Published in final edited form as:

Res Dev Disabil. 2013 April ; 34(4): 1190–1197. doi:10.1016/j.ridd.2013.01.002.

An Evaluation of the Aberrant Behavior Checklist for Children Under Age 5

Jonathan D. Schmidt,

Department of Behavioral Psychology, Kennedy Krieger Institute, Johns Hopkins University School of Medicine, Baltimore, MD

John M. Huete,

Department of Behavioral Psychology, Kennedy Krieger Institute, Johns Hopkins University School of Medicine, Baltimore, MD

Jill C. Fodstad,

Johns Hopkins University School of Medicine, Department of Behavioral Psychology, Kennedy Krieger Institute, Baltimore, MD

Michelle D. Chin, and

Department of Behavioral Psychology, Kennedy Krieger Institute, Baltimore, MD

Patricia F. Kurtz

Johns Hopkins University School of Medicine, Department of Behavioral Psychology, Kennedy Krieger Institute, Baltimore, MD

Abstract

Severe problem behaviors such as self-injury and aggression are frequently observed in young children under age 5 with intellectual and developmental disabilities (IDD). Although early identification of problem behavior is critical to effective intervention, there are few standardized measures available that identify severe problem behavior in this population. The Aberrant Behavior Checklist-Community (ABC-C; Aman & Singh, 1994) is a rating scale that measures the severity of a range of problem behaviors commonly observed in individuals with IDD. While it has been used with children under 5, investigations into the fit of the ABC-C for this population are sparse. The purpose of the present study was to report on ABC-C scores in a sample of 97 children under age 5 with problem behavior. Analyses included evaluating differences in scores between age groups, comparing sample norms to established norms for older children, and conducting a confirmatory factor analysis. Results indicated differences in mean scores based on age with younger children generally scoring higher on some subscales of the ABC-C. Furthermore, the original 5-factor structure of the ABC-C was not fully supported. In general, the ABC-C may over- or underestimate behavior problems in younger children; therefore more extensive investigation into the utility of the ABC-C for children under age 5 is warranted.

© 2013 Elsevier Ltd. All rights reserved.

Address correspondence to Jonathan D. Schmidt, Department of Behavioral Psychology, Kennedy Krieger Institute, 707 N. Broadway, Baltimore, MD 21205. Phone: 443-923-9209. schmidtj@kennedykrieger.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

problem behavior; young children; developmental disability; Aberrant Behavior Checklist

Young children with intellectual and developmental disabilities (IDD) are at risk for developing severe problem behavior such as self-injury, aggression, disruption, and tantrums. Prevalence estimates for severe behavior problems range from 13% to 30% in children with IDD, and are higher for children with an autism spectrum disorder (ASD) (Sturme, Seiverling, & Ward-Horner, 2008). These behavior problems can persist into adolescence and adulthood, often having deleterious effects on learning and social opportunities, and if severe enough, may result in extensive use of psychotropic medications or more restrictive environments. Early identification of certain behavior problems such as stereotypy and self-injurious behavior (SIB; e.g., head banging) can be complicated, though, as topographically-similar behaviors frequently are observed in typically-developing children (deLissovoy 1961). Therefore, assessment measures are needed to not only identify behavior problems in young children, but also classify and monitor the severity of these behaviors in order to distinguish children who are at greatest risk for developing chronic behavior problems. Unfortunately, there are few available measures that adequately assess the severity of problem behavior in children with or at risk for IDD under the age of 5 years.

The Aberrant Behavior Checklist (ABC) is one of the few empirically developed scales designed to measure psychiatric symptoms and behavioral disturbance exhibited by individuals with IDD across 5 domains: Irritability, Agitation, & Crying; Lethargy/Social Withdrawal; Stereotypic Behavior; Hyperactivity/Noncompliance; and Inappropriate Speech (Aman & Singh, 1986). The ABC originally was developed to assess the effectiveness of psychotropic medication, and has been used extensively in pediatric, as well as adult behavioral and psychiatric research due to its high reliability and validity (Aman, Singh, Stewart, & Field, 1985; Bihm & Poindexter, 1991; Paclawskyj, Matson, Bamburg, & Baglio, 1997; Matson, Cooper, Malone, & Maskow, 2008). Although the ABC was developed for use with individuals living in institutions and residential settings, revisions to the measure have made it more applicable for home and school settings; the commonly used measure is now the Aberrant Behavior Checklist-Community (ABC-C) version (Aman & Singh, 1994).

Perhaps due to the lack of available measures, the ABC-C has been utilized with children under the age of 5 (for this paper “young children” will refer to this age group), to evaluate behavioral and psychopharmacological treatment effects (Brown, Aman, & Haverkamp, 2002; Chadwick, Piroth, Walker, Bernard, & Taylor, 2000). The ABC-C also has been used to track the persistence of emerging problem behavior over several years for young children (Berkson, Tupa, & Sherman, 2001; Green, O’Reilly, Itchon, & Sigafos, 2005). For example, Green et al. (2005) administered the ABC-C every six months over a three-year time span to teachers of 13 preschool children with previously identified developmental disabilities. Each of the children emitted some form of problem behavior at the beginning of the study. Results of the repeated administration of the ABC-C showed that for nine of the children who exhibited some of the highest levels of problem behavior, their scores on the ABC-C remained elevated across the three years of measurement.

Despite its use with young children for assessment and treatment purposes, little is known regarding the (1) norms of the ABC-C, and (2) appropriateness or sensitivity of the ABC-C for this population, as the factor structure of the ABC-C has not been extensively evaluated with this population. When factor analyses have been conducted, sample sizes typically have been small (Sigafos, Pittendreigh, & Pennell, 1997) or included large age ranges (e.g., ages

3–23 years; Freund & Reiss, 1991; Rojahn & Helsel, 1991). This has made it difficult to draw more specific conclusions about the appropriateness of the ABC-C for these young children.

Thus, although the ABC-C is used with young children, there are several issues regarding its utility with this population. A primary concern is the developmental appropriateness of the items. The ABC-C was developed and normed with adults, therefore the items may only sample behavior problems commonly associated with adults with IDD. Behaviors which are concerning for young children may not be accounted for, and behaviors which are not relevant to young children may be overrepresented. Therefore, scores on the ABC-C for young children may be susceptible to false negative results. For example, it is likely that for younger children, lower scores may be obtained for certain subscales (e.g., Inappropriate Speech or Hyperactivity) because specific items (e.g., “talks excessively” or “inappropriately noisy and rough”) may not be appropriate due to the child’s developmental level. This may affect the sensitivity of the measure for tracking changes in problem behavior in this and other domains across time.

A secondary concern is that due to the potential developmental inappropriateness of certain items, the traditional 5-factor ABC-C structure may not be adequate for this population; research in this area has not yet provided conclusive results. Sigafos et al. (1997) evaluated the factor structure of the ABC-C for 32 children less than 6 years of age, as rated by parents and teachers. Results showed that the scores from both groups of raters were highly correlated, but more importantly, items loaded on the original 5-factors. Although these initial findings favor the appropriateness of the ABC-C with this population, there is a need for independent replication of these procedures to establish utility.

In summary, while there is a large body of research supporting the effectiveness of the ABC-C with a variety of populations, the research support for using the ABC-C with younger children is not sufficient. The current study had four goals. First, ABC-C mean scores were compared for 2 groups of children under the age of 5: one group were referred to a clinic for the treatment of multiple behavior problems while a second group included children from a research project investigating the emergence of self-injurious behavior. Second, the scores of the combined sample were compared to established norms for children ages 6–10 years (Brown et al., 2002) to evaluate differences. Third, a confirmatory factor analysis was conducted to determine the fit of the five-factor structure of the ABC-C with scores from the current study sample. Fourth, individual items were examined regarding the frequency of endorsement across different age groups.

Method

Participants

Demographic information for the study sample is summarized in Table 1. Ninety-seven children under the age of 5 years and their primary caregivers participated; 73 children (75.3%) were male, and 24 children (24.7%) were female. The study sample was comprised of two distinct groups. A majority of the sample, 65 children (Mean age = 3.08 years, range = 0.80 to 4.50), were referred to a hospital-based outpatient clinic for severe behavior problems over a 14-year time span (1997–2010). The other 32 children (Mean age = 2.49 years; range = 1.35 to 4.81) were recruited over a 5-year time span (2004–2008) for a clinic and home-based research study that specifically examined the emergence of SIB in young children (R01HD046722; Kurtz, Chin, Huete, & Cataldo, 2012). For the purposes of this paper, these two participant groups will be referred to as *clinic sample* and *research sample*, respectively. *Combined sample* will be used to refer to all participants (i.e., clinic sample plus research sample). Approximately 45.4% ($n = 44$) of the combined sample had some

previously documented level of developmental delay or intellectual disability; many had not been formally evaluated or diagnosed due to their young age. Approximately 13.4% ($n=13$) of participants were diagnosed with an autism spectrum disorder (ASD).

Data Collection

The ABC-C was provided to each child's primary caregiver at either the first appointment in the outpatient clinic or at the first research study visit. Completed forms typically were collected at or within 2 weeks of that initial visit. Trained examiners who were supervised by licensed psychologists scored all forms.

Aberrant Behavior Checklist-Community

The ABC-C is a 58-item questionnaire for caregivers that is designed to assess the presence and severity of various problem behaviors commonly observed in individuals diagnosed with IDD. Each item is scored as 0 (never a problem), 1 (slight problem), 2 (moderately serious problem), or 3 (severe problem). Items load onto one of five empirically derived subscales: Irritability, Agitation, & Crying (15 items); Lethargy/Social Withdrawal (16 items); Stereotypic Behavior (7 items); Hyperactivity/Noncompliance (16 items); and Inappropriate Speech (4 items). In addition, a Total Score can be calculated. In initial examinations of the utility of the measure, the ABC was shown to have sound psychometric properties with high internal consistency among subscales (mean $\alpha = 0.91$), excellent test-retest reliability (mean $r = 0.98$), acceptable inter-rater reliability (mean $r = 0.63$), and moderate correlations with measures of adaptive behavior (mean $r = 0.60$) (Aman et al., 1985).

Data Analysis

Preliminary analyses were conducted to determine if sample groups (i.e., clinic participants and the research participants) differed significantly with respect to age, IDD status, and ABC subscale scores. Independent sample t -tests were conducted for continuous data and a chi-square analysis was conducted for nominal data. It should be noted that the demographic variables of gender and ASD status were not able to be adequately assessed for differences across the two samples due to low endorsement rate (e.g., only 4 research participants were diagnosed with ASD at the time of data collection vs. 9 clinic participants, etc.).

Evaluation of study sample scores—During the first set of analyses, the clinic and research participants were evaluated as two separate groups. Within each group, participants were categorized by sex and by age (<1 year, 1 year, 2 years, 3 years, and 4 years). After being separated, the mean score for each subscale and a Total Score were calculated for each group. Means were evaluated to identify any significant differences due to age or sex.

For the remaining analyses, the *combined* sample of 97 participants was used for evaluation. Participants were grouped by sex and by age (<1 year, 1 year, 2 years, 3 years, and 4 years). After being separated, the mean score for each subscale and a Total Score were calculated for each group. However, due to the limited number of females, for the next analysis we ignored sex and grouped participants according to age only. A Welch t -test was conducted – due to the unequal sample sizes (and thus assumed unequal variances) – to determine if there were significant differences between the mean scores for each age group on each of the five ABC-C subscales, as well as for the Total ABC-C score (note: due to only having two children under the age of 1 a combined “under 2” age group was made which consisted of participants up to 1 year, 11 months of age). Additionally, mean scores were compared using a Welch t -test based on whether or not the participants had a diagnosis of an ASD, and whether the participants had a documented form of a developmental delay. To control for inflation of Type I error rate during these analyses, a Bonferroni correction was conducted

whereby alpha was adjusted to 0.01. Ultimately, the mean scores of the current sample were compared to the mean scores of 6–10 year-old children for comparison (Brown et al., 2002), as these scores were for children closest in age to the study sample.

Confirmatory factor analysis—A confirmatory factor analysis (CFA) was utilized to examine the fit between the original factor assignments of the ABC with the present data. Confirmatory factors were based on a Pearson inter-item covariance matrix; factors were allowed to correlate but measurement errors were not. Parameter estimations were made using Maximum Likelihood Estimation (MLE) with an oblique rotation. A second CFA was conducted where each item was coded dichotomously as *not occurring* (score of 0 or missing) or as *occurring* (score of 1, 2, or 3). This analysis was conducted in an attempt to account for infrequently endorsed items in a way that reflects item response instead of total item score; similar analyses have been employed with prior research using the ABC-C (Brown et al., 2002; Newton & Sturmey, 1988).

The AMOS program available for SPSS 20.0 was utilized for all CFA procedures. In addition to the normed chi-square value (χ^2/df), a close-fit index (Root Mean Square Error of Approximation [RMSEA]) and an incremental close-fit index were utilized (Comparative Fit Index [CFI]) to evaluate the goodness-of-fit between the measurement model and the data (Hu & Bentler, 1999). An adjusted RMSEA of .08 (due to small sample size; Browne & Cudek, 1993; Hu & Bentler, 1999) and close to .95 for the CFI (Hu & Bentler, 1999) were considered a good fit between the model and data. A χ^2/df value below 2.0 was used to determine fit acceptability (Carmines & McIver, 1981).

After the CFA, we investigated the frequency with which individual items were endorsed because of the potential differences between the mean scores of our study sample and those of the 6–10 year-old comparison; we compared frequencies across age groups as well as across the entire sample with the 6–10 year-olds normative data. The endorsement of individual items by care providers was also investigated to determine the frequency with which specific items were endorsed. A Welch t-test with an adjusted alpha of 0.01 was utilized for this evaluation.

Results

Evaluation of study sample score distribution

Results of preliminary *t*-test analyses indicated that there was a significant difference with respect to age, $t(1, 95) = 4.82, p < .001$ among the two samples. Age differences were expected, given that the research sample was selected to examine the emergence of SIB in very young children. Conversely, there were no significant differences between the research and clinic samples with subscale scores, all $p > .05$. Thus, given the limited amount of use with this measure, it would not be appropriate to conduct extensive analyses with the groups separate (e.g., the research sample only had 32 participants, there as a large difference in the number of males and females between samples). As such, for subsequent analyses data from the combined sample is reported, as opposed to reporting data from the clinic sample and research sample separately.

Analyses of the combined sample showed there were significant age differences in mean scores of the ABC-C subscales (Table 2). There was a significant difference in the scores of children under the age of 2 when compared to 2 year-olds on the Irritability ($-9.92, p = .001$) and Inappropriate Speech ($-2.477, p = .001$) subscales. Likewise, significant differences in mean scores were also found for children under the age of 2 when compared to 2, 3, and 4-year-olds on Lethargy ($-4.30, p = .004$; $-5.38, p = .016$; and $-6.37, p < .003$ respectively) and Hyperactivity ($-14.01, p < .001$; $-12.37, p = .002$; and $-14.07, p < .001$).

respectively) subscales, as well as Total score means (32.35, $p < .001$; -26.10, $p = .007$; and -28.75, $p = .002$ respectively).

Results of the chi-square indicated that there were no significant differences with respect to IDD status between the two samples, $\chi^2 (1, N = 97) = 2.35, p > .05$. Subsequent analyses were conducted to further investigate the impact of these differences. After controlling for a diagnosis of an ASD there were some small yet significant differences in mean scores on some subscales; however, these results must be interpreted with caution because of the uneven number of participants in each category, and the relatively small number of participants that had been diagnosed with an ASD (13 total for the sample; 4 research participants and 9 clinic participants). Outcomes from the Welch t -test analyses indicated that there were significant differences on the subscale of Lethargy, $t(95) = 2.91, p = .01$ and Stereotypy, $t(95) = 2.93, p = .01$, and in both instances those with ASD had higher mean scores on each subscale ($M = 12.69$, Lethargy; 7.83, Stereotypy) than participants not diagnosed as having an ASD ($M = 5.77$, Lethargy; 3.49, Stereotypy). Conversely, when similar analyses were conducted to investigate differences with respect to IDD status, no significant differences were found across any of the ABC-C subscales.

Comparison of sample scores to established norms

Table 3 displays the results of the t -test comparing the mean scores of the study sample to established mean scores of children 6–10 years of age (Brown et al., 2002). When looking at results across age groups, differences between the current sample and that of the preexisting norms from Brown et al. (2002) were somewhat consistent. Mean scores for the current combined sample were higher and significantly different for all age groups (children less than two years of age [$t(32) = 3.36, p = .002$], as well as children 2 years [$t(37) = 8.14, p < .001$], 3 years [$t(20) = 4.51, p < .001$], and 4 years of age [$t(16) = 3.69, p = .002$]). On the subscale of Lethargy/Withdrawal, significant differences were only found when children less than or equal to one year old ($M = 2.82$) were compared to the 6–10 years of age norms [$M = 5.47$; $t(50) = 3.22, p = .002$]. Significant differences on the Stereotypic Behavior subscale were found when children less than or equal to 1 year of age [$t(32) = 2.55, p = .01$] and when children 2 years old [$t(42) = 3.27, p = .002$] were compared to the normative sample, with both age groups having higher mean scores [$M = 5.75$ (less than or equal to 1 years old); 4.67 (2 years old)] than those 6–10 years old ($M = 2.71$). With respect to scores on the Hyperactivity subscale, children who were 2 years [$t(40) = 4.30, p < .001$], 3 years [$t(21) = 2.95, p = .007$], and 4 years of age [$t(17) = 3.68, p = .002$] were found to have statistically higher scores than the 6–10 year old normative sample. Finally, on the Inappropriate Speech subscale, children less than 2 years of age were found to have scores significantly lower than the normative sample [$t(56) = 3.38, p = .001$].

Results for the combined sample (without the inclusion of age categories as a variable) indicated that there was a significant difference between mean scores of the current sample and previously established scores for the Irritability [$t(144) = 9.04, p < .001$], Stereotypic Behavior [$t(180) = 2.64, p = .009$], and Hyperactivity subscales [$t(164) = 4.04, p < .001$]. There was no significant difference for the Inappropriate Speech subscale, even though differences in language were expected due to the age difference between the study and comparison samples.

Confirmatory factor analysis

The first CFA for the ABC-C yielded an RMSEA value of 0.12 (90% CI [0.11, 0.12]). The CFI value for the ABC was 0.55 and $\chi^2/df > 2.00 = 2.36$. These results suggest the data from the current sample do not adequately fit the original 5-factor structure of the ABC-C, indicating it is not an adequate measure of behavioral difficulties in this sample of

participants. Results for the second CFA yielded an RMSEA value of 0.09 (90% CI [0.08, 0.09]). The CFI value for the ABC-C was 0.60 and the $\chi^2/df < 2.00 = 1.76$. Although this second CFA yielded data that were a better fit than the first CFA, the RMSEA and CFI are both outside of an adequate range of acceptability.

To examine how frequently items were endorsed for each age group, we considered items endorsed for 70% of participants or more to be “*highly endorsed*” items (see Table 4). Within the Irritability subscale a total of seven items were highly endorsed and included items such as “injures self,” “temper tantrums,” and “needs demands met quickly.” On the Hyperactivity/Impulsivity subscale, a total of three items, were highly endorsed, “disobedient,” “does not attend to instructions,” and “will not sit still”; however, several items were endorsed between 60% and 70%. No additional items on any of the other subscales were considered highly endorsed.

Also of note is the frequency with which certain items were not endorsed on some of the ABC-C subscales. For the purposes of our evaluation we considered items endorsed for 20% of participants or less to be “*low endorsed*” items. For example, on the Irritability subscale, one item, “depressed,” met this criterion. Several items on the Lethargy subscale were considered low endorsed items including “inactive,” “listless/sluggish,” and “sits/stands in one position.” No additional items on any of the other subscales were considered low endorsed.

Discussion

The goals of this study were to (1) compare ABC-C mean scores for young children from clinic and research samples, (2) compare the scores of the combined sample to previously established norms for youth ages 6–10 years, (3) evaluate if the factor structure of the ABC-C is appropriate for this sample of young children, and (4) determine how frequently individual items were endorsed for participants. More consistent differences were found within the combined sample for this study when the means scores were compared across ages (i.e., children under age 2 when compared to children over the age of 2). Results indicated that when the current sample was categorized by more stringent age groups, and compared to previously established norms for 6–10 year-olds, each age group was different from the comparison sample on the Irritability subscale, but were more variable for the other subscales. When the age groups were collapsed and compared to the established scores for 6–10 year-olds, scores were significantly higher for the current sample on the Irritability, Stereotypic Behavior, and Hyperactivity subscales. Contrary to prior studies with younger children (e.g., Sigafoos et al., 1997), items did not load adequately on to the existing 5-factor structure. From these results, there are several interesting points that warrant further discussion.

Regarding the first goal, the comparison of the mean scores of the clinic and research samples, no general significant differences were found on the ABC-C subscales. However, after categorizing the sample by age groups, results clearly showed that there were significant differences for some subscales, but not others. Specifically, children under age 2 tended to have significantly lower mean scores, particularly on the Irritability, Hyperactivity, and Lethargy Subscales (see Table 2). Most likely, the differences in scores were due to the child’s developmental level. For example, on the Hyperactivity and Lethargy subscales, children under the age of 2 may be less physically mobile or active (e.g., walking or running) than children older than 2. Additionally, items within the Irritability subscale such as “yells inappropriately,” “bangs objects or slams doors,” and “depressed,” reference behaviors not likely to be intense enough to warrant enough attention due to the child’s physical limitations and abilities. The same is true for items on the Lethargy

(“listless/sluggish,” “seeks isolations,” and “prefers to be alone”) and Hyperactivity (“excessively active,” “impulsive,” and “constantly runs or jumps”) subscales. Results from additional analyses in this study that investigated the frequency with which items from these subscales were endorsed supports the claim that these items may not be appropriate for this age group (see Table 4). For these reasons, it becomes even more important to determine how well the current items on the ABC-C fit into the predetermined 5-factor structure.

When scores of this sample were compared to normative standards for 6–10 year-old children, our hypothesis that the scores would be comparable was partially supported, but not to the degree that was expected. Results show that the combined sample was significantly higher than the normative sample on the Irritability, Stereotypic Behavior, and Hyperactivity subscales. This is likely explained by the nature of the two different samples. That is, scores for the sample included in the Brown et al. (2002) study were from mail surveys that were returned for individuals who received special education services; the individuals did not have to emit problem behavior to be included. For the current study, all children were referred to a clinic for problem behavior that was deemed to be of considerable nature by their parents or a professional, or they were referred specifically for a study targeting the early emergence of SIB. Thus, differences would be expected due to the process in which each of the two studies collected data. Irrespective of this point, further analysis of scores for individual age groups provided evidence to support the hypothesis from the first objective; individuals under the age of 2 were significantly lower than the normative sample on the Inappropriate Speech and Lethargy subscales.

Regarding goals three and four, the current results suggest that the ABC-C factor structure may not be as informative for younger children as it is for older children, and may need to be revised for this young population. Outcomes of both CFA analyses of the ABC-C did not indicate there to be an adequate fit to the data. Thus, results of these analyses suggest that the five-factor structure proposed by Aman et al. (1985; 1994) may not accurately capture behavior problem domains in young children with behavioral difficulties. When items are not applicable, this produces the obvious effect of having fewer items that are available for discriminating between typical and atypical results in an age group. As a result, to build on a point referenced earlier, differences in scores may be due to a child’s developmental level, and not to an actual deficit. This makes it even more critical to assess the impact of the child’s communication ability, since this may have a dramatic effect on their ABC-C scores. Particularly, many items on the Inappropriate Speech subscale may subsequently be irrelevant (e.g., “talks excessively” or “talks loudly to self”). Additionally, further investigation of the frequency with which items were endorsed revealed several items that were endorsed below 20% of the time across the 58-items of the ABC-C.

While results of the current study are important in understanding the applicability of the ABC-C to younger children, there are several limitations that should be noted and used to guide future research on the use of the ABC-C with this age group. First, due to our small sample size, our goal was never to conduct an exploratory factor analysis, and in fact it would have been inappropriate to attempt to make a definitive claim with such a small sample. Future studies should evaluate the ABC-C with a larger sample to (1) implement a more powerful CFA, (2) conduct an exploratory factor analysis if necessary, (3) revise individual scoring items where applicable, and (4) ultimately, establish normative scores for this population. Additionally, the current sample consisted of a known clinical sample (outpatient referred participants) and non-clinical research study population which was identified as exhibiting emerging severe problem behavior (i.e., self-injury). Therefore, an adequate control group was not included for comparison. Future studies should include a neurotypical population for comparison, particularly because certain forms of problem

behavior are present at younger ages regardless of disability status (deLissovoy, 1961), and will thus be important for identifying atypical scores.

In addition, the convergent and divergent validity of the ABC-C to other similar measures that are relevant and commonly used with children under the age of 5 who have problem behavior was not evaluated. Since the item content of the ABC-C may not be reflective of the range of problem behaviors exhibited by young children with or at risk for IDD, it may miss certain behavior problems commonly observed in younger children and produce false negative findings. Comparing the outcomes of the ABC-C with other measures of problem behavior in younger children would aid in better understanding this.

Despite the aforementioned limitations, results of this study are a building block for future research to further evaluate the sensitivity and utility of the ABC-C for identifying and measuring the severity of problem behavior with younger children. Specifically, there is a dire need for evaluating measures more thoroughly if they are going to be used with younger children with IDD, since the presentation and severity of their problem behaviors may differ from that of older children as a result of developmental factors and more recent onset of the behavior problem. Results of this study indicate the ABC-C may have varying levels of sensitivity based on the age of the individual that is being rated, and thus may not allow for an accurate depiction of their level of problem behavior. For this reason, this study should serve as the impetus for further research regarding the sensitivity of the ABC-C.

Acknowledgments

This study was supported by a grant from the National Institute of Child Health and Human Development (R01HD046722).

References

- Aman, MG.; Singh, NN. Aberrant Behavior Checklist Manual. East Aurora, NY: Slosson Publications; 1986.
- Aman MG, Singh NN, Stewart AW, Field CJ. Psychometric characteristics of the Aberrant Behavior Checklist. *American Journal of Mental Deficiency*. 1985; 89:492–502. [PubMed: 3158201]
- Aman, M.; Singh, N. The Aberrant Behavior Checklist-Community. East Aurora, NY: Slosson Education Publications, Inc.; 1994.
- Berkson G, Tupa M, Sherman L. Early development of stereotyped and self-injurious behaviors: I. Incidence. *American Journal on Mental Retardation*. 2001; 106:539–547. [PubMed: 11708939]
- Bihm EM, Poindexter AR. Cross-validation of the factor structure of the Aberrant Behavior Checklist for persons with mental retardation. *American Journal on Mental Retardation*. 1991; 96(2):209–211. [PubMed: 1930950]
- Brown EC, Aman MG, Havercamp SM. Factor analysis and norms for parent ratings on the Aberrant Behavior Checklist-Community for young people in special education. *Research in Developmental Disabilities*. 2002; 23:45–60. [PubMed: 12071395]
- Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Beverly Hills, CA: Sage; 1993. p. 136-162.
- Carmines, EG.; McIver, JP. Analysing models with unobservable variables. In: Bohrnstedt, GW.; Borgatta, EE., editors. *Social measurement current issues*. Beverly Hills, CA: Sage; 1981. p. 65-115.
- Chadwick O, Piroth N, Walker J, Bernard S, Taylor E. Factors affecting the risk of behaviour problems in children with severe intellectual disability. *Journal of Intellectual Disability Research*. 2000; 44(2):108–123. [PubMed: 10898374]
- deLissovoy V. Head banging in early childhood: A study of incidence. *Journal of Pediatrics*. 1961; 58:803–805. [PubMed: 13720644]

- Freund LS, Reiss AL. Rating problem behaviors in outpatients with mental retardation: Use of the Aberrant Behavior Checklist. *Research in Developmental Disabilities*. 1991; 12(4):435–451. [PubMed: 1792366]
- Green VA, O'Reilly M, Itchon J, Sigafoos J. Persistence of early emerging aberrant behavior in children with developmental disabilities. *Research in Developmental Disabilities*. 2005; 26(1):47–55. [PubMed: 15590237]
- Hu L, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- Kurtz PF, Chin MD, Huete JM, Cataldo MF. Identification of emerging self-injurious behavior in young children: A preliminary study. *Journal of Mental Health Research in Intellectual Disabilities*. 2012; 5:260–285. [PubMed: 22844389]
- Matson JL, Cooper C, Malone CJ, Moskow SL. The relationship of self-injurious behavior and other maladaptive behaviors among individuals with severe and profound intellectual disability. *Research in Developmental Disabilities*. 2008; 29(2):141. [PubMed: 17350800]
- Newton JT, Sturmey P. The Aberrant Behavior Checklist: A British replication and extension of its psychometric properties. *Journal of Mental Deficiency Research*. 1988; 32:87–92. [PubMed: 3398039]
- Paclawskyj TR, Matson JL, Bamburg JW, Baglio CS. A comparison of the Diagnostic Assessment for the Severely Handicapped–II (DASH–II) and the Aberrant Behavior Checklist (ABC). *Research in Developmental Disabilities*. 1997; 18(4):289–298. [PubMed: 9216028]
- Rojahn J, Helsel WJ. The Aberrant Behavior Checklist with children and adolescents with dual diagnosis. *Journal of Autism and Developmental Disorders*. 1991; 21(1):17–28. [PubMed: 2037547]
- Sigafoos J, Pittendreigh N, Pennell D. Parent and teacher ratings of challenging behaviour in young children with developmental disabilities. *British Journal of Learning Disabilities*. 1997; 25(1):13–17.
- Sturmey, P.; Seiverling, L.; Ward-Horner, J. Assessment of challenging behaviors in people with autism spectrum disorders. In: Matson, JL., editor. *Clinical assessment and intervention for autism spectrum disorders*. Burlington, MA: Elsevier Inc.; 2008. p. 131-156.

Highlights

- Evaluated the utility of the ABC-C for 97 children with behavior problems.
- Younger children generally scored higher on several subscales of the ABC-C.
- Children under age 5 had significantly higher scores than 6–10 year olds.
- A confirmatory factor analysis did not support the factor structure of the ABC-C.

Table 1

Demographic Information for Study Sample

	Combined Sample (N = 97)	Clinic Sample (n=65)	Research Sample (n=32)
Sex			
Male	73	48	25
Female	24	17	7
Mean Age	2.79 years	3.08 years	2.49 years
Autism Spectrum Disorder	13.4%	13.8%	12.5%
Developmental Delay or Intellectual Disability	45.4%	43.1%	50.0%
Race/Ethnicity			
Caucasian	48	31	17
African-American	30	21	9
Hispanic	4	2	2
Asian	3	2	1
Other	12	9	3

Table 2

Mean Difference in ABC-C subscale and Total Scores by Age

Dependent Variable	Age 1	Age 2	Mean Difference	Sig
Irritability	<2	2	−9.92 *	.001
		3	−5.89	.083
		4	−4.02	.246
	2	3	4.03	.235
		4	5.91	.096
	3	4	1.88	.635
Lethargy	<2	2	−4.30 *	.004
		3	−5.38 *	.016
		4	−6.37 *	.003
	2	3	−1.11	.633
		4	−2.10	.333
	3	4	−0.98	.709
Stereotypy	<2	2	−1.67	.093
		3	−1.30	.315
		4	−3.00 *	.036
	2	3	0.37	.774
		4	−1.33	.342
	3	4	−1.70	0.294
Hyperactivity	<2	2	−14.01 *	<.001
		3	−12.37 *	.002
		4	−14.07 *	<.001
	2	3	1.64	.676
		4	−0.60	.987
	3	4	−1.70	.294
Inappropriate Speech	<2	2	−2.47 *	<.001
		3	−1.16	.057
		4	−1.29	.135
	2	3	1.31	.096
		4	1.18	.235
	3	4	−.14	.884
Total	<2	2	−32.35 *	<.001
		3	−26.10 *	.007
		4	−28.75 *	.002
	2	3	6.25	.513

Dependent Variable	Age 1	Age 2	Mean Difference	Sig
		4	3.60	.694
	3	4	-2.65	.803

Note: Scores under the “Age 1” column were subtracted from scores in the “Age 2” column to achieve the mean difference score, with negative scores indicating the score from the Age 1 column was lower.

*
 $p < .01$

Table 3

Comparison of Mean Scores of Study Sample to Established Norms

	2 (n=28)	2 (n=33)	3 (n=20)	4 (n=16)	Combined Sample (N=97)	Brown et al. (2002) (N = 195)	t (p)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	
Irritability	15.11* (9.93)	25.03* (11.18)	21.00* (12.14)	19.13* (11.30)	20.36 (11.59)	8.50 (8.11)	9.04 (< 0.001)*
Lethargy/Withdrawal	2.82* (3.70)	7.09 (6.98)	8.20 (8.73)	9.19 (7.00)	6.43 (6.98)	5.47 (6.05)	1.16 (= 0.25)
Stereotypic Behavior	4.67* (4.09)	5.75* (4.68)	4.05* (4.84)	2.50 (3.57)	4.08 (4.27)	2.71 (3.99)	2.64 (< 0.01)*
Hyperactivity	9.93 (10.99)	23.94* (13.91)	22.30* (13.68)	24.00* (11.57)	19.57 (13.96)	12.92 (11.70)	4.04 (< 0.001)*
Inappropriate Speech	0.89* (1.45)	3.36 (3.31)	2.05 (2.31)	2.19 (3.12)	2.19 (2.79)	2.01 (2.60)	0.53 (0.60)

Note: Mean scores from the study sample of Brown et al. (2002) were used for comparison to the current study across each age group and within combined sample scores. T-scores correspond to the combined sample mean scores with Brown et al. (i.e., ages 6–10 years mean scores).

* p < .01

Table 4

Percent Item Endorsement Across Age Groups

	2 (n=28)	2 (n=33)	3 (n=20)	4 (n=16)	Total (N=97)
<i>Irritability</i>					
Injures self	85.2	81.8	70.0	62.5	77.1
Aggressive to others	63.0	90.9	80.0	68.8	77.1
Screams inappropriately	53.8	71.9	65.0	62.5	63.8
Temper tantrums	96.4	93.8	95.0	81.3	92.7
Irritable	60.7	87.9	65.0	75.0	73.2
Yells inappropriately	28.6	59.4	65.0	56.3	51.0
Depressed	7.1	27.3	26.3	12.5	18.8
Needs demands met quickly	75.0	93.9	75.0	87.5	83.5
Cries over minor things	42.9	69.7	45.0	43.8	52.6
Mood lability	42.9	72.7	70.0	56.3	60.8
Cries/screams inappropriately	42.9	72.7	50.0	56.3	56.7
Stomps feet/banging on objects	34.6	75.8	70.0	75.0	63.2
Deliberately hurts self	71.4	72.7	75.0	50.0	69.1
Does physical violence to self	67.9	69.7	78.9	75.0	71.1
Throws temper tantrums	82.1	97.0	95.0	87.5	90.7
<i>Lethargy, Social Withdrawal</i>					
Listless, sluggish	0.0	9.1	15.0	18.8	9.3
Seeks isolations	14.3	30.3	35.0	50.0	29.9
Preoccupied	25.0	40.6	30.0	43.8	34.4
Withdrawn	14.3	12.1	35.0	37.5	21.6
Fixed facial expression(s)	7.1	27.3	30.0	18.8	20.6
Sits and watches others	3.6	21.2	10.0	25.0	14.4
Resists physical contact	3.6	33.3	30.0	18.8	21.6
Isolates self	10.7	15.2	30.0	50.0	22.7
Sits/stands in one position	14.3	12.1	10.0	6.3	11.3
Unresponsive to structured activity	17.9	31.3	50.0	31.3	31.3
Difficult to reach/contact	39.3	60.6	55.0	75.0	55.7

	2 (n=28)	3 (n=33)	4 (n=20)	Total (N=97)
Prefers to be alone	7.4	15.2	35.0	43.8
Communicates without words/gestures	21.4	30.3	50.0	56.3
Inactive	3.6	21.2	5.0	9.3
Responds negatively to affection	7.1	39.4	15.8	22.9
Shows few social reactions	10.7	33.3	47.4	33.3
<i>Stereotypic Behavior</i>				
Recurring body movements	32.1	43.8	40.0	68.8
Stereotypy	38.5	50.0	47.4	66.7
Odd/bizarre behavior	14.3	42.4	45.0	50.0
Moves or rolls head back and forth	21.4	28.1	10.0	31.3
Hand, body, or head stereotypy	32.1	50.0	35.0	62.5
Arm or leg stereotypy	23.1	28.1	31.6	53.3
Body rocking	25.0	21.2	20.0	22.7
<i>Hyperactivity, Noncompliance</i>				
Excessively active	24.0	71.9	68.4	93.3
Boisterous	40.7	75.8	63.2	61.1
Impulsive	29.6	66.7	78.9	75.0
Restless	50.0	75.8	75.0	81.3
Disobedient	57.1	78.8	80.0	87.5
Disturbs others	21.4	81.8	80.0	75.0
Uncooperative	46.4	72.7	70.0	81.3
Does not attend to instructions	51.9	75.8	85.0	71.9
Disrupts group activities	29.6	59.4	85.0	53.3
Does not stay in seat	40.7	84.8	75.0	70.8
Will not sit still	37.0	66.7	70.0	93.8
Easily distractible	44.4	75.0	70.0	80.0
Constantly runs or jumps	17.9	72.7	65.0	66.7
Pays no attention when spoken to	42.9	68.8	65.0	54.2
Excessively active	50.0	60.6	65.0	63.2
Deliberately ignores directions	46.4	69.7	70.0	86.7
<i>Inappropriate Speech</i>				63.5

	2 (n=28)	2 (n=33)	3 (n=20)	4 (n=16)	Total (N=97)
Talks excessively	14.8	36.4	30.0	26.7	27.4
Repetitive speech	14.3	46.9	25.0	26.7	29.5
Talks to self loudly	14.3	33.3	10.0	26.7	21.9
Repeats words/phrases	25.0	60.6	45.0	33.3	42.7

Note: Items in bold are either classified as “highly endorsed” (i.e., 70% or greater of the sample indicated the behavior occurred at some severity level) or “low endorsed” (i.e., 20% or less of the sample of the sample indicated the behavior occurred at some severity level).