


Adult ADHD Self-Report Scale Screener Ratings in College Students: Concurrent Validity and Test–Retest Reliability

Journal of Psychoeducational Assessment
2021, Vol. 39(8) 1015–1019
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/07342829211034367
journals.sagepub.com/home/jpa


Benjamin J. Lovett¹ , David E. Ferrier²,
Tina Wang¹, and Alexander H. Jordan^{3,4}

Abstract

The Adult ADHD Self-Report Scale screener is commonly used to determine if further evaluation of ADHD is required. However, despite use in college settings, the ASRS was not developed for this population, and limited relevant psychometric data exist. In this study, 190 college students completed the ASRS screener as well as a lengthier measure of the 18 DSM symptoms of ADHD. A subgroup ($n = 141$) completed the ASRS twice, with at least 1 week in between measures. Concurrent validity of ASRS scores was substantial ($r = 0.73$ with DSM inattentive symptoms and $r = 0.57$ with DSM hyperactive/impulsive symptoms) and test–retest reliability was adequate ($r = 0.69$ with an average interval of 42 days). However, the sensitivity of the screener was only moderate (66%), and approximately one-fifth of the students changed screening status (positive vs. negative) across the time interval. Implications for future research and practice are discussed.

Keywords

attention-deficit hyperactivity disorder, assessment, psychometrics, screening

Attention-deficit hyperactivity disorder (ADHD) is a relatively common psychiatric disorder, defined by high levels of inattention (distractibility, forgetfulness, etc.) and/or hyperactive and impulsive behavior. These symptoms must start by age 12, be present across settings, and cause significant interference with real-world functioning, while not being due to a different disorder (American Psychiatric Association, 2013). Originally thought to affect only children, ADHD has been recognized in adults for decades now (Doyle, 2004).

College students are often assessed for ADHD, but making the diagnosis in this population is complicated by several factors (cf. Nelson & Lovett, 2019). Apparent symptoms of ADHD can be due to the stresses of adjustment to college life, as well as problematic lifestyle behaviors related to sleep and substance use. Moreover, some students appear to seek a diagnosis to gain access to

¹Columbia University, New York, NY, USA

²University of Tennessee at Chattanooga, Chattanooga, TN, USA

³Harvard Medical School, Boston, MA, USA

⁴McLean Hospital, Belmont, MA, USA

Corresponding Author:

Benjamin J. Lovett, School Psychology Program, Teachers College, Columbia University, Box 120, 525 W. 120th Street, New York, NY 10027, USA.

Email: BL2799@tc.columbia.edu

stimulant medication or academic accommodations (Musso & Gouvier, 2014). Finally, college students may not reliably recall childhood symptoms or have easy and immediate access to historical records documenting early symptoms needed for the diagnosis (Suhr et al., 2009). For all of these reasons, a comprehensive evaluation is needed for accurate diagnosis.

Although self-reports of symptoms should not constitute the whole of an ADHD evaluation, they are one important component. Many college counseling centers and health offices use brief ADHD screeners to determine which students merit a full evaluation, and the Adult ADHD Self-Report Scale (ASRS) contains a very popular, six-item screening device (Kessler et al., 2005). The ASRS was developed by the World Health Organization to measure the presence of ADHD symptoms in adults and is now used widely. However, it was not designed for college students in particular, and most of the psychometric work on the scale has used general adult samples in primary care settings (for a review, see Anbarasan et al., 2020). Gray et al. (2014) investigated the psychometric properties of the ASRS in college students, but only in those who already had ADHD diagnoses, preventing a direct examination of the ASRS's clinical utility. Fuller-Kilgore et al. (2013) reported the screener total score's internal consistency in 111 college students ($r_{xx} = .67$) and found strong relationships between the score and students' responses on other self-report symptom measures. In the present brief report, we describe an evaluation of selected psychometric characteristics of the ASRS screener in a college student sample, looking in particular at its concurrent validity and test-retest reliability, and contrasting the features of its total score and dichotomous (positive/negative) interpretations. Concurrent validity was judged against a comprehensive self-report measure of ADHD symptoms. The concurrent validity data are also used to make a preliminary examination of the screener's clinical utility.

Method

Participants

The participants were 190 undergraduate students attending a medium-size public university in the Southeastern United States. Most of the participants were white (83.2%) and female (86.3%) and were distributed across the four grade levels fairly equally (between 20% and 30% of the sample at each year in college). No exclusion criteria were applied, and participants were not asked if they had been diagnosed with ADHD.

Measures

Attention-deficit hyperactivity disorder screener. The ASRS's six-item screener portion was used to screen participants for ADHD symptoms (see Kessler et al., 2005). Each item describes a symptom that is rated on a scale of 0–4, depending on how frequently it has been experienced in the previous 6 months. In addition, if at least four of the items are rated as being present to a clinical degree (defined as either “sometimes” or “often,” depending on the symptom), the screen is considered positive, which would suggest the need for a more thorough evaluation.

Comprehensive self-report measure. The Conners Adult ADHD Rating Scale (CAARS) long self-report form (Conners et al., 1999) was used as our criterion measure for concurrent validity. This scale consists of 66 items rated on a 0–3 scale based on how frequently various symptoms are experienced. The scale yields several norm-referenced index scores, two of which were used in this study: the Diagnostic and Statistical Manual of Mental Disorders (DSM) inattentive symptoms index and the DSM hyperactive/impulsive symptoms index. The DSM-based scores were used because the ASRS was developed to assess the DSM symptoms (Kessler et al., 2005).

The DSM index scores are each based on responses to items describing relevant DSM symptoms that are embedded within the CAARS (9 items for each score). The scores are scaled using a *T*-score distribution ($M = 50$, $SD = 10$), with scores above 65 representing clinical significance. These scores have acceptable reliability, and their validity evidence includes results from factor analyses on the normative sample (Conners et al., 1999).

Procedure

Participants were recruited from undergraduate psychology courses and given a small amount of course credit for their participation. They took the ASRS once as part of a pre-screening battery, as a requirement to participate in any study being offered by the psychology department. Later, if a student signed up for this particular study, they then took the ASRS a second time along with the CAARS. All testing was conducted in an online environment during the Fall 2019 semester (prior to the COVID-19 pandemic). “Catch” items (e.g., “select response 2 for this item to show that you are reading carefully”) were inserted into the measures to ensure effort and engagement while completing the measures, and participants who failed any catch item had already been excluded from data analyses.

Results

Concurrent Validity

Concurrent validity analyses were performed on the ASRS and CAARS that were completed at the same time, and for these analyses, the entire sample of 190 students was used. The ASRS total score correlated significantly with the CAARS DSM inattention ($r = .74$, $p < .001$) and DSM hyperactivity/impulsivity ($r = .57$) scores, and the former correlation was significantly greater than the latter, $z = 4.23$, $p < .001$.

Because the ASRS is expected to screen for ADHD regardless of subtype, a crosstabs analysis was performed looking at the relationship between a positive screen on the ASRS and having either CAARS DSM score above 65. The relationship was significant ($\phi = .46$, $p < .001$), and of 41 participants meeting the CAARS criterion, 27 screened positive on the ASRS (sensitivity of 66%), whereas of the 149 participants failing to meet the CAARS criterion, 125 screened negative on the ASRS (specificity of 84%).

The concurrent validity coefficients for the ASRS’s total score and dichotomous (positive/negative) judgments were compared, by correlating each with the binary variable of having either CAARS DSM score above 65. The ASRS total score showed a significantly higher correlation ($r = .55$) than the ASRS dichotomous judgment ($r = .46$), $z = 1.97$, $p < .05$.

Test–Retest Reliability

Reliability analyses were performed on a subsample of 141 students for whom at least 1 week elapsed between their two ASRS measurements. The mean test–retest interval was 42 days (range: 7–100 days). The ASRS total scores at the two points correlated significantly, $r = .69$, $p < .001$. Because the ASRS screener results are typically interpreted in a dichotomous fashion, a crosstabs analysis examined the relationship between screening positive at Time 1 and at Time 2. The relationship was significant ($\phi = .51$, $p < .001$). Of the 141 participants, 27 (19%) received a different screening status at the two time points, whereas 23 (16%) screened positive, and 91 (65%) screened negative, at both time points. The gap of time between the two ASRS measurements was unrelated to the change (absolute value of difference) in scores ($r = -.04$, $p = .65$).

Discussion

The present study examined selected psychometric properties of the ASRS screener in college students. First, we found significant relationships between the ASRS total score and both of the CAARS's DSM symptom scores. Using standard criteria for gauging effect sizes, both relationships were large in magnitude (Bosco et al., 2015). However, the ASRS showed a stronger relationship with CAARS inattentive symptoms than with hyperactive/impulsive items. This may be because inattention is the area of ADHD symptomatology most likely to continue into adulthood, or because of the mostly female composition of sample, and the fact that women are more likely to exhibit the predominantly inattentive type of ADHD (see Owens et al., 2015, for discussion of both of these issues). In any case, the strong relationships between the ASRS and CAARS (which were replicated with categorical cutoffs of clinical significance) constitute evidence of convergent validity.

The preliminary clinical utility analyses yielded more surprising results. Using the CAARS DSM scores as a criterion, the ASRS cutoff showed 84% specificity but only 66% sensitivity. Typically, screeners should have very high sensitivity, overidentifying cases that can later be eliminated through more thorough evaluation, whereas specificity is less important for screeners. Of course, a high score on the CAARS is not at all the same as a valid clinical diagnosis of ADHD (Harrison et al., 2019), but the low sensitivity may still be worrisome. Interestingly, even in its development sample, the ASRS showed much higher specificity than sensitivity (Kessler et al., 2005).

Finally, with regard to test–retest reliability, we found a .69 correlation between two ASRS scores obtained a month apart; Haynes et al. (2011) argue that this would be “adequate” evidence of test–retest reliability. However, a more appropriate index of reliability may be the phi coefficient between the screening statuses (positive vs. negative) at the two time points since the ASRS screener is typically interpreted in a simple dichotomous way. This coefficient was lower, .51, and may represent less reliability in the vicinity of the cut score. We found that approximately one in five of our participants changed screening statuses from one measurement to the other, a statistic to consider in applied use.

Limitations and Directions for Future Research

The present study had expected limitations, each of which suggests directions for future investigations. We did not conduct full clinical evaluations on the participants, and so we were unable to use an ultimate “gold standard” criterion to evaluate the ASRS against. Instead, as a first step, we wished to see if the ASRS would even relate strongly to a more comprehensive self-report symptom measure. Future studies involving full clinical evaluations would allow for more meaningful clinical utility indices, particularly positive and negative predictive values (Mayer, 2004). Full clinical evaluations would also allow for careful assessment of functional impairment, an often-overlooked construct whose relationship with symptom severity is far from perfect (Lewandowski et al., 2016). In addition, our sample was largely white and female, and the ASRS may have different properties in men and in college students with ethnic minority backgrounds. Future studies with more diverse samples are encouraged.

Conclusions

Despite the limitations noted above, our results nonetheless offer preliminary guidance to those using the ASRS in applied settings. They provide a reminder that screeners often err in both directions (i.e., false positives and negatives) and suggest that the ASRS may be particularly vulnerable to false negatives, given its relatively low (for a screener) sensitivity. These findings,

along with the test–retest reliability coefficient based on categorical cutoffs, lead us to recommend that clinicians consider interpreting the ASRS in a continuous manner. Higher ASRS scores do indicate a higher risk of stable, clinically significant levels of self-reported symptoms (even on a more comprehensive measure), but in particular, an ASRS score below the cutoff does not screen out ADHD in college students effectively.

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) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Benjamin J. Lovett  <https://orcid.org/0000-0003-4785-3070>

References

- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.
- Anbarasan, D., Kitchin, M., & Adler, L. A. (2020). Screening for adult ADHD. *Current Psychiatry Reports*, 22(12), 72-75.
- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100(2), 431-449.
- Conners, C. K., Erhardt, D., & Sparrow, E. (1999). *Conners' adult ADHD rating scales*. Multi-Health Systems.
- Doyle, R. (2004). The history of adult attention-deficit/hyperactivity disorder. *Psychiatric Clinics of North America*, 27(2), 203-214.
- Fuller-Killgore, M. D., Burlison, J., & Dwyer, W. (2013). Comparison of three ADHD screening instruments in college students of varying cognitive ability. *Journal of Attention Disorders*, 17(5), 449-454.
- Gray, S., Woltering, S., Mawjee, K., & Tannock, R. (2014). The adult ADHD self-report scale: Utility in college students with attention-deficit/hyperactivity disorder. *PeerJ*, 2, e324.
- Harrison, A. G., Nay, S., & Armstrong, I. T. (2019). Diagnostic accuracy of the conners' adult ADHD rating scale in a postsecondary population. *Journal of Attention Disorders*, 23(14), 1829-1837.
- Haynes, S. N., Smith, G. T., & Hunsley, J. D. (2011). *Scientific foundations of clinical assessment*. Routledge.
- Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, E., Howes, M. J., Jin, R., Secnik, K., Spencer, T., Ustun, T. B., & Walters, E. E. (2005). The world health organization adult ADHD self-report scale: A short screening scale for use in the general population. *Psychological Medicine*, 35(2), 245-256.
- Lewandowski, L. J., Lovett, B. J., & Gordon, M. (2016). Measurement of symptom severity and impairment. In S. Goldstein & J. A. Naglieri (Eds.), *Assessing impairment: From theory to practice* (2nd ed., pp. 229-245). Springer.
- Mayer, D. (2004). *Essential evidence-based medicine*. Cambridge University Press.
- Musso, M. W., & Gouvier, W. D. (2014). *Journal of Attention Disorders*, 18(3), 186-201.
- Nelson, J. M., & Lovett, B. J. (2019). Assessing ADHD in college students: Integrating multiple evidence sources with symptom and performance validity data. *Psychological Assessment*, 31(6), 793-804.
- Owens, E. B., Cardoos, S. L., & Hinshaw, S. P. (2015). Developmental progression and gender differences among individuals with ADHD. In R. A. Barkley (Ed.), *Attention-deficit hyperactivity disorder* (4th ed.). Guilford.
- Suhr, J., Zimak, E., Buelow, M., & Fox, L. (2009). Self-reported childhood attention-deficit/hyperactivity disorder symptoms are not specific to the disorder. *Comprehensive Psychiatry*, 50(3), 269-275.