Published in final edited form as:

J Atten Disord. 2020 October ; 24(12): 1775–1784. doi:10.1177/1087054716663632.

Initial Examination of the BRIEF2 in Clinically Referred Children With and Without ADHD Symptoms

Lisa A. Jacobson^{1,2}, Alison E. Pritchard^{1,2}, Taylor A. Koriakin³, Kelly E. Jones^{1,2}, E. Mark Mahone^{1,2}

¹Kennedy Krieger Institute, Baltimore, MD, USA

²Johns Hopkins University School of Medicine, Baltimore, MD, USA

³University of Connecticut, Storrs, CT, USA

Abstract

Objective—Clinical utility of the Behavior Rating Inventory of Executive Function, Second Edition (BRIEF2) was examined in 1,381 clinically referred youth (62.5% male, 5 to 18 years) with and without ADHD symptoms.

Method—Participants included children with restricted inattentive symptoms (IA only), restricted hyperactivity/impulsivity symptoms (HI only), symptoms in both domains (Combined), and non-ADHD clinical comparison.

Results—Greater hyperactivity (Combined, HI only) was associated with higher Behavior Regulation (BRI) and Emotion Regulation (ERI) scores, whereas greater inattentiveness (IA only, Combined) was associated with higher Cognitive Regulation scores. Effect sizes were largest for Inhibit, Working Memory, and Organization of Materials scales; these scales discriminated children with and without ADHD symptoms and restricted inattentive and hyperactive presentations.

Conclusion—The BRIEF2 distinguishes associated features of ADHD and the day-to-day executive impact. Sensitivity was consistently poorer than specificity. The referred nature of the sample and examination of restricted presentations suggest additional work is needed to examine whether ERI and BRI are dissociable.

Keywords

executive function; attention; self-regulation; rating scales; validity

The Behavior Rating Inventory of Executive Function, Second Edition (BRIEF2; Gioia, Isquith, Guy, & Kenworthy, 2015) is the recent revision of the original Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000), an

Reprints and permissions: sagepub.com/journalsPermissions.nav

Corresponding Author: Lisa A. Jacobson, Department of Neuropsychology, Kennedy Krieger Institute, 1750 E. Fairmount Ave., Baltimore, MD 21231, USA. jacobson@kennedykrieger.org

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

informant-report rating scale designed to assess executive behaviors in children and adolescents. The original BRIEF has been shown to be sensitive to clinical dysfunction across diagnostic groups. Specifically, considering the presence of persisting executive function deficits in many youth with ADHD (Barkley & Fischer, 2011; Seidman, 2006; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005), the BRIEF has been shown to discriminate between children with ADHD and typically developing controls (Mahone, Cirino, et al., 2002; Reddy, Hale, & Brodzinsky, 2011; Toplak, Bucciarelli, Jain, & Tannock, 2009) as well as between ADHD and clinically referred children without ADHD (McCandliss & O'Laughlin, 2007) or those with other clinical disorders (e.g., Gioia, Isquith, Kenworthy, & Barton, 2002; Mahone, Cirino, et al., 2002; Mahone, Zabel, Levey, Verda, & Kinsman, 2002; Vriezen & Pigott, 2002). Although subtypes of ADHD appear notably unstable over time (Lahey, Pelham, Loney, Lee, & Willcutt, 2005) suggesting that presentations may reflect the most severe current concerns rather than durable diagnostic differences, several studies have shown that the BRIEF also discriminates between (current) subtypes of ADHD; children with the combined presentation of ADHD tend to be rated by caregivers as higher on the Behavioral Regulation Index (BRI) and/or Inhibit scale of the BRIEF than are children with predominantly inattentive ADHD (Gioia et al., 2002; McCandliss & O'Laughlin, 2007), potentially reflecting the behavioral inhibition deficits characteristic of those with more hyperactive symptoms (e.g., Barkley, 1997; Fischer, Barkley, Smallish, & Fletcher, 2005). Likewise, children with the inattentive presentation of ADHD tend to be rated as showing a higher level of behaviors indicative of difficulty with working memory and cognitive control (e.g., the Metacognitive Index and/or Working Memory scale; Gioia et al., 2002; McCandliss & O'Laughlin, 2007; Reddy et al., 2011), though this finding has not been consistent across studies (e.g., McCandliss & O'Laughlin, 2007).

The recent revision of the measure, the BRIEF2 (Gioia et al., 2015), has resulted in a reduced number of items, the addition of a new index and renaming of the previous Metacognition Index, and rearrangement of scale content resulting in creation of two scales out of the previous Monitor scale items. Specifically, the new measure now consists of nine scales, with the Inhibit and Self-Monitor scales comprising the Behavior Regulation Index (BRI), the Shift and Emotional Control scales comprising the new Emotional Regulation Index (ERI), and the Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials scales comprising the Cognitive Regulation Index (CRI; the renamed Metacognition Index). The removal of the Shift and Emotional Control scales from the BRI and addition of the Self-Monitor scale to this index suggests that prior work supporting the utility of the BRI in discriminating between ADHD and non-ADHD groups and among ADHD subtypes may no longer be directly applicable. Validation of the new scale arrangement is needed to determine the degree to which the BRI continues to be useful for such purposes. Examination of the BRIEF2 in clinical samples is needed to determine the degree to which prior work remains useful and the confidence with which clinicians can use the revision to support diagnostic decision-making.

The original measure was designed to capture informant reports of children's day-to-day executive functioning in "real world" situations (Gioia, Kenworthy, & Isquith, 2010). Published work examining the BRIEF and other rating scales of executive behavior suggests

that traditional performance-based measures of executive function are not consistently correlated with ratings of everyday functioning provided by caregivers (e.g., McCauley, Chen, Goos, Schachar, & Crosbie, 2010). Given this limited pattern of association, rating measures like the BRIEF appear to measure unique constructs from those assessed by performance-based tests or other broad-band behavior rating scales (Jarratt, Riccio, & Siekierski, 2005) and add clinical value to assessment of everyday application of executive skills and their real-world behavioral impact (Barkley & Fischer, 2011; Goldberg & Podell, 2000; Isquith, Roth, & Gioia, 2013). Specifically, the BRIEF2 authors assert that the new measure was designed to show increased sensitivity to everyday executive dysfunction “in key clinical groups, such as ADHD” (Gioia et al., 2015, p. 3). The test manual describes BRIEF2 parent ratings in a sample of children with ADHD combined presentation (ADHD-C; $n = 218$) and inattentive presentation (ADHD-I; $n = 159$), relative to parent ratings of typically developing children drawn from the normative sample. Overall, the ADHD-C group showed significantly higher scores across all BRIEF2 scales and indices than controls, with the highest scale elevations on the Inhibit, Working Memory, and Shift scales and the highest index elevation on the BRI (Gioia et al., 2015, p. 140). Similarly, the ADHD-I group in this sample also showed significantly higher scores across scales and indices than controls, with highest scale elevations reported on Working Memory, Initiate, and Plan/Organize, and the CRI the most elevated index (Gioia et al., 2015). In a smaller research subsample (Isquith, Kenealy, Roth, & Gioia, 2015, reported in Gioia et al., 2015), parent ratings of children with ADHD-C were reported to be significantly higher than those of children with ADHD-I on Inhibit, Self-Monitor, Shift, and Emotional Control. Taken together, the authors assert that children with ADHD-C have “much higher average T scores on the Inhibit and Self-Monitor scales than children diagnosed with ADHD-I and ... both groups on average have much higher Working Memory, Plan/Organize, and Task Monitor scales than their typically developing peers” (Gioia et al., 2015, p. 140).

These preliminary data from the BRIEF2 authors are promising; however, further external validation of the new measure is needed to determine its utility for clinical purposes. As noted, given the rearrangements among the scale to index loading, prior work regarding the utility of the BRI in particular may not be directly applicable to the revision of the BRIEF. Additionally, most, if not all, of the prior studies investigating performance of either version of the measure in ADHD have examined only children with the inattentive or combined presentations of ADHD (e.g., Gioia et al., 2015; Gioia et al., 2002; McCandliss & O’Laughlin, 2007). Furthermore, meta-analytic and review studies (Doyle, 2006; Willcutt et al., 2005) have found a pattern of broad executive dysfunction among youth with ADHD, with notable heterogeneity, suggesting that more work is needed to characterize the heterogeneity in functioning among individuals with ADHD. To clarify the role of day-to-day inhibitory deficits as compared with attentional dysregulation, examination of patterns among children with all three presentations—including the predominantly hyperactive-impulsive presentation—is important.

The present study investigated the preliminary validity of the BRIEF2 in a large clinically referred sample of children and adolescents by examining patterns among the scales and indices of the BRIEF2 as well as sensitivity and specificity of selected scales in clinically referred youth with and without symptoms of ADHD (including individuals with symptoms

of all three presentations). We hypothesized that the revisions to the measure would provide evidence for discrimination between children with high levels of ADHD symptoms and referred children without ADHD, but also between ADHD symptom presentations. Furthermore, we hypothesized a “double dissociation” model: children with predominantly inattentive ADHD symptoms would show higher ratings on scales comprising the CRI relative to children with the predominantly hyperactive-impulsive presentation, whereas children with the predominantly hyperactive-impulsive presentation would show higher ratings on the scales comprising the BRI.

Methods

Procedures

As part of routine clinical practice at a large outpatient neuropsychology assessment center, parents of children referred to the clinic are typically asked to complete behavioral rating scales and all data are then entered into a clinical database by department clinicians via the hospital electronic health record. These data are securely maintained by the hospital's Information Systems Department. Data were collected over a 4-year period (October 2010–October 2014).

Following approval from the hospital Institutional Review Board, the clinical database was queried, and a limited data set was constructed of patients between 5 and 18 years of age for whom valid scores were available on all measures of interest (specified below). There were no exclusionary criteria beyond complete data on these measures, except as specified in procedures for group selection below.

Participants

Participants included parents of 1,969 youth, ages 5 to 18 years ($M = 10.45$, $SD = 3.22$), referred for outpatient neuropsychological assessment in a large outpatient neuropsychology clinic (see Table 1). Cases were selected for inclusion if complete caregiver ratings (i.e., not more than one missing item per scale) were available on the BRIEF2 and ADHD Rating Scale-IV (ADHDRS), and criteria for membership in one of the four groups were met as specified below. The majority of the children being rated were male (62.5%) and Caucasian (59.63%); 24.46% were African American, 5.45% were multi-racial, 2.3% were Asian American, and 4.8% were of unknown racial background, while 2.25% reported Hispanic ethnicity.

Group assignment—Children were categorized using modified Diagnostic and Statistical Manual of Mental Disorders (5th ed.; *DSM-5*; American Psychiatric Association, 2013) ADHD symptom criteria, including caregiver-report symptom count on the ADHD Rating Scale-IV (DuPaul, Power, Anastopoulos, & Reid, 1998) and ratings of impairment (rating of 4 or greater) in at least one domain of daily function on the Impairment Rating Scale (IRS; Fabiano et al., 2006). Participants in each of the ADHD symptoms groups were required to meet symptom count criteria (6 symptoms in the relevant domain(s) of inattention, hyperactivity/impulsivity, or both) with impairment reported in at least one functional domain. However, consistent with a recent latent class analysis suggesting that the severe

inattentive characterization is best predicted by at least six symptoms of inattention but fewer than three symptoms of hyperactivity/impulsivity (Volk, Todorov, Hay, & Todd, 2009), the IA only symptom group and HI only symptom group reflected *restrictive* symptom presentations (e.g., not simply subthreshold combined presentation) and were restricted to no more than three symptoms of the other symptom domain (i.e., children in the IA only group were reported to display at least six symptoms of inattention but no more than three symptoms of hyperactivity/impulsivity, whereas children in the HI only group were required to have at least six symptoms of hyperactivity/impulsivity, but no more than three symptoms of inattention). These groups are not diagnostic groups, per se, but rather reflect children exhibiting specific patterns of symptoms of ADHD. Four groups ($n = 1,381$) were created from the larger clinical sample, reflecting symptom presentations: a restricted inattentive group (IA only; $n = 395$), a restricted hyperactive/impulsive group (HI only; $n = 77$), a combined symptoms group (Combined; $n = 315$), and a non-ADHD clinical comparison ($n = 594$). The non-ADHD clinical comparison group was selected from the same referred cohort, but participants were only included if they were rated as exhibiting three or fewer symptoms of inattention and/or hyperactivity/impulsivity on the ADHDRS. Children were not included in the analyses if they did not meet the above group assignment criteria ($n = 588$), either because they were rated by parents as showing substantial subthreshold symptoms (five symptoms of inattention, hyperactivity, or both), because at least six symptoms were endorsed within one domain *and* four or five symptoms were endorsed in the other, or because clinical symptoms were endorsed without impairment (IRS).

Measures

ADHDRS—Caregivers completed the ADHDRS (DuPaul et al., 1998), an 18-item measure, reflecting the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; *DSM-IV*; American Psychiatric Association, 1994) and *DSM-5* ADHD diagnostic criteria. Item content reflects *DSM-IV* symptom criteria and ratings are based on the child's behavior over the past 6 months, using a 4-point scale (0 = *not at all*; 1 = *sometimes*; 2 = *often*; 3 = *very much*). Subscales correspond to the *DSM-IV* Inattentive and Hyperactive/ Impulsive criteria. Total subscale scores were obtained by adding item ratings (range: 0-27 for each). The ADHDRS has been shown to demonstrate adequate reliability and validity (DuPaul et al., 1998); internal consistency estimates for the parent-report version ranged from .86 for the Inattention scale to .88 for the Hyperactivity/Impulsivity scale, with test-retest reliability over short periods of time ranging from .78 for Inattention to .86 for Hyperactivity/ Impulsivity. In the current sample, internal consistency for the ADHDRS subscales was good (Inattention $r_\alpha = .859$; Hyperactivity/Impulsivity $r_\alpha = .856$).

IRS—The IRS (Fabiano et al., 2006) is an informant report measure of a youth's impairment in several domains of functioning. The parent version of the measure is comprised of seven items (six domain items and a summary rating); five of these items were collected as part of the present study. Parents were asked to rate, on a 7-point Likert scale, the extent to which their child's problems are impacting their academic progress, peer relationships, relationships with caregivers, self-esteem, and home life. For the purposes of

the present study, a score at or above the midpoint of 4 on this 7-point scale was considered an endorsement of functional impairment in a given domain.

BRIEF2 parent-report form—The BRIEF2 (Gioia et al., 2015) is a revision of the BRIEF, a rating scale assessing everyday behaviors reflecting executive functions across the school-age span (ages 5-18). The BRIEF2 parent-report form contains 63 items within nine theoretically derived clinical scales: Inhibit, Self-Monitor, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials. The nine clinical scales form three Composite or Index scores: the BRI, ERI, and CRI as well as the overall Global Executive Composite (GEC) summary score. All Parent Form coefficient alpha values for index scores were reported to fall above .90, with coefficients for individual scales ranging from .80 (Monitor) to .91 (Emotional Control) in the standardization sample (Gioia et al., 2015, p. 101). The manual reports test-retest reliability estimates from a sample of 163 parent ratings; the mean test-retest correlation coefficient for clinical scales was .79 (range = .67-.92) over an average time interval of 2.9 weeks (Gioia et al., 2015, p. 111). According to the test manual (Gioia et al., 2015, p. 144), the Working Memory scale reportedly showed the largest effect size in comparisons between typically developing controls and youth with ADHD (either inattentive or combined presentation).

For the present study, item level data used for analysis were obtained via the original BRIEF parent rating form. As all of the item content for the revised BRIEF2 is derived from existing items on the original BRIEF, it was possible to extract the items present on the BRIEF2 from original BRIEF data stored in the computer scoring program. The BRIEF2 publisher (PAR, Inc.) then arranged these extracted items into the new BRIEF2 clinical scales and indices. These new clinical scale and index raw scores were compared with the BRIEF2 normative sample to obtain age- and sex-based *T* scores for clinical scales, indices, and the overall composite. Correlations between corresponding BRIEF and BRIEF2 scales were strong, with all values above .90 with exception of Organization of Materials (.83; due at least in part to substantive item changes between versions) and Monitor (.85 with Self-Monitor, .78 with Task Monitor) scales. The overall GEC correlated highly across measures ($r = .975$), with the original Metacognition Index correlating .96 with the CRI and the original BRI correlating slightly lower with the revised BRI (.85).

Analysis Plan

Following group assignment, detailed above, examination of BRIEF2 scale internal consistency was completed via calculation of Cronbach's alpha. Next, two multivariate analyses of variance (MANOVAs) were conducted to examine (a) performance of the BRIEF2 index scores across groups, and (b) performance of the clinical scales across groups, while correcting for multiple simultaneous comparisons. Post-hoc tests, correcting for multiple comparisons (Bonferroni correction), were examined to identify scales that discriminated between clinical groups. Although BRIEF2 scores were age normed, given the broad school-age range of the sample, an additional follow-up multivariate analysis of covariance (MANCOVA) was conducted to determine whether the pattern of between-group differences among clinical scales remained consistent when controlling for age.

Next, as it is important to know the likelihood of specific impairments in clinical groups relative to typically developing children (e.g., Chelune, 2010), clinical base rates of elevated scores across specific scales were calculated across groups, to identify approximate prevalence of such elevations in a mixed clinical sample. Finally, discriminative value for accurately classifying symptom presentations was examined via receiver operator characteristic (ROC) curve analyses.

Results

Reliability

In this mixed clinical sample, the BRIEF2 exhibited excellent internal consistency (Cronbach's $\alpha = .965$) for the items contributing to the GEC. Scale reliabilities (α) ranged from .772 (Initiation, five items) to .923 (Emotional Control, eight items; see Table 2).

Clinical Group Profiles

Index score profiles—As expected, there was a large overall effect of group on the BRIEF2 indices, $F(12, 4128) = 139.84, p < .0001; \eta_p^2 = .289$. Comparison of BRIEF2 Index ratings revealed that children rated as showing both inattentive and hyperactive symptoms (Combined group) were rated as showing greater executive dysfunction than those children without substantial ADHD symptoms (non-ADHD, all $p < .01$; see Table 1). On average, children in the Combined symptoms group exhibited elevations of $2 SD$ above the mean of the normative sample (i.e., $T = 50$) across all index scores. Children with high levels of hyperactivity (HI only and Combined groups) showed equally elevated BRI scores, whereas those children rated as primarily inattentive (IA only group) exhibited less severe behavioral dysregulation (BRI T scores $< 1 SD$ above the mean). In contrast, children with greater inattention symptoms (IA only group) showed more elevated CRI scores ($2 SD$ above the mean) than those without (HI only group, CRI T scores = $1 SD$ above the mean). Although the new ERI composite score was elevated in all groups with clinical symptoms, children with more symptoms of hyperactivity (HI only, Combined groups) were rated with greater dysfunction in this domain than children without substantial hyperactivity.

Clinical scale profiles—There was a significant overall effect of group across the BRIEF2 clinical scales, $F(27, 4113) = 67.11, p < .0001; \eta_p^2 = .306$ (see Figure 1). As depicted in Figure 1, post-hoc tests indicated that children in groups with high levels of hyperactivity (HI only and Combined symptoms groups) did not differ from each other on mean Inhibit or Emotional Control scale scores, but were rated significantly higher on these scales than those without hyperactivity symptoms (IA only and non-ADHD groups). Examining the Self-Monitor and Shift scales, the Combined group again showed higher scores than the IA only and non-ADHD groups, although this group was rated as significantly higher than the HI only group on Self-Monitor. The greatest elevations on the scales comprising the CRI were observed in the Combined and IA only groups, and there were no differences among mean scores between these two groups. Both groups were rated as showing significantly greater cognitive dysregulation than the HI only group, which had higher mean scores on the CRI scales than the non-ADHD group.

Among children in the HI only group, ratings of more than 2 *SD* above the mean were observed only on the Inhibit scale, with mean ratings of >1.5 *SD* on Shift and Emotional Control. Ratings on Initiate, Plan/Organize, and Organization of Materials were largely within normal limits (on average not more than 0.5 *SD* above the mean). Among children with greater inattentive symptomatology (IA only group), elevations of >2 *SD* were seen only on the Working Memory scale, with elevations of >1.5 *SD* on Initiate, Plan/Organize, Organization of Materials, and Task Monitor. Ratings on the Inhibit and Emotional Control scales fell within 1 *SD* of the mean for this group. As expected then, children with high levels of *both* inattentive and hyperactive symptoms (Combined group) showed elevated mean scores (>1.5 *SD*) across all subscales, with mean ratings of >2 *SD* on Inhibit, Shift, and Working Memory. Effect sizes for across group comparisons (η_p^2) were largest for the Inhibit (.495), Working Memory (.461), and Organization of Materials (.368) scales, followed by Plan/Organize (.355), Initiate (.346), Task Monitor (.260), Self-Monitor (.242), Emotional Control (.226), and Shift (.145).

Examining the consistency of findings after controlling for age (MANCOVA), both group ($\eta_p^2 = .298$) and age ($\eta_p^2 = .122$) showed a significant overall effect on BRIEF2 clinical scale scores. The pattern of between-group differences described above remained largely consistent after controlling for age, with the exception of slight differences in the pattern of findings for the Shift, Working Memory, and Organization of Materials scales. Whereas there was previously not a significant difference between the Combined group and the IA only group on the Working Memory and Organization of Materials scales, all groups were significantly different after controlling for age. In contrast, on the Shift scale, scores for the HI only group were significantly higher than those of the IA only group after controlling for age (with a persisting pattern of no difference between the HI only group and the Combined group).

Clinical Utility of Scale Elevations

Base rates of elevated scores are important to consider in evaluating performance for clinical purposes (Chelune, 2010) and the proportion of children in each group with clinical scale and index *T* scores ≥ 70 is presented in Table 3. Children who were reported to exhibit clinically elevated levels of ADHD symptoms were consistently more likely to exhibit elevations of ≥ 2 *SD* above the normative mean across BRIEF2 scales than those without (non-ADHD; all $p < .01$), with the exception of the Working Memory and Organization of Materials scales (HI only group not different from the non-ADHD group; see Table 3). With regard to within symptom group comparisons, referred children showing high levels of hyperactivity/impulsivity symptoms (either HI only or Combined groups) were more likely than those with primarily inattentive symptoms (IA only) to show elevations greater than 2 *SD* ($T \geq 70$) on the Inhibit scale (both $p < .001$; see Table 3). In contrast, elevations ≥ 2 *SD* on the Working Memory scale were more likely in children with greater inattention (IA only and Combined presentation vs. HI only; both $p < .001$). Consistent with these findings, the BRI composite score was more frequently elevated ($T \geq 70$) in youth with hyperactive symptoms (HI only or Combined vs. IA only; both $p < .001$), whereas the CRI was more frequently elevated in youth with greater inattentiveness (IA only or Combined vs. HI only; both $p < .001$).

The scales associated with the new ERI showed a less consistent pattern: children rated as displaying high levels of inattention *and* hyperactivity (i.e., Combined group) were more likely than both the IA only and HI only groups to show clinically significant elevations on the Shift scale (both $p = .009$; IA only and HI only groups not significantly different from each other), with the Combined and HI only groups more likely than the IA only group to show elevations on the Emotional Control scale ($p = .009$). As such, the new ERI was more likely to be elevated in children with greater hyperactivity (i.e., Combined and HI Only groups vs. the IA only group; both $p < .001$).

Between-Group Discrimination

The overall GEC composite accurately discriminated children with elevated symptoms of ADHD (any presentation) from those without (area under the curve [AUC] = .888; see Table 4). Discriminative power of specific clinical scales for identifying any symptom presentation versus the non-ADHD referred group and HI only versus IA only was examined across those scales showing the largest between group (i.e., ADHD symptoms vs. non-ADHD) effect size: Inhibit, Working Memory, and Organization of Materials.

ADHD vs. non-ADHD—With regard to discriminating children in any of the ADHD symptom groups from the non-ADHD group, the Inhibit scale showed a good AUC (.806) and excellent specificity (see Table 4); over 96% of children with Inhibit scores below 70 were not reported to show clinically significant ADHD symptoms (i.e., were in the non-ADHD group). However, the Inhibit scale correctly classified only 63.07% of the sample at $T = 70$ and only 69.8% of the sample at $T = 65$ (sensitivity = 53.88, specificity = 90.91).

In our clinical sample, the Working Memory scale showed a very good AUC of .872 and correctly classified 75.67% ($T = 70$) and 79.58% ($T = 65$) of the sample as showing elevated symptoms of ADHD or not. Examining specificity, almost 88% of those with Working Memory scores below 70 were not rated as having clinically elevated ADHD symptoms.

The Organization of Materials scale exhibited the third largest effect size in this clinical sample, with a good AUC (.834) and excellent specificity (95.79% at $T = 70$) for discriminating between children rated with elevated ADHD symptoms and those without.

HI only vs. IA only—With regard to discriminating children with predominantly hyperactive symptoms from those with predominantly inattentive symptoms, the BRIEF2 Inhibit scale showed good discrimination (AUC = .874) and accurately classified 83.9% of these youth at a T score of 70. Sensitivity was marginal with stronger specificity, and sensitivity improved at a T score of 65 (sensitivity = 80.52; specificity = 74.68; 75.64% correctly classified). Discriminating between restricted symptom presentations, the Working Memory scale accurately classified 72.25% ($T = 70$) and 81.36% ($T = 65$), with an AUC of .844. In contrast, the Organization of Materials scale exhibited a somewhat lower AUC (.784) and classification accuracy (49.79%) when discriminating between children with a Hyperactive only versus IA only presentation. However, specificity remained high (94.81%), suggesting that almost 95% of those with Organization of Materials scale scores ≥ 70 were in the IA only group.

Discussion

These findings provide very promising initial evidence of construct validity and clinical utility of the BRIEF2 in clinically referred children. Caregiver ratings on the BRIEF2 scales were specifically associated with reports of inattentive and hyperactive ADHD symptoms, such that the Inhibit scale was most likely to be elevated in children with greater hyperactivity, whereas the Working Memory scale was more likely to be elevated in children with high levels of inattention. Likewise, children with combined symptoms were likely to have elevated scores on both Inhibit and Working Memory. Notably, these scales were less likely to be elevated by high levels of symptoms in the opposite domains, suggesting specificity of prediction as well as evidence of convergent and discriminant validity of the revised BRIEF2 scales. As hypothesized, the BRIEF2 scales showed a pattern of double dissociation between the restricted presentations of ADHD symptoms—highlighted by the differences between the IA and hyperactive only groups and complemented by the combined group in each case. As few, if any, existing studies on the original measure included samples of youth with a hyperactive only presentation, the present study also provides a unique examination of the BRIEF2 profile in children with the restricted hyperactive presentation.

These data further suggest that the BRIEF2 discriminates well between specific clinical groups. The overall composite score (GEC) is sensitive to the everyday executive dysfunction associated with ADHD, and discriminates well between referred children with high levels of ADHD symptoms and those without. However, the GEC does not discriminate well between ADHD presentations, and the profile across the clinical scales appears more useful for describing the differential impact of inattentive versus hyperactive symptoms. These findings are consistent with preliminary validity evidence for the measure from the test manual (Gioia et al., 2015), but extend those initial findings into a large sample of clinically referred children including those with a pattern of restricted hyperactive-impulsive symptoms.

Although some recent work suggests that ratings on the BRIEF may reflect broad behavioral impairment rather than executive dysfunction per se (McAuley et al., 2010), the present findings highlight a degree of precision in the pattern of overlap between the BRIEF2 scales and behavioral ratings of ADHD symptomatology. Specifically, the Inhibit and Organization of Materials scales show excellent specificity for discriminating between children with and without ADHD, whereas the Organization of Materials scale shows excellent specificity for discriminating between restricted inattentive and hyperactive ADHD presentations. Taken together, these findings suggest a pattern of shared functional impairment between the BRIEF2 and parent ratings of ADHD symptoms that is consistent with findings observed using the original measure in McAuley et al. (2010). It should be noted, however, that some of the item content on the BRIEF2 Inhibit and Working Memory scales corresponds closely to the *DSM* diagnostic criteria for ADHD hyperactive-impulsive and inattentive presentations, respectively. As such, the classification accuracy may simply reflect a degree of rater consistency across similar items. However, the present findings also highlight a pattern of dissociation in day-to-day behavioral impact of ADHD symptoms as measured by specific BRIEF2 scales. Although the pattern of elevations on the Inhibit and Working Memory scales in part reflects the core weaknesses thought to characterize the hyperactive

and inattentive presentations of ADHD, respectively, this examination adds to our understanding of the relationships between ADHD symptoms and functional impact on informant reports of day-to-day executive behavior.

In the present sample of clinically referred youth, sensitivity of the BRIEF2 scales was consistently poorer than specificity, suggesting that the clinical scales may more accurately help to *rule out* ADHD symptomatology than rule it in. Furthermore, the overall classification accuracy across scales examined did not reach substantively above 75% (examining a cut score of $T = 70$) for discriminating those with any elevated ADHD symptoms from the non-ADHD comparison group. These findings are generally consistent with those presented in the test manual (Gioia et al., 2015, p. 260) using a cut score of $T = 70$ on the Inhibit scale to discriminate between children with inattentive versus combined type ADHD (classification accuracy: 68.4%-72.7%), but somewhat lower than the classification accuracy presented for the Working Memory scale in discriminating between children with ADHD and typical controls ($T = 65$; classification accuracy: 83.1%-84.8%). The lower sensitivity and classification accuracy in the present sample for discriminating between children with high levels of ADHD symptomatology and those without may reflect the referred nature of the present sample, in which a variety of symptom presentations, in addition to ADHD, may disrupt effective daily executive behavior. While the BRIEF2 may show greater sensitivity and classification accuracy in discriminating between typical controls and youth with ADHD, the purpose of clinical assessment is not usually so discrete.

Interestingly, in spite of the revisions to the BRIEF2 scale and index structure, the BRI and ERI do not appear different from one another in terms of what they appear to be measuring in this clinically referred sample. Within these groups, it is not clear that the modifications to the BRIEF2 index structure have resulted in better articulation of the constructs of behavioral self-regulation versus affective self-regulation—at least, in children presenting with symptoms of ADHD. This differentiation, and the discriminatory power of the ERI, may be more evident in other clinical groups such as children with anxiety disorders or Autism Spectrum Disorders, given movement of the Shift scale to the ERI (e.g., Gioia et al., 2002; Hovik et al., 2014). Further work in other clinical groups will be needed to clarify the clinical utility of this new Index and how well it can be differentiated from the BRI.

This study is one of the first to examine the construct validity of the new BRIEF2. Strengths of the study include the large sample size and use of a clinically referred cohort to examine ADHD symptom severity including a clinically referred, non-ADHD comparison group. Furthermore, examination of restrictive inattentive and hyperactive symptom presentations allowed for a more careful examination of the association between BRIEF2 scales and specific types of ADHD symptomatology. At the same time, limitations of the study include the clinical nature of the sample, which limited the availability of data from teachers/school settings for each participant. As such, group assignment was made based upon parent ratings of symptoms and impairment, as information about cross-setting symptom presentation from teachers was not consistently available. Additionally, we recognize that the use of restricted presentations is not entirely consistent with clinical diagnostic practice (e.g., exclusion of children with six symptoms in one domain and five in the other from analyses), which can be expected to impact the reported classification accuracy values. Furthermore, the clinical

nature of the sample precluded comparison of the measure's predictive power in discriminating referred from non-referred (e.g., typically developing) children. As such, additional examination of the BRIEF2's validity in large samples of referred and typically developing children, with cross-informant symptom reports of both ADHD symptoms and executive dysfunction, will help to support and replicate these initial findings.

Taken together, the present study provides initial evidence for reliability and validity of the revised BRIEF2 for assessing behavioral symptoms of childhood impairment in day-to-day settings. While not designed to directly assess symptoms of ADHD, the BRIEF2 appears well able to accurately characterize the pattern of self-regulatory weaknesses in children with clinically elevated ADHD symptoms, providing an important source of complimentary clinical information to ADHD symptom ratings. Given these findings, the additional benefits of a shorter length and reduced administration time appear to have made the BRIEF2 a useful improvement over the original.

Acknowledgments

The authors gratefully acknowledge the assistance of Peter Isquith, PhD, and Jennifer Greene at PAR, Inc., in transcription of existing BRIEF item data onto the BRIEF2 item structure and provision of normed BRIEF2 scores from these original data.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors acknowledge the support of the Kennedy Krieger Institute's Intellectual and Developmental Disabilities Research Center U54-HD-079123.

Biographies

Lisa A. Jacobson, PhD, is board certified in clinical neuropsychology by the American Board of Professional Psychology and a licensed psychologist in Maryland. She is also an assistant professor of Psychiatry and Behavioral Sciences at the Johns Hopkins University School of Medicine.

Alison E. Pritchard, PhD, is a licensed clinical psychologist and research scientist in the Department of Neuropsychology at the Kennedy Krieger Institute, where she manages the Executive Function Clinic, the Neuropsychology Research Lab, and the Neuropsychology Research Training Program. Dr. Pritchard is also an assistant professor of Psychiatry and Behavioral Sciences at the Johns Hopkins University School of Medicine.

Taylor A. Koriakin, MA, is currently a school psychology doctoral student at University of Connecticut's Neag School of Education. Prior to graduate school, she worked as a research assistant in the Department of Neuropsychology at the Kennedy Krieger Institute from 2011 to 2015.

Kelly E. Jones, PhD, is a graduate of Drexel University and postdoctoral fellow in pediatric neuropsychology at the Kennedy Krieger Institute and Johns Hopkins School of Medicine. She was awarded the Student Liaison Committee's Student Research Award at the 44th

Annual meeting of the International Neuropsychological Society for her abstract based upon a portion of this work.

E. Mark Mahone, PhD, is board certified in clinical neuropsychology by the American Board of Professional Psychology and a licensed psychologist in Maryland. He is the immediate past president of the American Academy of Clinical Neuropsychology. He is a professor of psychiatry at the Johns Hopkins School of Medicine and director of the Department of Neuropsychology at the Kennedy Krieger Institute.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th. Author; Washington, DC: 1994.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th. American Psychiatric Publishing; Arlington, VA: 2013.
- Barkley RA. Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*. 1997; 121:65–94. [PubMed: 9000892]
- Barkley RA, Fischer M. Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology*. 2011; 36:137–161. [PubMed: 21347918]
- Chelune G. Evidence-based research and practice in clinical neuropsychology. *The Clinical Neuropsychologist*. 2010; 24:454–467. [PubMed: 18821179]
- Doyle AE. Executive functions in attention-deficit/ hyperactivity disorder. *Journal of Clinical Psychiatry*. 2006; 67(Suppl. 8):21–26.
- DuPaul, GJ, Power, TJ, Anastopoulos, AD, Reid, R. ADHD Rating Scale-IV. Guilford Press; New York, NY: 1998.
- Fabiano GA, Pelham WE Jr, Waschbusch DA, Gnagy EM, Lahey BB, Chronis AM, Burrows-MacLean L. A practical measure of impairment: Psychometric properties of the Impairment Rating Scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child & Adolescent Psychology*. 2006; 35:369–385. [PubMed: 16836475]
- Fischer M, Barkley RA, Smallish L, Fletcher K. Executive functioning in hyperactive children as young adults: Attention, inhibition, response perseveration, and the impact of comorbidity. *Developmental Neuropsychology*. 2005; 27:107–133. [PubMed: 15737944]
- Gioia, GA, Isquith, PK, Guy, SC, Kenworthy, L. BRIEF: Behavior Rating Inventory of Executive Function. Psychological Assessment Resources; Lutz, FL: 2000.
- Gioia, GA, Isquith, PK, Guy, SC, Kenworthy, L. BRIEF2: Behavior Rating Inventory of Executive Function. Second. Psychological Assessment Resources; Lutz, FL: 2015.
- Gioia GA, Isquith PK, Kenworthy L, Barton RM. Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*. 2002; 8:121–137. [PubMed: 12638065]
- Gioia GA, Kenworthy L, Isquith PK. Executive function in the real world: BRIEF lessons from Mark Ylvisaker. *Journal of Head Trauma Rehabilitation*. 2010; 25:433–439. [PubMed: 21076244]
- Goldberg E, Podell K. Adaptive decision making, ecological validity, and the frontal lobes. *Journal of Clinical and Experimental Neuropsychology*. 2000; 22:56–68. [PubMed: 10649545]
- Hovik KT, Egeland J, Isquith PK, Gioia G, Skolgli EW, Andersen PN, Øie M. Distinct patterns of everyday executive function problems distinguish children with Tourette syndrome from children with ADHD or autism spectrum disorders. *Journal of Attention Disorders*. 2014
- Isquith PK, Roth RM, Gioia GA. Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child*. 2013; 2:125–132. [PubMed: 23442015]
- Jarratt KP, Riccio CA, Siekierski BM. Assessment of attention deficit hyperactivity disorder (ADHD) using the BASC and BRIEF. *Applied Neuropsychology*. 2005; 12:83–93. [PubMed: 16083397]

- Lahey BB, Pelham WE, Loney J, Lee SS, Willcutt E. Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry*. 2005; 62:896–902. [PubMed: 16061767]
- Mahone EM, Cirino PT, Cutting LE, Cerrone PM, Hagelthorn KM, Hiemenz JR, Denckla MB. Validity of the Behavior Rating Inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*. 2002; 17:643–662. [PubMed: 14591848]
- Mahone EM, Zabel TA, Levey E, Verda M, Kinsman S. Parent and self-report ratings of executive function in adolescents with myelomeningocele and hydrocephalus. *Child Neuropsychology*. 2002; 8:258–270. [PubMed: 12759823]
- McCandliss S, O’Laughlin L. The clinical utility of the Behavior Rating Inventory of Executive Function (BRIEF) in the diagnosis of ADHD. *Journal of Attention Disorders*. 2007; 10:381–389. [PubMed: 17449837]
- McAuley T, Chen S, Goos L, Schachar R, Crosbie J. Is the Behavior Rating Inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*. 2010; 16:495–505. [PubMed: 20188014]
- Reddy LA, Hale JB, Brodzinsky LK. Discriminant validity of the Behavior Rating Inventory of Executive Function parent form for children with attention-deficit/hyperactivity disorder. *School Psychology Quarterly*. 2011; 26:45–55.
- Seidman LJ. Neuropsychological functioning in people with ADHD across the lifespan. *Clinical Psychology Review*. 2006; 26:466–485. [PubMed: 16473440]
- Toplak ME, Bucciarelli SM, Jain U, Tannock R. Executive functions: Performance-based measures and the Behavior Rating Inventory of Executive Function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*. 2009; 15:53–72. [PubMed: 18608232]
- Volk HE, Todorov AA, Hay DA, Todd RD. Simple identification of complex ADHD subtypes using current symptom counts. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2009; 48:441–450. [PubMed: 19318883]
- Vriezen ER, Pigott SE. The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology*. 2002; 8:296–303. [PubMed: 12759826]
- Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*. 2005; 57:1336–1346. [PubMed: 15950006]

Figure 1. Profiles of mean BRIEF2 subscale and index scores across clinical groups

Note. BRIEF2 = Behavior Rating Inventory of Executive Function, Second Edition; BRI = Behavioral Regulation Index; ERI = Emotion Regulation Index; CRI = Cognitive Regulation Index; SelfMon = Self-Monitor; EmoCtrl = Emotional Control; WM = Working Memory; PlanOrg = Plan/Organize; TaskMon = Task Monitor; OrgMat = Organization of Materials.

Table 1

Sample Characteristics.

	Non-ADHD (<i>n</i> = 594)		IA only (<i>n</i> = 395)		HI only (<i>n</i> = 77)		Combined (<i>n</i> = 315)		η_p^2
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Sex (% male)	60.6 ^a		65.8 ^b		75.3 ^c		70.8 ^c		.050
Age	10.56 ^a	3.32	11.29 ^b	3.15	8.71 ^c	2.68	9.65 ^c	2.88	
Parent ratings									
ADHDRS IA	7.85 ^a	3.27	20.12 ^b	3.16	11.79 ^c	2.31	22.06 ^d	3.43	.799
ADHDRS HI	4.68 ^a	3.64	7.13 ^b	3.51	19.16 ^c	2.96	20.66 ^c	3.51	.783
BRIEF2 BRI	52.18 ^a	10.17	58.74 ^b	10.38	70.92 ^c	8.37	74.14 ^c	8.55	.452
BRIEF2 ERI	55.63 ^a	11.45	60.61 ^b	11.57	68.30 ^c	10.54	71.19 ^c	11.07	.233
BRIEF2 CRI	56.38 ^a	8.49	71.17 ^b	6.94	60.16 ^c	7.12	71.90 ^b	6.65	.487
BRIEF2 GEC	56.48 ^a	8.80	68.54 ^b	7.66	67.61 ^b	7.49	76.30 ^c	7.27	.497

Note. Superscripts indicate significant differences between groups (identical letters mean groups are not statistically different [a, a; group means do not differ]; different letters indicate significant differences between groups [a, b, c; all groups are different]). IA only = inattentive symptoms group (e.g., 6 symptoms of inattention, 3 symptoms of hyperactivity/impulsivity); HI only = hyperactive-impulsive symptoms group (e.g., 6 symptoms of hyperactivity/impulsivity, 3 symptoms of inattention); comb = Combined symptoms group (e.g., 6 symptoms of both inattention and hyperactivity/impulsivity); ADHDRS = Attention Deficit Hyperactivity Disorder Rating Scale; ADHDRS IA = Inattention scale sum; ADHDRS HI = Hyperactive/Impulsive scale sum; BRIEF2 = Behavior Rating Inventory of Executive Function, Second Edition; BRI = Behavioral Regulation Index; ERI = Emotion Regulation Index; CRI = Cognitive Regulation Index; GEC = Global Executive Composite.

Table 2

Internal Consistency (Cronbach's Alpha) by BRIEF2 Scale.

BRIEF2 scale	Alpha	Items
Emotional Control	.923	8
Inhibit	.913	8
Working Memory	.892	8
Shift	.851	8
Organization of Materials	.858	6
Plan/Organize	.842	8
Self-Monitor	.838	4
Task Monitor	.821	5
Initiate	.772	5

Note. BRIEF2 = Behavior Rating Inventory of Executive Function, Second Edition.

Table 3Base Rates of Scale Elevations $T \geq 70$.

	Non-ADHD ($n = 594$)		IA only ($n = 395$)		HI only ($n = 77$)		Combined ($n = 315$)	
	n	%	n	%	n	%	N	%
Inhibit	23	3.9 ^a	46	11.6 ^b	47	61.0 ^c	207	65.7 ^c
Shift	110	18.5 ^a	141	35.7 ^b	32	41.6 ^b	178	56.5 ^c
EmoCtrl	72	12.1 ^a	76	19.2 ^b	37	48.1 ^c	169	53.7 ^c
Initiate	66	11.1 ^a	216	54.7 ^c	16	20.8 ^b	180	57.1 ^c
WM	72	12.1 ^a	274	69.4 ^b	10	13.0 ^a	239	75.9 ^c
Plan/Org	36	6.1 ^a	174	44.1 ^b	10	13.0 ^c	156	49.5 ^b
OrgMaterials	25	4.2 ^a	162	41.0 ^b	4	5.2 ^a	139	44.1 ^b
Self-Monitor	77	13.0 ^a	90	22.8 ^b	21	27.3 ^b	169	53.7 ^c
Task Monitor	49	8.2 ^a	151	38.2 ^b	8	10.4 ^c	119	37.8 ^b
BRI	36	6.1 ^a	68	17.2 ^b	48	62.3 ^c	230	73.0 ^d
ERI	76	12.8 ^a	93	23.5 ^b	38	49.4 ^c	182	57.8 ^c
CRI	41	6.9 ^a	259	65.6 ^b	10	13.0 ^c	207	65.7 ^b

Note. Superscripts indicate significant differences between groups (identical letters mean group proportions are not statistically different [a, a; group proportions do not differ]; different letters indicate significant differences between group proportions [a, b, c; all groups are different, with ^a representing the lowest proportion]). IA only = inattentive only group; HI only = hyperactive/impulsive only group; Comb = combined group; EmoCtrl = Emotional Control; WM = Working Memory; Plan/Org = Plan/Organize; OrgMaterials = Organization of Materials; BRI = Behavioral Regulation Index; ERI = Emotional Regulation Index; CRI = Cognitive Regulation Index.

Table 4

Classification Accuracy Measures for Discriminating Between Groups With Selected Scales at $T = 70$.

$T = 70$	Any ADHD symptoms vs. non-ADHD						IA only vs. HI only					
	Sens	Spec	CA	PPV	NPV	AUC	Sens	Spec	CA	PPV	NPV	AUC
Inhibit	38.12	96.13	63.07	92.88	53.97	.806	61.04	88.35	83.90	50.54	92.08	.874
WM	66.45	87.88	75.67	87.90	66.41	.872	69.37	87.01	72.25	96.48	35.64	.844
Org	38.75	95.79	63.29	92.42	54.14	.834	41.01	94.81	49.79	97.59	23.86	.784
GEC	61.88	93.27	75.38	92.41	64.87	.888	46.84	58.44	48.73	85.25	17.65	.535

Note. IA only = inattentive only group; HI only = hyperactive/impulsive only group; Sens = sensitivity; Spec = specificity; CA = classification accuracy; PPV = positive predictive value; NPV = negative predictive value; AUC = area under the curve; WM = Working Memory; Org = Organization of Materials; GEC = Global Executive Composite.