

## ARTICLES

---

# Assessment of Children's Social Skills: Recent Developments, Best Practices, and New Directions

Kenneth W. Merrell

*Division of Psychological & Quantitative Foundations  
University of Iowa*

This article provides an overview of some of the recent developments in assessing social skills of children and youth, as well as a discussion of “best practices” in conducting assessment and linking assessment to effective intervention. Naturalistic behavioral observation and behavior rating scales are proposed as the 2 assessment methods that should be considered primary or “first-line” choices for social skills assessment. A review of some specific tools that have been developed within these 2 assessment methods is provided. Interviewing and sociometric techniques, although not necessarily primary methods of assessment for children's social skills, sometimes may be an important part of an assessment design, or “second-line” choices. Projective–expressive techniques or objective self-report instruments for assessing children's social skills sometimes may help illuminate the overall assessment results but should never be used as primary assessment methods for social skills; thus, these methods are considered “third-line” methods for this purpose. Six best-practices recommendations, all of which are supported by previous empirical research, are offered for improving social skills assessment efforts. Finally, conclusions are drawn regarding the current state of the art in social skills assessment, as well as implications for future directions in this area. Clearly, the challenge facing researchers and practitioners who are concerned with this area will be to make assessment “functional” and link it to intervention more effectively.

Screening and assessment are essential foundations for effective intervention in social–behavioral problems of children and youth. Without the careful identification, classification, and selection that should be a part of good assessment, social behavior inter-

ventions are likely to be haphazard and disorganized at best and ineffective at worst. Clearly, practitioners and researchers who are serious about understanding and dealing effectively with children's social skills and related social-behavioral issues must provide the proper foundation for their work through carefully planned assessment activities.

There have been considerable advances in research and practice in social skills assessment during the past 2 decades. This article provides a brief overview of some of these advances, highlights recommendations for "best practices" in assessment, and discusses some of the issues pertinent to future directions in this area. Because of space constraints, this article is not intended to be a comprehensive guide to assessment practices. Rather, it illustrates some emerging ideas and practices that show a great deal of promise. Readers who are interested in more comprehensive discussions about assessing children's social behavior are referred to other recent sources (e.g., Merrell, 1999, 2000; Merrell & Gimpel, 1998).

## METHODS OF ASSESSMENT

Most experts in social-emotional assessment of children and youth would agree that there are six primary methods of gathering this type of assessment information: behavioral observation, behavior rating scales, interviewing, self-report instruments, projective-expressive techniques, and sociometric techniques. Although these methods have their own unique set of advantages and disadvantages and are useful for various assessment purposes in their own right, they are clearly not equal when it comes to usefulness for assessing social skills of children and adolescents (Merrell, 1999).

Projective-expressive techniques such as thematic approaches to assessment, sentence completion techniques, and drawings may be useful for gaining rapport with a child and in developing hypotheses to guide further assessment, but one would be hard pressed to find even a shred of evidence that these techniques can be even minimally useful in identifying and classifying specific social skills.

Self-report instruments, which consist primarily of objective psychometric measures that compare children's responses on groups of self-report items to the responses of a normative group, increasingly have been found useful for evaluating internalizing problems such as depression and anxiety, self-concept, and general personality development. In the area of social skills assessment, however, very little has been done to demonstrate the effectiveness of self-report assessment. The only well-researched instrument of this type designed specifically for assessing social skills is the innovative self-report form of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990). As is evident from the data in the SSRS examiner's manual and in subsequent published studies, there is a great deal of divergence between self-reported social skills of children and youth and more direct measures of social skills, such as behavioral observation and behavior rating scales. At this point, it is unclear what part self-report evaluation should play in social skills assessment and whether youth who have significant social behavior problems can report accurately on their own social skills. Thus, self-report assessment of children's social skills should be considered an experimental method.

Sociometric assessment techniques, such as peer nomination, peer rating, peer ranking, and alternative procedures, have a long and rich history in the child development literature. It has been demonstrated convincingly that sociometric procedures have very high levels of reliability and validity and that they may be powerful predictors of future social outcomes (e.g., McConnell & Odom, 1986). However, two issues limit sociometric assessment of social skills. The first issue has to do with the constructs that are measured through sociometric assessment. It generally is understood that such procedures measure peer acceptance or social reputation and do not actually measure social skills, notwithstanding the obvious relation between social skills and peer acceptance (Landau & Milich, 1990). The second issue is pragmatic. Sociometric assessment is done in social groups (i.e., classrooms) rather than individually. It generally is acknowledged that obtaining parental and administrative consent to conduct sociometric assessments in schools is very difficult, because of concerns about the possibility of further social rejection of some students as a result of participating in the assessment. Additionally, effective sociometric assessment, to be valid, requires that the entire group (or almost all of it) be included in the assessment, thus necessitating separate informed consent from the parents of all students in a group (Merrell, 1999), a daunting task if the goal is social skills assessment of a select few students or an individual. In short, sociometric assessment can provide a tremendous wealth of information regarding children's social functioning, but because of the difficulties involved in this process, it should be used occasionally for specific purposes rather than as a first line of action for assessment.

Interviewing is perhaps the oldest and most widely used social-emotional assessment method. Interviewing is highly adaptable and can range from the use of highly structured interview schedules to free-flowing, unstructured, "stream of consciousness" interviews. Interviews may occur directly with the child or youth who is the target of assessment or with the child's parents and teachers. A particular advantage of interviewing to assess social-behavioral problems is that it is one of the few methods capable of providing relevant and functional information about the environment in which the behavioral problems occur, and thus the method may be linked effectively to intervention. Additionally, the flexibility of interviewing provides the possibility of role playing with children and thus may be useful in conducting analogue observations of children actually engaging in important social behaviors. However, the specific applications of interviewing in assessing children's social skills have not been articulated, researched, or condensed clearly into formats that lend themselves directly to consistent use by practitioners. Thus, interviewing, like sociometrics, should be considered an important secondary method of assessing social skills.

Thus, the first two methods that were discussed, projective-expressive techniques and self-report instruments, should be considered "third-line" methods for assessing children's social skills because of concerns regarding how directly and reliably these methods tap the construct in question. The third and fourth methods discussed, sociometric techniques and interviewing, may be very useful for assessing children's social skills in given situations but have inherent pragmatic problems that limit how broadly they should be used for this purpose. These two methods therefore should be considered "second-line" assessment methods for children's social skills. The obvious implication is that the remaining two methods, direct behavioral observation and behavior rating scales, should be consid-

ered the “first-line” methods for assessing children’s social skills. Applications of these two methods are illustrated in the following two sections.

### NATURALISTIC BEHAVIORAL OBSERVATION

Several years ago, Elliott and Gresham (1987) proposed that “analyzing children’s behavior in natural settings ... is the most ecologically valid method of assessing children’s social skills” (p. 96). Presumably, the analysis of behavior in naturalistic settings must rely heavily on the use of naturalistic direct behavioral observation, which has long been the assessment method of choice for behaviorally oriented practitioners and researchers. By definition, naturalistic behavior observation includes three key components: (a) observation and recording of behaviors at the time of occurrence in their natural settings; (b) the use of trained, objective observers; and (c) a behavioral description system that requires a minimal level of subjective inference by the observer–coders (Jones, Reid, & Patterson, 1979). There are other forms of behavioral observation in addition to the naturalistic type, such as analogue observation, which seeks to simulate the conditions of the natural environment in a controlled clinic or laboratory setting. However, the best evidence to date indicates that naturalistic observation is the preferred method of observation for most social skills assessment purposes.

For the assessment of children’s social skills using naturalistic behavioral observation, the most appropriate settings are those in which children commonly interact with peers. Thus, schools are particularly relevant places in which to assess social skills through naturalistic observation. Experts have long focused on the school settings in which there is more peer interaction and less structure as the best places for naturalistic observation of social skills. For younger students, recess times and the playground are ideal assessment settings. Finding such ideal assessment settings in secondary schools is somewhat more challenging, because of the usual lack of a general break or recess period and the pragmatic issues associated with conducting naturalistic observation in settings where the numbers in the student population sometimes might be overwhelming. One relatively unexplored school setting for naturalistic social observation of adolescents is the cafeteria or lunch period. Outside of school-based settings, there are certainly other possible settings for naturalistic behavioral observation of social skills, such as the home, community, and peer group. However, there is a practical reality to conducting naturalistic observations of children and youth outside of school settings. Although some researchers might be able to successfully tap into these nonschool settings for some very specific purposes, the vast majority of practitioners are unlikely to find such settings accessible, practical, or useful. Thus, for naturalistic behavioral observation assessment of the social behavior of children and adolescents, school settings are clearly the preferred choice on several counts.

To conduct a naturalistic behavioral observation of social skills, one first must determine an appropriate coding system to use in recording social behaviors. For all but the most esoteric purposes, assessment experts generally agree that there are five observational coding and recording procedures from which to choose. These include *event recording*, recording the number of times a specific behavior occurs during the entire

length of the observational period; *interval recording*, dividing the observational period into time intervals and recording the specified behaviors that occur at any point during the interval (partial-interval method) or during the entire interval (whole-interval method); *time-sampling recording*, dividing the observational period into intervals and recording specified behaviors that occur momentarily at the interval; *duration recording*, recording how long a particular behavior lasts; and *latency recording*, recording the amount of time from the end of one behavior to the beginning of another behavior. Although any of these five coding and recording procedures may be used in naturalistic observational assessment of social skills, some clearly have been used more widely for this purpose. Practitioners tend to rely on event recording and, to a lesser extent, interval recording. Researchers concerned with children's social skills have tended to rely more frequently on interval recording, particularly the partial-interval procedure, which is less likely to underestimate lower frequency behaviors than the whole-interval method.

Naturalistic behavioral observation, unlike some other assessment methods, does not require the use of specific instruments or tests. Instead, it relies on observer-constructed observational protocols, which are built on the basic empirical principles of naturalistic observation and which are designed specifically to meet the assessment needs of the situation. Thus, there are relatively few published assessment tools for naturalistic behavioral observation of children's social skills. An exception that is worth illustrating in more detail because it may serve as an excellent exemplar for observer-constructed tools is the Peer Social Behavior Code (PSBC), which is part of Walker and Severson's (1992) innovative *Systematic Screening for Behavior Disorders* (SSBD), a multiple-gating screening system for use with children in Grades 1–6. The PSBC consists of a series of 10-sec intervals. The actual number of intervals used in an observation varies depending on the situation. The recording forms include spaces for 40 different intervals. Observations are always conducted during free-play situations (e.g., during recess), and a typical observation period might last 15 min. Five recording categories are included in the PSBC: Social Engagement, Participation, Parallel Play, Alone, and No Codable Response. Behavior in the first two categories may be coded as either positive (+) or negative (-). Behavior in the Parallel Play and Alone categories is coded by simply checking the appropriate box. The No Codable Response category is coded with a check when the child is out of view and a dot when the child is interacting with an adult rather than a peer. After the observation, the observational data are transferred from the recording forms to an observational summary sheet. The number of intervals recorded is entered for each category, and then the percentage of time spent for each category is calculated by dividing the total number of intervals in the session into the intervals recorded under different categories and multiplying by 100. The SSBD manual contains thorough directions for interpreting PSBC observation data by using the normative tables provided. Extensive research went into the development of the PSBC, and the validity evidence and technical properties reported in the manual and in other sources are impressive. An excellent observer training tape is provided as part of the SSBD kit, which also includes an audio timing tape for accurately using the 10-sec intervals. The PSBC is an exemplary interval-based coding procedure for direct observation of children's social behavior and also may serve as a model for constructing similar coding systems for more specific purposes.

Despite the many advantages of and the strong empirical basis for using naturalistic behavioral observation to assess social skills of children and youth, there are also some significant potential problems. One of the most practical problems in using direct behavioral observation of any kind is the extensive time that it requires. The actual time spent conducting the observation is only one part of the time requirement. Successful use of naturalistic behavioral observation also requires extensive time “up front” in targeting the most important social behaviors, selecting and refining an appropriate coding and recording system, and training observers. A second potential concern for effective use of behavioral observation assessment involves some of the psychometric problems that may be caused by poor planning and implementation. I (Merrell, 1999; Merrell & Gimpel, 1998) identified seven such potential threats to the validity of behavioral observations, ranging from observer reactivity to biased expectations of observers. These potential validity threats, along with potential consequences and possible solutions, are summarized in Table 1. A third potential concern for effective use of naturalistic behavioral observation in assessing children’s social skills is the issue of how many observations are required for reliable and useful measurement. This issue has surfaced only

TABLE 1  
Some Potential Threats to the Validity of Behavioral Observations, and Possible Solutions to These Threats

<i>Problem</i>	<i>Potential Consequences</i>	<i>Possible Solutions</i>
Poorly defined observational domains	Observational recording system is either too cumbersome or too vague.	Carefully define and select behaviors to be observed based on assessment problem and intervention goals.
Unreliability of observers	Observers drift from original definitions; interrater reliability decreases.	Provide high quality initial training; conduct periodic reliability checks and retraining.
Lack of social comparison data	Interpretations of behavior are not based on a normative perspective; deviancy may be under- or overestimated.	Include typical or randomly selected participants in the same setting for behavioral comparison.
Observer reactivity	Participant behavior is influenced by the presence of the observer.	Select and participate in observational settings in a discrete, unobtrusive manner.
Situational specificity of behavior	Interpretations of observational data may not represent the larger picture.	Conduct observations in multiple settings; do not overgeneralize from limited data.
Inappropriate recording techniques	Behaviors are not adequately depicted; inappropriate conclusions are reached.	Select recording systems to carefully match the behavioral domain.
Biased expectations of the observer	Borderline behaviors may be systematically coded in a biased manner.	Resist pressure to confirm expectations of persons with vested interests; remain scrupulously objective in coding behavior.

*Note.* Adapted from *Behavioral, Social, and Emotional Assessment of Children and Adolescents* (p. 68), by Kenneth W. Merrell. Copyright © 1999 by Lawrence Erlbaum Associates, Inc. Reprinted with permission of publisher.

recently, based primarily on the findings of an intriguing study by Doll and Elliott (1994), who found that for young children, several observations (as many as six) over time may be necessary to measure social behavior reliably. The apparent reason that several observations may be required for reliable measurement is the situational specificity and reactivity of social behavior of young children. Although this was only one study limited to young children, it raised some important questions. The standard practice of having one observer conduct one 20- to 30-min observation of a student's social skills, no matter how elaborate and internally stable the observational system is, may yield data that are unreliable and therefore of questionable usefulness. Practitioners and researchers who desire to use naturalistic behavioral observation should understand that this time-honored assessment method, like other methods, is subject to problems that may render the obtained data useless and that a great commitment of planning, training, and implementation is required to combat these potential threats effectively.

### BEHAVIOR RATING SCALES

During the last decade or 2 of the 20th century, the fields of psychological and educational assessment witnessed tremendous advances in the availability and technical sophistication of rating scales designed for use by parents and teachers in assessing social-emotional behaviors of children and adolescents. Today, behavior rating scales are one of the most widely used methods of assessing child behavior, and new developments in behavior rating scale technology continue to build on previous advancements. Assessment of social skills (and social behavior in general) of children and adolescents is one area in which behavior rating scales have made a particularly strong impact in recent years and in which an impressive body of supportive empirical evidence has accrued (Merrell & Gimpel, 1998).

Behavior rating scales offer many advantages for clinicians and researchers in assessing social skills of children and adolescents. The following six points, which I have identified previously (Merrell, 1999), represent the most apparent advantages of behavior rating scales. First, in comparison with direct behavioral observation, behavior rating scales are less expensive in terms of professional time involved and amount of training required to use the assessment system. Second, behavior rating scales are capable of providing data on low-frequency but important behaviors that might not be seen in a limited number of direct observation sessions. Third, behavior rating scales constitute an objective assessment method that yields data that are more reliable than data provided through unstructured interviews or projective-expressive techniques, such as drawing tests, thematic techniques, sentence completion procedures, and so on. Fourth, behavior rating scales can be used to assess individuals who cannot readily provide information about themselves, such as children with limited verbal skills or youth who are extremely uncooperative. Fifth, behavior rating scales capitalize on observations over a period of time in a child or adolescent's natural environment (i.e., school or home settings). Sixth, behavior rating scales capitalize on the judgments and observations of persons who are highly familiar with the child or adolescent's behavior, such as parents or teachers, and thus are considered to be "expert" informants. These six advantages of using rating scales show

why this method has become so popular—rating scales provide a “big picture” of the assessment problem in a short amount of time, at moderate cost, and with a good deal of technical precision and practical utility.

Behavior rating scales have been developed to assess a broad array of social–emotional behavior of children and youth, ranging from general screening of behavioral concerns to attention deficit hyperactivity disorders to conduct problems. Assessment of social skills is one of the emerging targets of behavior rating scale technology. The first nationally standardized and psychometrically sound social behavior rating scales began to emerge in the 1980s, and there are currently several good choices available. Three of these instruments are discussed very briefly in this section, as examples, including the School Social Behavior Scales, SSRS, and Walker–McConnell Scales of Social Competence and School Adjustment. The major features of these three instruments are summarized in Table 2. Other sources have a more comprehensive review of the psychometric properties and specific applications of these three measures and similar instruments (e.g., Merrell, 1999; Merrell & Gimpel, 1998).

### School Social Behavior Scales

Designed specifically for use in educational settings, the School Social Behavior Scales (SSBS; Merrell, 1993) is a teacher rating scale aimed at evaluating social behaviors of students in kindergarten through Grade 12. The SSBS includes 65 items on two separate conormed scales—Scale A: Social Competence (32 items) and Scale B: Antisocial Behavior (33 items). All items are rated on a scale ranging from 1 (*never*) to 5 (*frequently*). The empirically derived structure for each scale includes three subscales. Scale A includes the Interpersonal Skills, Self-Management Skills, and Academic Skills subscales, as well as a total score. Scale B includes the Hostile–Irritable, Antisocial–Aggressive, and Demanding–Disruptive subscales, as well as a total score. All subscale and total scores are converted to *social functioning levels*, which serve as general indicators of the normative range of skill deficits or problem excesses. In addition, the total scores for each scale are converted to standard scores and percentile ranks. SSBS items and scales are scored by using a very simple scoring key that is printed on the rating form. Both scales of the SSBS were standardized with a sample of 1,855 students in kindergarten through Grade 12 from several U.S. states.

Internal consistency reliability of the SSBS total scores and subscales has been shown to range from .91–.98. Test–retest reliability at 3-week intervals is reported at .76–.83 for the Social Competence scores and .60–.73 for the Antisocial Behavior scores. Interrater reliability between resource room teachers and paraprofessional aides ranges from .72–.83 for the Social Competence scores and .53–.71 for the Antisocial Behavior scores. Validity of the SSBS has been supported through research documented in the manual and through several published studies. Convergent and discriminant construct validity of the SSBS has been demonstrated through correlational studies with several other behavior rating scales. Extensive research documenting the discriminating power and sensitivity to group differences with various educational groups has been published, including students in special education, gifted students, regular education students, and

TABLE 2  
Overview of Three Child Behavior Rating Scales Designed Specifically  
for Assessing Social Behavior

<i>Instrument</i>	<i>Publisher</i>	<i>Purpose</i>	<i>Norm Sample</i>	<i>Items and Subscales</i>
School Social Behavior Scales	Assessment-Intervention Resources, <a href="http://www.assessment-intervention.com">http://www.assessment-intervention.com</a>	Assessment of social competence and antisocial behavior; a parent report form is under development.	1,858 students ages 5–18	32 social competence items with 3 subscales; 33 antisocial behavior items with 3 subscales
Social Skills Rating System	AGS, 4201 Woodland Road, Circle Pines, MN 55014–1796; 800–328–2560	Assessment of social skills and academic competence, with a brief problem behavior screen; parent and self-report forms are also available.	Varies based on specific form; over 4,000 students for all versions combined.	Number of items varies slightly depending on form and age range—elementary teacher rating form includes 57 items with 6 subscales.
Walker–McConnell Scales of Social Competence and School Adjustment	Singular Publishing, 4101 West A Street, Suite 325, San Diego, CA 92101–7904; 800–521–8545	Assessment of teacher- and peer-preferred social competencies	Approximately 2,000 students in Grades 5–18	4-item elementary-age version includes 3 subscales; 53-item adolescent version includes 4 subscales.

at-risk students without identified disabilities. Additionally, the subscale structure of the SSBS has been supported through additional confirmatory factor analyses, which strengthen the theoretical foundation for the scale design.

The SSBS has the advantage of being focused exclusively on social behavior and providing a comprehensive screen of both social competence and antisocial behavior of students in school settings. It also is very easy to administer and score, is relatively brief, and focuses on routine social competencies and problems of youth rather than low-rate clinical behaviors. Because the SSBS is focused exclusively on social behavior, it should be supplemented with problem-specific assessment tools if problems such as attention deficit hyperactivity disorder or depression are a major concern. A version of the SSBS for use by parents and other community-based informants is currently under development (Merrell & Caldarella, 1999; Robbins & Merrell, 1998) and will be available soon.

### Social Skills Rating System

The SSRS (Gresham & Elliott, 1990) is a multicomponent social skills rating system aimed at behaviors that affect parent-child relations, teacher-student relations, and peer acceptance. The entire system includes separate rating scales for teachers and parents, as well as a self-report form for students. The teacher rating form of the SSRS, the specific focus of this overview, includes separate forms for ages 3-5, kindergarten through Grade 6, and Grades 7-12. The three forms are similar in organization, but some of the content differs according to developmental considerations of that age range. The elementary-level version includes 57 items divided in three scales: Social Skills (30 items), Problem Behaviors (18 items), and Academic Competence (9 items). For Social Skills and Problem Behaviors items, teachers respond to descriptions using a 3-point response format based on how often a given behavior occurs (0 = *never*, 1 = *sometimes*, and 2 = *very often*). On the Social Skills items, teachers are also asked to rate how important a skill is (on a 3-point scale) to success in the classroom. The importance rating is used not to calculate ratings for each scale but for planning interventions. On the Academic Competence scale, teachers rate students as compared with other students on a 5-point scale. All raw scores are converted to standard scores ( $M = 100$ ,  $SD = 15$ ). The combined SSRS forms were standardized by using a nationwide sample of more than 4,000 cases. The specific standardization groups for each of the three teacher rating forms are considered to be adequate to good.

For the teacher forms of the SSRS, reliability was measured using internal consistency (i.e., alpha coefficients ranged from .74 to .95), interrater, and test-retest (i.e., .75-.93 correlations across the three scales) procedures. Criterion-related and construct validity of these scales has been demonstrated through significant correlations between the SSRS and other rating scales. The empirically derived subscales for Social Skills (Cooperation, Assertion, and Self-Control) and Problem Behavior (Internalizing Problems, Externalizing Problems, Hyperactivity) are based on sound analytic procedures. The external research base on the SSRS is growing, and additional validity evidence continues to accrue.

The primary advantage of the SSRS is that it is an integrated multicomponent system of instruments for use by teachers, parents, and students, and it is one of the few social skills rating scales that include a comprehensive parent report version. The manual is very well written, and the rating instruments are easy to understand and use. Because the Problem Behavior scale items are few in number, this scale should be considered as a brief screen, and if possible problems are detected, a more thorough assessment with comprehensive problem behavior assessment tools is advised.

### Walker–McConnell Scales of Social Competence and School Adjustment

Like the SSBS, the Walker–McConnell Scales of Social Competence and School Adjustment (SSCSA; Walker & McConnell, 1995a, 1995b) are social behavior rating scales designed specifically for teachers and other school-based professionals. However, the SSCSA focuses exclusively on the domain of social competence and does not include an antisocial behavior screen. Two versions of the scale are available—an elementary version for use with students in kindergarten through Grade 6 and an adolescent version for use with students in Grades 7–12. The elementary version contains 43 positively worded items that reflect adaptive social–behavioral competencies within the school environment and includes three empirically derived subscales (Teacher-Preferred Social Behavior, Peer-Preferred Social Behavior, and School Adjustment Behavior). The adolescent version, which includes the same three subscales, includes the 43 items from the elementary version (although 9 are reworded to reflect adolescent-specific concerns), plus 10 additional items that comprise a fourth subscale designed to measure self-related social adjustment and empathy. The items are rated by using a 5-point scale ranging from 1 (*never occurs*) to 5 (*frequently occurs*). The scale yields standard scores on three subscales ( $M = 10$ ,  $SD = 3$ ) as well as a total score ( $M = 100$ ,  $SD = 15$ ), which is a composite of the three subscales. The scales were standardized with ratings of approximately 2,000 students representing all four U.S. geographical regions.

Studies undertaken during the development of the SSCSA that are cited in the technical manuals indicate adequate to excellent psychometric properties. Reliability of the scales was established by using test–retest (e.g., correlations of .88–.92 over a 3-week period with 323 participants), internal consistency (e.g., alpha coefficients ranging from .95 to .97) and interrater (e.g., a .53 correlation between teachers' and aides' ratings on the total score in a day treatment facility) procedures. Validity of the scales was evaluated by using a variety of procedures. Sensitivity of the scales to theory-based group differences was indicated in studies that found the scales to differentiate among groups of students who would be expected to differ behaviorally (behavior disordered and normal, antisocial and normal, behaviorally at risk and normal, and with and without learning problems). Criterion-related validity was demonstrated through significant correlations between the scales and various criterion variables, such as other rating scales, sociometric ratings, academic achievement measures, and a systematic behavioral screening procedure. Construct validity of the scales has been demonstrated by strong correlations between evaluative comments of students from their peers and teacher rat-

ings on the scales, and by the finding that low social skills ratings were strongly associated with the emergence of antisocial behavior in a longitudinal study of at-risk boys. The factor structure of the two scales has been shown to be robust.

Both versions of the SSCSA are brief, are easy to use, and contain items that are highly relevant for assessing social skills in educational settings. The research base behind the scales is exemplary. Because neither version of the SSCSA was designed to measure problem behaviors, these instruments should be supplemented with an appropriate problem behavior assessment if warranted by the referral issues.

## BEST PRACTICES

Effective assessment of children's social skills requires more than an understanding of the basic characteristics and the availability of good assessment methods and instruments. Although there have been significant advances in social skills assessment in recent years, there is still a great need to identify the specific ways in which assessment can be used more effectively regarding children and adolescents' social skills and related concerns. Perhaps more important, there is still a great need to find effective ways of linking social skills assessment to the solution of social-behavioral problems. The following list of best-practice suggestions has been derived from research and clinical experience in the application of social skills assessment processes (e.g., Caldarella & Merrell, 1997; Merrell, 1999, 2000; Merrell & Gimpel, 1998). It reflects an attempt to develop a conservative framework for effective social skills assessment that focuses primarily on the two methods emphasized in this article: naturalistic behavioral observation and behavior rating scales. These suggestions should not be considered absolute. It is quite possible that within another two decades, the field will have advanced to the point where more specific practices for improved assessment of children's social skills can be identified.

### Use Behavior Rating Scales Routinely for Early Screening

Effective screening practices involve systematically identifying with a high degree of accuracy children who may be in the early stages of developing behavioral, social, or emotional problems. Those children identified as such then are evaluated more thoroughly (i.e., by using naturalistic behavioral observation and other time-intensive assessment methods) to determine whether their problems warrant special program eligibility, intervention services, or both. The purpose of screening for social-emotional problems is usually for *secondary prevention*, preventing existing problems from becoming worse. Screening for early intervention is one of the best uses of behavior rating scales, given that they cover a wide variety of important behaviors and take very little time to administer and score. For general screening purposes, it is recommended that children or youth whose rating scale scores are one or more standard deviations above instrument normative means in terms of social problem excesses or social skill deficits be evaluated in more detail (Merrell, 1999). This practice will narrow the screening pool down to approximately 16% of the overall population, and this selected group then can be evaluated more

comprehensively. This sort of screening criterion typically will result in some false positive errors, or identification of some students who do not require further evaluation or services, but such errors should be easy to detect on further evaluation and consideration. Moreover, a screening criterion of 1 standard deviation seldom will result in false negative errors, or failure to identify children who are truly at risk. For screening purposes, false positive error is more tolerable than false negative error (Merrell, 1999).

### Use Behavioral Observation for Later Stages of Screening and Identification

Naturalistic behavioral observation of children's social behavior should be considered an absolutely essential assessment method when decisions are being made that involve classification, service eligibility, and specific intervention planning. However, as already discussed, this method is very time intensive and may require multiple applications across time and settings for optimum reliability. Therefore, a recommended practice is to use lower cost assessment methods (i.e., behavior rating scales) earlier in the screening and identification process and use behavioral observation later in the process, after the population and specific concerns have been narrowed. This practice is similar to the multiple-gating process employed in Walker and Severson's (1992) SSBD, wherein teacher ratings and rankings are used widely early in the process, and social behavior observations are used later in the process but on a much smaller number of youth who have been screened through the first two gates in the process.

### Use the "Aggregation Principle," But Carefully

The so-called aggregation principle involves obtaining assessment data from a variety of sources, across a variety of settings, and using a variety of methods (Merrell, 1999). Because of source and setting variance, each piece of information may present a slightly different picture. When observations and rating scales are used for purposes other than routine screening (in which one brief rating scale may be sufficient), obtaining assessment data that is aggregated in this manner is recommended to identify the particular behavioral excesses or deficits that are most consistent and troublesome across settings and from various perspectives. In other words, one should expect some differences in how various teachers rate the same student, even if the teachers are using the same tool for their ratings. Likewise, behavioral observations conducted across different school settings are likely to produce data that show different patterns of behavior. However, when there are significant behavioral problems or deficits, they likely will be consistent across settings, sources, and measures. In other cases, where a behavioral concern is truly situation specific, an aggregated assessment design may be of great help in determining which behaviors may be most problematic under specific conditions. Using an aggregated multimethod, multisource, multisetting assessment approach can help to identify the concerns that emerge the most consistently but also may pose some challenges for interpretation. Some practitioners will not have the skills to sufficiently integrate and interpret aggregated assessment data, and,

as a result, faulty conclusions may be reached. Therefore, it is important to use the aggregated data collection approach with a great deal of care and caution.

### Design Interventions to Match Specific Problems

A best practice in providing interventions for students who exhibit significant social skills problems and other challenging behaviors is to specifically match the intervention to the most critical problem excesses or skill deficits (Peacock Hill Working Group, 1991). This matching process is the opposite of a “one-size-fits-all” mentality for designing interventions, wherein any student who is identified as having significant behavioral concerns is recommended to receive a generic behavioral intervention of some kind. Such a global (if not vague) approach to intervention typically will fail to target the most important problems specifically enough, and the result typically will be weak or inconsistent intervention gains. A simple way to link observation and rating scale data to intervention in this problem-matched manner is to carefully review the assessment data and the specific clusters of skill deficits and problem excesses that emerge as the most significant concerns. If any of these clusters of concerns are found consistently across observers, raters, and settings, it is particularly important to specify them in designing the intervention. Once a narrowed list of key behaviors has been identified and targeted, specific intervention techniques that precisely match those behaviors can be selected. This particular method is very consistent with the Keystone Behavior Strategy, which has been touted as a promising way to effectively link assessment data to intervention planning (Nelson & Hayes, 1986).

### Always Consider the Environment in Planning Interventions

One of the advantages of using naturalistic behavioral observation as an assessment method is that it allows the practitioner to make inferences about the “function” of the problem behaviors, or the environmental conditions that elicit, shape, and maintain such problems. This sort of “ecobehavioral” approach to assessment moves the locus of the problem from residing only within the child to occurring within an environmental context that may be part of the problem. Thus, a primary advantage of naturalistic behavioral observation is that it is one of the few assessment methods that provide a clear avenue for identifying variables outside of the child that may be important in developing hypotheses about possible interventions. Although this process is often complex, certain applications of it are relatively simple, such as using event recording during the observation and then dividing all observation data into “A-B-C” columns on a sheet of paper, denoting their presumed function as antecedents (A), problem behaviors (B), or consequences of those behaviors (C; Merrell, 1999).

### Assess Progress During and After Intervention

Continuous assessment and monitoring of student progress after the initial assessment and during treatment are important in successfully implementing behavioral interventions (Kerr & Nelson, 1989). With respect to this type of progress monitoring, I believe that naturalistic behavioral observation will be very time intensive and difficult to use.

However, one easily might evaluate progress toward behavioral intervention goals by identifying weekly or biweekly problems using social behavior rating scales. Such continuous measurement may help in modifying the intervention if it does not produce the desired effect (Merrell & Gimpel, 1998). In reality, full-length rating scale measures may be too long and time consuming for frequent and repeated use. Under such circumstances, it is easy and reasonable to select a few critical behavioral items from the rating scales that were targeted for intervention after the initial assessment and use these few items as a brief informal measure of intervention progress. An additional comprehensive assessment after the intervention (including observation and rating scales) also can be a useful tool. The primary reason for conducting a follow-up assessment is to determine how well the intervention effects have been maintained over time (e.g., after 3 months) and how well the behavioral changes have generalized to other settings (e.g., other classrooms). Information gathered from a follow-up assessment may help determine whether additional interventions or “booster sessions” are needed.

## CONCLUSIONS AND FUTURE DIRECTIONS

With the advances in understanding, assessing, and treating children’s social skills have developed certain heuristics for research and practice. In addition, the advances that have occurred during the previous 2 decades have created a pressure for additional advances, particularly in linking research and practice. The five general methods of assessing children’s social and emotional behavior are not equal when it comes to viability for effective social skills assessment. Naturalistic behavioral observation and behavior rating scales should be considered first-line choices for assessing children’s social skills. Their ecological and social validity, coupled with the practical considerations in conducting a good assessment, clearly makes some combination of the two approaches the most defensible for both research and clinical applications. Following the best practice of using a multimethod, multisource, multisetting assessment design whenever feasible for evaluating child and adolescent social and emotional behavior, practitioners who are moving beyond basic screening processes should strive for social skills assessment designs that, at a minimum, include naturalistic behavior observations in at least two settings (or over time if multiple settings are not possible) and behavior rating scales completed by at least two informants. Ideally, the behavior rating scales should include the parents and teachers of the child as informants. The second-line assessment methods of interviewing and sociometrics may provide excellent additional information for the assessment of social skills and, in some cases, may even be essential. The third-line assessment methods of projective–expressive techniques and objective self-report instruments may shed some additional light on children’s social skills and related concerns but should never be used as a primary assessment source because of some technical and practical concerns.

A reasonable array of reliable and valid social skills assessment tools have been developed for both research and clinical purposes. Of course, additional improvements in this area are always needed and will be welcome additions to what is currently available. However, the major assessment challenge, as far as children’s social skills are concerned, is no longer the paucity of appropriate methods and tools for evaluation. Rather,

the present and future challenge in assessment is to find meaningful ways to make assessment results “functional,” in the sense of tying specific results to important social outcomes and to the development of effective instructional and therapeutic programs. The solution to this problem appears to be more elusive than was the solution to the absence of psychometrically sound assessment tools 2 or 3 decades ago. Moving the field to the point where we can easily link assessment processes to effective intervention surely will require a concerted and sustained effort by research and development specialists and practitioners.

## REFERENCES

- Caldarella, P., & Merrell, K. W. (1997). Common dimensions of social skills of children and adolescents: A taxonomy of positive behaviors. *School Psychology Review, 26*, 265–279.
- Doll, B., & Elliott, S. N. (1994). Representativeness of observed preschool social behaviors: How many data are enough? *Journal of Early Intervention, 18*, 227–238.
- Elliott, S. N., & Gresham, F. M. (1987). Children’s social skills: Assessment and classification practices. *Journal of Counseling and Development, 66*, 96–99.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Jones, R. R., Reid, J. B., & Patterson, G. R. (1979). Naturalistic observation in clinical assessment. In P. McReynolds (Ed.), *Advances in psychological assessment* (Vol. 3, pp. 42–95). San Francisco: Jossey-Bass.
- Kerr, M. M., & Nelson, C. M. (1989). *Strategies for managing behavior problems in the classroom* (2nd ed.). Columbus, OH: Merrill.
- Landau, S., & Milich, R. (1990). Assessment of children’s social status and peer relations. In A. M. LaGreca (Ed.), *Through the eyes of the child* (pp. 259–291). Boston: Allyn & Bacon.
- McConnell, S. R., & Odom, S. L. (1986). Sociometrics: Peer-referenced measures and the assessment of social competence. In P. Strain, M. J. Guralnick, & H. M. Walker (Eds.), *Children’s social behavior: Development, assessment, and modification* (pp. 215–284). New York: Academic.
- Merrell, K. W. (1993). *School Social Behavior Scales*. Iowa City: Assessment-Intervention Resources. Retrieved April 4, 2001 from the World Wide Web: <http://www.assessment-intervention.com>
- Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children and adolescents*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Merrell, K. W. (2000). Informant report: Rating scale measures. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessment of child and adolescent behaviors* (pp. 203–234). New York: Guilford.
- Merrell, K. W., & Caldarella, P. (1999). Social-behavioral assessment of at-risk early adolescent students: Validity of a parent report form of the School Social Behavior Scales. *Journal of Psychoeducational Assessment, 17*, 36–49.
- Merrell, K. W., & Gimpel, G. A. (1998). *Social skills of children and adolescents: Conceptualization, assessment, treatment*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Nelson, R. O., & Hayes, S. C. (Eds.). (1986). *Conceptual foundations of behavioral assessment*. New York: Guilford.
- Peacock Hill Working Group. (1991). Problems and promises in special education and related services for children and youth with emotional or behavioral disorders. *Behavioral Disorders, 16*, 299–313.
- Robbins, R., & Merrell, K. W. (1998). Cross-information comparisons of the Home and Community Social Behavior Scales and the School Social Behavior Scales. *Diagnostic, 23*, 204–218.
- Walker, H. M., & McConnell, S. (1995a). *Walker–McConnell Scale of Social Competence and School Adjustment, Adolescent Version*. San Diego, CA: Singular.
- Walker, H. M., & McConnell, S. (1995b). *Walker–McConnell Scale of Social Competence and School Adjustment, Elementary Version*. San Diego, CA: Singular.
- Walker, H. M., & Severson, H. (1992). *Systematic Screening for Behavior Disorders*. Longmont, CO: Sopris West.

Copyright of Exceptionality is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.