

Informant Discrepancies in Externalizing and Internalizing Symptoms and Adaptive Skills of High-Functioning Children With Autism Spectrum Disorder

Christin A. McDonald,
Christopher Lopata, James P. Donnelly,
Marcus L. Thomeer,
and Jonathan D. Rodgers
Canisius College

Allyson K. Jordan
University at Buffalo, The State University
of New York

Assessment of clinical symptoms requires information from multiple informants. Discrepancies between informants' ratings can have significant implications in school settings (e.g., access to services, treatment planning, progress monitoring). This study examined parent–teacher discrepancies for ratings of internalizing and externalizing symptoms, and adaptive skills of high-functioning children with autism spectrum disorder. A total of 236 Behavior Assessment System for Children–2nd Edition ratings of children with high-functioning children with autism spectrum disorder from 2 informant groups (parents and teachers) were analyzed. Each informant pair ($n = 118$ parents/caregivers and $n = 118$ teachers) rated the same child. Scores on the Internalizing Problems, Externalizing Problems, and Adaptive Skills Composites were examined for mean differences, level of agreement, linear relationship, and moderators of discrepancies. There were no significant mean differences between raters for the Internalizing and Externalizing Composites or their constituent scales (except Hyperactivity). Parent–teacher ratings on these composites and scales were significantly correlated (generally moderate), and the discrepancies were not moderated by the included child or parent variables. In contrast, teacher ratings were significantly higher than parents for the Adaptive Skills Composite and several of its constituent scales. Correlations between informants on the Adaptive Skills Composite were significant (low-to-moderate), with notable variability in the correlations among its constituent scales. The degree of parent–teacher discrepancy differed significantly across the Adaptive Skills Composite score range, but it was not moderated by the included child or parent variables. This study suggests a reduced likelihood of informant discrepancies for externalizing and internalizing symptoms, with larger discrepancies expected when assessing adaptive skills.

Keywords: ASD, informant discrepancies, externalizing and internalizing symptoms, adaptive skills

Assessing the clinical symptoms and adaptive functioning of children with autism spectrum disorder (ASD) is a significant challenge. This is due

to the multicomponent nature of the diagnostic features and significant variability in functional levels (American Psychiatric Association, 2013).

Christin A. McDonald, Christopher Lopata, James P. Donnelly, Marcus L. Thomeer, and Jonathan D. Rodgers, Institute for Autism Research, Canisius College; Allyson K. Jordan, Department of Counseling, School, and Educational Psychology, University at Buffalo, The State University of New York.

The research reported in this article was supported by Department of Education, Institute of Education Sciences Grants R324A130216 and R324A080136, and a

research grant from the Organization for Autism Research (OAR). Findings and conclusions are those of the authors and do not necessarily reflect the views of the funding agencies.

Correspondence concerning this article should be addressed to Christin A. McDonald, Institute for Autism Research, Canisius College, Science Hall 1016A, 2001 Main Street, Buffalo, NY 14208. E-mail: mcdona37@canisius.edu

One subgroup that has drawn increased attention is high-functioning children with ASD (HFASD). Children with HFASD differ from children with ASD and concomitant intellectual disability (ID) based on cognitive ability, with a cut score of 70 commonly used to differentiate between the two groups (Lecavalier, 2014). The most recent data indicated that children with HFASD comprise approximately half of children with ASD and they may account for the reported increase in ASD prevalence (Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators and the Centers for Disease Control and Prevention, 2014). Beyond the diagnostic features, many children with HFASD exhibit significant externalizing and internalizing symptoms (Lecavalier, 2006; Volker et al., 2010). For example, Volker et al. (2010) compared Behavior Assessment System for Children–2nd Edition–Parent Rating Scales (BASC-2-PRS) profiles of 62 children with HFASD to profiles of matched typically developing (TD) controls. Results indicated significant differences between the HFASD and TD samples on the majority of scales (exceptions being Aggression, Somatization, and Conduct Problems), with the largest mean difference on the Behavioral Symptoms Index scale. For the HFASD sample, scores were in the clinically significant range on the Behavior Symptoms Index, Atypicality, Withdrawal, and Developmental Social Disorders scales and in the at-risk range for Hyperactivity, Attention Problems, and Depression scales.

In addition, many children with HFASD exhibit adaptive behavior deficits (McDonald et al., 2015; Volker et al., 2010). Adaptive functioning has been defined as the “ability to translate capacities into consistent, habitual behaviors fostering self-sufficiency in naturalistic settings” (Klin, Saulnier, Tsatsanis, & Volkmar, 2005, p. 2005). A broad range of factors/behaviors are subsumed under this definition, including communication, socialization, self-care, and independence (Perry, Flanagan, Dunn Geier, & Freeman, 2009). The presence of adaptive behavior deficits is notable, given the cognitive strengths of children with HFASD (McDonald et al., 2015). McDonald et al. (2015) examined the adaptive functioning of 71 children with HFASD, ages 7–12, using the Vineland Adaptive Behavior Scales–2nd Edition (Sparrow et al., 2005). Results indicated that these children displayed overall deficits on

the Adaptive Behavior Composite, as well as deficits in all three adaptive domains (Communication, Daily Living Skills, and Socialization). Further, all adaptive areas were significantly lower than population estimates and in comparison to the sample’s average cognitive ability. These reported adaptive functioning deficits, as well as a wide range of affected areas of behavioral concern, suggest that clinical evaluators, including school psychologists, should include a broad assessment of externalizing and internalizing symptoms, as well as adaptive functioning for children with HFASD.

Rating scales are often used to assess the clinical symptoms and adaptive functioning of students with ASD (Stratis & Lecavalier, 2015). As part of a comprehensive assessment, rating scales offer a number of advantages including the fact that they are brief, inexpensive, and easy to administer, and allow for ratings from multiple informants in naturalistic settings (Norris & Lecavalier, 2010). In contrast to diagnostic instruments that measure symptoms dichotomously (i.e., present or absent), rating scales often use continuous scaling that allows for measurement of aspects of gradation and whether/where a child falls within a clinical category which is useful for monitoring response to interventions and/or overall degree of impairment (Achenbach, 2011; Reynolds & Kamphaus, 2004).

A central principle of clinical assessments is that they consider information from multiple sources/informants (Stratis & Lecavalier, 2015). Parents and teachers are regarded as essential in clinical assessments as they know the child’s symptoms and skills in different but naturalistic settings (De Los Reyes, 2011; Norris & Lecavalier, 2010). Information provided by parents and teachers can help identify major deviations in functioning and/or how symptoms and skills are expressed across settings (Achenbach, 2011). School psychologists gathering rating scale information from parents and teachers of children with HFASD are challenged to understand and integrate potentially disparate information into their clinical conclusions. Prior research has generally yielded low-to-moderate correlations between parent and teacher ratings for TD youth (e.g., see meta-analysis by Achenbach, McConaughy, & Howell, 1987) and indicated that easily observable behaviors are likely to produce greater agreement than those requir-

ing greater inference (Achenbach, 2011). The overall findings of low-to-moderate correlations among informant ratings for TD youth have led to some assertions about the source(s) of the discrepancies. Specifically, that the discrepancies are a function of differences in the informants' understanding of the items, differences in the informants' perceptions of the symptoms/behaviors, and/or actual differences in the children's behaviors across settings, rather than a result of measurement error (Achenbach, 2011; De Los Reyes, 2011). These measurement challenges led De Los Reyes (2011) to recommend further studies of informant discrepancies and the reason(s) that they occur. Such research is needed as informant discrepancies have significant implications for identifying co-occurring externalizing and/or internalizing symptoms, adaptive levels, and/or treatment planning and monitoring (De Los Reyes & Kazdin, 2004).

Despite their common use in assessing emotional and behavioral symptoms and adaptive skills, little is known about parent-teacher informant agreement for rating scales for students with ASD/HFASD (Stratis & Lecavalier, 2015). Given the paucity of studies specifically addressing informant discrepancies for ASD, Stratis and Lecavalier (2015) investigated this phenomenon by examining the results of studies ($N = 49$) that included ratings by multiple informants of 2- to 22-year-olds with ASD or ID. Their meta-analysis specifically examined informant agreement and moderators of agreement for scales measuring externalizing and internalizing symptoms and social skills of youth with ASD or ID. Results indicated an overall mean correlation of .36 across all raters and behaviors. Of particular relevance to the current study, the average correlations between parents and teachers were .38 for externalizing symptoms, .25 for internalizing symptoms, and .27 for social skills. The lack of information in many studies limited the moderator analyses, however age and IQ (as continuous variables) were found to be significant moderators of agreement among some raters for internalizing symptoms and the aggregate of all behaviors. According to Stratis and Lecavalier (2015), results for youth with ASD or ID were generally similar to prior meta-analyses of informant discrepancies for TD youth. They cautioned, however, that the moderator analyses were limited as less than half of the studies included data on

average IQ of their samples and both the ASD and ID samples had to be combined for this analysis due to limitations in the published data. The authors identified the need for ongoing studies of discrepancies among specific rater pairs (e.g., parent-teacher), in addition to analyses of an expanded number of moderators (in addition to age and IQ). The study was also limited to an examination of externalizing and internalizing symptoms and social skills, and did not examine informant discrepancies for ratings of adaptive skills.

This review strongly supports the need for studies specifically examining informant discrepancies for children with HFASD. These examinations should be comprehensive and include intraclass correlations coefficients (ICCs; indicating the variance of ratings of an individual child by two informants and the variance across the entire sample of children) and simple correlations (linear relationships between informant groups; Stolarova, Wolf, Rinker, & Brielmann, 2014), as well as between-groups comparisons of mean scores of the rated children (Achenbach, 2011). In addition, the studies should use standard scores to increase the interpretability of findings (Achenbach, 2011; De Los Reyes & Kazdin, 2004). This investigation was conducted to comprehensively analyze informant discrepancies for parent and teacher ratings of externalizing symptoms, internalizing symptoms, and adaptive skills for children with HFASD. Based on the current review, this was the first study to examine parent-teacher discrepancies for these symptom and adaptive areas for children with HFASD using between-groups mean comparisons, ICCs, Pearson correlations, a Bland-Altman plot, and moderator analyses.

Method

Participants

This study included a total of 236 BASC-2 ratings of children with HFASD from two informant groups (parents and teachers). The majority of parent raters were mothers ($n = 115$; 97.5%). Each parent-teacher pair ($n = 118$ parents/caregivers and $n = 118$ teachers) rated the same child with HFASD. The children, ages 6 to 11 years, with HFASD were recruited from school-based or community samples from sev-

eral clinical trials assessing the effectiveness of psychosocial treatments for children with HFASD. Child participants in those clinical trials had a prior clinical diagnosis of autism, Asperger's, or pervasive developmental disorder—not otherwise specified, a Wechsler Intelligence Scale for Children—4th Edition (WISC-IV; Wechsler, 2003) short-form IQ >70 (and short-form Verbal Comprehension Index [VCI] or Perceptual Reasoning Index [PRI] score ≥80), and Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999) short-form Expressive or Receptive language score >70. The short-form WISC-IV consisted of the Block Design, Similarities, Vocabulary, and Matrix Reasoning subtests; the short-form CASL consisted of the Antonyms, Synonyms, Syntax Construction, and Paragraph Comprehension subtests. All children also had their diagnosis confirmed using the Autism Diagnostic Interview—Revised (Rutter, LeCouteur, & Lord, 2003). No exclusionary criteria were applied for comorbid conditions (e.g., internalizing, externalizing, etc.). The majority of the child sample ($n = 97$; 82.2%) received special education services, while only about one third ($n = 39$; 33.1%) were reported to have taken at least one psychotropic medication within the past year. The child sample for the current study was predominantly male ($n = 106$; 89.8%) and Caucasian ($n = 115$; 97.5%). The participants' measured short form IQ ranged from 71 to 140 ($M = 103.12$; $SD = 14.35$), VCI ranged from 67 to 146 ($M = 102.22$; $SD = 14.07$), and PRI ranged from 63 to 140 ($M = 103.51$; $SD = 16.38$). Mean language levels of the child sample were in the average range and average parent education (used as a proxy measure of socioeconomic status) was 13.8 years (for a detailed description of the child sample, see Table 1).

Measure

The BASC-2 PRS and Teacher Rating Scale (BASC-2 TRS; Reynolds & Kamphaus, 2004) assess a range of clinical symptoms and adaptive skills to assist with differential diagnosis, treatment planning, and progress monitoring. The BASC-2 includes three age-specific forms that were designed to be consistent across the parent and teacher versions to allow cross-informant comparisons; this study used the

Table 1
Demographic Characteristics of Child Sample

Characteristic	Child participants ($n = 118$) M (SD)
Age in years	8.74 (1.32)
WISC-IV	
Short-Form IQ	103.12 (14.35)
Short-Form VCI	102.22 (14.07)
Short-Form PRI	103.51 (16.38)
CASL	
Short-Form Expressive Language	98.34 (15.58)
Short-Form Receptive Language	101.40 (16.96)
ADI-R	
Impairment in Social Interaction	19.20 (5.79)
Impairment in Communication	15.64 (4.74)
Restricted Repetitive Behavior	5.92 (2.39)
	n (Percentage of total sample)
Gender	
Male	106 (89.8)
Female	12 (10.2)
Ethnicity	
Caucasian	115 (97.5)
Latino	2 (1.7)
Other	1 (.8)

Note. WISC-IV = Wechsler Intelligence Scale for Children—4th Edition; VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; CASL = Comprehensive Assessment of Spoken Language; ADI-R = Autism Diagnostic Interview—Revised. Cutoff scores for the ADI-R for a diagnosis of ASD are Impairments in Social Interaction (cutoff = 8), Impairments in Communication (cutoff = 10), and Restrictive Repetitive Behavior (cutoff = 3).

Child form (ages 6–11) for parents (PRS-C) and teachers (TRS-C). The PRS assesses behavior in the community and home settings and the TRS assesses behavior in the school setting. Individual items are rated on a 4-point frequency scale ranging from 0 (*never*) to 3 (*almost always*). Item scores are summed and converted to T scores ($M = 50$, $SD = 10$) for interpretation. For the clinical scales, higher scores represent more problematic symptoms/behaviors (i.e., 60–69 at-risk and ≥70 clinically significant) and for the adaptive behavior scales, lower scores indicate more severe deficits (i.e., 31–40 at-risk and ≤30 clinically significant).

The primary focus of this study was on the Externalizing, Internalizing, and Adaptive Skills composites, with the individual scales that comprise each composite used to assist with interpretation. Across the PRS-C and TRS-C

forms, the Externalizing Composite is derived from the Hyperactivity, Aggression, and Conduct Problems scales and the Internalizing composite is derived from the Anxiety, Depression, and Somatization scales. For the Adaptive Skills Composite, the PRS-C and TRS-C forms share the Adaptability, Social Skills, Leadership, and Functional Communication scales, with the PRS-C also including the Activities of Daily Living scale and the TRS-C including the Study Skills scale. Because the Activities of Daily Living and Study Skills scales both assess performance of basic expected tasks in the home (performance of everyday tasks in a safe and acceptable manner) or school (study and organizational skills for academic success) settings, they were paired for the analyses in this study. In order to insure that the two differing scales were not accounting for the between-groups difference in the Adaptive Skills Composite, the same between-groups comparison was conducted using an unweighted Adaptive Skills Composite (created using the mean of the shared scales; i.e., Adaptability, Social Skills, Leadership, and Functional Communication). Results of that comparison yielded the same outcome for the Adaptive Skills Composite. Therefore, the pairing of home and study skills for comparison was kept for the current study.

Psychometric properties of the PRS-C and TRS-C forms are reported in the manual (Reynolds & Kamphaus, 2004). Reported internal consistency reliability estimates for the composites ranged from .90 to .95 for the PRS-C and .88 to .97 for the TRS-C, and for the individual scales ranged from .73 to .88 for the PRS-C and .78 to .95 for the TRS-C. Adjusted interrater reliabilities for the PRS-C (correlations of ratings by two parents) ranged from .70 to .77 for the composites and .53 to .80 for the scales and for the TRS-C (correlations of ratings by two teachers) ranged from .52 to .65 for the composites and .41 to .74 for the scales (with the exception of Anxiety, which was lower at .23). Validity evidence is presented in factor analytic data supporting the derivation of scales, as well as in moderate-to-high intercorrelations among scales within the same composite. Concurrent validity is supported in moderate-to-high correlations between the PRS-C and TRS-C scales and composites and similar scales and composites on other established rating

scales (for additional psychometric information, see Reynolds & Kamphaus, 2004).

Procedures

Institutional Review Board approval was received for the treatment trials from which the cases were drawn, and each of the trials was conducted according to the approved procedures, including attainment of consent and assent for the participants. In each of the studies, parents and teachers completed the BASC-2 as part of a pretreatment (baseline) battery of measures. Each parent-teacher pair rated one child with HFASD. Each BASC-2 protocol was checked upon return for completeness and any errors (e.g., omissions, multiple endorsements for an item) were promptly corrected with the rater. All completed protocols were scored by graduate-level research assistants using the BASC-2 ASSIST Plus computer scoring software which includes a second entry check for accuracy. To insure the accuracy of data entry, the BASC-2 scores were entered into a database and independently checked by a second research assistant, with any discrepancies resolved by a third member of the research team.

Data Analyses

The data analysis plan was constructed to comprehensively examine discrepancies in parent and teacher reports. As previously indicated, this is a secondary analysis of data from several treatment trials which were powered to meet the primary questions of those trials. In the current study, the sample size of 118 pairs of raters was sufficient to detect a correlation of .30 at an alpha of .05 with power of .78 (ESCI, Power-Paired; Cumming, 2012). The correlation of .30 was identified as the minimally important value based on prior parent-teacher reliability studies of youth with ASD (Stratis & Lecavalier, 2015). As previously noted, the primary focus of this study was on the Externalizing, Internalizing, and Adaptive Skills composites, with the individual scales that comprise each used to assist with interpretation. Data analyses included (a) parent versus teacher means (paired *t* tests), (b) interrater agreement and consistency (ICCs, Pearson correlations, Bland-Altman plot, and regression), and (c) examination of potential moderators of parent-teacher discrepancies (Pearson correlations and scatterplots). Given

the need for more studies examining a broad range of moderators (in addition to age and IQ; Stratis & Lecavalier, 2015), parent education (a proxy measure of socioeconomic status) was included as a potential moderator in this analysis. Parent education, as a proxy of socioeconomic status, was included as family financial status has been positively associated with the likelihood of being classified with ASD (Lecavalier, 2014). This has implications for access to services, which could influence parent ratings. Effect sizes and confidence intervals were calculated for all estimates. For the parent versus teacher mean score comparisons, family wise alpha (for each composite and its constituent scales) was maintained at .05 by utilizing an adjusted critical $p < .008$ (family-wise critical value based on the Adaptive Skills Composite that had the largest number of comparisons, i.e., six comparisons; .05/6). Data analysis was accomplished with SPSS 22 (IBM, 2014) and ESCI (Cumming, 2012).

Results

Cross-Informant Group Comparisons

Cross-informant group comparisons were conducted for each of the BASC-2 composites, as well as the scales that comprised each com-

posite (see Table 2). Results were nonsignificant for the Externalizing Composite and the Aggression and Conduct Problems scales (ds .02–.25). Only the Hyperactivity scale within the Externalizing composite yielded a significant between-groups difference, with parent ratings being significantly higher than teacher ratings ($d = .45$). For the Internalizing composite and its scales, all the between-groups comparisons were nonsignificant ($ds \leq .17$).

In contrast to the Externalizing and Internalizing composites, the between-groups comparison was significant for the Adaptive Skills Composite ($d = .73$), as well as the constituent scales of Leadership ($d = .48$), Activities of Daily Living/Study Skills ($d = 1.06$), and Functional Communication ($d = .55$). The between-groups comparisons for Adaptability and Social Skills were nonsignificant (ds .20 and .30, respectively). In all adaptive comparisons, parent ratings were lower than the teacher ratings.

Parent–Teacher Reliability Estimates

Two sets of coefficients were calculated in order to examine parent–teacher reliability in terms of agreement and linear relationship (see Table 3). The column of ICC values shows significant and moderate parent–teacher agreement for the Externalizing composite and its constituent scales. The

Table 2

Parent and Teacher BASC-2 Scores, Tests of Significance, Effect Sizes, and Confidence Intervals

Composite/Scale	Parent <i>M</i> (<i>SD</i>)	Teacher <i>M</i> (<i>SD</i>)	<i>t</i> (119)	<i>p</i>	Cohen's <i>d</i> [95% CI]
Externalizing Problems	56.81 (10.86)	54.23 (9.62)	2.43	.017	-.25 [-.46, -.05]
Hyperactivity	64.16 (12.84)	58.68 (11.24)	4.62	<.001*	-.45 [-.66, -.25]
Aggression	54.18 (10.54)	53.53 (11.47)	.53	.600	-.06 [-.28, .16]
Conduct Problems	49.82 (10.71)	49.66 (8.20)	.61	.873	-.02 [-.22, .19]
Internalizing Problems	55.14 (11.59)	56.86 (11.19)	-1.38	.171	.15 [-.07, .37]
Anxiety	55.43 (12.59)	57.57 (12.49)	-1.61	.109	.17 [-.04, .38]
Depression	57.58 (12.05)	58.72 (11.15)	-.90	.370	.10 [-.12, .31]
Somatization	49.42 (11.55)	50.10 (10.97)	-.56	.575	.06 [-.15, .27]
Adaptive Skills	34.58 (8.63)	39.96 (5.75)	-6.21	<.001*	.73 [.48, .98]
Adaptability	37.31 (9.24)	39.09 (8.52)	-1.63	.106	.20 [-.04, .44]
Social Skills	38.11 (10.10)	40.76 (7.50)	-2.44	.016	.30 [.06, .54]
Leadership	37.15 (7.43)	40.25 (5.04)	-4.35	<.001*	.48 [.26, .71]
Activities of Daily Living/Study Skills	35.47 (10.16)	44.82 (7.10)	-8.81	<.001*	1.06 [.79, 1.34]
Functional Communication	36.20 (9.30)	40.87 (7.59)	-4.95	<.001*	.55 [.32, .78]

Note. BASC-2 = Behavior Assessment System for Children–2nd Edition; CI = confidence interval. Paired-samples *t* tests calculated as two-tailed tests. All comparisons based on parent and teacher ratings of $n = 118$ children with high-functioning autism spectrum disorder. Bonferroni family-wise adjusted critical value based on the Adaptive Skills Composite, which had the largest number of comparisons (i.e., six comparisons; .05/6).

* $p < .008$.

Table 3
Correlations Between Parent and Teacher BASC-2 Ratings

Composite/Scale	Intraclass correlations			Pearson correlations		
	Coefficient	<i>p</i>	95% CI	Coefficient	<i>p</i>	95% CI
Externalizing Problems	.54	<.001***	[.34, .68]	.37	<.001***	[.20, .52]
Hyperactivity	.60	<.001***	[.43, .72]	.43	<.001***	[.27, .57]
Aggression	.40	.003**	[.14, .59]	.25	.006**	[.07, .41]
Conduct Problems	.52	<.001***	[.30, .66]	.36	<.001***	[.19, .51]
Internalizing Problems	.44	.001**	[.20, .61]	.28	.002**	[.11, .44]
Anxiety	.51	<.001***	[.30, .66]	.34	<.001***	[.17, .49]
Depression	.46	<.001***	[.23, .63]	.30	.001**	[.13, .46]
Somatization	.47	<.001***	[.24, .63]	.31	.001**	[.14, .46]
Adaptive Skills	.30	.028*	[−.01, .51]	.19	.039*	[−.01, .36]
Adaptability	.18	.140	[−.18, .43]	.10	.281	[−.08, .28]
Social Skills	.21	.099	[−.13, .45]	.12	.181	[−.06, .29]
Leadership	.41	.002**	[.16, .59]	.28	.002**	[.11, .44]
Activities of Daily Living/Study Skills	.24	.068	[−.09, .47]	.15	.114	[−.03, .32]
Functional Communication	.43	.001**	[.18, .60]	.29	.002**	[.12, .45]

Note. BASC-2 = Behavior Assessment System for Children–2nd Edition; CI = confidence interval. All calculations based on parent and teacher ratings of *n* = 118 children with high-functioning autism spectrum disorder. Intraclass correlation coefficients calculated in SPSS 22 as two-way mixed effects model using absolute agreement definition of reliability (Shrout & Fleiss, 1979).

* *p* < .05. ** *p* < .01. *** *p* < .001.

degree of linear relationship in parent–teacher ratings on these scales, represented by the Pearson correlations, were also significant and reflected moderate agreement overall. For the Internalizing composite and its scales, significant ICCs and Pearson correlations were found and these were generally moderate in strength. A significant ICC and Pearson correlation were also found for the Adaptive Skills composite, with the magnitude of the coefficients being generally moderate and low, respectively. Significant ICCs and Pearson correlations were found for two of the adaptive scales including Leadership and Functional Communication; these were moderate-to-low in strength. The ICCs and Pearson correlations for the remaining three adaptive scales were low-to-negligible and nonsignificant.

Bland-Altman plots of the Externalizing, Internalizing, and Adaptive Skills composite scores were created to further examine the degree of agreement between parents and teachers, and to determine whether there were any systematic trends across the ranges of scores. The vertical (*y*) axis is the difference between parent and teacher ratings; the horizontal (*x*) axis is the mean of the parent–teacher pairs of ratings (see Figure 1). The solid horizontal line references the mean difference score, and the dotted lines above and below represent the 95%

confidence interval for the differences. The teacher score was subtracted from the parent score; thus, a positive difference score indicates that the parent submitted a higher rating than the teacher for the particular child. Bland and Altman (1986) recommended regression analysis to examine the possibility of systematic trends across the scale range by regressing the mean onto the difference score. Results of the regression analyses showed no significant relationship between the difference scores and the means for the Externalizing composite scores or between the difference scores and the means for the Internalizing composite scores. In contrast, the regression analysis for the Adaptive Skills composite showed a significant positive relationship between the difference scores and the means (*B* = .656, *t* = 4.585, *p* < .001; Figure 1).

Analysis of Potential Moderators of Parent–Teacher Discrepancies

Several variables were analyzed for their potential association with parent–teacher rating differences. In order to explore whether such variables might moderate parent–teacher difference scores, scatterplots and correlations were examined. The variables included child age,

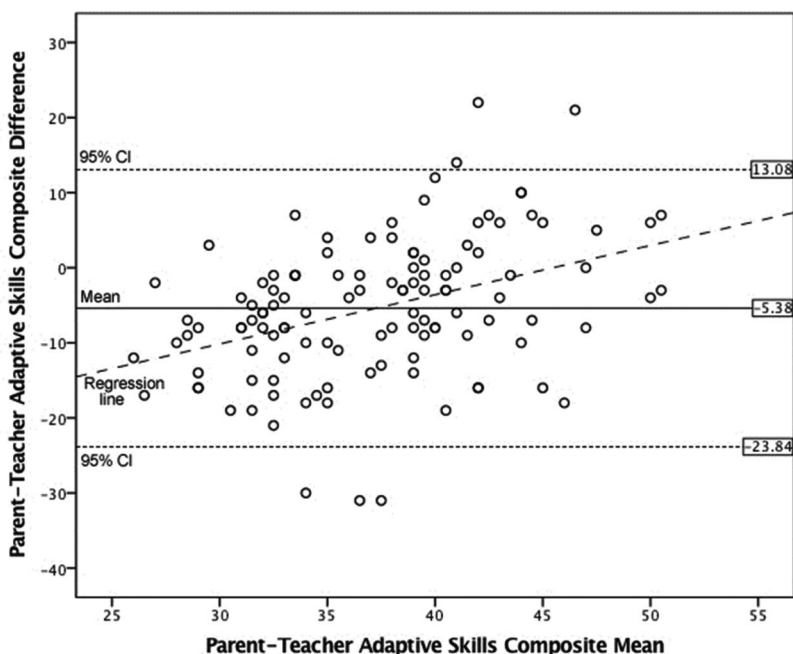


Figure 1. Bland-Altman plot of Adaptive Skills Composite by parent-teacher pairs.

WISC-IV (short form) VCI, PRI, and Full-Scale IQ, and CASL (short form) Expressive and Receptive language scores, as well as average parent education. There were no significant correlations between any of the potential moderators and the parent-teacher difference scores. The scatterplots did not indicate nonlinear relationships for any pair of variables, nor did they reveal nonnormal distributions or outliers that might impact the correlation. Gender and ethnicity may be associated with parent and teacher ratings, but the distribution of these variables was skewed such that a reasonable analysis could not be completed.

Discussion

Assessment of clinical symptoms requires information from multiple informants (Renk & Phares, 2004; Stratis & Lecavalier, 2015). However, informant discrepancies may have significant repercussions when performing evaluations in schools, such as access to services, treatment planning, and progress monitoring (De Los Reyes & Kazdin, 2004). Understanding informant discrepancies may be considered even more important for children with ASD, who often have diffi-

culty with generalization of skills across settings (Stratis & Lecavalier, 2015). Despite the possible effects on access to services and intervention planning, there is a paucity of research that has specifically and comprehensively examined informant discrepancies for scales assessing externalizing, internalizing, and adaptive behaviors for children with HFASD. This study was conducted to comprehensively examine parent-teacher discrepancies in BASC-2 ratings for children with HFASD.

Results revealed little difference between informants in the areas of internalizing and externalizing symptoms. For the Internalizing and Externalizing Composite scores, both informant groups rated the children in the average range, with no significant differences between the parent and teacher raters. This was also the case for each of the internalizing and externalizing scales, with the exception of only one externalizing scale (i.e., hyperactivity parent mean in the at-risk range and teacher mean in the average range). The reason for the difference between raters on the Hyperactivity scale is unknown. For example, it may be that the teachers reported lower scores due to increased structure in the classroom environment, relative to less structure in the home environment.

Parent–teacher ratings on the Internalizing and Externalizing composites and scales were also significantly correlated and mainly moderate in magnitude, with higher levels of agreement/association generally observed for the externalizing behaviors than internalizing behaviors. These findings are similar to those found by others in samples of TD children or children with ASD (e.g., Achenbach et al., 1987; Stratis & Lecavalier, 2015), and also support the contention that behaviors that are more easily observable may result in higher agreement levels (Achenbach, 2011). Additional analyses for the Externalizing and Internalizing Composites indicated that the discrepancies in informant-pairs' ratings did not differ significantly across the range of scores and the discrepancies were not moderated by child age, IQ, or language level, or parent education level. This is in contrast to the findings of Stratis and Lecavalier (2015), who found several factors (e.g., age, IQ) were significant moderators of agreement for some symptoms. Overall, the findings suggested that parent–teacher ratings of externalizing and internalizing symptoms were largely commensurate on average, with a moderate degree of agreement/association.

In contrast to the results for the Externalizing and Internalizing Composite scores, the mean parent rating was significantly different (lower) than the mean teacher rating on the Adaptive Skills Composite, with both falling in the at-risk range (mean parent–teacher discrepancy of 5.38 *T* score points). The correlations between parent and teacher ratings on this composite were significant, but small-to-moderate in magnitude. Additional analysis indicated that the discrepancies in informant-pair's ratings on the Adaptive Skills Composite differed significantly across the range of scores, with a larger difference between parent and teacher scores found as the pair's mean score increased. Specifically, 15% of the variance in difference scores was explained by the parent–teacher pair means. Consistent with the Internalizing and Externalizing Composite results, the moderator analyses yielded no significant relationship between parent–teacher discrepancies and child or parent variables.

Overall, findings on the Adaptive Skills Composite indicated greater discrepancy and lower correlations between parent and teacher ratings than observed on the Internalizing and Externalizing Composites. The individual adaptive scales

provide some explanation for the greater discrepancy on the Adaptive Skills Composite. Across the scales that comprise the Adaptive Skills Composite, the findings were inconsistent. To illustrate, the average teacher ratings were significantly higher than parent ratings for three of the five scales (i.e., Leadership, Activities of Daily Living/Study Skills, and Functional Communication). The inconsistent pattern was also observed in the correlations that were significant and generally moderate for two of the scales (i.e., Leadership and Functional Communication) and non-significant and largely negligible for the other three scales.

The reason(s) for the inconsistent pattern of findings between the Externalizing and Internalizing Composites and scales, and the Adaptive Skills Composite and scales is unknown but may involve the nature of the adaptive skills construct. For example, the content of the adaptive skills scales may be more diverse and abstract than the symptom scales that comprise the externalizing and internalizing areas. Social Skills, Activities of Daily Living/Study Skills, and Functional Communication scales are all included in the adaptive skills area but may be capturing highly diverse skills. The diverse nature of adaptive functioning is reflected in the definition of adaptive behaviors established by the American Association on Intellectual and Developmental Disabilities (1992, 2002), which includes nine areas of adaptive functioning. Alternatively, internalizing and externalizing problems may be more salient to raters than those adaptive behaviors requiring social skills, resulting in greater variability in the assessment of adaptive level (Renk & Phares, 2004). Setting differences might also be implicated. For example, ratings of adaptive skills, especially those with a social component, could be affected as parents may have less opportunity to observe their children in social situations with multiple peers over an extended period (Murray et al., 2009; Renk & Phares, 2004) or there may be inconsistent opportunities to display nonsocial adaptive behavior between settings. These differences in opportunities across settings may result in more dependence on context, variability in performance, and greater disparity between raters (Murray et al., 2009). Finally, it is also possible that the greater discrepancies in the adaptive skills area are a function of differences in the raters' understanding of the items that comprise these scales (Achenbach, 2011; De Los Reyes, 2011), al-

though that level of discrepancy was not observed for the externalizing and internalizing areas.

These results have several implications for school psychologists. One implication involves the reduced likelihood of large parent–teacher rating discrepancies for externalizing and internalizing symptoms. In this study, parents and teachers exhibited similar ratings overall, moderate agreement, no systematic differences across the ranges of scores, and no moderation of the discrepancies. This suggests that ratings between these informants will likely yield similar determinations about the presence of and need to treat externalizing and internalizing symptoms in students with HFASD. In contrast, this study suggested that school psychologists are more likely to confront parent–teacher discrepancies when assessing adaptive skills and these may have screening and treatment implications. For example, at the composite level a significant discrepancy could result in a student not receiving appropriate intervention if both of the informants do not identify the need (e.g., when the scores fall into different clinical ranges of severity). Poor agreement can also have significant treatment and progress monitoring implications including difficulties identifying what to target, establishing accurate baseline levels, and assessing intervention efficacy. The positive relationship between mean and difference scores of parent and teacher ratings (depicted on the Bland-Altman plot) suggests that school psychologists may encounter larger discrepancies between parent–teacher ratings as the reported scores increase. Given these clinical implications, school psychologists will be compelled to gather additional information to determine the reason(s) for the disagreement (e.g., raters’ perceptions of the behavior, understanding of the items, real variability in the behavior across settings; Achenbach, 2011; De Los Reyes, 2011) and use this to inform screening and intervention decisions. Objective, clinical observations of the child in multiple settings may be beneficial, in addition to the information derived from the rating scale (Norris & Lecavalier, 2010).

This study had a number of strengths, including being the first to comprehensively examine informant discrepancies for children with HFASD using the BASC-2. This study also included a large homogeneous, well-characterized sample of children with HFASD, as recommended by Norris

and Lecavalier (2010), and examined parent–teacher discrepancies using standard scores, which increases the interpretability of findings (De Los Reyes & Kazdin, 2004). Despite these strengths, several limitations should be noted. The majority of this sample was male and Caucasian, and all were high-functioning, which limits the generalizability of the findings. Further, although examination of rating scales in more functionally homogeneous ASD subgroups is needed, the inclusion criteria for this study may have restricted the variability of the sample’s characteristics (e.g., age, IQ, language level) used in the moderator analysis. Given these limitations, future studies on informant discrepancies should include more diverse samples with ASD/HFASD (i.e., age, ethnicity, functional level, etc.). Additionally, future studies should consider examining a broader range of informant differences including parent, teacher, and child report.

References

- Achenbach, T. M. (2011). Commentary: Definitely more than measurement error: But how should we understand and deal with informant discrepancies? *Journal of Child and Adolescent Psychology*, 40, 80–86. <http://dx.doi.org/10.1080/15374416.2011.533416>
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213–232. <http://dx.doi.org/10.1037/0033-2909.101.2.213>
- American Association on Intellectual and Developmental Disabilities. (1992). *Mental retardation: Definition, classification, and systems of supports* (9th ed.). Washington, DC: Author.
- American Association on Intellectual and Developmental Disabilities. (2002). *Mental retardation: Definition, classification, and systems of support*. Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307–310. [http://dx.doi.org/10.1016/S0140-6736\(86\)90837-8](http://dx.doi.org/10.1016/S0140-6736(86)90837-8)
- Carrow-Woolfolk, E. (1999). *Comprehensive assessment of spoken language*. Circle Pines, MN: American Guidance Service.

- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.
- De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 40, 1–9. <http://dx.doi.org/10.1080/15374416.2011.533405>
- De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, 16, 330–334. <http://dx.doi.org/10.1037/1040-3590.16.3.330>
- Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators, & the Centers for Disease Control and Prevention. (2014). Prevalence of autism spectrum disorder among children aged 8 years: Autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR*, 63, 1–21.
- Klin, A., Saulnier, C. A., Tsatsanis, K., & Volkmar, F. R. (2005). Clinical evaluation in autism spectrum disorders: Psychological assessment within a transdisciplinary framework. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Vol 2. Assessment, interventions, and policy* (3rd ed., pp. 772–798). Hoboken, NJ: Wiley.
- Lecavalier, L. (2006). Behavioral and emotional problems in young people with pervasive developmental disorders: Relative prevalence, effects of subject characteristics, and empirical classification. *Journal of Autism and Developmental Disorders*, 36, 1101–1114. <http://dx.doi.org/10.1007/s10803-006-0147-5>
- Lecavalier, L. (2014). Phenotypic variability in autism spectrum disorder: Clinical considerations. In T. E. Davis, S. W. White, & T. H. Ollendick (Eds.), *Handbook of autism and anxiety* (pp. 15–29). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-06796-4_2
- McDonald, C. A., Thomeer, M. L., Lopata, C., Fox, J. D., Donnelly, J. P., Tang, V., & Rodgers, J. D. (2015). VABS-2 ratings and predictors of adaptive behavior in children with HFASD. *Journal of Developmental and Physical Disabilities*, 27, 235–247. <http://dx.doi.org/10.1007/s10882-014-9411-3>
- Murray, D. S., Ruble, L. A., Willis, H. W., & Molloy, C. A. (2009). Parent and teacher reports of social skills in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, 40, 109–115. [http://dx.doi.org/10.1044/0161-1461\(2008/07-0089\)](http://dx.doi.org/10.1044/0161-1461(2008/07-0089))
- Norris, M., & Lecavalier, L. (2010). Screening accuracy of Level 2 autism spectrum disorder rating scales. A review of selected instruments. *Autism*, 14, 263–284. <http://dx.doi.org/10.1177/1362361309348071>
- Perry, A., Flanagan, H. E., Dunn Geier, J., & Freeman, N. L. (2009). Brief report: The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism and Developmental Disorders*, 39, 1066–1078. <http://dx.doi.org/10.1007/s10803-009-0704-9>
- Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review*, 24, 239–254. <http://dx.doi.org/10.1016/j.cpr.2004.01.004>
- Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children* (2nd ed.). Circle Pines, MN: AGS.
- Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism Diagnostic Interview-Revised*. Los Angeles, CA: Western Psychological Services.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428. <http://dx.doi.org/10.1037/0033-2909.86.2.420>
- Sparrow, S. S., Cichetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Minneapolis, MN: NCS Pearson, Inc.
- Stolarova, M., Wolf, C., Rinker, T., & Briemann, A. (2014). How to assess and compare inter-rater reliability, agreement and correlation of ratings: An exemplary analysis of mother-father and parent-teacher expressive vocabulary rating pairs. *Frontiers in Psychology*, 5, 509. <http://dx.doi.org/10.3389/fpsyg.2014.00509>
- Stratis, E. A., & Lecavalier, L. (2015). Informant agreement for youth with autism spectrum disorder or intellectual disability: A meta-analysis. *Journal of Autism and Developmental Disorders*, 45, 1026–1041. <http://dx.doi.org/10.1007/s10803-014-2258-8>
- Volker, M. A., Lopata, C., Smerbeck, A. M., Knoll, V. A., Thomeer, M. L., Toomey, J. A., & Rodgers, J. D. (2010). BASC-2 PRS profiles for students with high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40, 188–199. <http://dx.doi.org/10.1007/s10803-009-0849-6>
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th ed.). San Antonio, TX: The Psychological Corporation.

Received October 28, 2015

Revision received January 15, 2016

Accepted January 26, 2016 ■