Reliability and Validity of the BRIEF-A for Assessing Deaf College Students’ Executive Function

Peter C. Hauser¹, Jennifer Lukomski¹, and Vince Samar¹

Abstract
This study investigated the reliability and validity of the Behavior Rating Inventory of Executive Functions—Adult Form (BRIEF-A) when used with deaf college students. The BRIEF-A was administered to 176 deaf and 184 hearing students of whom 25 deaf students and 56 hearing students self-identified as having an Attention Deficit Hyperactivity Disorder (ADHD) diagnoses. Cronbach’s alpha internal consistency reliabilities for the deaf participants ranged from .55 to .96, similar to the published BRIEF-A normative sample. Differential Item Function analysis revealed that only 3 of the 75 items showed evidence of item bias for deaf students. The participants with ADHD had significantly higher scores on all nine scales. Discriminant analysis revealed comparable sensitivity and specificity of the BRIEF-A for discriminating ADHD from non-ADHD individuals for deaf and hearing groups. Finally, the deaf and hearing ADHD groups exhibited score profiles across the nine BRIEF-A scales that followed the pattern of the BRIEF-A ADHD clinical sample profile. The results suggest that the BRIEF-A is a reliable, largely unbiased diagnostic tool for deaf college students, with comparable discriminant and predictive validity for ADHD.

Keywords
BRIEF-A, executive function, ADHD, deaf, hard of hearing, psychometrics

Executive functions refer to cognitive processes that allow for future, goal-oriented behavior (Barkley, Murphy, & Fischer, 2010; Morgan & Lilienfeld, 2000) that are generally considered necessary for socially appropriate behavior and underlie productive educational and occupational performance (Biederman et al., 2004; Katz & Hartman-Maeir, 2005). The assessment of executive functions has therefore become central to psychoeducational assessment, clinical neuropsychological evaluation, and neuroforensic assessment of children, adolescents, and adults (e.g., Barkley et al., 2010; Biederman et al., 2004; Gioia, Isquith, Kenworth, & Baron, 2002; McCloskey, Perkins, & Divner, 2009).

¹Rochester Institute of Technology, Rochester, NY, USA

Corresponding Author:
Peter C. Hauser, National Technical Institute for the Deaf, Rochester Institute of Technology, 52 Lomb Memorial Drive, Rochester, NY 14623, USA
Email: peter.hauser@rit.edu.
No executive function assessment instruments have been developed specifically for the deaf and hard-of-hearing population (henceforth, *deaf*). It is common for deaf individuals to be assessed in educational and clinical settings with tests that are administered in English and that have been normed only on the hearing population (Hauser, Wills, & Isquith, 2006). Many deaf individuals represent a culturally distinct minority group, use sign language as their primary language, and vary widely in their English language skills (see Hauser, O’Hearn, McKee, Steider, & Thew, 2010 for discussion). Many have limitations of English literacy due mostly to limited access to English language input during childhood. Therefore, differences between deaf and hearing cultural norms and English literacy limitations are population factors that can undermine the reliability and validity of English language–based psychological assessment instruments (Hauser et al., 2006; Oberg & Lukomski, 2011).

This study investigated the reliability and validity of a widely used clinical screening inventory of executive function, the Behavior Rating Inventory of Executive Function—Adult Version Self-Report Form (*BRIEF-A*; Roth, Isquith, & Gioia, 2005), when used with deaf college students. The BRIEF-A is potentially a valuable screening instrument because it has been shown to be selectively sensitive to disorders of executive functions that influence cognitive skills and behaviors important for academic achievement (e.g., Rabin et al., 2006). The child version of this inventory (BRIEF) has been widely used in clinical practice and research (see Baron, 2000 for review) and is often used to evaluate Attention Deficit Hyperactivity Disorders (ADHD; e.g., McCandless & O’Laughlin, 2007; Pratt, 2000). The BRIEF-A is a new version of that inventory that is a reliable and valid measure of executive function in adults (see Roth et al., 2005 for review) and has also been shown to be sensitive as a screening tool for ADHD (e.g., Rotenberg-Shpigelman, Rapaport, Stern, & Hartman-Maerir, 2008).

We administered the BRIEF-A to samples of deaf and hearing students with and without self-reported prior diagnoses of ADHD. ADHD was selected as the criterion for testing predictive validity because it is widely recognized as a signature developmental disorder of executive functions (Barkley, 1997; Barkley et al., 2010; Roth et al., 2005). We included hearing ADHD and non-ADHA control groups in these analyses to verify that our methods replicate the results of the normative national sample on the same college campus as our deaf groups.

**Method**

**Participants**

This study included a total of 360 (176 deaf and 184 hearing) college students from the Rochester Institute of Technology in Rochester, New York. Twenty-five of the deaf students and 56 of the hearing students self-identified as having an ADHD diagnosis. The deaf participant sample included students who had various levels of audiometric hearing ability, may or may not have had cochlear implants, may have used sign language, spoken language, or a combination of both, came from different educational settings, and had either deaf or hearing parents. The deaf participants’ reading and academic skills ranged from precollege to college levels. The ratio of men to women in the deaf groups (56.8% male) was significantly different, $\chi^2(1) = 4.31$, $p < .05$, than that of the hearing sample (71.7% male). Furthermore, the deaf sample ($M_{age} = 22.0$, $SD_{age} = 3.03$) was approximately 1 year older on average than the hearing sample ($M_{age} = 21.1$, $SD_{age} = 2.7$), $F(1, 356) = 4.20$, $p < .05$. There was a higher percentage of White, non-Hispanic among hearing participants (86%) than among deaf participants (72%) although this difference was not significant in chi-square group comparisons across the six race/ethnicity categories.
Measures

Behavior Rating Inventory of Executive Function-Adult Version Self-Report Form (BRIEF-A). The BRIEF-A is composed of 75 items that comprise a total of nine clinical scales: Inhibit, Shift, Emotional Control, Self-Monitor, Initiate, Working Memory, Play/Organize, Task Monitor, and Organization of Materials. The first four scales are summed to provide a Behavioral Regulation Index (BRI), while the last five scales are summed to provide a Metacognitive Index (MI). In addition all nine scales are summed to provide an overall summary score called the Global Executive Composite (GEC). Three additional validity scales are designed to assess the overall validity of an individual BRIEF-A administration (see Roth et al., 2005, for the psychometric properties of the test).

Attention Deficit Hyperactivity Disorder Self-Report (ADHD Self-Report). Participants were classified as having self-reported ADHD if they circled ADHD in response to the question “Have you been diagnosed with (please circle): Learning Disability, Depression, ADD/ADHD, Anxiety, Bipolar, None, or Other.” They were classified as non-ADHD otherwise.

Conners Adult ADHD Rating Scales-Self-Report: Long Version (CAARS-S:L-S:L). The CAARS-S:L (Conners, Erhardt, & Sparrow, 1998) is a reliable and valid measure of ADHD symptoms for use with hearing adults. It provides three DSM-IV symptoms measures and an ADHD index (see Conners et al., 1998, for psychometric information).

Procedure

All of the participants were recruited via flyers that were posted around campus and handed out at student events. All participants who volunteered were included in the study and there was no exclusion criteria as the authors wanted to maximize the generalizability to deaf college students. Participants were tested individually and in groups, and were paid US$10. They were first administered a demographic questionnaire, then the BRIEF-A. Participants who reported that they had ADHD and a randomized subgroup of non-ADHD participants were administered the CAARS-S:L in a counterbalanced order with the BRIEF-A. The participants’ BRIEF-A and CAARS-S:L raw scores were converted to T-scores using normative data provided in the test manuals.

Results

Predictive Validity of ADHD Self-Report

The CAARS-S:L was used as a convergent assessment instrument to validate the ADHD Self-Report. The CAARS-S:L provides two overall index scores, the ADHD Index that best distinguishes individuals with ADHD from nonclinical individuals and the DSM-IV ADHD Symptoms Total, which assesses how strongly participants meet the overall DSM-IV criteria for ADHD diagnosis. Eighty-one of the participants took the CAARS-S:L (deaf: 17 ADHD and 22 non-ADHD; hearing: 23 ADHD and 19 non-ADHD). Logistic regressions of these participants’ self-reported ADHD diagnoses onto these two CAARS-S:L indexes were used to derive Resource Operating Characteristic (ROC) curves for deaf and hearing groups separately. The area under the ROC curve provides a measure of the ability of a Connors index to discriminate self-reported ADHD participants from non-ADHD participants.

The following values were obtained: Deaf Group ADHD Index = .833; Hearing Group ADHD Index = .805; Deaf Group DSM-IV Symptoms Total = .916; Hearing Group DSM-IV Symptoms Total = .928. Values of .8 to .89, as is the case for the ADHD index, reflect good diagnostic
classification and values of .9 to 1.0, as is the case for the DSM-IV Symptom Total index, reflect excellent diagnostic classification (Rosenberg, Joseph, & Barkun, 2000). These results confirm that self-reported ADHD diagnoses in deaf and hearing participants meaningfully reflect the presence of ADHD related behavioral patterns and correlate well with the clinical criteria established by the DSM-IV for ADHD diagnosis. Therefore, the ADHD Self-Report measure can be taken as a valid criterion measure for use with the full participant sample in this study.

Reliability

The BRIEF-A includes validity scales designed to assess the overall validity of an individual BRIEF-A administration. These are Negativity (whether the client is biased toward a negative response pattern), Infrequency (whether the client has produced an atypical response pattern), and Inconsistency (whether the client has answered related questions in an inconsistent manner). Deaf and hearing participants displayed elevations on the three validity scales rarely—8 out of 176 deaf students and 3 out of 184 hearing—and there were no significant differences between deaf and hearing groups in these proportions (Fischer’s Exact Test, \( p = .1324 \)).

Table 1 displays Cronbach’s alpha for each of the nine scales and the three indexes for each of the four participant groups (i.e., deaf non-ADHD, hearing non-ADHD, deaf ADHD, hearing ADHD) and for the normative sample from the BRIEF-A manual.

The reliability coefficients for the scales and index scores across the participant groups fell into the moderate to high range (i.e., .55 to .96). Hakstian-Wallen tests (Hakstian & Whalen, 1976) revealed no evidence of significant mean alpha coefficient differences among the four experimental groups and the normative sample for five of the nine scales. However, the Shift, Working Memory, Plan/Organize, and Organization of Materials scales did show evidence of
significant differences among the five groups. Significant Hakstian-Wallen tests on these latter four scales justify post hoc tests for pairwise alpha coefficients to determine if any of the experimental groups’ mean alpha coefficients differed reliably from the normative sample alpha coefficients. Post hoc tests were conducted using the Fisher-Bonett statistic (Bonett, 2003) with a Šidák correction (Šidák, 1967) for multiple post hoc comparisons. Only two comparisons remained significant. For the Working Memory scale, the hearing ADHD group showed a significantly lower reliability coefficient than the normative sample. For the Plan/Organize scale, the non-ADHD deaf group showed a significantly lower reliability coefficient than the normative sample.

A Hakstian-Wallen test was also conducted on the BRI, MI, and, GEC. The BRI displayed significant differences in reliability among the five groups. Fisher-Bonett post hoc tests revealed that this effect was due to the reliability coefficient for the hearing ADHD group being significantly lower than the normative sample on the BRI. The Hakstian-Wallen test on the GEC was also significant. Fisher-Bonett post hoc tests revealed that this difference was again due to the hearing ADHD group having a significantly lower GEC than the normative sample GEC.

Overall, the reliability coefficients of the BRIEF scales and indexes for the non-ADHD deaf and hearing groups are statistically equivalent to the reliability coefficients for the manual’s normative sample. This is also the case for the reliability coefficients of the deaf ADHD group. For the hearing ADHD group, there was a tendency for their reliability coefficients to be numerically, but typically not significantly, lower than the other groups on individual scales. This tendency was reliably captured by the lower BRI and GEC coefficients for this group. Nevertheless, the latter coefficient was still robust at values of .84 and .93 respectively. Collectively, these data suggest that the BRIEF-A has comparable and acceptable reliability for deaf and hearing college students.

Validity

Differential item functioning. An item bias is present when two groups display different response distributions on an item despite no difference between the groups in the underlying executive function trait that the item was designed to measure. Differential item functioning (DIF) analysis allows us to detect evidence of such biases in each item of a given BRIEF-A scale. If biases occur for deaf individuals, the predictive validity (sensitivity and specificity) of the BRIEF-A for that population will be compromised. DIF works by examining the relative group distributions of responses to a given item when the groups are equated on their performance on the underlying trait that the item measures. Similar to hearing individuals, we assume that deaf individuals with poor executive function on a BRIEF-A scale trait will respond that they often exhibit a specific negative behavior (as stated in a scale item), whereas others with better executive function on the scale trait will respond that they sometimes or never exhibit that behavior. Cultural or linguistic factors may shift the probability with which a deaf individual will endorse a particular frequency of behavior on an item compared with a hearing peer who has exactly the same underlying level of the specific executive function trait measured by that item.

DIF was computed on the non-ADHD group participants and Zumbo’s (1999) ordinal logistic regression (OLR) method was used to perform DIF analyses. Due to the significant differences in the age and gender characteristics of the deaf and hearing samples, age and gender were included in the OLR model as the first two covariates to control for any potential effect of these variables on DIF-related factors. We adopted Jodoin and Gierl’s (2001) standards here and required a minimum $R^2$ value of .035 to declare DIF on the BRIEF-A items to ensure greater power to detect relatively small but still potentially meaningful biases for deaf participants.
Only three items met the criteria for DIF: Item 15 (Plan/Organization Scale: “I have trouble prioritizing activities”); Item 2 (Task Monitor Scale: “I make careless errors when completing tasks”); and, Item 31 (Organization of Materials scale: “I lose things, such as keys, money, wallet, homework, etc.”). Item 15, which displayed moderate uniform DIF, accounting for 4.2% of the OLR variance after equating the groups on their levels of the Plan/Organization scale, was almost entirely accounted for by the variance associated with the uniform DIF Group effect. The additional variance associated with the non-uniform DIF interaction term was negligible. Items 2 and 31 displayed strong uniform DIF, accounting for 31.3% and 20.7% of the OLR variance, respectively, with negligible additional variance contributed by the non-uniform DIF interaction term. For Items 15 and Item 2 more deaf students endorsed “1,” meaning that the specific behavior has never been a problem for them in the past. For Item 31 more hearing students endorsed “3,” meaning that the specific behavior has often been a problem for them, whereas more deaf students endorsed “2” rather than “3,” meaning that the specific behavior has sometimes been a problem for them rather than often been a problem for them. We do not have any hypotheses as to why these three items appear to be biased with the deaf college student population. In general, these results indicate that item biases for the deaf students were rare (< 4% of the items showed evidence of item bias) and, because the three biased items were distributed over different scales, respectively, they had minimal effects on their total scale scores.

Predictive validity of BRIEF-A for ADHD Self-Report. The predictive validity of the BRIEF-A was examined by using discriminant analysis to determine the relative sensitivity and specificity of the instrument for discriminating between participants who reported either a history of diagnosis of ADHD or no history of diagnosis of ADHD, for both deaf and hearing groups. If the BRIEF-A is valid for use with deaf people, discriminant analysis should successfully classify deaf individuals as having ADHD who report previous ADHD diagnoses (sensitivity), and should successfully classify individuals as not having ADHD who report no previous ADHD diagnoses (specificity). To test this hypothesis a 2 Group (deaf, hearing) × 2 Diagnosis (ADHD, non-ADHD) × 9 Scale repeated measures ANOVA was conducted using T-scores on the nine BRIEF-A scales as dependent variables. Age and Gender were included as covariates. The main effect for Diagnosis was significant, $F(1, 354) = 90.35, p < .0001$. Simple effects ANOVAs on the individual scales revealed that participants with ADHD diagnoses had higher mean scale scores over all nine scales, $F(1, 354) = 3.72$ to 9.97, $p < .0002$ to $p < .0001$. This effect held for both deaf and hearing groups separately, deaf: $F(1, 172) = 4.12$ to 29.38, $p < .04$ to .0002 and hearing: $F(1, 180) = 2.13$ to 9.38, $p < .035$ to .0001. The difference is shown in Figure 1 by the separation of the solid (ADHD) and dashed (non-ADHD) line graphs.

There was also a Scale main effect, $F(8, 347) = 6.05, p < .0001$, a Scale × Diagnosis interaction, $F(8, 347) = 6.84, p < .0001$, a Scale × Group interaction, $F(8, 347) = 9.52, p < .0001$, and a three-way interaction of Scale × Diagnosis × Group, $F(8, 347) = 2.51, p < .015$. The data patterns that are responsible for these significant effects can be appreciated by inspection of Figure 1. The figure shows that the deaf and hearing ADHD groups showed more dramatic profiles than their non-ADHD counterparts of mean scale scores across the nine BRIEF-A scales (approximately a twofold larger range of scores). These impressions were confirmed by the significantly larger variances across scale means for the ADHD groups compared with the non-ADHD groups, deaf: variance ratio = 5.6, $F(8, 8) = 5.59, p < .013$; hearing: variance ratio = 3.8, $F(8, 8) = 3.79, p < .04$.

It is apparent from Figure 1 that there were both similarities and differences in the profiles of scores across the nine scales for the deaf ADHD group compared with the hearing ADHD group. On the BRI scales, the deaf and hearing groups both tended to follow a similar and overlapping U-shaped pattern. The two groups differed significantly only on the Self-Monitor scale, with the deaf group showing higher mean scale scores, $F(1, 77) = 7.11, p < .01$. On the MI scales, both groups followed a nearly parallel M-shaped pattern, with the hearing groups showing
systematically higher mean scale scores across the five component scales: Initiate, $F(1, 77) = 10.44, p < .0018$; Working Memory, $F(1, 77) = 8.25, p < .006$; Planning, $F(1, 77) = 3.44, p < .068$; Task Monitoring, $F(1, 77) = 11.43, p < .0011$; and Organization, $F(1, 77) = 5.30, p < .025$. The non-ADHD groups differed significantly on two BRI scales, the Inhibit, $F(1, 275) = 7.49, p < .007$, and Self-Monitor scales, $F(1, 275) = 12.54, p < .0005$. They also differed significantly on two MI scales, the Working Memory, $F(1, 275) = 4.21, p < .04$, and Task Monitoring scales, $F(1, 275) = 7.84, p < .006$.

Figure 2 displays scatterplots of the relationship between the mean $T$-scores for the nine BRIEF-A scales for deaf and hearing groups versus the mean $T$-scores for the normative and clinical samples presented in the manual. The mean values used for the ADHD manual sample were the average of the medicated and unmedicated subgroups. For both the hearing and deaf ADHD groups, their $T$-score scale means correlated significantly with those of the manual’s sample. These results indicate that the profiles of scale scores for both the hearing and deaf ADHD groups tend to follow the same profile that has been found in the BRIEF-A reference clinical population although the correlation was weaker for the deaf participant group. By contrast, the range of mean scale scores for the deaf and hearing non-ADHD groups was much smaller and was uncorrelated with the scale scores of the BRIEF-A clinical reference group.

Sensitivity and specificity. Sensitivity and Specificity were determined by conducting stepwise discriminant analyses on the BRIEF-A scales and index scores for deaf and hearing groups
separately. For each group, the self-reported diagnosis (ADHD vs. non-ADHD) was used as the classification variable in three discriminant analyses, each successive analysis using scores at a higher level of trait scoring. For the first analysis, the nine scale’s T-scores were entered into the discriminant analysis in a stepwise manner. For deaf and hearing participants separately, the nine scale covariates competed for stepwise entry into the discriminant equation until their p-to-enter exceeded 0.25, an average criterion based on Monte Carlo simulation studies of the optimal stopping criterion for multivariate 2-valued classification procedures (Hosmer & Lemeshow, 2000; Lee & Koval, 1997). Sensitivity and specificity were calculated for each hearing status group based on the success of classification using the subset of scale variables selected by each group’s discriminant analysis. For the second analysis, the BRI and MI scores were entered jointly to predict the proportions of individuals correctly classified as members of the ADHD and as non-ADHD groups, respectively. Sensitivity and specificity were determined for each hearing status group based on the success of those classifications. For the third analysis, the GEC scores were used to predict the proportions of individuals correctly classified as ADHD and as non-ADHD, respectively. Sensitivity and specificity were determined for each hearing status group based on the success of those classifications. Typically, sensitivities and specificities of approximately .7 or higher are considered satisfactory for psychometric tests.

Table 2 shows the results of three alternative analyses to estimate the sensitivity and specificity of the BRIEF-A for the participant groups. The first analysis used the T-scores from the nine BRIEF-A scales as covariates to determine participants’ predicted ADHD diagnostic classification. For both deaf and hearing participants, the Inhibit, Working Memory, and Initiate scales entered as significant covariates. In addition, for the deaf participants, the Shift and Self Monitor scales contributed further to the discriminant equation. Sensitivity and specificity had similar magnitudes for the deaf and hearing participants and compared reasonably well with the BRIEF-A published values for ADHD diagnosis (see selected scale covariates solution section in Table 2). All groups had sensitivity and specificity above .7 except for the deaf ADHD group who had a sensitivity of .680.
The second analysis used the BRI and MI indexes as covariates. For both deaf and hearing participants, BRI and MI both entered as significant covariates. The joint BRI & MI solution section in Table 2 shows that the sensitivity and specificity values for this analysis were comparable for deaf and hearing participants. Sensitivity values were somewhat lower than those produced by the discriminant equations based on the individual scale covariates (Deaf ADHD = .680 and Hearing ADHD = .698).

The third analysis used the overall GEC index as the sole classifier variable. The global score solution section in Table 2 shows that the sensitivity value for hearing participants increased relative to the value produced by the joint BRI and MI analysis with the Deaf ADHD group, again, being the only one with a value below .7 (sensitivity = .680). For deaf participants, however, the sensitivity value was identical to that of the joint BRI and MI analysis. The specificity values decreased somewhat for both groups in the GEC analysis compared with the joint BRI and MI analysis. In general, the various approaches above to estimating the sensitivity and specificity of the BRIEF-A suggest that these metrics of predictive validity for the deaf participants typically produce values comparable to those of the hearing participants.

### Table 2. Estimated Sensitivity and Specificity From Scale and Index-Based Discriminant Analyses

<table>
<thead>
<tr>
<th>Predicted diagnosisa</th>
<th>Selected scale covariates solution</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADHD</td>
<td>ADHD</td>
</tr>
<tr>
<td>Deaf</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>non-ADHD</td>
<td>31</td>
<td>120</td>
</tr>
<tr>
<td>Hearing</td>
<td></td>
<td>47</td>
</tr>
<tr>
<td>ADHD</td>
<td>27</td>
<td>101</td>
</tr>
<tr>
<td>non-ADHD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Joint BRI &amp; MI solution</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADHD</td>
<td>ADHD</td>
</tr>
<tr>
<td>Deaf</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>non-ADHD</td>
<td>37</td>
<td>114</td>
</tr>
<tr>
<td>Hearing</td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>ADHD</td>
<td>24</td>
<td>104</td>
</tr>
<tr>
<td>non-ADHD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Global score solution</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADHD</td>
<td>ADHD</td>
</tr>
<tr>
<td>Deaf</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>non-ADHD</td>
<td>41</td>
<td>110</td>
</tr>
<tr>
<td>Hearing</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>ADHD</td>
<td>27</td>
<td>101</td>
</tr>
<tr>
<td>non-ADHD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aScales appear in the order of their entry into discrimination equation: (a) Deaf: Inhibit, Working Memory, Shift, Self-Monitor, Initiate; and, (b) Hearing: Working Memory, Inhibit, Initiate.
Discussion

For deaf college students, the BRIEF-A reliability was high, the differential item functioning was similar between deaf and hearing college students, and the BRIEF-A had good predictive validity for classifying individuals as having ADHD or not. Deaf college students’ BRIEF-A’s scales, indexes, and composite scores were found to have moderate to high internal consistency reliability with the exception of the Plan/Organize scale. Overall, the reliability coefficients of the BRIEF-A scales and indexes for the non-ADHD deaf and hearing groups were statistically equivalent to the reliability coefficients for the manual’s normative sample. This was also the case for the reliability coefficients of the deaf ADHD group (although not the hearing ADHD group). Collectively, these data suggest that the BRIEF-A has comparable and acceptable reliability for deaf and hearing college students.

With regard to differential item functioning, only three of the 75 BRIEF-A items demonstrated different item functioning and the three items were distributed across different scales suggesting that the impact on the overall score is minimal. Furthermore, the BRIEF-A results demonstrated discriminant validity when comparing ADHD and non-ADHD groups. Similar to hearing individuals with ADHD, deaf individuals with ADHD reported higher scores on all of the nine scales compared to the respondents without ADHD. The only scale that was more elevated for deaf respondents with ADHD than hearing respondents with ADHD was the Self-Monitor scale. In comparison, the deaf students with ADHD had lower scale scores compared to hearing people with ADHD on the Initiate, Working Memory, Planning, Task Monitor, and Organization scales.

The profiles of scale scores for both the hearing and deaf ADHD groups tended to follow the same profile that has been found in the BRIEF-A reference clinical population although the correlation was weaker for the deaf group. By contrast, the range of scale scores for the deaf and hearing non-ADHD groups was much smaller and was uncorrelated with the scale scores of the BRIEF-A reference clinical group. This result is expected since variation in nonclinical groups should only represent random sampling error across scales, not signature patterns of relative deficits that comprise a stable profile. The similar profile for the deaf ADHD group with the clinical population is evidence for substantial construct validity of the BRIEF-A because it replicates the global pattern of relative deficits on specific executive function constructs found in the hearing normative sample. It remains for future research to determine whether the weaker correspondence to the normative sample profile displayed by the deaf participants compared with the hearing participants reflects stable deaf-related deviations from the hearing profile pattern or simply reflects greater sampling error due to the smaller number of ADHD deaf participants in this study.

The sensitivity and specificity had similar magnitudes for deaf and hearing participants and compared reasonably well with the BRIEF-A published values for ADHD diagnosis. The ADHD vs. non-ADHD sensitivity and specificity were comparable across deaf and hearing groups for the BRI, MI, and GEC. Overall, the present findings provide robust evidence that the BRIEF-A is a useful screening inventory for deaf college students.

Conclusion

This study provides substantial support for the construct validity of the BRIEF-A for use with deaf college students. The deaf college student community, although highly heterogeneous, represents a subgroup of the young adult deaf population with cognitive abilities, academic achievement, and opportunity to obtain a college degree. Even though this sample is not representative of the whole deaf population, it is estimated that 45% of the deaf population attend college (Schroedel, Watson, & Ashmore, 2005), suggesting that our findings can be generalized to a significant proportion of the deaf population.
There are some limitations of this study. Although the number of non-ADHD participants is sizeable, factor analytic studies with a larger sample should be undertaken to more reliably examine the construct validity of the BRIEF-A for the deaf young adult population. Quantitative estimates of the reliability, sensitivity, and specificity parameters in this study are mathematically limited by the somewhat small sample size for the deaf ADHD participants. Additionally, we relied on self-report of ADHD diagnosis to form the groups used in this study since authoritative documentation of ADHD diagnoses (e.g., medical reports) were not available to us. Nevertheless, this limitation was mitigated by our inclusion of a widely used ADHD self-assessment scale, the CAARS-S: L-S: L, to provide convergent evidence of the validity of the self-reported ADHD diagnosis question. Considering the limitations of the study, the results provide crucial validation for the use of the BRIEF-A to support evidence-based evaluation practices with deaf college going individuals.

Acknowledgments
We thank Josh Allmann, Corey Clark, Wyatte Hall, Tara Hillman, Jason Listman, Rebecca Mowell, Tiffany Panko, and Alex Pouliot for their assistance with this project.

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.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by grants to P.H. and J.L. from National Science Foundation Science of Learning Center on Visual Language and Visual Learning (SBE-0541953) and to Daphne Bavelier and P.H. From the National Institute on Deafness and Other Communication Disorders (DC04418).

References


