Title

Differences in profiles of emotional behavioral problems across instruments in verbally fluent versus language impaired children with Autism Spectrum Disorder

Short Running Tile

EBP Measure Differences in ASD

Authors

Megan Fok, Graduate School of Applied and Professional Psychology, Rutgers University-New Brunswick

Vanessa H. Bal, PhD, Graduate School of Applied and Professional Psychology, Rutgers University-New Brunswick

Acknowledgements

This work was supported by K23MH115166-01 to VHB. We are grateful to all of the families at the participating Simons Simplex Collection sites, as well as the principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, E. Wijsman). We appreciate obtaining access to phenotypic data on SFARI Base.

Lay Summary

Some emotional and behavioral problem scales include items that may be inappropriate for children with ASD with limited language. The present study found that there was a tendency for children with language impairment to have lower internalizing scores on the Child Behavior Checklist, but higher scores lethargy and irritability scores on the Aberrant Behavior Checklist,

relative to verbally fluent children. This suggests that each of these instruments may underestimate emotional and behavioral problems in certain subsets of children.

Abstract

There has been increasing attention to the assessment of minimally verbal (MV) children with Autism spectrum disorder (ASD). Previous research has begun to examine the relationship between verbal abilities and emotional and behavioral problems (EBP). The current study compared parent-reported EBP in children of differing language levels on two instruments commonly used in ASD research and clinical practice, the Child Behavior Checklist (CBCL) and Aberrant Behavior Checklist (ABC). The study consisted of 1,937 6-18 years old children with ASD from the Simons Simplex Collection. Children were divided into three language groups, by ADOS module (Module 1 = MV, 2 = phrase speech (PS), 3 and 4 = verbally fluent (VF)) and then compared on CBCL and ABC subscales. The ABC and CBCL showed different patterns of elevations across the language groups. MV children were reported to have more impairment than VF children on the ABC Irritability, Lethargy and Hyperactivity scales. Children with less language (MV and PS) exhibited less impairment on the CBCL Internalizing domain than VF children, but did not differ on the Externalizing domain. Post hoc comparisons showed that Internalizing differences were driven by fewer children with less language exhibiting clinically elevated Anxious/Depressed scores compared to VF children. The present study underscores the significance of considering language when assessing EBP. Results have implications for the psychiatric screening of children with ASD, particularly those with language impairments. Researchers should exercise caution when applying EBP instruments designed for use with different populations and purposes to broad samples of children.

Keywords

Language, Affect/emotion, Children

Introduction

Autism spectrum disorder (ASD) is widely acknowledged as a heterogeneous syndrome which manifests in different ways depending on the individual's age, language and cognitive abilities. Recent studies suggest that approximately 30% of children with ASD remain minimally verbal (MV; Howlin, Savage, Moss, Tempier, & Rutter, 2014; Pickles, Anderson, & Lord, 2014), though precise estimates may vary by how MV is defined (Bal, Katz, Bishop, & Krasileva, 2016). It is widely acknowledged that consideration of language level is essential in the assessment of children with ASD. A dearth of appropriate measures to assess MV individuals challenges clinical assessment and impedes inclusion of this population in research (Jack & Pelphrey, 2017; Kasari, Brady, Lord, & Tager-Flusberg, 2013; Tager-Flusberg & Kasari, 2013). Recent studies have begun to explore the relationship between emotional and behavioral problems (EBP) and verbal abilities (Lerner et al., 2017; Williams, Siegel, & Mazefsky, 2017). The present study aims to extend this literature by comparing profiles from two measures commonly used to assess EBP in studies of children with ASD: the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) and the Aberrant Behavior Checklist (ABC; Aman, Singh, Stewart, & Field, 1985).

Many studies have reported high rates of co-occurring EBP and psychiatric disorders in children with ASD (Anderson, Maye, & Lord, 2011; Chandler et al., 2016; Gotham, Brunwasser, & Lord, 2015; Kanne, Abbacchi, & Constantino, 2009; Lecavalier, 2006; Leyfer et al., 2006; Rosen, Mazefsky, Vasa, & Lerner, 2018; Simonoff et al., 2008). Only a handful of studies, however, have explored how EBP vary in children with different language abilities. Williams and colleagues (2017) examined EBP across multiple measures in an inpatient sample of children with ASD. After controlling for age and NVIQ, there were no significant differences

between MV (i.e., children receiving an ADOS-2 Module 1 or 2) and verbal children (i.e., administered a Module 3 or 4) on the ABC Stereotypy and Irritability subscales. A lack of language group differences could reflect a tendency for children to be referred to an inpatient unit for EBP. On the Vineland Adaptive Behavior Scales-II (Sparrow, Cicchetti, & Balla, 2005), however, verbal children had significantly higher externalizing subscale scores than MV children.

In a mixed sample of children with ASD drawn from inpatient and outpatient clinics, Lerner and colleagues (2017) investigated the differences between different verbal ability groups (also defined by the ADOS-2) on the Child and Adolescent Symptom Inventory (CASI; Gadow & Sprafkin, 2005). Verbal children reportedly had more severe symptoms than the MV children across the CASI Depression, Oppositional Defiant Disorder, and Generalized Anxiety Disorder domains. In contrast, a greater proportion of the MV group fell above the clinical cut-off on the ADHD-Hyperactive/Impulsive domain. It is possible that some psychiatric disorders (e.g., Anxiety, Depression) manifest differently in individuals of varying intellectual and language levels (Einfeld & Aman, 1995). Alternatively, group differences could be due to some symptoms (particularly related to internalizing problems) being more difficult to assess in children who cannot verbally express their thoughts or feelings (Lerner et al., 2017).

In recognition of the need for measures appropriate for use with minimally verbal children, Mazefsky and colleagues (2018) recently developed the Emotion Dysregulation Inventory (EDI). Using Item Response Theory analyses to inform item selection, the initial validation lends support for the use of the EDI as a clinical and research tool with children across the full range of cognitive and verbal abilities. While this may be a useful tool for future use, there remains a need to understand instruments that have been used to assess EBP to date, as

these are commonly included in large public datasets that are being used for secondary data analyses.

The Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) is one of the most commonly used measures to assess EBP in children with ASD. The CBCL was developed as a screening tool for primary care physicians to assess for psychiatric problems in children to inform what services would best next serve them (Achenbach & Ruffle, 2000). However, it includes many items that are directly dependent on verbal language abilities and, therefore, may not be appropriate for MV individuals. These items reference thoughts that may be otherwise difficult to assess without the child directly stating them, such as feeling guilty, worthless, or that he/she might do something bad. Although there have been some studies studying the use of the CBCL in children with ASD (Duarte, Bordin, Oliveira, & Bird, 2003; Mazefsky, Anderson, Conner, & Minshew, 2010; Pandolfi, Magyar, & Dill, 2012), none have systematically investigated the validity of its use in MV children.

On the other hand, the Aberrant Behavior Checklist (ABC; Aman et al., 1985) is a frequently used measure that may be well-suited for measuring EBP in children who have ASD (Kaat, Lecavalier, & Aman, 2014). Unlike the CBCL, the ABC was created to assess pharmacologic treatment effects in individuals with severe intellectual disability and therefore symptom ratings do not rely on spoken language ability. In a study assessing the validity of the ABC in a heterogeneous sample of children with ASD, there were negligible to small (r<.15) associations between ABC subscales and both age and IQ (Kaat et al., 2014). Correlations between scales assumed to measure similar constructs on the ABC and CBCL subscales supported the convergent validity. While this study provided initial support for the use of ABC in

children with ASD across developmental levels, it did not specifically investigate potential effects of varying language levels.

Taken together, both empirical evidence and the specific characteristics of different instruments suggest a need for more systematic investigation of the presentation and assessment of EBP in children with varying language abilities. Such information is critical to inform clinical practices and promote accurate screening for EBP and psychiatric comorbidities, particularly in MV children with ASD. Assessment of EBP is also increasingly important in many domains of ASD research. In clinical studies, it is often necessary to "control for" potential confounding effects of EBP on measures of ASD severity (Havdahl, von Tetzchner, Huerta, Lord, & Bishop, 2016; Hus, Bishop, Gotham, Huerta, & Lord, 2013). Understanding how EBP profiles may differ by instrument across children of differing language abilities has implications for researchers seeking to identify associations between genetic disorders or high-risk genes and clinical phenotypes, as well as large-scale studies combining heterogeneous samples to explore genetic susceptibility across a range of neurodevelopmental and psychiatric conditions. Thus, the aim of the current study was to investigate how EBP profiles from the CBCL and ABC compared across school-aged children with ASD of different language levels in order to gain insight about how each measure may capture differences in the presentation of EBP.

Methods

Participants

Participants were selected from the Simons Simplex Collection, which includes 2,658 children with ASD (Fischbach & Lord, 2010). Children too young to receive the CBCL 6-18 parent report form were excluded from the present study (n=721). Study procedures were

approved by Institutional Review Boards at each university site. The resulting sample comprised 1,937 children between the ages of 6 and 18 (*M*=10.1 years old, *SD*=2.96, 25% at least 12 years old); 87% male, 79% white, 60% maternal education of a bachelor's degree or higher, and 25% with a nonverbal IQ less than 70.

Measures

Language groups. Language groups were based on the Autism Diagnostic Observation Schedule (Lord et al., 2000; Lord et al., 2012). Children administered a module 1 were categorized as Minimally Verbal (MV), module 2 as Phrase Speech (PS) and module 3 as Verbally Fluent (VF). Table 1 provides additional breakdown of demographics per language group.

Aberrant Behavior Checklist (Aman et al., 1985). The ABC is a 58-item scale created to measure treatment-related change in individuals with developmental disabilities. Caregivers rate behavior over the previous 4 weeks on a 4-point Likert scale (ranging from 0= "Not at all a problem", 1= "The behavior is a problem but slight in degree", 2= "The problem is moderately serious", and 3= "The problem is severe in degree"). The ABC includes five subscales: Irritability, Lethargy, Stereotypy, Hyperactivity, and Inappropriate Speech. The Stereotypy and Inappropriate Speech subscales were excluded from the present analysis due to overlap with ASD symptoms and inapplicable items for MV children, respectively. Confirmatory factor analyses in two independent samples suggest marginal to moderate fit for the original five-factor structure that was deemed acceptable from a theoretical and practical perspective (Kaat et al., 2014). While some studies have suggested variation in item placement across scales (Brinkley et

al., 2007), the present study used the original scales for comparability to the majority previous studies using the ABC in ASD research.

Child Behavior Checklist 6-18 (Achenbach & Rescorla, 2001). The Child Behavior Checklist 6-18 (CBCL) is a 118-item diagnostic screening tool. Caregivers rate their child's behavior over the previous 6 months on a three-point Likert scale (0 = "Not True", 1 = "Somewhat or Sometimes True", or 2 = "Very True or Often True"). The CBCL contains two broad domain scales (Internalizing and Externalizing). The Internalizing domain is comprised of three syndrome scales: Anxious/Depressed, Somatic Complaints, and Withdrawn/Depressed. The Externalizing domain measures includes Rule Breaking Behavior and Aggressive Behavior syndrome scales. Raw scores for each syndrome scale are converted to norm-referenced T scores (M = 50, SD = 10). Gender-based norms are provided for 6–11 and 12–18-year age ranges. Domain T scores above 63 are considered clinically significant, and T scores 60-63 fall in the borderline clinical range. T scores $\geq$70 on the syndrome scales indicate clinically significant levels of symptoms. Confirmatory factor analyses support the original factor structure (i.e., domain and syndrome scales) in children with ASD (Pandolfi et al., 2012).

Statistical Analysis

First, Pearson correlations were used to assess convergent validity of the CBCL and ABC across language groups. Correlation coefficients less than .3 were considered weak, between .3 and .6 are moderate, and correlation coefficients greater than .7 were strong (Cohen, 1988). Next, Analysis of Variance with post-hoc comparisons were used to compare language groups (MV, PS, VF) on the ABC subscales and the CBCL Internalizing and Externalizing domains. Effect sizes for the overall F-test (η_p^2) and post-hoc comparisons (Cohen's *d)* were also computed.

Effect sizes are reported as partial eta squared (η_p^2=.01-.05, η_p^2=.06-.13, and $\eta_p^2 \geq$.14 were considered small, medium, and large, respectively) and Cohen's *d* (*d*= .2-.4, *d*=.5-.7, and *d*≥.8 were considered small, medium, and large, respectively (Cohen, 1988). Next, Chi-square analyses were used to investigate whether distribution of clinical classifications (i.e., clinically significant, borderline, non-clinical) on the two CBCL domains differed across language groups. Effect sizes (Cramer's *V*) were also computed and *V*=.1, *V*=.3, and *V*=.5 were considered small, medium, and large, respectively (Cohen, 1988). Finally, post hoc comparisons with Bonferroni correction of the CBCL syndrome scales were conducted to better understand language group differences on the Internalizing domain. Results were highly similar when covarying nonverbal IQ (NVIQ); therefore, for comparability to other samples, statistics are reported without covariates unless noted. A significance level of p<.001 was set across all analyses to reduce type I error due to multiple comparisons.

Results

Correlations between CBCL and ABC scales, by language

As seen in Table 2, ABC-CBCL correlations were highly similar across language groups. Consistent with previous studies (e.g., Kaat et al., 2014), moderate correlations were found between the ABC Hyperactivity subscale and CBCL Externalizing domain, as well as between the ABC Lethargy subscale and the CBCL Internalizing domain. The ABC Irritability subscale was strongly correlated with the CBCL Externalizing domain and moderately correlated with the Internalizing domain.

Aberrant Behavior Checklist (ABC)

Subscale scores on the Aberrant Behavior Checklist (ABC) differed across language groups: Irritability ($F(2,1934)=11.9$, $p<.001$, $\eta_p^2=.012$); Lethargy ($F(2,1934)=32.3$, $p<.001$, $\eta_p^2=.033$); and Hyperactivity ($F(2,1934)=26.3$, $p<.001$, $\eta_p^2=.027$) (see Table S1). As indicated in Figure 1, while children with less language tended to have higher scores across all subscales, each subscale showed a different pattern of language group differences. On the Lethargy subscale, MV children had higher scores than both the PS ($p<.001$; $d=.308$) and VF ($p<.001$; $d=.494$) groups; there was not a significant difference between the PS and VF groups ($p=.01$; $d=.184$). On the Hyperactivity subscale, both the MV ($p<.001$; $d=.391$ and PS ($p<.001$; $d=.306$) groups had higher scores than VF children, but the MV and PS groups did not differ significantly ($p=.71$; $d=.093$). While the MV group had higher Irritability scores than the VF ($p<.001$; $d=.284$), these differences were no longer significant when NVIQ was covaried ($F(2, 83.13)=1.10$, $p=.33$). The PS group did not differ from the MV ($p=.44$; $d=.116$) or VF ($p=.02$; $d=.171$) groups on the Irritability scale.

Child Behavior Checklist (CBCL)

Language groups differed significantly on the CBCL Internalizing (Int) domain ($F(2,1934)=46.16$, $p<.001$, $\eta_p^2=.046$; Figure 2A, Table S1). Bonferroni comparisons showed that the MV ($p<.001$; $d=.504$) and PS ($p<.001$; $d=.459$) groups had significantly lower Int scores than the VF group and showed no differences between MV and PS ($d=.031$). The language groups showed no significant difference in Externalizing (Ext) domain scores ($F(2,1934)=.72$, $p=.49$, $\eta_p^2=.001$; Figure 2B, Table S1).

The distribution of clinical classifications on the Int domain differed significantly by language groups ($X^2(4) = 74.48$, $p<.001$, $V=.148$); fewer children in the MV and PS group scored

within the clinical range of concern compared to the VF group (Figure 2C). The distribution of clinical classifications on the Ext domain was not statistically significant across language groups ($X^2(4) = 13.83$, p=.008, *V*=.063; Figure 2D).

Post hoc CBCL Syndrome Scales

To better understand language group differences on the Int domain, ANOVA and Chi-square analyses were repeated for the three syndrome scales corresponding to the Int domain. Pearson correlations between the three syndrome scales and the ABC Lethargy scale were also examined to inform interpretation of profile differences across instruments.

Language groups differed on both the Anxious/Depressed (AnxD; F(2,1933)=120.93, p<.001, η_p^2=.111) and Somatic Complaints syndrome scales (SomC; F(2,1934)=14.78, p<.001, η_p^2=.015) (see Table S1). As seen in Figure 3A and 3B, the MV and PS groups had lower AnxD (p<.001; MV *d*=.967; PS *d*=.680) and SomC (p<.001; MV *d*=.313; PS *d*=.228) scores than the VF group, but did not differ from each other (p=.03, *d*=.320; p=.99 *d*=.090, respectively). Distributions of clinical classifications on the syndrome scales differed for the AnxD syndrome scale ($X^2(4) =164.86$, p<.001, *V*=.206). Consistent with the mean score distributions, fewer children with language impairment scored in the clinically significant range compared to the VF group (Figure 3D). However, as shown in Figure 3E, distributions of clinical classifications on the SomC scale did not differ significantly across language groups (p=.02, *V*=.057), suggesting mean score differences on that syndrome scale may not be as clinically meaningful. There were no significant differences between the language groups for the Withdrawn/Depressed syndrome scale (WithD; F(2, 1934)=4.09, p=.02, η_p^2=.004; Figure 3C) nor were there were significant differences in the distributions of clinical classifications (p=.006, *V*=.061; Figure 3F).

Across language groups, the ABC Lethargy was differentially correlated with the three CBCL internalizing syndrome scales. Table 3 indicates that ABC Lethargy was weakly to moderately correlated with the CBCL AnxD and SomC syndrome scales, but moderately to strongly correlated with the WithD scale.

Discussion

Previous studies (e.g., Lerner et al., 2017 and Williams et al., 2017) have suggested that different instruments may yield different EBP profiles when used with children of varying language levels. The current study adds to this literature by demonstrating that children with impaired language show different domain profiles of EBP on two commonly used instruments in ASD research and clinical practice: the ABC and the CBCL.

Consistent with a previous study by Lerner and colleagues (2017), verbal children were reported to have higher levels of internalizing symptoms (on the CBCL) than children with language impairment (i.e., MV and PS groups). In addition to medium-sized effects of language group, differences on the Internalizing domain were also clinically meaningful, such that verbal children, on average, fell in the Borderline Clinical range, whereas both language impaired groups fell below the range of clinical concern (Figure 2). In contrast, on the ABC Lethargy scale, which was moderately correlated with the Int scale, MV children had higher scores than both the PS and VF groups, with effect sizes in the small-to-moderate range (Figure 1). ABC Lethargy differences were somewhat more difficult to interpret in the absence of clinical cut-offs or ranges; however, the difference of approximately 3.5 points (MV vs VF) out of a total 48 for that subscale may be less clinically meaningful than the group differences on the CBCL affecting clinical classification. The reason for different patterns of symptoms on the CBCL vs. ABC is likely due to the differences on the internalizing CBCL being driven by the medium-to-large

effects of language on the AnxD scale (Figure 3A), whereas the Lethargy scale correlated most strongly with the WithD scale. Consistent with the less clinically significant Lethargy scale difference, language groups differences on the WithD scale were negligible to small and not statistically significant (Figure 3C).

Patterns of externalizing symptoms were somewhat different across instruments. There was a small overall effect of language on ABC Hyperactivity scores. Children with impaired language exhibited more impairment on the ABC Hyperactivity scale than children who were VF (effect sizes in the small-to-moderate range; Figure 1C). In contrast, mean CBCL Ext scores did not differ across language groups (Figure 1B). In contrast, Williams and colleagues (2017) reported that verbal children exhibited more impairment on the Vineland Externalizing scale than MV children, but that language groups did not differ on the ABC subscales. This could be due to differences in definitions (i.e., their MV group included children with phrase speech, whereas the current study separated MV and phrase speech groups) or sample ascertainment (i.e., children from an inpatient unit vs. research participants recruited to a genetic study).

Although effect sizes on the ABC were smaller than those observed for the CBCL, the difference in profiles on each measure across languages groups suggest a need to exercise caution when interpreting results from these instruments across samples of children with mixed language abilities. On one hand, the present results suggest that the CBCL may underestimate EBP in language-impaired children with ASD, whereas the ABC may underestimate EBP in more verbal children. Differences in profiles of EBP on the CBCL and ABC may be expected considering differences in the instruments themselves. The CBCL was developed as a psychiatric screener for use in primary care settings, where most 6-18 year olds may be expected to have spoken language and therefore contains items that would be difficult for parents of MV children

to assess and report. The ABC was designed to assess treatment-related change in problem behaviors in studies of individuals with severe intellectual impairment. Therefore, items were likely intentionally written to assess observable behaviors and may miss internal experiences that might only be apparent through verbal report. Although differences may be expected, these instruments are commonly used in clinical and research settings without careful attention to possible language effects.

Based on the present analyses, we cannot rule out the possibility that the differences in profiles reflect true differences in the manifestation of EBP in children of varying language and ability levels. Unfortunately, the specific item level data necessary to demonstrate that measurement properties of these instruments differ when applied to children of varying language levels was not available through this study. Future studies including item-level data are needed to further explore whether language group differences are due to aspects of the measures themselves or whether results provide further evidence that EBP manifest differently in individuals of varying ability levels (e.g., Einfeld & Aman, 1995).

While this is a critically important distinction to be made, the current study further contributes to the literature emphasizing a need for greater attention to consideration of verbal abilities in assessment of EBP. Although the CBCL and ABC have historically been the most common assessments of EBP in ASD research, The Emotion Dysregulation Inventory (EDI; Mazefsky et al., 2018) was recently developed using Item Response Theory analyses to calibrate items to remove biases related to IQ and verbal ability (as well as gender). In the initial validation study, authors report that the EDI was not only appropriate for use with youth with ASD of any cognitive or language ability, but also provided comparable or more information than corresponding ABC or CBCL subscales. Although additional independent validation of the

EDI is warranted, these results suggest the EDI may be a more valid alternative for assessment of EBP in studies including children with ASD of varying cognitive and language levels.

Limitations

The Simons Simplex Collection is not a representative sample and may underrepresent minimally verbal children (Bal et al., 2016); however, as the goal of this study was to compare instrument performance, rather than estimate prevalence of co-occurring EBP, ascertainment bias seems unlikely to account for results. It is possible that the different time frames that the CBCL and ABC ask parents to report on their child's behavior (i.e., previous six months and four weeks, respectively) could have influenced results. This has not been raised in other studies comparing these two instruments and it is difficult to imagine how timing could have systematically influenced parent reporting to produce the observed pattern of language group differences. As noted above, the most significant limitation of this study is that analyses were limited by the lack of availability of CBCL item-level data. Thus, our interpretation of language group differences relating to different foci of items (i.e., observable behavior vs. verbal reports by the child) remains to be explored in studies with item level data for both instruments. Such datasets may also allow application of statistical techniques to combine information across instruments (e.g., Integrative Data Analysis; Curran & Hussong, 2009), which may also identify constructs or dimensions of psychopathology that can be assessed across developmental levels.

Perhaps the biggest impact the results of this study have is for studies that examine larger groups that examine individuals with ASD and more broad range of abilities. The mixed pattern of EBP reported across the CBCL and ABC for verbal and MV children suggests that larger studies that examine mixed participant samples, such as studies of the genetic consortia, should

consider the use of different instruments to measure specific constructs and combine the overall results instead of using a single measure to continuously measure across a broader group. Alternatively, they may consider measures designed for use with children with ASD across varying verbal and cognitive levels, such as the EDI (Mazefsky et. al, 2018).

Conclusion

Recent studies have suggested a need for considering language levels when assessing EBP in children with ASD. Findings from the present study suggest that two instruments commonly used to assess EBP in ASD research may yield different patterns of EBP for children of varying language levels. The present study cannot differentiate whether these differences are attributable to differences in the measurement properties of the questionnaires or reflect true differences in manifestation of EBP across language groups. Nonetheless, as more attention is given to the clinical assessment of MV children and research increasingly calls for large samples that often include children with varied language abilities, these findings further underscore the critical need to systematically investigate the instruments being used to assess and quantify any behavioral construct. Evaluation of EBP is highly important across clinical and research settings where underestimates have serious implications on multiple levels – from missed opportunities for treatment to misleading investigations of pathophysiology.

References

Achenbach, T. M., & Ruffle, T. M. (2000). The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in Review, 21*(8), 265–271.

Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985). The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency*, *89*(5), 485–491.

Anderson, D. K., Maye, M. P., & Lord, C. (2011). Changes in maladaptive behaviors from midchildhood to young adulthood in autism spectrum disorder. *American Journal on Intellectual and Developmental Disabilities*, *116*(5), 381–397.

Bal, V. H., Katz, T., Bishop, S. L., & Krasileva, K. (2016). Understanding definitions of minimally verbal across instruments: Evidence for subgroups within minimally verbal children and adolescents with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *57*(12), 1424–1433.

Brinkley, J., Nations, L., Abramson, R. K., Hall, A., Wright, H. H., Gabriels, R., … Cuccaro, M. L. (2007). Factor analysis of the aberrant behavior checklist in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(10), 1949–1959.

Chandler, S., Howlin, P., Simonoff, E., O'Sullivan, T., Tseng, E., Kennedy, J., … Baird, G. (2016). Emotional and behavioural problems in young children with autism spectrum disorder. *Developmental Medicine and Child Neurology*, *58*(2), 202–208.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods, 14*(2), 81.

Duarte, C. S., Bordin, I. A. S., de Oliveira, A., & Bird, H. (2003). The CBCL and the Identification of Children with Autism and Related Conditions in Brazil: Pilot Findings. *Journal of Autism and Developmental Disorders*, *33*(6), 703–707.

Einfeld, S. L., & Aman, M. (1995). Issues in the taxonomy of psychopathology in mental retardation. *Journal of Autism and Developmental Disorders*, *25*(2), 143–167.

Fischbach, G. D., & Lord, C. (2010). The Simons simplex collection: A resource for identification of autism genetic risk factors. *Neuron*, *68*(2), 192–195.

Gadow, K. D., & Sprafkin, J. (2005). *Child and adolescent symptom inventory*. Stony Brook, NY: Checkmate Plus.

Gotham, K., Brunwasser, S. M., & Lord, C. (2015). Depressive and anxiety symptom trajectories from school age through young adulthood in samples with autism spectrum disorder and developmental delay. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*(5), 369-376.e3.

Havdahl, K. A., Bal, V. H., Huerta, M., Pickles, A., Øyen, A. S., Stoltenberg, C., Lord, C. & Bishop, S. L. (2016). Multidimensional influences on autism symptom measures: Implications for use in etiological research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *55*(12), 1054-1063.

Howlin, P., Savage, S., Moss, P., Tempier, A., & Rutter, M. (2014). Cognitive and language skills in adults with autism: A 40-year follow-up. *Journal of Child Psychology and Psychiatry, 55*(1), 49–58.

Hus, V., Bishop, S., Gotham, K., Huerta, M., & Lord, C. (2013). Factors influencing scores on the Social Responsiveness Scale. *Journal of Child Psychology and Psychiatry*, *54*(2), 216–224.

Jack, A., & Pelphrey, K. (2017). Annual Research Review: Understudied populations within the autism spectrum–current trends and future directions in neuroimaging research. *Journal of Child Psychology and Psychiatry*, *58*(4), 411–435.

Kaat, A. J., Lecavalier, L., & Aman, M. G. (2014). Validity of the Aberrant Behavior Checklist in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(5), 1103–1116.

Kanne, S. M., Abbacchi, A. M., & Constantino, J. N. (2009). Multi-informant ratings of psychiatric symptom severity in children with autism spectrum disorders: The importance of environmental context. *Journal of Autism and Developmental Disorders*, *39*(6), 856–864.

Kasari, C., Brady, N., Lord, C., & Tager-Flusberg, H. (2013). Assessing the minimally verbal school-aged child with autism spectrum disorder. *Autism Research*, *6*(6), 479-493.

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*(8), 1101–1114.

Lerner, M. D., Mazefsky, C. A., Weber, R. J., Transue, E., Siegel, M., & Gadow, K. D. (2017). Verbal ability and psychiatric symptoms in clinically referred inpatient and outpatient youth with ASD. *Journal of Autism and Developmental Disorders, 48*(11), 3689-3701.

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., … Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, *36*(7), 849–861.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., … Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule: ADOS-2*. Los Angeles, CA: Western Psychological Services.

Mazefsky, C. A., Anderson, R., Conner, C. M., & Minshew, N. (2010). Child behavior checklist scores for school-aged children with autism: Preliminary evidence of patterns suggesting the need for referral. *Journal of Psychopathology and Behavioral Assessment*, *33*(1), 31–37.

Mazefsky, C. A., Day, T. N., Siegel, M., White, S. W., Yu, L., & Pilkonis, A. (2018). Development of the Emotion Dysregulation Inventory: A PROMIS®ing method for creating sensitive and unbiased questionnaires for autism spectrum disorder. *Journal of Autism and Developmental* Disorders, 48(11), 3336-3746.

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2012). An initial psychometric evaluation of the CBCL 6–18 in a sample of youth with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *6*(1), 96–108.

Pickles, A., Anderson, D. K., & Lord, C. (2014). Heterogeneity and plasticity in the development of language: A 17-year follow-up of children referred early for possible autism. *Journal of Child Psychology and Psychiatry, 55*(12), 1354–1362.

Rosen, T. E., Mazefsky, C. A., Vasa, R. A., & Lerner, M. D. (2018). Co-occurring psychiatric conditions in autism spectrum disorder. *International Review of Psychiatry, 30*(1), 40–61.

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(8), 921–929.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales – 2nd edition (VABS-II)*. Livonia, MN: Pearson Assessments.

Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research, 6*(6), 468–478.

Williams, D. L., Siegel, M., & Mazefsky, C. A. (2017). Problem behaviors in autism spectrum disorder: Association with verbal ability and adapting/coping skills. *Journal of Autism and Developmental Disorders, 48*(11), 3668-3677.

Table 1. Descriptive breakdown by language group

	MV	PS	VF
N	296	303	1338
Age in years (*M*, SD)	9.84, 3.11	9.54, 2.97	10.28, 2.90
Age (% 12)	24%	20%	27%
Gender (% Male)	85%	82%	88%
Race (% White)	62%	74%	83%
Maternal Education (% BS)	55%	59%	61%
NVIQ (% <70)	85%	41%	9%

Note: % 12 =at least 12 years old; BS=Bachelor's degree or higher; NVIQ=Nonverbal IQ

Table 2. Pearson correlations of CBCL domains and ABC subscales by language

	MV			PS			VF		
ABC Subscales	Irrit.	Leth.	Hyper.	Irrit.	Leth.	Hyper.	Irrit.	Leth.	Hyper.
CBCL Internalizing	0.41*	0.50*	0.28*	0.34*	0.44*	0.20*	0.50*	0.50*	0.32*
CBCL Externalizing	0.72*	0.28*	0.64*	0.73*	0.21*	0.61*	0.76*	0.25*	0.64*

Note: *$p < 0.001$; Irrit.=Irritability; Leth.=Lethargy; Hyper.=Hyperactivity

Table 3. Pearson correlations of CBCL Int syndrome scales and ABC Lethargy by language

	MV	PS	VF
CBCL Anxious/Depressed	0.35*	0.25*	0.32*
CBCL Somatic Complaints	0.23*	0.11**	0.29*
CBCL Withdrawn/Depressed	0.51*	0.63*	0.63*

*p<.001; **p<.05

Figure Legends

Clinical

Borderline

Not Clinical

*p<.001

Figure 1. Language group differences of mean ABC subscale scores.
Note: *p<.001; Bars represent +/- 1 SE; [1]group differences were no longer significant when controlling for NVIQ; MV=Minimally Verbal; PS=Phrased Speech; VF=Verbally Fluent.

Figure 2. (A-B) Language group differences of mean CBCL domain scores; (C-D) Language group differences of distribution of mean CBCL domain score distributions in the clinical, borderline, and not clinical range.

Note: *p< .001; Bars represent + /- 1 SE; MV=Minimally Verbal; PS=Phrased Speech; VF=Verbally Fluent.

Clinical Borderline Not Clinical

Figure 3. (A-C) Language group differences of mean CBCL syndrome scale scores; (D-F) Language group differences of distribution of mean CBCL syndrome scale score distributions in the clinical, borderline, and not clinical range.

Note: **p*< .001; Bars represent + /- 1 SE; MV=Minimally Verbal; PS=Phrased Speech; VF=Verbally Fluent.

Clinical Borderline Not Clinical