Published in final edited form as:

Semin Speech Lang. 2019 August ; 40(4): 272–290. doi:10.1055/s-0039-1692963.

Redmond (2002) Revisited: Have Standardized Behavioral Rating Scales Gotten Better at Accommodating for Overlapping Symptoms with Language Impairment?

Sean M. Redmond, Ph.D.[1], **Kirsten M. Hannig, B.A.**[1], **Amy Wilder, M.S.**[1]

[1]Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, Utah.

Abstract

Seventeen years ago, Redmond reviewed five standardized behavioral rating scales and identified several aspects of their design that made them prone to mischaracterize language impairments as socioemotional behavioral disorders. The purpose of this report is to provide an update and extension of the original audit. We consulted test manuals to evaluate: (1) representation of children with language impairments in their standardization samples; (2) presence of language, or academic items within their inventories; (3) accommodations for administering the measure to children with language impairments; and (4) procedures for identifying inordinately punitive ratings. Overlapping language and academic symptoms continued to be a problem across current behavioral rating scales. Improvements since Redmond occurred in the representation of children with language impairments in standardization samples and in procedures for identifying inordinately punitive ratings. We discuss implications for clinical assessment, research programs, and instrument development.

Keywords

differential diagnosis; specific language impairment; ADHD; behavioral rating scales; reliability and validity

Suspected socioemotional and behavioral disorders, such as attention-deficit hyperactivity disorder (ADHD), social anxiety disorder, and conduct disorder, represent some of the most common reasons children are referred for clinical evaluations.[1] These disorders rarely occur in isolation. For example, ADHD in the presence of comorbid disorders is considerably more prevalent than the presentation of ADHD alone.[2] Studies provide evidence that language impairments in particular commonly co-occur in children with socioemotional and behavioral disorders, although there has been considerable variability across study samples.[3–7] Clear phenotypic boundaries and psychometrically robust measurement systems

Address for correspondence: Sean M. Redmond, Ph.D., Department of Communication Sciences and Disorders, University of Utah, 390 South 1530 East BEHS 1201, Salt Lake City, UT 84105 (sean.redmond@health.utah.edu).

Executive Function Skills in Children with Language Differences and Disorders; Guest Editor, Margarita Kaushanskaya, Ph.D.

NONFINANCIAL DISCLOSURES
No relevant nonfinancial relationships exist for any of the authors.

CONFLICT OF INTEREST
None declared.

represent prerequisites for examining associations between language impairments and socioemotional and behavioral disorders. The presence of imprecision in clinical tools compromises differential diagnosis, the identification of comorbidity, and the evaluation of theoretical models of co-occurrence.

Seventeen years ago, Redmond[8] presented a critical review of assessment practices associated with the identification of socioemotional and behavioral disorders in children. The central thesis of the review was that, relative to other methods, standardized behavioral rating scales represented the best choice for identifying co-occurring socioemotional and behavioral disorders in children with language impairments. For example, unlike clinical interviews, self-reports, and projective techniques, standardized behavioral rating scales do not require children to be fluent in their retrieval of words or to be able to produce complex sentences in response to interview questions that they might not understand. However, even these instruments were prone to mischaracterize language impairments as potential symptoms of socioemotional and behavioral disorder.

To support his conclusion that standardized behavioral rating scales were prone to potential language bias, Redmond[8] reviewed five commonly used standardized behavioral rating scales across four dimensions: (1) the representation of children with speech and language impairments in their standardization samples; (2) the presence of speech, language, or academic items within their inventories; (3) accommodations for administering the instrument to children with language impairments; and (4) procedures for identifying inordinately punitive ratings. This last item was included in the review for two reasons. First, the tendency of raters to assign pejorative ratings because the individual they are rating possesses a negative or undesirable characteristic (e.g., obesity or membership in a particular ethnic or cultural community), represents a widely recognized source of potential bias in socioemotional and behavioral rating scales.[9] Item validity checks and other procedures have been developed by rating scale designers to guard against both inordinately negative and positive ratings. Second, communication disorders in general and language impairments in particular represent negative traits that appear to trigger these kinds of biases in raters. For example, when asked to judge demographic, intellectual, personality, and moral attributes of speakers with communication disorders based solely on exposure to brief samples of connected speech, raters have consistently provided pejorative attributions—even when the raters were speech–language pathologists.[10–12]

Across the five rating scales evaluated, Redmond[8] found that children with language impairments and/or learning disabilities were either entirely excluded from or underrepresented in the standardization samples. This was particularly troublesome for language impairments given their relatively high prevalence rates, as confirmed by epidemiological study samples, and yet low levels of identification.[13–17] In addition, all five standardized rating scales reviewed included several items in their inventories that could be characterized as either symptomatic of a primary language impairment or representative of a secondary academic consequence (e.g., “Can’t talk”; “Doesn’t speak clearly”; “Has trouble following directions”; “Difficulty doing homework”). These items typically loaded onto ADHD, Social Problems, or Immaturity subscales across these clinical instruments. None of the standardized rating scales provided accommodations for the presence of speech

and/or language impairments. Two of the standardized rating scales, however, the *Revised Behavior Problem Checklist* (RBPC)[18] and the *Behavior Assessment System for Children* (BASC),[19] provided assessment teams with disaggregated norms for children with identified learning disabilities. The BASC was the only standardized rating scale reviewed to utilize a validity check, referred to in the instrument as the "faking bad" or F index, which measured a respondent's tendency to assign excessively negative ratings. The presence of internal checks on rating scales would help assessment teams interpret the meaning of divergent evaluations collected across raters, which is likely to occur with children who have language impairments. For example, in a longitudinal investigation of students from K-2nd grade, Redmond and Rice[20] found significant differences between the socioemotional and behavioral ratings collected on children with specific language impairment (SLI) and typically developing (TD) children from parents and teachers using the *Child Behavior Checklist* (CBCL)[21] and the *Teacher Report Form* (TRF)[22] from the Achenbach System of Empirically Based Assessment (ASEBA). Specifically, teachers, but not parents, identified significantly more children with SLI than children with TD as being in the clinical range across multiple behavioral subscales. Redmond and Rice[20] also found inconsistencies across ratings collected from different teachers over different grades. It was common, for example, in the dataset for one teacher to assign ratings within the clinical range for a student with SLI and then for another teacher to assign values within the normal range for the same student. The inclusion of something like the BASC's F index into the CBCL and TRF rating scales would have helped identify potential sources of variability between the parent and teacher groups and variability within the group of teachers.

In the spirit of offering suggestions toward improving diagnostic practices, Redmond[8] provided assessment teams with a protocol to apply to socioemotional behavioral rating scales when children either have or are suspected of having language impairment. A key element of the protocol involved the removal of individual items from rating scale inventories that overlapped with language and academic symptoms, prior to the calculation of clinical scores. At the time, this accommodation could only appeal to the logic behind reasonable accommodation. Data were not available to examine directly the extent to which adjusting rating scales for language bias would compromise them. For example, would removing overlapping symptoms from ADHD syndrome scales decrease their sensitivity for identifying ADHD? To address this question, Redmond and Ash[23] examined the consequences of removing language and academic items from the CBCL and the parent version of the *Conners Rating Scales-Revised* (CRS-R)[24] to their sensitivity to detect cases of ADHD. The study sample consisted of 20 children with SLI, 20 children with ADHD, and 20 children with TD. Results indicated that removal of the language and academic items resulted in improved specificity for discriminating cases of ADHD from SLI (especially for the CRS-R) but had very little impact on either instrument's sensitivity for discriminating cases of ADHD from TD. In other words, removal of language and academic items from the inventories of these two commonly used rating scales generally improved rather than compromised their diagnostic accuracy.

A lot has changed in both communication disorders and child psychopathology since 2002. Both disciplines have experienced taxonomic, diagnostic, as well as therapeutic realignments. To keep up with these changes, test developers have provided new editions for

some of the socioemotional behavioral rating scales Redmond[8] reviewed, bringing them into stronger alignment with current practices. Other rating scales included in the Redmond [8] review, the *Louisville Behavior Checklist*[25] and the *Revised Behavior Problem Checklist*,[18] have fallen out of favor and exist now primarily as historical examples.

The purpose of this report is to examine the extent to which these changes have translated into improvements on the issue of differential diagnosis of language impairments and socioemotional behavioral disorders. Have standardized socioemotional behavioral rating scales gotten better at accommodating for overlapping symptoms with language impairment? To address this question, we brought in current editions of three of the original five behavioral rating scales, replicating the Redmond[8] audit of the BASC, CBCL, and CRS. We added four new socioemotional behavioral rating scales into our review to more accurately reflect current practices in research studies and clinical settings. These included a general behavioral rating scale widely used in research studies as well as three scales designed to target specific clinical designations. Because Redmond[8] identified ADHD as a condition particularly prone to overlapping symptoms with language impairments, we included rating scales targeting symptoms of ADHD, the related attention disorder of sluggish cognitive tempo (SCT),[26] and children's executive function deficits (the Strengths and Difficulties Questionnaire [SDQ][27]; the ADHD Rating Scale-5 [ADHD RS-5][28]; Barkley Sluggish Cognitive Tempo Scale [BSCTS][29]; and the Behavior Rating Inventory of Executive Function, Second Edition [BRIEF-2][30]).

We used a consensus process and arrived at an agreed upon set of inventory items across the standardized rating scales that we considered potentially representative of primary language symptoms or their secondary academic consequences. Table 1 summarizes our reviews for the ASEBA (CBCL and TRF), BASC-3, and CRS-3 rating scales. Table 2 provides a summary for the SDQ, ADHD RS-5, BSCTS, and BRIEF-2.

UPDATED AUDITS ON NEW VERSIONS OF THE CHILD BEHAVIOR CHECKLIST/TEACHER REPORT FORM, BEHAVIORAL ASSESSMENT SYSTEM FOR CHILDREN, AND CONNERS RATING SCALES

Child Behavior Checklist and Teacher Report Form: The Achenbach System of Empirically Based Assessment (ASEBA) is a battery of behavioral observations, behavioral ratings, self-reports, and clinical interviews. The ASEBA is a well-regarded international standard of clinical assessment that regularly appears in research reports (https://aseba.org/bibliography/). Translations of the ASEBA elements are available to assessment teams in more than 75 languages.[31,32] The current version of the ASEBA system includes a preschool parent scale, a preschool teacher scale (1½ to 5 years), a school-age parent scale, and a school-age teacher scale for children[33] (6–11 and 12–18 years). The preschool version of the CBCL (1½ to 5) was updated to include the co-normed *Language Development Survey* (LDS).[34] The LDS screens for the presence of developmental delays by tracking children's language milestones and includes a vocabulary checklist. Items on the LDS are scored separately from the other syndrome scales and do not load onto other clinical scales. With the exception of the LDS on the CBCL 1½ to 5, all other clinical scales use a three-point

scale to describe the frequency of problematic behavior: 0 = *not true,* 1 = *somewhat or sometimes true*, 2 = *very true or often true.*

Standardization of the CBCL and TRF scales was updated in 2001 using a larger, more representative national sample of children from geographically, ethnically, and socioeconomically diverse populations. Data from nonreferred children within this sample were used to create norms for the Adaptive and Competence scales, the Syndrome scales, DSM-oriented scales, and the LDS. Minor changes in wording occurred for all forms and an expansion to the scoring process for Adaptive (TRF) and Competence (CBCL) scores was added to the current versions for greater clarity. Overlapping items across multiple syndrome scales, present in previous versions, were eliminated. Thresholds of risk for syndromes for all ages were also lowered. These adjustments mean, relative to older versions of these rating scales, a smaller number of symptom endorsements is needed to reach clinical levels in the current versions. In terms of the syndrome scales, the 1991 versions and current versions are very similar. Second-order factors of internalizing and externalizing symptoms for both the CBCL and the TRF did not change from their 1991 factor structure. Some of the individual inventory items were changed; however, according to ASEBA manuals, most children would receive the same syndrome ratings whether they were using the 2001 or 1991 scales.

Adjustments and supplements to the 2001 CBCL and TRF continue to occur. In 2007, multicultural norms were published as a supplement to the current manual[32] to reflect differences in cultures across the world.[31] Grouped by country, commonwealth, and, sometimes, partial territories within a country, these data allow assessment teams to more directly compare children's ratings to children from similar backgrounds. The multicultural supplement also includes norms for additional problem scales including obsessive-compulsive problems (OCP), posttraumatic stress problems (PTSP), SCT, and positive qualities (PQ). With the release of the DSM-5, alterations to the DSM-oriented scales were made to reflect the newly published criteria.[34] These adjustments were minor; the Autism Spectrum Problems scale replaced Pervasive Development Problems, and items on the Anxiety Problems and Somatic Problems scales were updated.

Test–retest reliability for the 2001 versions of the CBCL and TRF for children 1½ to 5 years ranged from adequate to excellent (*r*-value range: 0.57–0.92), while the forms for children aged 6 to 18 reported good to excellent coefficients (*r*-value range: 0.85–0.95) across all scales. Reports of inter-rater reliability were poor for the TRF across all ages (range: 0.20–0.79), whereas the CBCL inter-rater reliability was adequate, ranging from 0.48 to 0.88, with lower values on the range generally reported for the preschool CBCL form. Cross-informant agreement reported between the CBCL and TRF for ages 6 to 18 was poor, ranging from 0.29 to 0.35 across all scales. According to the manuals, follow-up interviews are recommended to resolve rating differences when cross-informant agreement is low. The manuals provide little guidance, however, on how to conduct these interviews in a way that identifies sources of disagreement.

Items from the 1991 versions of the CBCL and TRF that represented potentially overlapping language symptoms were retained in the current versions ("Acts young," "Speech problems," "Won't talk," "Difficulty following directions," "Has difficulty learning"). New

items added to the current CBCL and TRF that represent potentially overlapping language symptoms include “Poor schoolwork,” “Talks out of turn,” “Messy work,” “Talks too much,” and “Fails to carry out assigned tasks.” These items loaded onto the Anxious/Depressed, Withdrawn/Depressed, Social, and Attention Problems syndrome scales. Of importance, consistent with earlier versions, the current protocols do not provide assessment teams with methods for identifying potential rater bias nor do they provide guidelines for adjusting syndrome scores when language impairments are expected to be a contributing factor.

Behavioral Assessment System for Children, Third Edition: The BASC-3[36] is also a comprehensive assessment battery composed of many components: a teacher, parent, and self-report rating scale, parent relationship questionnaire, behavioral and emotional screenings, a structured developmental history, a behavioral observation form, and tools for intervention. The parent and teacher report scales (PRS and TRS) of the BASC-3 are used for children aged 2 to 18 years. Specific forms for assessment vary according to age and include preschool (PRS-P/TRS-P; ages 2–5), child (PRS-C/TRS-C; ages 6–11), and adolescent (PRS-A/TRS-A; ages 12–18) versions. The BASC-3 uses a four-point rating scale. Possible responses include *0 = Not very true at all, 1 = Just a little true, 2 = Pretty much true,* and *3 = Very much true.* The number of items on the PRS ranges between 139 and 175, depending on the age of the child, and TRS forms include between 105 and 165 items. The current edition of the BASC includes five scales: clinical, adaptive, content, composite, and probability indexes. Probability indexes were added to the current edition to facilitate identification of difficulties associated with specific conditions such as ADHD, emotional/behavioral disorders, autism spectrum disorders, functional impairment, and general clinical probability.

Validity scales, present on the earlier version of the BASC, were expanded in the BASC-3. The F index, carried over from earlier versions, provides assessment teams with a method for identifying negative biases in either parental or teacher ratings. The response pattern index identifies the possibility of inattentiveness on the part of the respondent, which if excessive would invalidate their responses. The consistency index detects circumstances of inconsistency for paired items. In cases for which these indices identify bias, it is recommended that assessment teams follow up with the respondents to determine the validity of responses. Test–retest reliability for the parent scales on the BASC-3 ranged from good to excellent (*r*-value range: 0.80–0.94) which represents an improvement over previous editions. Inter-rater reliabilities were modest and ranged from 0.47 to 0.87, with higher values associated with the PRS-A, reflecting increasing reliability as children age.

The third edition of the BASC includes gender-specific and gender-combined norms for three types of groups: the general population, the clinical population, and the ADHD population. General population norms included children with specific learning disorder (~ 5%) and children classified as having “other” difficulties (~ 1%; predominantly speech and/or language disorders). The inclusion of disaggregate norms for various clinical populations makes the BASC-3 unique, compared with other behavioral assessments. Clinical norms for 4- to 18-year-olds were based on children who were already classified as having behavioral and/or emotional problems and/or those receiving special education

services. Children with specific learning disorder represented 2 to 14% of the clinical norms included across different age bands, and 10 to 43% of children were reported as having “other difficulties” (predominantly speech and language disorders). Although the disaggregated norms provided are less than optimal for the purposes of directly comparing individuals with language impairments to a comparison group of other children with language impairments, the level of representation on the BASC-3 is a considerable improvement over previous editions.

Several items on the parent and teacher forms of the BASC-3 were language related. Many of these items fell within the area of Functional Communication scale which was designed to screen for problems in communication (“Provides full name when asked,” “Provides home address when asked,” “Starts conversations,” “Is able to describe feelings accurately,” “Has trouble getting information when needed,” “Likes to talk about his or her day,” “Responds appropriately when asked a question,” “Communicates clearly,” “Accurately takes down messages,” “Is unclear when presenting ideas,” and “Speaks in short phrases that are hard to understand”). This collection of language symptoms is scored separately from socioemotional and behavioral disorder in a manner that is similar to how the preschool version of the CBCL incorporates items in the LDS. However, beyond the Functional Communication Scale, language-related items appear on the BASC-3: Hyperactivity, Aggression, Attention, Atypicality, Withdrawal, Social Skills, and Executive Function Scales. Items on these scales include “Pays attention,” “Listens to directions,” “listens carefully,” “Pays attention when being spoken to,” “Speech is confused or disorganized,” “Says things that make no sense,” “Begins conversations appropriately,” Answers telephone properly,” “Politely asks for help,” “Communicates clearly,” “Compliments others,” “Says all the letters of the alphabet when asked,” “Readily starts up conversations with new people,” “Interrupts others when they are speaking,” “Talks over others,” “Has difficulty explaining rules of games to others,” “Is effective when presenting information to a group,” “Starts conversations,” “Gives good suggestions for solving problems,” “Tracks down information when needed,” “Is able to describe feelings accurately,” and “Babbles to self.” Inclusion of these items could unduly penalize children with speech and/or language problems and lead to misclassification.

Conners Rating Scales 3rd Edition: The CRS-3[37] represents a culmination of research on ADHD and its comorbid disorders in children and adolescents, aged 6 to 18. A 2008 update of the CRS-R,[24] the CRS-3 was designed to reflect the DSM-IV-TR diagnostic criteria for ADHD and its most closely comorbid disorders, including disruptive behavior disorders, anxiety, depression, and severe conduct problems.[37] The CRS-3 was updated again in 2014 to align more closely with DSM-5 criteria.[38] The CRS-3 uses cross-informational data from parents, teachers, and self-report forms (long and short versions) to assist in the identification of attentional problems across contexts (home, community, academic).

The CRS-3 boasts a normative sample of more than 2,300 ratings collected from parents and teachers. Its sampling distribution reflects an ethnically diverse set of ratings collected on children across several regions of the United States. The CRS-3 is normed separately for males and females, aged 6 to 18. Children with learning disorders (i.e., reading, mathematics, written expression, and multiple learning disorders) represented approximately

5% of the normative sample for both forms, and 16 to 17% of the clinical sample for both parent and teacher forms.

The CRS-3 includes symptom scales of ADHD inattentive, ADHD hyperactive-impulsive, conduct disorder, and oppositional defiant disorder. New features include validity scales, executive function assessment, as well as additional items on the impairment scales. Three new validity scales are designed to identify excessively positive or negative responses and other inconsistencies that could indicate rater bias: the *Positive Impression* scale (PI), *Negative Impression* scale (NI), and *Inconsistency Index* (IncX). The PI and NI indicate the possibility that the rater demonstrated an inordinately positive or negative view of the child's behavior. The PI and NI are available on both parent and teacher forms. The IncX is found only on the parent long form and identifies instances of inconsistencies across similar items. When these indices suggest potential rater bias, the manual recommends follow-up interview and observation to determine the validity of the ratings.

The addition of new content items to the parent and teacher forms of the CRS-3 allows for the evaluation of executive functioning in children. Additionally, new content items on the impairment scales are designed to identify contexts in which attention and behavior problems exist (e.g., classroom, playground, home) and the degree to which they may interfere with children's daily functioning. The index scales for ADHD and global scores have been renamed to reflect updated norms. Otherwise, the information found within these scales is the same as previous editions. The CRS-3 was also adjusted to align better with commonly used eligibility requirements for educational intervention. The assessment of ADHD inattentive symptoms separately from learning disorders within the content scales represents an important adjustment in the CRS-3. The category of emotional issues, found in previous additions, was removed from the parent and teacher questionnaires of the CRS-3 and placed in the Conners Comprehensive Behavior Rating Scales (CBRS). DSM-5 symptom scales are included in the long forms only. ADHD and Global Index scales are included in the long forms but are also available on their own forms. Alignment of content items across parent and teacher forms was updated and improved for the current edition of the Conners rating scales, to facilitate more effective comparison of scores across informants. Screener items for anxiety and depression were also included.

The CRS-3 forms include 110 and 115 content items for the parent and teacher long forms, respectively. The short forms contain 45 and 41 content items for parent and teacher forms. The ADHD and Global Index scales each contain 10 content items. To more closely align the parent and teacher long and short forms, additional content items were added to the current version of the Conners. However, these additions increased the number of language-related items, such as "Spelling is poor," "Has trouble reading," "Does not understand what he/she reads," and "Needs extra explanation of instructions." Language-related items from the previous version, such as "Forgets things he/she has learned," "Does not seem to listen," and "Fails to finish things they start" remain on the current versions. The majority of the language-related items on the inventories load on to the Learning Problems Content scale. However, some language-related items load on to other content scales, such as executive functioning, family relations, hyperactivity/impulsivity, and inattention. These items include "Does not seem to listen to what is being said to him/her,"

"Fails to complete schoolwork, chores or tasks," "Forgets instructions quickly," and "Needs extra explanation of instructions." The CRS-3 does not provide assessment teams with accommodation guidelines for accommodating for the presence of language impairments in children's ratings.

Reliability of the Conners-3 improved significantly relative to previous editions. Test–retest coefficients ranged from good to excellent (*r*-value range: 0.72–0.98) for all forms and scales. Inter-rater reliability examining parent consistency ranged from good to excellent (*r*-value: 0.74–0.94), while teacher consistency ranged from poor to good (*r*-value: 0.55–0.82). Parent to teacher agreement also ranged from poor to good (*r*-value: 0.52–0.67), indicating moderate agreement can be expected across home and school contexts.

In sum, all three standardized rating scales previously audited in Redmond[8] demonstrated marked psychometric improvements over the course of their revisions. These improvements were primarily in the areas of normative sampling, reliability, and stronger alignment with the DSM-5 taxonomy.[39] The BASC-3 and the CRS-3 increased their representations of children with language impairments and/or learning disorders considerably. When they are available, assessment teams should use the disaggregate norms provided by these instruments to make more informed clinical decisions in cases of suspected or known language impairment. The BASC-3 and the CBCL have extended their coverage of clinical symptoms to include separate language scales, reflecting a recognition that language represents a separate domain. However, all three rating scales continued to include several items in their socioemotional behavioral scales that potentially overlap with either primary language impairments or their secondary academic consequences. The CRS-3 joined the BASC-3 in providing assessment teams with internal checks to guard against inordinately negative ratings, a noted concern when asking adults to rate children with language impairments and other communication disorders. Manuals of the newer versions of the ASEBA, BASC, and CRS each provided more discussion of the need for assessment teams to consider divergent ratings from different informants relative to earlier editions. However, there is still room for improvement. None of the revised versions offered assessment teams much in the way of explicit guidance on how to interpret divergent ratings. Furthermore, the ASEBA, BASC, and CRS remain silent on the need to make accommodations for children's receptive or expressive language difficulties.

CONSIDERATION OF ADDITIONAL STANDARDIZED RATING SCALES: ADHD RATING SCALE-5, BARKLEY SCT SCALE, BRIEF-2, SDQ

ADHD Rating Scale-5: The ADHD RS-5[28] measures symptoms and functional impact associated with ADHD in children aged 5 to 17 years. This scale is widely used to screen for ADHD, diagnose ADHD, and evaluate treatments of ADHD. The scale includes four forms, a child form (ages 5–10) and an adolescent form (ages 11–17) for both the home version, which is completed by parents, and the school version, which is completed by teachers. The rating scale items are based on the diagnostic and DSM-5 criteria for ADHD. Each form contains essentially the same 18 behavioral items, with slightly different wording, across the

child and adolescent forms. Likert scales of *never or rarely, sometimes, often,* or *very often* are used to describe children's behavior over the previous 6 months.

ADHD RS-5 items load onto two subscales: the Inattention Subscale and the Hyperactive-Impulsive subscale. A third scale, the Functional Impairment scale, is completed twice: first after rating the nine inattention items and then again after the nine hyperactive-impulsive items. The Functional Scale measures the impact of specific ADHD behaviors across six domains including relationships with family members, or teachers for the school version; relationships with other children/teenagers; homework functioning; academic functioning; behavioral functioning; and self-esteem. These domains are rated using a four-point scale of *no problem, minor problem, moderate problem, severe problem.*

The normative sample for the home version included 2,079 randomly selected participants aged 5 to 17 years with parent/guardian respondents aged 20 to 77 years. The sample was selected to approximate the 2010 U.S. census data for ethnic group, region, and income distributions. The inclusion of children with language impairments or learning disabilities in the home version normative sample is not reported. The school version normative sample included 2,140 participants, aged 5 to 17 years, selected to approximate the 2010 U.S. census data for region and ethnic group. The respondents included 1,070 predominately Caucasian teachers. Children who were receiving special education services comprised 16.4% of the sample.

Internal consistency among the behavioral items and the impairment items is high with comparative fit indices ranging from 0.92 to 0.98. Test–retest reliability ranges from moderate to strong (Pearson's correlations = 0.61–0.87) for the behavioral scales and from weak to strong (Spearman's Rho correlations = 0.14–0.90) for the functional impact scales, with lower reliability on the adolescent home version. Interrater agreement between parents and teachers ranges from weak to moderate (Pearson's correlations = 0.01–0.77) for the behavioral scales (Spearman's Rho correlations −0.06–0.77) for the impact scale. Four of the behavioral items on the ADHD RS-5 overlap with speech, language, or learning disabilities, including "Fails to give close attention to details or makes careless mistakes in schoolwork," "Does not seem to listen when spoken to directly," "Does not follow through on instructions and fails to finish work," and "Avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (e.g., schoolwork or homework)." All four of these items are part of the Inattention Subscale. These items also influence values assigned to the Functional Impact scale. The manual does not provide any guidelines for making accommodations for children with language or learning disabilities.

Barkley Sluggish Cognitive Tempo Scale: The BSCTS[29] is a parent rating scale for measuring attention deficits in children aged 6 to 17 years. The clinical designation SCT, first coined by Lahey and colleagues,[40] is meant to capture a second type of attention disorder, conceptualized as relatively distinct from but often co-occurring with ADHD. SCT, also referred to in the literature as concentration deficit disorder (CDD), is characterized by a combination of symptoms including hypoactivity, daydreaming, trouble staying awake/alert, mental fogginess, and staring a lot.[29] The BSCTS contains 12 behavioral items that parents/guardians rate using a Likert scale of *never or rarely, sometimes, often,* or *very*

often. The BSCTS provides two subscales, SCT Daydreamy and SCT Sluggish, as well as a composite SCT total score.

The normative sample included 1,800 parents (900 fathers and 900 mothers) of children aged 6 to 17 years. Variability across geographical regions, educational attainment, and socioeconomic status of parents in the normative samples was similar to the 2000 U.S. census. The sample included parental ratings of children with language delay (5%), children receiving special education services (9.4%), and children receiving speech and language therapy (11.4%). The manual provides disaggregated raw score means and standard deviations for each disordered group included in the normative sample. The presence of speech, language, or learning disorder, however, does not require an adjustment in children's standard scores or the clinical cutoffs used.

Internal consistency for the BSCTS is high with α-coefficients ranging from 0.87 to 0.93 across the two subscales and the total score composite. Test–retest reliability coefficients are strong with *r*-values ranging from 0.79 to 0.84. One item on the BSCTS, "Doesn't seem to understand or process questions or explanations as quickly or accurately as others," overlaps with the primary symptoms of language impairment. This item is part of the SCT Sluggish subscale. The BSCTS manual provides an alternate parent interview format of the scale to be used when parents experience difficulty reading the form.

The Behavior Rating Inventory of Executive Function, Second Edition: The BRIEF-2[30] assesses impairments of executive function in children aged 5 to 18 years. The BRIEF-2 includes a parent rating form and a teacher rating form. The BRIEF-2 is divided into nine clinical subscales: Inhibit, Self-Monitor, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Task-Monitor, and Organization of Materials. The BRIEF-2 also includes four index scores: behavior regulation, emotion regulation, cognitive regulation, and global executive composite. Shorter screening versions are also available.

The BRIEF-2 includes both a typical normative sample (n = 1,400) and a clinical sample (n = 3,007). The clinical sample includes 113 children with learning disabilities (3.8%). The manual provides comparisons between clinical groups and TD children to identify expected scores for clinical populations. Means, standard deviations, and base rates of T-scores for all scales, indices, and composite scores for each clinical population are provided to assist assessment teams with whether obtained scores are within the expected range associated with a comparable clinical group.

Several items on both the parent and teacher forms overlap with symptoms of speech, language, or learning disorders, including "Talks at the wrong time", "Has trouble getting started on homework or tasks", "When given three things to do, remembers only the first or last"; "Has trouble with chores or tasks that have more than one step", "Has trouble finishing tasks (chores, homework, etc.)", "Has trouble concentrating on tasks, schoolwork, etc."; "Has good ideas but cannot get them on paper"; "Written work is poorly organized", "Makes careless errors", and "Tests poorly even when knows correct answers." These items affect the inhibit, shift, working memory, plan/organize, and task-monitor subscales.

Three additional BRIEF-2 subscales may be used to identify potential rater bias. The Inconsistency Scale indicates the level of contradictory responses as acceptable, questionable, or inconsistent. The Negativity Scale determines whether the respondent exhibits an overly pejorative perception regarding the child. If a high Negativity Scale score is obtained, the manual recommends additional observation and clinical judgment to determine whether the parent or teacher ratings reflect an attributional bias or whether the child does, indeed, have an executive function disorder. A high score on the Infrequency Scale indicates responses as either haphazard or extreme, or that the respondent may have had reading difficulties.

Internal consistency for the items on the BRIEF-2 ranges from acceptable to excellent across the subscales and forms (α-coefficients: 0.76–0.97). Test–retest reliability ranges from moderate to high across the subscales and forms (Pearson's correlations: 0.67–0.92). Interrater reliability between parents and teachers ranges from low to moderate (Pearson's correlations: 0.24–0.72) with lower correlations in the clinical sample and higher correlations in the typical sample.

The Strengths and Difficulties Questionnaire: The SDQ[27] standardized behavioral rating scale is designed to collect both positive and negative ratings from adults on children's behaviors (age range: 4–16 years). The SDQ is widely used in research studies and contains 25 items that may be completed by either parents or teachers. The 25 items load onto five subscales, with 5 items each: Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Relationship Problems, and Prosocial Behavior. Alternatively, items from the SDQ may be pooled into three composite scales: the prosocial scale (five items), the internalizing problems (emotional and peer symptoms, 10 items), and the externalizing problems (conduct and hyperactivity symptoms, 10 items). These last two composites have parity with the internalizing and externalizing scales of the CBCL and TRF scales.

A normative sample for the SDQ was obtained as part of the National Health Interview Survey (NHIS) in 2001.[41] A parent (biological, adoptive, or step; 92%) or grandparent (4.4%) completed the SDQ for 9,878 children between ages 4 and 17 years. The extent to which children with speech/language impairments or learning disorders were included in the normative sample is unavailable. Other key demographic variables on the SDQ, such as racial/ethnic community, socioeconomic status, and geographical region, are missing as well. Factor analysis of the items on the SDQ showed that internal consistency ranged from poor to adequate across the five subscales (α-coefficients: 0.53–0.76).[42] Test–retest reliability, calculated using parent responders, ranged from moderate to high across the five subscales (Pearson's correlations: 0.57–0.76).[42] Interrater agreement between parents and teachers is low (*r* range: 0.26–0.47).[42]

One item on the SDQ, "Good attention span, sees chores or homework through to the end," if not endorsed positively by a parent or teacher, could reflect the presence of a language impairment or learning disability. This item is included in the SDQ Hyperactivity Subscale. Guidelines for making accommodations for children with speech/language impairments or learning disorders are not provided.

SUMMARY OF REVIEW OF SELECTED BEHAVIORAL RATING SCALES

Tables 1 and 2 summarize key points associated with our review of these seven clinical instruments. As a collection, the standardized behavioral rating scales were prone to many of the original criticisms raised by Redmond[8]—with some notable exceptions. One key area of improvement in the current review was the increased representations of children with language impairments and/or learning disabilities across the different rating scales. The BASC-3, CRS-3, BSCTS, and the BRIEF-2 provide assessment teams with disaggregated norms. Another bright spot from our review was the increased incorporation of validity checks across the different rating scales we reviewed (BASC-3, CRS-3, and BRIEF-2). Validity checks designed to capture inordinately negative evaluations would be particularly helpful for assessment teams trying to interpret divergent ratings collected on children with language impairments.

The total number of items within the rating scales we reviewed varied from 12 (BSCTS) to 173 (BASC-3). Language and academic items regularly appeared within their inventories as they did in Redmond.[8] Although the representation of language and academic items within a particular rating scale's inventory could be small relative to the total number of items, given the manner in which clinical thresholds for individual subscales are determined, the influence of these items on assessment decisions could still be substantial. Test designers iteratively develop their behavioral inventories through factor analyses and other reduction techniques with the goal of arriving at the smallest number of items needed for each subscale. Consequently, the endorsement of even a small number of items on some of these subscales could be sufficient to place the rating of otherwise TD children's behavior within clinical thresholds. As Redmond[6] pointed out, comorbidity in cases of overlapping symptoms starts to become meaningless when we cannot trust clinical measures to reliably differentiate between disorders. Without a clear understanding of the source of individual causes of comorbidity, faulty assumptions guide clinical management. This could then lead to contraindicated intervention approaches and wasted resources.[43]

Reasonable accommodations for children with language impairments—such as removing language and academic items prior to the calculation of clinical scores—were systemically overlooked across the seven rating scales. There is some evidence adjusting scales for overlap with language symptoms can improve their specificity without compromising their sensitivity.[23] Additional research is needed, however, to determine whether this characterization holds across individual rating scales.

It seems, then, the answer to the question "Have standardized behavioral rating scales gotten better at accommodating for overlapping symptoms with language impairment?" arrives at "They have, but there is still considerable room for improvement."

RECOMMENDATIONS FOR CLINICAL PRACTICE AND RESEARCH PROGRAMS

Almost two decades have passed since Redmond[8] reported on the presence of overlapping symptoms and other psychometric limitations within commonly used standardized rating

scales that compromised the integrity of these instruments to identify socioemotional and behavioral disorders in children who have language impairments. In this section, we revisit Redmond's[8] four recommendations regarding clinical practice in light of our updated audit. We also offer additional suggestions for clinical practice, research programs, and instrument development.

1. Collect standardized measures of socioemotional integrity from multiple informants.

The collection of ratings from multiple informants was widely recognized as an important aspect of multidisciplinary assessment when Redmond[8] conducted his review, and doing so continues to align with recognized best practices.[9] For children with either known or suspected language impairments, standardized ratings from both parents and teachers provide assessment teams with important information about the range of situations in which symptoms are present. Because parents of children with language impairments may themselves have language and literacy limitations, assessment teams should also be prepared to offer alternative formats for collecting ratings, such as the structured interview format provided by the BSCTS.

2. Consider discrepancies between informants in light of the differences across situations.

Most of the socioemotional and behavioral disorders captured by standardized rating scales require their symptoms to appear in multiple settings. For example, criteria for ADHD provided by the DSM-5 are very clear on the requirement that symptoms of inattention must be impairing in nonacademic settings. DSM-5 differentiates inattentiveness that could result from "frustration, lack of interest, or limited ability" associated with language and learning disabilities from symptoms attributable to ADHD.[39] Over the course of a multidisciplinary assessment using socioemotional and behavioral rating scales, ADHD symptoms reported by a teacher that are not endorsed by parents should be tested against the possibility that the symptoms reported by the teacher are due to either a language impairment or a learning disability. In addition, whenever possible, preference should be given to those socioemotional and behavioral rating scales that provide assessment teams with validity checks to guard against potentially inordinately negative evaluations of children with language impairments.

3. Consider the reported behavior problems in light of instrument bias.

For the most part, the widespread presence of overlapping symptoms on socioemotional and behavioral rating scales was as true for the scales assessed in the current review as it was for those reviewed by Redmond.[8] The results of Redmond and Ash[23] highlight the value of removing language and learning items from scales prior to calculating clinical scores for the purpose of differential diagnosis. Speech language pathologists should review rating scales used by assessment teams for potential language bias and, when appropriate, suggest adjusting clinical scores to accommodate for the presence of language and learning items. There is nothing particularly groundbreaking or controversial behind this suggestion. The removal of potentially overlapping symptoms from standardized behavioral ratings and clinical checklists prior to differential diagnosis represents a commonly suggested strategy in child psychopathology.[44]

4. Collect local norms.

Redmond[8] suggested those assessment teams who regularly work with families of children with language impairments and have the resources to do so should collect local norms. This represented one way of addressing the very limited representation of children with language impairments and learning disabilities across normative samples associated with the standardized socioemotional and behavioral rating scales Redmond[8] reviewed. Fortunately, since the original Redmond[8] review, the inclusion of disaggregated norms within current behavioral scale manuals has, become more common.

Our general recommendations regarding the use of standardized socioemotional rating scales by assessment teams in clinical settings extend to researchers and their programs as well. Estimates offered for the co-occurrence of language impairments and ADHD across reports have been remarkably unstable. Redmond[6] characterized the situation as “theoretically generous” in that individual reports could be selected out of the literature to support a variety of claims about the underlying nature of ADHD + LI comorbidity. Put differently, this area of research has suffered from a reproducibility problem. Some of the highest rates associated with estimates of ADHD + LI comorbidity have been based on teacher ratings. Regularly incorporating parental ratings into research studies would provide an important check against potential inflations brought in by teacher ratings. Combining parental and teacher ratings, and using validity indices, aligns with recognized best practices. These practices, however, have rarely been incorporated into empirical studies of ADHD + LI comorbidity. Likewise, adjusting clinical scores by removing overlapping symptoms prior to running tests of significance, although rarely done, provides researchers with more valid estimates of their observed effects. Reliable and valid estimates of socioemotional and behavioral symptoms are needed to examine potentially shared mechanisms/linkages between language impairments and ADHD. Their absence could lead to theoretical “red herrings,” reproducibility problems, and wasted resources. A stronger empirical base is well worth the effort. Language impairments and ADHD affect millions of students. A better understanding of the manner in which comorbidity is established between these two common, and potentially commonly co-occurring disorders, could lead to earlier detection and eventually the discovery of mitigating factors for associated academic and social risks.

Going forward, our review of these seven rating scales has potential implications for the next wave of clinical indices. For example, developers of standardized behavioral rating scales should regularly include input from experts in child language disorders during the process of item generation and selection, to guard against potential language bias. Input from experts in child language disorders should also extend over the course of data reduction to the processes of interpreting item factors loadings and assigning scale structure. The inclusion of communication and academic scales on the BASC-3 and the LDS language milestone checklist into the CBCL, where these language symptoms are separated from other clinical scales, represents an important development in rating scale design. Continued progress in this and other aspects of behavioral rating scales moves us closer to the effective differential diagnosis of language impairments from socioemotional and behavioral disorders.

FINANCIAL DISCLOSURES

S.M.R. receives a salary from the University of Utah. His research is funded by the National Institute on Deafness and Other Communication Disorders (NIDCD). He serves as Editor in Chief for the language section of the *Journal of Speech, Language, and Hearing Research.* K.M.H. is a doctoral student at the University of Utah and receives funding through Sean Redmond's NIDCD projects.

A.W. is a doctoral student at the University of Utah and receives funding through Sean Redmond's NIDCD projects.

REFERENCES

1. Ogundele MO. Behavioural and emotional disorders in childhood: a brief overview for paediatricians. World J Clin Pediatr 2018;7(01):9–26 [PubMed: 29456928]
2. Larson K, Russ SA, Kahn RS, Halfon N. Patterns of comorbidity, functioning, and service use for US children with ADHD, 2007. Pediatrics 2011;127 (03):462–470 [PubMed: 21300675]
3. Beitchman J, Brownlie E. Language development and its impact on children's psychosocial and emotional development. In: Rvachew S, Tremblay R, Bovin M, Peters R, eds. Encyclopedia on Early Childhood Development [online]. Montreal, Quebec: Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development; 2010:1–8. Available at: http://www.child-encyclopedia.com/documents/Beitchman-BrownlieANGxp_rev.pdf Accessed April 24, 2019
4. Cohen NJ, Vallance DD, Barwick M, et al. The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. J Child Psychol Psychiatry 2000;41(03):353–362 [PubMed: 10784082]
5. Snowling MJ, Bishop DV, Stothard SE, Chipchase B, Kaplan C. Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. J Child Psychol Psychiatry 2006; 47(08):759–765 [PubMed: 16898989]
6. Redmond SM. Language impairment in the attention-deficit/hyperactivity disorder context. J Speech Lang Hear Res 2016;59(01):133–142 [PubMed: 26502026]
7. Sciberras E, Mueller KL, Efron D, et al. Language problems in children with ADHD: a community-based study. Pediatrics 2014;133(05):793–800 [PubMed: 24753530]
8. Redmond SM. The use of rating scales with children who have language impairments. Am J Speech Lang Pathol 2002;11:124–138
9. Whitcomb SA. Behavioral, social, and emotional assessment of children and adolescents, 5th ed. Milton Bridge, Oxon, United Kingdom: Routledge; 2018
10. Allard ER, Williams DF. Listeners' perceptions of speech and language disorders. J Commun Disord 2008;41(02):108–123 [PubMed: 17586521]
11. DeThorne LS, Watkins RV. Listeners' perceptions of language use in children. Lang Speech Hear Serv Sch 2001;32(03):142–148 [PubMed: 27764405]
12. Rice M, Hadley P, Alexander A. Social biases toward children with speech and language impairments: a correlative causal model of language limitations. Appl Psycholinguist 1993;14:445–471
13. Beitchman JH, Nair R, Clegg M, et al. Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. J Speech Hear Disord 1986;51(02):98–110 [PubMed: 3702369]
14. Norbury CF, Gooch D, Wray C, et al. The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. J Child Psychol Psychiatry 2016;57(11):1247–1257 [PubMed: 27184709]
15. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. J Speech Lang Hear Res 1997;40(06):1245–1260 [PubMed: 9430746]
16. Wittke K, Spaulding TJ. Which preschool children with specific language impairment receive language intervention? Lang Speech Hear Serv Sch 2018;49 (01):59–71 [PubMed: 29188274]

17. Zhang X, Tomblin J. The association of intervention receipt with speech-language profiles and social-demographic variables. Am J Speech Lang Pathol 2000;9(04):345–357
18. Quay H, Peterson D. Manual for the Revised Behavior Problem Checklist. Coral Gables, FL: Department of Psychology, University of Miami; 1987
19. Reynolds CR, Kamphaus RW. Behavior Assessment System for Children. Circle Pines, MN: American Guidance Service, Inc.; 1992
20. Redmond SM, Rice ML. Stability of behavioral ratings of children with SLI. J Speech Lang Hear Res 2002;45(01):190–201 [PubMed: 14748648]
21. Achenbach T Manual for the Child Behavior Checklist/4–18. Burlington, VT: University of Vermont Press; 1991
22. Achenbach T Manual for the Teacher Report Form. Burlington, VT: University of Vermont Press; 1991
23. Redmond SM, Ash AC. A cross-etiology comparison of the socio-emotional behavioral profiles associated with attention-deficit/hyperactivity disorder and specific language impairment. Clin Linguist Phon 2014;28(05):346–365 [PubMed: 24456478]
24. Conners C Conners' Rating Scales-Revised: User's Manual. New York: Multi-Health Systems, 2002
25. Miller L Louisville Behavioral Checklist-Revised. Los Angeles, CA: Western Psychological Services; 1984
26. Barkley RA. Distinguishing sluggish cognitive tempo from ADHD in children and adolescents: executive functioning, impairment, and comorbidity. J Clin Child Adolesc Psychol 2013;42(02):161–173 [PubMed: 23094604]
27. Goodman R The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry 1997;38(05):581–586 [PubMed: 9255702]
28. DuPaul G, Power T, Anastopoulos A, Reid R. ADHD Rating Scale-5 for Children and Adolescents. New York, NY: Guilford Press; 2016
29. Barkley R Barkley Sluggish Cognitive Tempo Scale: Children and adolescents. New York, NY: Guilford Press; 2013
30. Gioia G, Isquith P, Guy S, Kenworthy L. BRIEF-2: Behavior Rating Inventory of Executive Function: Professional Manual. Lutz, FL: Psychological Assessment Resources; 2015
31. Achenbach T, Rescorla L. Multicultural Supplement to the Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2007
32. Achenbach T, Rescorla L. Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2001
33. Achenbach T Manual for the Teacher Report Form. Burlington, VT: University of Vermont Press; 1991
34. Achenbach T, Rescorla L. Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2000
35. Scoring Update DSM-V. Pearson Education Ltd. Available at: https://www.pearsonclinical.co.uk/Psy-chology/ChildMentalHealth/ChildADDADHD-Behaviour/Conners3rdEdition(Conners3)/ForThisProduct/conners-3-dsm-v-update.aspx. Accessed April 24, 2019
36. Reynolds C, Kamphaus R. BASC-3 Behavioral and Emotional Screening System manual. Circle Pines, MN: Pearson; 2015
37. Conners CK. Conners Third Edition. Los Angeles, CA: Western Psychological Services; 2008
38. Conners 3. MHS Assessments. Available at: https://www.mhs.com/MHS-Assessment?prodname=conners3. Accessed April 24, 2019
39. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington, VA: American Psychiatric Publishing; 2013
40. Lahey BB, Pelham WE, Schaughency EA, et al. Dimensions and types of attention deficit disorder. J Am Acad Child Adolesc Psychiatry 1988;27(03): 330–335 [PubMed: 3379015]

41. Normative SDQ. Data from the USA. SDQ info. Available at: https://www.sdqingo.org/norms/USNorm.html. Accessed April 24, 2019

42. Stone LL, Otten R, Engels RC, Vermulst AA, Janssens JM. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. Clin Child Fam Psychol Rev 2010;13(03):254–274 [PubMed: 20589428]

43. Redmond S, Timler G. Addressing the social concomitants of developmental. In: Apel K, Kamhi A, Masterson J, eds. Clinical Decision Making in Developmental Language Disorders. Baltimore, MD: Brookes; 2007

44. Milberger S, Biederman J, Faraone SV, Murphy J, Tsuang MT. Attention deficit hyperactivity disorder and comorbid disorders: issues of overlapping symptoms. Am J Psychiatry 1995;152(12):1793–1799 [PubMed: 8526248]

Learning Outcomes:

As a result of this activity, the reader will be able to (1) list potential sources of measurement bias when behavioral rating scales are used with children who have language impairments; (2) discuss key areas of improvement in the design of standardized behavioral rating scales that have occurred over the past 20 years; (3) describe accommodations for children's language impairments that assessment teams can implement to identify potential concomitant socioemotional behavioral disorders.

Table 1

Overview of Current Versions of Scales Reviewed in Redmond (2002)

Instrument	Age	Normative sample		Scales	Language/Academic Items
		N	Representation of LI and LD children		
ASEBA CBCL[34] LDS[34]	1 ½ to 5 y	700 parent ratings 278 parent ratings	Syndrome scales and LDS: children in special education and head start programs included DSM scales: none	Aggressive behavior Anxious/depressed Attention problems Emotionally reactive Sleep problems Somatic complaints Withdrawn	Acts too young; doesn't answer when people to talk to them; speech problem LDS (vocabulary checklist)
ASEBA CBCL[21]	6–18 y	4,994 DSM and syndrome scales 1,753 competence scale	No information provided	Aggressive behavior Anxious/Depressed Attention problems Rule-breaking behavior Social problems Somatic complaints Thought problems Withdrawn/depressed	Acts too young; fails to finish things; poor school work; refuses to talk; speech problems
ASEBA TRF[34]	1 ½ to 5 y	1,192 teacher ratings	Syndrome scales: included children from 11 clinical settings DSM: none	Aggressive behavior Anxious/depressed Attention problems Emotionally reactive Somatic complaints Withdrawn	Acts too young; doesn't answer when people talk to them; difficulty following directions; speech problem
ASEBA TRF[21]	6–18 y	4,437 DSM and syndrome scales 2,319 adaptive functioning scale	No information provided	Aggressive behavior Anxious/depressed Attention problems Rule-breaking behavior Social problems Somatic complaints Thought problems Withdrawn/depressed	Acts too young; fails to finish things; difficulty following directions; has difficulty learning; talks out of turn; poor school work; messy work; refuses to talk; speech problem; talks too much; fails to carry out assigned tasks

Instrument	Age	Normative sample		Scales	Language/Academic Items
		N	Representation of LI and LD children		
BASC-3 PRS-P36	2–5 y	600 parent ratings	Clinical sample: specific learning disorder 2.4% Other (predominantly speech/ language disorder) 43.4%	Clinical: aggression, anxiety, attention problems, atypicality, conduct problems, depression, hyperactivity, learning problems, somatization, withdrawal Adaptive: activities of dally living, adaptability, functional communication, social skills Content: anger control, bullying, developmental social disorders, emotional self-control, executive functioning, negative emotionality, resiliency Composites: externalizing, Internalizing, adaptive skills, behavioral symptoms index, functional impairment index, executive function index	Begins conversations appropriately; provides full name when asked; listens to directions; answers telephone properly; politely asks for help; pays attention when being spoken to; provides home address when asked; communicates clearly; compliments others; responds appropriately when asked a question; Is unclear when presenting ideas; starts conversations; says things that make no sense; says all the letters of the alphabet when asked; readily starts up conversations with new people; listens carefully; is clear when telling about personal experiences; interrupts others when they are speaking; speaks in short phrases that are hard to understand; babbles to self; is able to describe feelings accurately
BASC-3 PRS-C[36]	6–11 y	600 parent ratings	General sample: Specific learning disorder 5.2% Other (predominantly speech/ language disorder) 1.3% Clinical sample: Specific learning disorder 13.8% Other (predominantly speech/ language disorder) 15.7%	Same as above plus Adaptive: leadership, study skills	Responds appropriately when asked a question; listens to directions; accurately takes down messages; interrupts others when they are speaking; answers telephone properly; has difficulty explaining rules of games to others; communicates clearly; listens carefully; has trouble getting information when needed; likes to talk about his or her day; starts conversations; says things that make no sense; pays attention when being spoken to: Is clear when telling about personal experiences; babbles to self; speech is confused or disorganized; is able to describe feelings accurately; is unclear when presenting ideas
BASC-3 PRS-A[36]	12–18 y	600 parent ratings	General sample: Specific learning disorder 4.7% Other (predominantly speech/ language disorder) 1.0% Clinical sample: Specific learning disorder 9.5% Other (predominantly speech/ language disorder) 10.1%	Same as above	Pays attention; likes to talk about his or her day; talks over others; accurately takes down messages; listens to directions; has difficulty explaining rules of games to others; communicates clearly; has trouble getting information when needed; listens carefully; responds appropriately when asked a question; is unclear when presenting ideas; starts conversations; Is effective when presenting information to a group; says things that make no sense; pays attention when being spoken to; gives good suggestions for solving problems; tracks down information when needed; is able to describe feelings accurately; babbles to self; is clear when telling about personal experiences; interrupts others when they are speaking
CRS-3[36]	6–18 y	2,300 parent ratings 2,437 teacher ratings	Children with learning disorders represent 16–17% of parent- and teacher-rated clinical sample Children with learning disorders represent 5% of	Inattention Hyperactivity/impulsivity Learning problems Executive	Spelling is poor; does not understand what he/she reads; is good at memorizing facts; forgets things already learned; makes mistakes; fails to finish things he/she starts; does not seem to listen to what is being said to him/her; has trouble reading; has to struggle to complete hard tasks; needs extra explanation of instructions; learns information as separate facts—does

Instrument	Age	Normative sample		Scales	Language/Academic Items
		N	Representation of LI and LD children		
			parent- and teacher-rated total sample	functioning Defiance/aggression Peer relations Family relations Conduct disorder Oppositional defiant disorder ADHD index Global index	not "get the big picture"; reads slowly and with a lot of effort; fails to complete schoolwork, chores or tasks (even when he/she understands and is trying to cooperate); does not remember what they read; forgets instructions quickly; needs extra explanation of instructions

Table 2

Overview of Selected Socioemotional and Behavioral Rating Scales

Instrument	Age	Normative sample		Scales	Language/Academic items
		N	Representation of LI and LD		
ADHD Rating Scale-5[28]	5–17 y	2,079 parentv ratings 2,140 teacher ratings	No information provided for parent-rated sample Children receiving special education services represent 16.4% of the teacher-rated sample	Hyperactlvlty-lmpulslvlty Inattention Total score Impairment ratings	Makes careless mistakes in schoolwork; does not seem to listen when spoken to directly; does not follow through on instructions and fails to finish work; avoids, dislikes, or is reluctant to engage In tasks that require sustained mental effort
Barkley SCT Scale[29]	6–17 y	1,800 parent ratings	Children with language delay represent 5% of the sample; children receiving speech/ language therapy represent 11.4% of the sample	Emotionally reactive Anxious/ depressed Somatic complaints Withdrawn Attention problems Aggressive behavior Other problems	Doesn't seem to understand or process questions or explanations as quickly or accurately as others
BRIEF-2[30]	4–18 y	TD 3,007 clinical parent ratings TD 1,825 clinical teacher ratings	Children with learning disabilities represent 3.8% of the parent-rated clinical sample Children with learning disabilities represent 5.5% of the teacher-rated clinical sample	Inhibit Self-monitor Shift Emotional control Initiate working memory Plan/organize Task-monitor Organization of materials	Talks at the wrong time; has trouble getting started on homework or tasks; when given three things to do, remembers only the first or last; has trouble with chores or tasks that have more than one step; has trouble finishing tasks (chores, homework, etc.); has trouble concentrating on tasks, schoolwork, etc.; has good ideas but cannot get them on paper; written work Is poorly organized; makes careless errors
SDQ[27]	3–16 y	9,878 parent ratings	No information provided	Emotional problems Conduct problems Hyperactivity Peer problems Prosocial Internalizing Externalizing	Good attention span, sees chores or homework through to the end