victor have more internalizing problems than other boys his age?” (Internalizing composite). Another instance is a composite score exceeding T of 69, with lower scores on the subscales—a noteworthy combination suggesting that the sum of this child’s problems clearly exceeds its parts. For the clinical purposes of diagnosis, classification, or treatment planning, the individual scales provide the specificity necessary to answer these questions; the composite scores do not.

57
The TRS and the PRS subscales show good levels of test specificity (defined as unique and reliable variance associated with scores on a particular scale), although parent ratings (the PRS) have been found to have greater specificity across scales than teacher ratings (TRS; Daniel, 1993). This finding means that parents' ratings will produce profiles with greater variability than teachers' and will denote less influence by the higher order factors (e.g., internalizing and externalizing). Daniel (1993) contends that these findings indicate that parents may be more sensitive to specific aspects of their child’s behavior. This is a reasonable view; given parents’ greater exposure to their children, they can likely make finer distinctions regarding behavior and may be influenced less by any positive or negative halo. These interpretations are consistent with TRS and PRS profiles for various diagnostic groups reported in the BASC manual as well.

In the sample tables that follow, all potentially important T scores are identified in boldface type for the case of “Victor,” discussed in Chapter 4.

<table>
<thead>
<tr>
<th></th>
<th>Teacher</th>
<th>Mother</th>
<th>Father</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperactivity</td>
<td>83</td>
<td>58</td>
<td>52</td>
</tr>
<tr>
<td>Aggression</td>
<td>57</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>Externalizing Problems composite</td>
<td>57</td>
<td>48</td>
<td>45</td>
</tr>
<tr>
<td>Anxiety</td>
<td>50</td>
<td>47</td>
<td>42</td>
</tr>
<tr>
<td>Depression</td>
<td>45</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>Somatization</td>
<td>43</td>
<td>49</td>
<td>37</td>
</tr>
<tr>
<td>Internalizing Problems composite</td>
<td>45</td>
<td>43</td>
<td>36</td>
</tr>
<tr>
<td>Atypicality</td>
<td>74</td>
<td>37</td>
<td>41</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>41</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>Attention Problems</td>
<td>78</td>
<td>73</td>
<td>73</td>
</tr>
<tr>
<td>Behavioral Symptoms Index</td>
<td>69</td>
<td>48</td>
<td>47</td>
</tr>
</tbody>
</table>

On the Adaptive scales, scores below 30 are considered clinically low, but scores from 31 to 70 are also representative of high-risk areas.

<table>
<thead>
<tr>
<th></th>
<th>Teacher</th>
<th>Mother</th>
<th>Father</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>44</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Social Skills</td>
<td>42</td>
<td>66</td>
<td>57</td>
</tr>
<tr>
<td>Adaptive Skills composite</td>
<td>42</td>
<td>65</td>
<td>60</td>
</tr>
</tbody>
</table>

58
**Step 2. For each scale listed, identify supportive and nonsupportive evidence of a significant problem or competency.**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Supportive evidence</th>
<th>Nonsupportive evidence</th>
</tr>
</thead>
</table>
| Hyperactivity | - Parents and teacher report an inability to learn basic academic concepts, such as numbers and letters, a high level of activity, and an inability to focus on academic tasks.  
- Father reports his own lifelong problem of sitting still.  
- Teacher reports that he cannot focus on task at hand and, as a result, his problem is worsening.  
- Current teacher reports that he is always on the run, talks excessively, is unable to stay seated for any period of time, and frequently interrupts class activities. | - Parents report some difficulty with fine and gross motor skills. He is not yet able to catch a ball or tie his shoes, although he is able to ride a bike. His teacher also reports that he has a difficult time writing his letters, coloring, or drawing (could be a case of general developmental delay or mental retardation).  
- Teacher reports that he is unable to retain information over even short periods of time (further evidence of developmental delay). |
<p>| Atypicality | - Parents report some difficulty with fine and gross motor skills. He is not yet able to catch a ball or tie his shoes, although he is able to ride a bike. His teacher also reports that he has a difficult time writing his letters, coloring, or drawing (could be a case of general developmental delay or mental retardation). | (continued) |</p>
<table>
<thead>
<tr>
<th>Scale</th>
<th>Supportive evidence</th>
<th>Non-supportive evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention Problems</td>
<td>Parents and teacher report an inability to learn basic academic concepts, such as numbers and letters, a high level of activity, and an inability to focus on academic tasks. Teacher reports that he cannot focus on task at hand and, as a result, his problem is worsening. During testing, his inability to focus on task at hand or sustain attention for even small lengths of time caused him to miss many items he would otherwise have gotten correct.</td>
<td>Parents report some difficulty with fine and gross motor skills. He is not yet able to catch a ball or tie his shoes, although he is able to ride a bike. This teacher also reports that he has a difficult time writing his letters, coloring, or drawing (could be a case of general developmental delay or mental retardation). Teacher reports that he is unable to retain information over even short periods of time (further evidence of developmental delay).</td>
</tr>
<tr>
<td>Adaptability</td>
<td>Parents report that he is responsible for keeping his room clean, is very polite, well-mannered, and able to dress and bathe himself.</td>
<td></td>
</tr>
<tr>
<td>Social Skills</td>
<td>Parents report that he is responsible for keeping his room clean, is very polite, well-mannered, and able to dress and bathe himself.</td>
<td></td>
</tr>
</tbody>
</table>

60
Step 3. For each scale remaining, draw conclusions regarding diagnosis, subsyndromal problems, and competencies. Victor has significant problems with inattention and hyperactivity, as indicated by BASC T scores and other measures (see test results in Chapter 4). Although his PAIS Hyperactivity scale T scores are not significant, these results reflect underestimations of his activity level at home when compared to the parents' verbal descriptions of their management of his activity level at home (see Chapter 4). There is overwhelming evidence for the diagnosis of ADHD, combined type. Moreover, the potential alternative cause of developmental delay or mental retardation is effectively ruled out by the various cognitive test results, which were consistently average to low-average. He tended to score lower on tests requiring spatial and visual motor skills, which could be attributed to a mild developmental delay.

Although his parents report adaptive strengths at home and they clearly have a very favorable opinion of his behavior, noting that he has many friends, this is not the case at school. Qualitative information and behavior ratings from his teacher do not support this conclusion. Hence, it is probably premature to conclude that he has above-average adaptive competencies in comparison to other children his age (see full report in Chapter 4).

In this case, the Atypicality scale functions as another indicator of hyperactivity, as demonstrated in Chapter 2. Therefore, this result can be subsumed under the Hyperactivity scale results.

No noteworthy subsyndromal problems (T = 60–69) were identified for Victor. This conclusion suggests that his ADHD symptomatology is not accompanied by any current evidence of emerging comorbidities, with the possible exception of a learning disability, for which he will be observed (see Chapter 4).

While we recognize that these three steps represent a simplified version of the complex set of heuristics typically used by clinicians, they do provide a structure for initiating BASC scale interpretation. The report for Victor, presented in Chapter 4, includes more of the nuance that characterizes clinical case formulation. To facilitate understanding and interpretation of various score levels on all standardized subscales of the BASC, we have provided rapid summary tables, one for each BASC subscale previously mentioned in Chapter 2. Now we turn to explore additional interpretive steps that may enhance the BASC's utility.
Critical Items

Individual test items provide narrow views of behavior, and item reliabilities pale in comparison to the reliability of a composite of items (a subscale). However, reviewing item content and following up with questions can prove extremely useful in gaining a deeper understanding of a pattern’s etiology. Some items cry out for follow-up by virtue of their content (e.g., “Sometimes I want to hurt myself”). Such items are tagged as Critical Items on all levels of the BASC TRS, PRS, and SRP on both the hand-scored and computer-scored options. Examples of Critical Items from the SRP-A are:

- I can’t seem to control what happens to me.
- I don’t seem to do anything right.
- Someone wants to hurt me.
- No one understands me.
- I hear voices in my head.
- I have fainting spells.
- I just don’t care anymore.
- Sometimes I want to hurt myself.
- Nothing goes my way.
- Nothing about me is right.
- Nobody ever listens to me.
- I cannot control my thoughts.
- Other kids hate to be with me.
- Sometimes voices tell me to do bad things.
- I get into fights at school.
- I cannot stop myself from doing bad things.

As is apparent, these items deal principally with issues of self-harm, loss of control, bizarre symptoms, and rejection/rejection—all critical issues for adolescents that may denote more severe problems than have otherwise been detected or revealed. Follow-up questioning and sometimes confrontation may be necessary although many times a clear, rational explanation for the behavior or feeling is given by the child.

Examples of Critical Items from the rating scales are also instructive. The following are designated as Critical Items on the PRS-A:

- Tries to hurt self.
- Sleeps with parents.
- Uses medication.
- Uses illegal drugs.
- Says, “I want to die,” or “I wish I were dead.”
- Has muscle spasms.
- Has a hearing problem.
- Says, “I’m afraid I’ll hurt someone.”
- Has seizures.
- Has eye problems.
- Is in trouble with the police.
- Says, “I want to kill myself.”
- Stutters.
- Plays with fire.
- Drinks alcoholic beverages.
- Says, “I hate myself.”
- Threatens to hurt others.
Common themes of violence, poor sense of control, and odd behaviors, in conjunction with well-established developmental flags (e.g., "Plays with fire") and somatic complaints (e.g., muscle spasms, hearing or eye problems, seizures) should prompt a medical referral if the child has not been seen lately by his or her primary care physician. Critical items on these and all other BASC standardized scales are clearly identified when endorsed, regardless of whether the clinician uses the hand score BASC forms or any of the BASC computer scoring programs.

According to Butcher (1990), and reiterated by Graham (2000), the primary and most appropriate use of Critical Items is to detect highly specific problems or behaviors that are important but might not be reflected in clinical profile elevations. Highly salient themes in the child’s life may be illuminated. However, as single items, Critical Items have limited reliability, and the types of problems depicted are restricted in range. They are thus a supplement to subscale interpretation, not a substitute or an independent interpretive process. Endorsement of Critical Items should lead the examiner to inquire further into areas designated by the items.

**INTERVENTION PLANNING AND EVALUATION**

The interpretive steps described are primarily useful for case formulation, classification, and diagnostic purposes. The following discussion offers principles intended to enhance the utility of BASC results for treatment of behavioral and emotional difficulties in children and adolescents.

**Specifying Treatment Objectives and Target Behaviors**

Results from the BASC may form the basis for establishing treatment objectives. Examiners are cautioned to avoid establishing unrealistic objectives—and therefore expectations—for change. Indeed, it may be inappropriate to offer treatment objectives for all BASC constructs. It may not be warranted, for example, to expect change on the Withdrawal or Anxiety scale if somatic therapy for Hyperactivity and Attention Problems is the only treatment being used. Similarly, pharmacological treatments for ADHD (e.g., amphetamines) have not been shown to affect internalizing problems in predictable ways (discussed more later). Therefore, it seems unwise to communicate the expectation that these scales will change when, in fact, there is no treatment aimed at changing the behavior designated by the construct. On the other hand, if
Comorbid depression is being treated through medication and cognitive-behavior therapy; then change in the Depression scale may be an appropriate treatment objective. Treatment objectives should always be linked to actual treatments delivered in order to ensure that parents, teachers, and others have realistic expectations regarding change.

Any of the four BASC components (PRS, TRS, 8RP, SOS) can then be used to identify target behaviors for intervention. Erhardt and Conners (1996) support their suggestion that deviant scales and items be used to set target behaviors by observing that “it is reasonable in drug studies to use the most elevated scales on rating measures as predictors of drug treatment outcome, as well as primary target behaviors for medication effects” (p. 130).

The BASC Plus computer scoring program offered by American Guidance Service provides recommended target behaviors that are useful for prioritizing interventions. The selection of these behaviors was heavily influenced by surveys of teachers and more than 500 school-age children during the development of the BASC. These teachers and students listed, in prioritized order, behaviors they most disliked in others. Frequency and saliency were determined from these ratings; those most likely to result in a child being ostracized were retained on the BASC. The BASC Plus identifies these behaviors when marked by a teacher or a parent. These are often behaviors associated with aggressive tendencies, which typically cause children to be excluded from activities with normally behaving, nonaggressive children. By targeting these behaviors for intervention—unless self-handicap is noted, which would take priority—children are less likely to experience exclusion and ostracization. Children who display inappropriate behavior benefit from being with normal and well-behaving peers who model appropriate behavior. Due to the level of saliency among peers and teachers and the resultant social validity and potential for positive impact, these target behaviors should receive careful, due consideration as a starting point whenever they are detected.

Recording Treatment Data

The major prescriptive inference made by the BASC user is a link between treatment and outcomes. Careful treatment records are necessary in order to draw such inferences. The reverse side of the Parent and Teacher Monitor forms that are part of the BASC ADHD Monitor, for example, provides space for documenting the types of various treatments, including their onset, cessation, and adherence. Although adherence is probably the most difficult to document, attempts should be made to do so. It may be necessary to use telephone interviews with parents, teachers, and/or clinicians, or conduct reviews of collateral assessment data.
of clinical records in order to assess adherence to treatment regimens. The collection of treatment information allows clinicians to be confident in their assessments of change.

Collecting at Least Three Data Points

The BASC forms and software are designed to facilitate the collection of three or more data points so as to encourage frequent follow-up and a more accurate assessment of a child’s underlying change trajectory. Francis, Fletcher, Stuebing, Davidson, and Thompson (1991) observed that individual growth is difficult to measure accurately with only two data points, such as a pretest and posttest.

Assessing Changes: Clinical Significance

The assessment of change is characterized by numerous methodological confounders and issues. Jacobson and Truax (1991) criticize the use of meta-analyses for assessing change by observing that statistical significance may be unrelated to the clinical significance of change. They cite the example of weight loss where a loss of 2 or 3 pounds may be deemed statistically significant yet hardly satisfying to many of those undergoing treatment. In relationship to change subsequent to psychotherapy, they note that there are three potential indicators of clinical significance. We believe these indicators apply to all types of psychological interventions and have rephrased them here to reflect their more general applicability:

1. The level of functioning subsequent to behavioral intervention should fall outside the range of the dysfunctional population, where the range is defined as extending 2 SD’s beyond (in the direction of functionality) the mean of the population.

2. The level of functioning subsequent to behavioral intervention should fall within the range of the functional or normal population, where the range is defined as within 2 SD’s of the mean of that population.

3. The level of functioning subsequent to behavioral intervention places the client closer to the mean of the functional population than it does to the mean of the dysfunctional population.

To these we would add an important fourth:

4. The response to treatment makes a positive difference in the day-to-day life of the patient.
Consider, for example, a child who has a severe case of ADHD and responds to psychopharmacotherapy to a degree that reduces his score on the Hyperactivity scale from 90 to 75. This child is still more overactive than 99% of children at his age. Nevertheless, if he can now attend a regular classroom with peers who have good social skills, as opposed to attending a segregated classroom with children who have behavioral problems and less developed social skills, then the treatment would be successful.

Whether or not such goals are realistic for ADHD and other developmental disorders, however, remains in question. If ADHD is conceptualized as a developmental disorder, there may be cause for modest expectations for change, whereas we would not expect children with mental retardation to ever test within the normal range. However, in accordance with criterion (4), an intervention allowing a person with mental retardation to handle activities of daily living independently would make a huge difference in quality of life and be judged successful. Similarly, expectations for change due to treatment of children with syndromes of reading disabilities or autism might not include a return to the normal range of functioning for many. For a child with autism to be able to engage in interactive play, or for a child with dyslexia to develop positive self-esteem, a high level of self-acceptance, and compensatory skills are all big successes that do not attain "normalcy" but should be recognized and encouraged. Other potential indicators of acceptable change could include change that is recognizable by peers, teachers, parents, or others, and/or reduced risk for various health problems (Jacobson & Trux, 1991). Certainly the limitations of statistical indicators of change should not be overlooked.

For the above reasons, it is essential that clinical significance be considered when assessing change. In part, clinical significance will eventually become evident through research on the outcomes of children with various disorders. Until then, the clinical assessor of change must decide on a measure of clinical significance that is appropriate for each child. Perhaps the best way to proceed is to define clinical significance by setting objectives that are commensurate with the treatment objectives for the child, parents, and/or caregivers. In fact, setting appropriate goals for change may be an important topic to discuss with parents and others prior to the implementation of treatment. If treatment objectives are not specified and/or realistic for a given child, the effectiveness of treatment will remain ambiguous. Therefore, the establishment of treatment goals is a key step in RASC interpretation. Detailed information on development of treatment plans and objectives can be found in Hersen and Bellak (1998), Hughes (1999), Marrons, Witt, Daly, and Vollmer (1999), and Leahy and Holland (2000).
Assessing Change: Other Methodological Issues

Francis et al. (1991) have offered statistical procedures and formulae for individual growth models of change. These authors also discuss some of the measurement prerequisites that are necessary for the accurate assessment of change. Francis et al. (1991) concluded that measures of change, such as the BASC, should (1) have interval scales of measurement, (2) be relatively free of ceiling or floor effects, (3) encourage collection of three or more observations in order to assess an individual growth trajectory, and (4) not use age-based standard scores for measuring change. Each of these criteria will be discussed as they apply to the BASC.

The BASC strives to produce an interval scale of measurement through the use of T scores. An interval scale is necessary for assessing change because a scale with unequal units would not allow for the precise computation of change indices. The methods of scaling the BASC (Reynolds & Kamphaus, 1992) ensure the existence of an equal-interval scale of T scores.

Because ceiling and floor effects would detract from the usefulness of a scale for assessing change, the BASC T-score range is not artificially restricted to a prespecified range. The T scores were scaled based on the distributional properties of the construct, as sampled from the population (see Reynolds & Kamphaus, 1992). Ceiling and floor effects are rarely encountered with the BASC, and when they do occur, are almost exclusively present on the SRP.

The BASC does use age-based standard scores (T scores) as the featured score for measuring change. For many children, however, the effects of age standard scores on the interpretation of change are nonexistent, because raw scores are not significantly different across age groups, as is indicated by the use of the same norm tables for large age groups of children. In the case of the PRS, the raw scores did not differ significantly for ages 8–11, allowing for the use of one norm table for this age range. For example, if a child is diagnosed initially at age 8, the same raw score to T-score conversion table is used until the child reaches age 11. In the case of a 7-year-old child, however, different norm tables will be used for the computation of T scores at age 7, which could result in an inaccurate assessment of change. When norm tables change for children, it may be wise for clinicians to consider raw scores, in addition to T scores, when assessing change over this time period. Raw scores can be used to assess change for a child or a group of children in a common age range but should not be used for classification or diagnostic purposes.

Another methodological issue to consider is the reliability of different (pretest-posttest) scores. Since an inverse relationship exists between the reliability of a difference score and the correlation between the pretest and
posttest, unreliable difference scores may reflect the possibility that the trait being measured changes little over time or in response to treatment (Francis et al., 1991). The long-term stability of the BASC scales is currently established only for a 7-month period (see Reynolds & Kamphaus, 1992), but additional information, when it becomes available through future research, may shed light on the malicability that may be detected by the scales. However, when groups of children are considered, regression effects and unreliability of change scores are easily accommodated.

**Informing Treatment Providers and Caregivers**

Due to the complexity of symptoms commonly associated with children’s disorders, children often receive multimodal therapies, including somatic, behavioral, and educational interventions. The involvement of multiple treatment providers makes communication critical. BASC output is designed to foster communication among parents, physicians, teachers, psychologists, and other clinicians by offering straightforward graphs. T scores and percentile ranks are also provided to enhance quantitative interpretation.

We recommend that BASC results be shared generously among all individuals involved with a case. Results, however, should always be accompanied by competent interpretation in order to ensure that they are used appropriately.

**ADDITIONAL INTERPRETIVE ISSUES**

**Child-Adult Agreement**

We expect disagreement between adult and child raters, even when common items and scales are utilized. Handwerk, Larzelere, Soper, and Trisman (1999) compared parent Child Behavior Checklists (CBCL) and adolescent Youth Self Report (YSR) ratings and found that the adolescent rates significantly under-reported their problem behaviors whether they were being treated in acute-care psychiatric shelter, psychiatric inpatient, residential placement settings. In addition, the adolescents were able to describe their internalizing symptoms with only a minimal level of accuracy. Specifically, parent reports exceeded clinical cut scores 48% of the time, whereas children exceeded cut scores with 22% frequency for the Internalizing scale; the corresponding values for the Externalizing scale were 78% and 30%. In light of these and other data, the authors concluded that, in this specialized treatment setting, parents tended to overrate the prevalence of externalizing symptoms compared to ex-
ternal criteria, such as incident reports. Hence, both child and parent ratings showed evidence of bias regarding the reporting of externalizing problems. It should be noted, however, that these samples were clinical in nature, with the residential sample including adolescents who had a minimum of two and a half psychiatric diagnoses per case.

In another study, Kamphaus et al. (2000) found that the mean scores for children in normative samples were similar for parent and teacher ratings. We compared the parent and teacher ratings of four groups of children: U.S. Anglo, U.S. African American, U.S. Hispanic, and metropolitan Medellín, Colombia. We found that parents tended to rate their children as having more adaptive skills and significantly more problem behaviors than teachers did.

The Kamphaus et al. (2000) study is also the largest comparison study of mean gender and ethnic group differences. Generally speaking, the results can be summarized as follows (although there were a few exceptions):

- Child gender produced the most consistent trends for the total U.S. and Colombian samples. In the Colombian sample teachers rated boys as having more aggression, atypicality, conduct problems, and hyperactivity, whereas parents rated boys as having more problems in all those areas as well as attention. Girls were not rated by teachers as having more problems than boys; however, parents viewed girls as having more problems with anxiety (ages 6–18), withdrawal (ages 4–11), depression (ages 12–18), and somatization (ages 12–18). (These results also support Handwerk et al.'s conclusion that parents were more severe raters than teachers.) Finally, both parents and teachers generally agreed that girls have better adaptive skills than boys (indicated for most age levels on all of the BASC Adaptive scales). This finding is consistent with the BASC standardization data.

- Few of the differences between the four ethnic groups were as consistent as the gender differences. For example, the four groups did not differ on ratings of aggression by teachers. Parents of Colombian and U.S. Anglo children rated their children as more aggressive than did parents of U.S. African American and U.S. Hispanic children. There were, however, some differences at individual age levels. For example, parents of U.S. Hispanic children, ages 4–5, rated their children as being as aggressive as parents of Colombian and U.S. Anglo children rated theirs. These data reveal that the ethnic and linguistic differences are much less interpretable than the clear cut gender differences, which are much more consistent across rater and ethnic group.

- Some of the ethnic group differences were more consistent within rater categories. For example, although teachers did not observe significant differ-
ences in somatization, parents consistently identified their U.S. African American and Colombian children as having more somatization problems than did parents of U.S. Hispanic and Anglo children (see Kamphaus et al., 2000, for a full report of these findings).

- Analyses of socioeconomic status (SES) for the Colombian sample revealed that teachers rated fewer differences, with the exception of low-SES children having more problems with somatization and conduct. Many differences were noted among parents by SES groups, as groups in the lower SES strata rated their children as having more aggression, physicality, depression, hyperactivity, and conduct problems. Other differences were quite variable by age group.

Our interpretation of these findings are that (1) child gender is the most important determinants of both teacher and parent ratings, and (2) child ethnicity and SES are more likely to affect parent than teacher ratings. Of course, the significance of these findings is tempered by the fact that these were nonclinical samples. Nevertheless, results may have implications for research and theory development regarding the interplay of child gender, family SES, and ethnic and language group influences on child behaviors and adult perceptions of these behaviors.

Teacher and parent ratings of young children have been found to be good predictors of behavioral maladjustment over a 6-year period (Verhulst, Koot, Van der Ende. 1994). Verhulst and colleagues used parent and teacher ratings on the CBCL to predict behavioral maladjustment, based on interviews with parents, six years later. Interestingly, although both teacher and parent ratings were predictive, teacher ratings were significantly better predictors, despite the fact that the criteria data were collected from parents. This finding underscores our belief that both parent and teacher reports are necessary in each case we evaluate.

Kamphaus and Frick (2002) emphasize an interpretive approach that considers all information, regardless of the source, to be of potential value. Take the case of Victor, for example. If he were old enough to take the SBP, his results would also be submitted to the same three steps of interpretation. We agree with this approach and advise our trainees similarly—to consider all sources of information as potentially valuable. Case studies in the next chapter demonstrate the value of child/adolescent ratings as well. In addition, we wish to point out that we developed the SOS and SDH for one reason: observations and interview findings are of considerable importance for case conceptualization.

70
ADHD Comorbid with Mental Retardation

The most recent diagnostic manual of the American Association on Mental Retardation (1992) suggests that all individuals with mental retardation should receive a full social-emotional evaluation because they are at greater risk of depression, ADHD, schizophrenia, and a long list of other psychiatric disorders. The prevalence of ADHD is, indeed, higher in this group than in the remainder of the population. Unfortunately, due to inclusion of a codical to this effect in the DSM-IV, clinicians may consider ADHD symptomatology to reflect a developmental delay that is associated with the child's mental age. Well, it is conceivable that the DSM-IV is simply incorrect on this point. Pearson and Annan (1994) conclude that this is the case:

Findings suggest that chronological age should be taken into consideration when behavior ratings are used to assess cognitively delayed children for ADHD. However, the results do not support guidelines stating that mental age must be used to determine what norms should be applied when such children are evaluated clinically. (p. 395)

The use of mental age as a consideration in making the ADHD diagnosis for children with mental retardation may result in the denial of somatic and behavioral treatments that are known to have demonstrated efficacy.