J Autism Dev Disord (2017) 47:90–100
DOI 10.1007/s10803-016-2933-z

ORIGINAL PAPER

Evaluation of the ADHD Rating Scale in Youth with Autism

Benjamin E. Yerys[1,2] · **Jenelle Nissley-Tsiopinis**[3] · **Ashley de Marchena**[1,4,8] · **Marley W. Watkins**[5] · **Ligia Antezana**[1,7] · **Thomas J. Power**[2,3,6] · **Robert T. Schultz**[1,2,6]

Published online: 13 October 2016

Abstract Scientists and clinicians regularly use clinical screening tools for attention deficit/hyperactivity disorder (ADHD) to assess comorbidity without empirical evidence that these measures are valid in youth with autism spectrum disorder (ASD). We examined the prevalence of youth meeting ADHD criteria on the ADHD rating scale fourth edition (ADHD-RS-IV), the relationship of ADHD-RS-IV ratings with participant characteristics and behaviors, and its underlying factor structure in 386, 7–17 year olds with ASD without intellectual disability. Expected parent prevalence rates, relationships with age and externalizing behaviors were observed, but confirmatory factor analyses revealed unsatisfactory fits for one-, two-, three-factor models. Exploratory analyses revealed several items cross-loading on multiple factors. Implications of screening ADHD in youth with ASD using current diagnostic criteria are discussed.

Keywords Autism · Attention deficit/hyperactivity disorder · Comorbidity · Screening · Factor analysis

Electronic supplementary material The online version of this article (doi:10.1007/s10803-016-2933-z) contains supplementary material, which is available to authorized users.

✉ Benjamin E. Yerys
Yerysb@email.chop.edu

1 Center for Autism Research, The Children's Hospital of Philadelphia, 3535 Market Street, Ste 860, Philadelphia, PA 19104, USA

2 Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

3 Center for the Management of ADHD, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

4 Intellectual and Developmental Disabilities Research Center, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

5 Department of Educational Psychology, Baylor University, Waco, TX, USA

6 Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

7 Present address: Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

8 Present address: Department of Behavioral and Social Sciences, University of the Sciences in Philadelphia, Philadelphia, PA, USA

Introduction

The 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association 2000), unlike the 4th edition, permits the concurrent diagnosis of autism spectrum disorder (ASD) and attention deficit/hyperactivity disorder (ADHD). In the context of an evaluation clinicians attempt to distinguish whether ASD and ADHD symptom endorsement reflects the true presence of both disorders, whether parents endorse a particular behavior as supporting symptoms for both disorders, or whether the presence of one disorder leads to expression of symptoms in the other (behavioral phenocopy). However, there is no empirical evidence demonstrating that existing ADHD rating scales measure the same constructs in the ASD population as in the general population, and thus that they are valid for assessing ADHD in those with ASD.

Although the wording of individual symptoms for inattention and hyperactivity/impulsivity do not overlap with those for ASD, it is important to note that the vast majority of ADHD symptoms occur in a social context. Thus, symptoms might be endorsed on a questionnaire due to

impairments in social skills and not to underlying difficulties with inattention and impulsivity/hyperactivity. For example, ADHD symptoms of "Is easily distracted," "Has difficulty sustaining attention to activities or play," or "Does not seem to listen when spoken to directly" may be observed by a caregiver or a teacher because the individual is inattentive or hyperactive/impulsive, or because the youth with ASD lacks social skills to engage and disengage from an interaction appropriately. Likewise, respondents contextually interpret questionnaires (Schwarz 2010). Due to the complexity of the relationship between ADHD and ASD, it is imperative for clinicians and scientists to possess clinical tools for screening/diagnosing ADHD that exhibit the expected two-factor structure in youth with ASD. Clinical tools with this level of validity will improve clinicians' confidence in the validity of the scores.

One of the most commonly used screening measures for ADHD is the ADHD Rating Scale-Fourth Edition (ADHD-RS-IV). It has been normed with a community sample in both home and school settings, and the two-factor model (Inattention and Hyperactivity/Impulsivity) has been validated in both community (DuPaul et al. 2016, 1998) and ADHD samples (Willcutt et al. 2014). Although the ADHD Rating Scale has been updated to reflect DSM-5 criteria, the core items of the disorder have not changed; modifications in symptom items in the updated version of the scale (ADHD-RS-5) are relatively minor and pertain primarily to the wording of items in the adolescent version of the measure (DuPaul et al. 2016). The ADHD-RS-IV was selected in the present study because it includes all 18 ADHD symptoms for making a diagnosis whereas many other screeners do not (e.g., Child Behavior Checklist). Scores on the hyperactivity/impulsivity subscale of the ADHD-RS-IV decrease with age in the normative sample (DuPaul et al. 1998). The ADHD-RS-IV's construct validity has been established in school-age youth with ADHD by demonstrating large positive correlations between Inattention and Hyperactivity/Impulsivity scores with the Behavioral Regulation ($r = .57$–$.70$) and Metacognition Indices ($r = .59$–$.85$), of the Behavior Rating Inventory of Executive Functions–Parent Form (BRIEF; Mahone et al. 2002). Furthermore, consistent with the high comorbidity of ADHD and disruptive behaviors, elevations in both the Inattentive and Hyperactive/Impulsive Subscales of the ADHD-RS-IV have been significantly associated with elevations in the Externalizing Problems subscale relative to the Internalizing Problems subscale of broad screeners (Humphreys et al. 2012; Reiersen and Todorov 2013). Thus, validation of the ADHD-RS-IV in ASD would require a demonstration of these same relationships across age and measures.

The ADHD-RS-IV is used to inform diagnoses of ADHD in ASD for determining treatment options, school-based services, and research on shared and unique features of the two disorders. Thus, there is a critical need for establishing its validity in ASD populations. It is important to know if elevated scores on the ADHD-RS-IV stem from different sources in those with ADHD and those with ASD; in other words, does this scale measure the same underlying constructs in both populations or are the ADHD symptoms exhibited by those with ASD different in important ways. This study characterizes caregiver and teacher ADHD-RS-IV ratings in a relatively large ASD sample ($N = 386$). We tested whether our sample was consistent with others that found ~30–50 % of youth with ASD exhibited clinically elevated ADHD symptoms (Leyfer et al. 2006; Sinzig et al. 2009). In addition, we tested factorial validity with confirmatory and exploratory factor analyses to determine whether the ADHD-RS-IV captures the same underlying constructs of Inattention, Hyperactivity, and Impulsivity in youth with ASD across caregiver and teacher ratings, as it does in community-based and ADHD samples. Consistent with the ADHD literature, we also sought to confirm a relationship between greater ADHD symptoms and greater executive function impairments (Doyle et al. 2005; Willcutt et al. 2005) and greater externalizing behavior problems (Reiersen and Todorov 2013). We measured executive function and externalizing behaviors with the BRIEF (Gioia et al. 2000) and the Behavior Assessment System for Children-2nd Edition (Reynolds and Kamphaus 2004), respectively. Demonstrating this set of relationships would suggest that the behaviors captured by the ADHD-RS-IV might stem in part from impaired attention and hyperactivity/impulsivity in youth with ASD, just as it does in community-based and ADHD samples. Our null hypothesis was that the two-factor model described in DSM-5 would be the best fit in the ASD sample, just as it is in community and ADHD populations (DuPaul et al. 2016, 1998; Willcutt et al. 2014). We also tested whether common findings in the ADHD literature, such as decreasing Hyperactivity/Impulsivity behaviors with age, were present in ASD.

Method

Participants

ADHD-RS-IV caregiver ratings were available for 386 youth with ASD, and teacher ratings were available for a subset of 203 (see Table 1). The 203 with both caregiver and teacher reports did not differ in age, sex-ratio, or IQ from the 183 with caregiver reports only (see Supplementary Information—Table S1). There were small but significant differences in current ASD presentation, such that those with both caregiver and teacher versions of the ADHD-RS-IV scale had slightly higher levels of symptoms (Cohen's $d \leq 0.25$). To ensure that these small differences did not significantly influence results, all analyses of the caregiver version were conducted twice:

Table 1 Participant characteristics

	Total sample	Subset with caregiver and teacher report
Sample size	386	203
Chronological age		
Years [*M* (SD)]	10.54 (3.02)	10.40 (3.11)
Min–Max	6.00–18.83	6.08–18.83
Race[a]		
Asian	6 (1.55%)	3 (1.48%)
Black	14 (3.63%)	5 (2.46%)
Caucasian	341 (88.34%)	182 (89.66%)
Native American	1 (0.26%)	0 (0.00%)
Biracial	19 (4.92%)	9 (4.43%)
Missing/not reported	5 (1.30%)	4 (1.97%)
Ethnicity[a]		
Hispanic	26 (6.74%)	15 (7.39%)
Not Hispanic	333 (86.27%)	174 (85.71%)
Not reported	27 (6.99%)	14 (6.90%)
GCA		
M (SD)	100.38 (18.10)	99.81 (17.55)
Min–Max	61–162	67–158
Verbal reasoning		
M (SD)	100.80 (18.91)	100.29 (18.03)
Min–Max	60–161	62–161
Nonverbal reasoning		
M (SD)	100.83 (17.66)	100.26 (16.84)
Min–Max	61–166	64–166
Spatial reasoning		
M (SD)	99.24 (16.05)	99.06 (16.18)
Min–Max	62–150	63–141
Sex-ratio (M:F)	342:44	185:17
ADI-R		
Reciprocal social interaction	18.63 (6.14)	18.70 (6.51)
Communication	14.93 (4.70)	14.98 (4.86)
Repetitive behaviors	5.79 (2.46)	5.93 (2.49)
ADOS		
Total [*M* (SD)]	12.24 (4.95)	12.81 (4.92)
Social affect [*M* (SD)]	9.08 (4.10)	9.44 (4.16)
Repetitive behavior [*M* (SD)]	3.17 (1.93)	3.38 (1.95)
Calibrated severity score [*M* (SD)]	6.84 (2.31)	7.07 (2.19)
Age-sex[a]		
6- to 7-year-old males	85 (22.02%)	51 (25.12%)
6- to 7-year-old females	2 (0.52%)	0 (0.00%)
8- to 10-year-old males	119 (30.83%)	71 (34.98%)
8- to 10-year-old females	26 (6.74%)	9 (4.43%)
11- to 13-year-old males	84 (21.76%)	38 (18.72%)

Table 1 (continued)

	Total sample	Subset with caregiver and teacher report
11- to 13-year-old females	11 (2.85%)	5 (2.46%)
14- to 17-year-old males	49 (12.69%)	26 (12.81%)
14- to 17-year-old females	5 (1.30%)	3 (1.48%)

ADI-R autism diagnostic interview-revised; Cut-off scores: reciprocal social interaction = 10; communication = 8; repetitive behaviors = 3

ADOS-2 autism diagnostic observation schedule-2nd edition; ASD cut-off score total = 7

[a]Raw count (percent sample)

once with the whole group (i.e., those only with caregiver report) and once with the subset that had both caregiver and teacher report. Youth were recruited for studies at the Center for Autism Research at the Children's Hospital of Philadelphia from the Philadelphia metropolitan region of the United States by means of flyers, online postings for parent groups and ASD groups, and setting up booths at local events for individuals with ASD and their families.

Youth in the ASD group met DSM-IV-TR criteria for either autism, asperger's syndrome, or pervasive developmental disorder—not otherwise specified (American Psychiatric Association 2000); the Autism Diagnostic Observation Schedule (Lord et al. 2000) with the revised diagnostic algorithm that parallels the 2nd edition's algorithm (Gotham et al. 2007) and the Autism Diagnostic Interview—Revised (Lord et al. 1994) were used by clinicians to inform their judgment when completing a DSM-IV-TR checklist. Parent report was used to determine the presence of comorbid genetic or neurological disorders, extreme premature birth (i.e., <32 weeks), or other medical conditions that may affect neural development; youth whose caregivers reported such conditions were excluded from the sample. Eighty-six youth with ASD (22% of the total sample) were prescribed one psychoactive medication, and 94 (24%) were prescribed multiple medications. Regarding ADHD medications, 94 youth (24%) were prescribed a stimulant, 46 (12%) were prescribed an alpha-2A agonist, and 17 (4%) were prescribed a selective norepinephrine reuptake inhibitor. Many youth treated for ADHD symptoms were prescribed more than one ADHD medication. Because the ADHD-RS-IV was developed and validated in a community sample of youth from general education classrooms, very few youth in the normative sample had comorbid Intellectual Disability, as is often seen in ASD. It is thus unclear whether the ADHD-RS-IV would validly diagnose ADHD in youth with significantly impaired IQ. To avoid confounding ASD with Intellectual Disability in our analyses of the ADHD-RS-IV, we only included participants with IQ > 60[1] on the Verbal, Nonverbal, and Spatial Reasoning domains of the Differential Ability Scales—Second Edition (Elliott 2007). Informed consent and assent were obtained from participants in accordance with the Children's Hospital of Philadelphia IRB guidelines.

[1] A total of 9 children had IQ scores below 70, the traditional cut-off for establishing an Intellectual Disability diagnosis. Excluding these children had no significant effect on our results. For example, on all confirmatory factor analyses changes in fit indices was 0.002 or less. For this reason, we elected to include these children in our final sample.

Measures

ADHD-Rating Scale IV (ADHD-RS-IV; DuPaul et al. 1998) home and school versions assess severity of inattention and hyperactivity/impulsivity symptoms according to caregiver or teacher report, respectively. This 18-question scale yields two domains: inattention and hyperactivity/impulsivity. For each question, caregivers or teachers use a 0–3 scale to rate the participant. A higher score indicates more symptom severity. A score of 2 or 3 is considered a significant symptom; six or more significant symptoms in the inattention and/or hyperactivity/impulsivity domains is used to determine whether an individual meets criteria for an ADHD diagnosis according to that reporter. Dependent variables included the group means of each item (0–3), inattention and hyperactivity/impulsiveness total scores (0–27), group means of total symptoms endorsed in each domain (0–9), and percentages of youth meeting DSM-IV-TR criteria for different presentations of ADHD. Data were collected from both caregivers and teachers.

Behavior Rating Inventory of Executive Functions—Parent Form (BRIEF) (Gioia et al. 2000) is an informant report of executive function in everyday situations comprised of eight scales that are collapsed into two broad indices: Behavioral Regulation and Metacognition. Results are reported as T-scores. Higher scores indicate greater impairment; T-scores ≥65 (i.e., 1.5 SDs ≥ the mean) indicate clinically significant ratings. Dependent variables include the two broad indices. Data was collected from caregivers only.

Behavior Assessment System for Children–Second Edition (BASC-2) (Reynolds and Kamphaus 2004) is an informant report of both adaptive and problematic behaviors in everyday settings. Dependent variables include T-scores (Mean = 50, SD = 10) from the two broad problem domains: Externalizing and Internalizing Problems. Data was collected from caregivers only.

Procedures

These measures were completed as part of a larger test battery from multiple studies at the Center for Autism Research (Antezana et al. 2016; Chevallier et al. 2014, 2015; Granader et al. 2014; Herrington et al. 2016; Parish-Morris et al. 2013; Pugliese et al. 2015). The hospital's institutional review board approved the research protocol. Prior to participation, consent was obtained from all legal guardians and assent was obtained from all youth ≥7 years.

Analysis Plan

We examined the percentage of youth above clinical cut-offs within each domain (i.e., six or more symptoms) and each rating form. We used both the "And" and "Or" rule for counting symptoms, which represents the most and least stringent (but most commonly used) criteria for diagnosis (Lahey et al. 1994). Using the "And" rule, youth meet criteria for a symptom only if it is endorsed (i.e., score of 2 or 3) by both the caregiver *and* the teacher. With the "Or" rule youth meet criteria for a symptom if it is endorsed (i.e., score of 2 or 3) by either the caregiver or the teacher. We examined the distribution of data, means, and standard deviations for the Inattention and Hyperactivity/Impulsivity scales, as well as individual items.

To establish factorial validity, confirmatory factor analyses were conducted using one-factor [all 18 items (see Table S2) load on a single "ADHD" factor]; two-factor [nine items (odd numbers) that correspond to inattention symptoms load on one factor labeled "Inattention," and the six items related to hyperactivity (items 2, 4, 6, 8, 10, 12) and three items related to impulsivity (items 14, 16, 18) load on a second factor labeled "Hyperactivity/Impulsivity"]; and three-factor models (nine items load on an Inattention factor, six items load on a "Hyperactivity" factor and three items load on a distinct "Impulsivity" factor) for caregiver and teacher report. Given the nonparametric nature of the data, we used robust weighted least squares estimation with polychoric correlations (Flora and Curran 2004) as implemented in Mplus version 7.3 (Muthén and Muthén 2015). We used the Chi square likelihood ratio test of exact fit, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA) to evaluate the approximate fit for all models (Brown 2015). CFI represents the degree of improvement over the worst-fitting model and is scaled from 0 to 1 with higher values indicating better fit. The TLI is similar to the CFI and is also scaled from 0 to 1, but it is more sensitive to model complexity than the CFI. RMSEA represents the degree of model misfit and is also scaled from 0 to 1, but lower values indicate better fit. For further explanation of these indices and how they are applied to evaluate goodness of fit when using CFA, see Brown (2015). Hu and Bentler (1999) suggested a combination rule whereby CFI/TLI values ≥.95 *and* RMSEA values ≤.06 would best control both Type-I and Type-II errors. If none of the models meet the Hu and Bentler (1999) combination rule for an acceptable fit, we will conduct an exploratory factor analysis (EFA). In such situations, Gorsuch (1997) recommended that a follow-up EFA be conducted because parameters that were constrained in a CFA (i.e., zero loadings of items onto non-salient factors) are not restricted in an EFA, allowing identification of alternative models or reasons for poor fit (Wegener and Fabrigar 2000). Further, Bentler and Wu (2002) noted that EFA is superior to CFA modification indices in these measurement models. One potential reason for poor fit is when items have salient associations with more than one factor (i.e., "cross-load"). For example, item 12 ("Talks Excessively") may have a moderate association with both factors of Hyperactivity and Impulsivity. This cross-loaded item is easy to identify in an EFA. Items were considered to saliently load on a single factor when pattern coefficients ≥.30 (Child 2006).

Finally, we explored two-tailed correlations across caregiver and teacher forms, as well as commonly observed relationships of ADHD symptoms with age and (caregiver) ratings of executive function and behavior problems to establish construct validity across forms and with established measures. Because no differences were observed in results between the full sample with the caregiver form and the subset that also had teacher forms, we reported correlation analyses on the full sample.

Results

Table 2 provides means and standard deviations of the raw scores within each domain (0–27) and a Total combined score (0–54) for caregiver and teacher ratings. This table also provides estimates of the total number of symptoms endorsed in each domain (0–9), and the percentages of youth receiving scores above the clinical cut-off (six or more symptoms). Caregiver ratings showed that 39 % of youth in the sample received scores above the clinical cut-offs for Inattention symptoms and 27 % for Hyperactivity/Impulsivity symptoms, whereas teacher ratings showed only 24 % of youth receiving scores above the clinical cut-offs for Inattention symptoms and 13 % for Hyperactivity/Impulsivity symptoms. Table 2 also provides further breakdowns by DSM-IV-TR ADHD subtypes by each informant, as well as use of the "And" and "Or" rules that combines endorsed symptoms from either rater. The item endorsed most often as clinically elevated by caregivers and teachers was "Is easily distracted" (65 % caregiver and 53 % teacher). Youth with ASD demonstrated a normal distribution for Inattention and Hyperactivity/Impulsivity scales, as well as the Total score for caregiver and teacher ratings; detailed breakdown of item-level analyses from caregiver and teacher settings can

Table 2 ADHD Rating Scale summary scores and symptoms endorsed

Summary scale	Mean	SD	*N* (%) above cut-off
ADHD caregiver version			
Inattention score	14.13	6.38	–
H/I score	11.65	6.50	–
Total score	25.78	11.75	–
Inattention symptoms endorsed	4.51	2.95	151 (39%)
H/I symptoms endorsed	3.62	2.75	103 (27%)
ADHD caregiver version			
Predominantly inattentive	–	–	69 (18%)
Predominantly H/I	–	–	23 (6%)
Combined	–	–	82 (21%)
Did not meet	–	–	212 (55%)
ADHD teacher version			
Inattention score	10.98	6.32	–
H/I score	8.00	6.13	–
Total score	18.92	11.14	–
Inattention symptoms endorsed	3.15	2.87	49 (24%)
H/I symptoms endorsed	2.18	2.43	27 (13%)
ADHD teacher version			
Predominantly inattentive	–	–	30 (15%)
Predominantly H/I	–	–	8 (4%)
Combined	–	–	19 (9%)
Did not meet	–	–	146 (72%)
ADHD "And" rule			
Predominantly inattentive	–	–	19 (9%)
Predominantly H/I	–	–	3 (1%)
Combined	–	–	10 (5%)
Did not meet	–	–	171 (85%)
ADHD "Or" rule			
Predominantly inattentive	–	–	47 (23%)
Predominantly H/I	–	–	15 (7%)
Combined	–	–	64 (32%)
Did not meet	–	–	77 (38%)

Caregiver ratings n = 386

Teacher ratings n = 203

"And" and "Or" rule ratings n = 203

H/I hyperactivity/impulsivity

be found in Tables S2 and S3 and Figs. S1–S18 shows the distribution plot by item for each Informant.

None of the CFA models for either caregiver or teacher ratings from the ADHD-RS-IV were exact fits to the data (i.e., Chi square values were all highly significant). Likewise, no model met the Hu and Bentler (1999) combination rule that required CFI ≥ .95 *and* RMSEA ≤ .06, or more liberal fit criteria whereby RMSEA is >.06 but the upper boundary of the 90% confidence interval is <.08 (see Table 3 for details). The three-factor model for teachers was closest to meeting this criterion, but its RMSEA value was excessive (.083; upper limit of confidence interval =.095). For comparison, McCallum et al. (1996) labeled RMSEA values of .08 to .10 as mediocre and >.10 as poor. Additionally, the correlation between the second and third factors for teachers was extremely high (.911) and exceeded their alpha reliability (.82 and .88, respectively). Similarly, the correlation between the second and third factors for caregivers (.82) indicated poor discriminant validity (Brown 2015). In sum, CFA failed to identify an acceptable structure for these data.

Accordingly, we followed-up the unsatisfactory CFA with EFA using best-practice criteria (Goldberg and Velicer 2006; Gorsuch 1997) with the *psych* package in the **R** statistics system (Revelle 2015). Beginning with polychoric correlation matrices, principal axis factor extraction with an oblique oblimin rotation was applied in all models. Given that it is better to overextract than underextract (Gorsuch 1997), a three-factor model was identified, although both parallel analysis and minimum average partial correlation criteria indicated that two factors might be sufficient (Velicer et al. 2000).

The three-factor solution for caregiver ratings identified several problematic items on the ADHD-IV-RS. Of the nine items designed to measure Inattention, items 3, 5, and 15 had salient associations (cross-loaded) with the factor intended to measure Hyperactivity. Likewise, items that were designed to measure Hyperactivity (items 8 and 12) and Impulsivity (item 16) had salient associations with the other factor. In all, items 3, 5, 8, 12, 15, and 16 did not perform well in a three-factor structure. The two-factor solution was less problematic, as the results were more similar to what we would expect based on prior research; however, three inattentive items continued to cross-load within this structure: notably, inattention items 3, 5, and 15 had salient associations with both the Inattention and Hyperactivity/Impulsivity factors (see Tables S4 and S5).

The three-factor solution for teacher ratings also identified problematic items. Its Inattention factor was marked by weak loadings (associations) of items with Inattention and cross-loadings of inattention items with the factor intended to measure Impulsivity. A third factor had two of the inattention items and three of the hyperactivity items loaded on it but none of the intended impulsivity items (See Table S6). This disorganization was clarified by a two-factor solution (see Table S7) whereby all nine items of the presumed Inattention factor cohered, and eight of the presumed Hyperactivity/Impulsivity factor items fused into the second factor. Only item 2 (Hyperactivity/Impulsivity) was not performing

Table 3 Confirmatory factor analysis goodness-of-fit statistics

	df	χ^2	CFI	TLI	RMSEA (90% CI)
Caregiver ($n=386$)					
One-factor	135	1058.6	.872	.855	.133 (.126, .141)
Two-correlated factors	134	625.0	.932	.923	.097 (.090, .105)
Three-correlated factors	132	570.1	.939	.930	.093 (.085, .101)
Teacher ($n=203$)					
One-factor	135	710.5	.874	.857	.145 (.134, .155)
Two-correlated factors	134	372.8	.948	.940	.093 (.082, .105)
Three-correlated factors	132	316.8	.959	.953	.083 (.071, .095)

To determine the acceptability of the model fit we used Hu and Bentler's (1999) combination rule whereby CFI/TLI values ≥.95 *and* RMSEA values ≤.06 would best control both Type-I and Type-II errors

CFI comparative fit index, *CI* confidence interval, *RMSEA* root mean square of approximation

well in this structure, as it was cross-loaded on both the Inattention and Hyperactivity/Impulsivity factors.

Caregiver and teacher ratings were moderately correlated within and across ADHD symptom domains ($r=.25$–$.43$; Table S8). Caregiver ratings of Inattention and H/I symptoms had moderate to large positive correlations with the Behavior Regulation and Metacognition Indices, and Global Executive Composite from the BRIEF ($r=.44$–$.75$; Table S9). Caregiver ratings of Inattention and Hyperactivity/Impulsivity symptoms also had moderate to large positive correlations with Externalizing Problems on the BASC-2 ($r=.49$–$.67$), but small correlations with Internalizing Problems ($r<.18$; Table S10). Caregiver and teacher ratings of Hyperactivity/Impulsivity symptoms both decreased with age (r's $=-.31$, $-.28$, respectively), and teacher ratings of Inattention symptoms decreased with age ($r=-.21$; Table S8).

Discussion

The present study takes a new and critical step for improving our measurement of ADHD among youth with ASD. Researchers have routinely used the ADHD-RS-IV (or similar DSM-based screeners like the Vanderbilt ADHD Diagnostic Parent/Teacher Rating Scale) to assess ADHD symptoms in youth with ASD, as well as the association of ADHD symptoms with maladaptive behaviors, and attention and executive function task performance (Andersen et al. 2013; Corbett et al. 2009; Johnston et al. 2013; Sikora et al. 2012; Yerys et al. 2013); however, this work has not considered whether ADHD screeners are valid in ASD samples. While the overall factor structure of the scale generally corresponds with an expected two-factor solution, the present study shows that the ADHD-IV-RS fails to meet goodness-of-fit criteria for factorial validity in youth with ASD. Our results suggest that the scale does not adequately separate the constructs of inattention and hyperactivity/impulsivity in ASD. An important next step would be to modify the cross-loaded items on the ADHD-IV-RS and re-assess the scale's factorial validity. Ideally, minor wording changes will help informants minimize the influence of ASD on ratings of target behaviors. Until this modification occurs, the ADHD-RS-IV remains a useful clinical tool for assessing ADHD in the context of ASD, with additional clinical interviewing focused on separating inattention and hyperactivity/impulsivity symptoms from ASD symptoms. In what follows we will step through the interpretation of our factor analyses in greater detail before returning to issues of convergent validity, estimates of ADHD prevalence, and research and clinical recommendations based on our findings.

Confirmatory factor analyses demonstrated an unacceptable fit for the latent variables of Inattention and Hyperactivity/Impulsivity in youth with ASD. They also highlight the need for future research to improve measurement precision in assessing these constructs in youth with ASD. This failure to fit the data to any model, particularly the two-factor model, contrasts with studies of community populations and those with ADHD (Baumgaertel et al. 1995; DuPaul et al. 2016, 1998; Pillow et al. 1998; Willcutt et al. 2014). Follow-up investigation with EFA found that three Inattention items from caregiver ratings (numbers 3, 5, and 15) were cross-loaded, as was one Hyperactivity/Impulsivity item from teacher ratings (number 2). The complex items for caregivers ask about the following behaviors: (3) Difficulty sustaining attention in tasks or play; (5) Doesn't listen when spoken to directly; and (15) Is easily distracted. The complex item for teachers asks if the youth "Fidgets or squirms in seat." Some item cross-loadings may be the result of item wording that leads caregivers to endorse a symptom because of poor social skills or odd behaviors rather than inattention or hyperactivity (i.e., construct irrelevant variance). Or they may be the result of an ASD diagnosis leading to the expression of ADHD symptoms in social settings (a behavioral phenocopy of ADHD in ASD). Therefore, an important next step is to revise the wording of these problematic items, and

evaluate if a modified ADHD-IV-RS captures Inattention and Hyperactivity/Impulsivity constructs in the ASD population (i.e., factorial validity). This suggestion may raise concern that we are calling for a change in the diagnostic criteria for ADHD in youth with a diagnosis of ASD. This is not the goal. ASD and ADHD symptoms certainly *can* be disentangled with careful clinical interviewing. Instead, the goal is to refine a well-established measure of ADHD symptoms so that clinicians and researchers can have confidence that elevated scores represent symptoms of Inattention and Hyperactivity/Impulsivity and not simply a behavioral phenocopy. That is, that the endorsement of ADHD symptoms is not better explained by core ASD symptoms.

The ADHD-RS-IV showed construct validity for caregiver report, because we found expected relationships with executive function, externalizing behavior problems, and age in youth with ASD. Caregiver report demonstrated positive correlations between Inattention and Hyperactivity/Impulsivity symptoms and both executive function and externalizing behaviors, but not internalizing behaviors. This pattern suggests that caregivers are not biased to rate *all* problem behaviors as elevated in these youth; instead, caregivers are endorsing problem behaviors most associated with ADHD symptoms. Hyperactivity/Impulsivity symptoms decreased with age in caregiver and teacher ratings, which is also seen in the normative sample (DuPaul et al. 2016, 1998); teacher ratings of Inattention also decreased with age (DuPaul et al. 2016). We note that because this is a cross-sectional, and not a longitudinal study, age effects should be interpreted with caution. Nevertheless, these findings establish construct validity as they replicate known relationships observed in youth with ADHD without ASD (Doyle et al. 2005; Reiersen and Todorov 2013).

Despite the limitations in factorial validity, using the ADHD-RS-IV we observed prevalence rates of ADHD in an ASD sample similar to those established via in-depth semi-structured clinical interviews (Leyfer et al. 2006; Sinzig et al. 2009). Prevalence rates of ADHD within ASD are comparable overall; however, our findings using the ADHD-IV-RS, along with an unmodified psychiatric interview (Sinzig et al. 2009), yield the highest reported rates. Reported ADHD prevalence rates are lowest when clinicians use a psychiatric interview (i.e., KSADS) specifically adapted to disentangle primarily social symptoms from primarily inattentive or hyperactive/impulsive symptoms (Leyfer et al. 2006). It is important to note that the clinicians in this study ignored the DSM-IV criteria to not diagnose ASD and ADHD concurrently (American Psychiatric Association 2000). The current study is also the first to report teacher ratings independently from caregiver ratings; teachers identified 24 % of youth with ASD meeting ADHD diagnostic criteria. This lower rate for teachers is also seen in the normative community sample (DuPaul et al. 2016), and may result in part from their experience of child behaviors in a more restricted (and demanding) range of settings. The "Or" rule, which may be a more accurate estimate than single-informant or conservative "And" rule approaches (Schwarz 2010), revealed a prevalence rate of 62 %. This higher rate is expected as symptoms endorsed in either setting are counted.

This study informs future research that will use such screeners to parse groups into ASD versus ASD + ADHD or using ADHD-RS-IV scores as continuous measures of ADHD symptom severity. Although these scores may yield expected relationships based on what is seen in prior research, the origins of these behaviors—most notably on the Inattention scale—are potentially capturing variability related to ASD symptoms in addition to inattention. Furthermore, in future clinical trials treating ADHD symptoms in ASD we may need to exercise caution with these types of screeners to measure treatment-related change. In its current form, the origin of some symptoms on the ADHD-IV-RS may differ in ASD, and it is not yet known how this difference may or may not relate to treatment outcome measurement. In our own work, we will continue to use screeners like the ADHD-IV-RS, but we will use statistical approaches to account for the role of ASD symptoms when examining relationships between ADHD symptoms and task performance or other dependent variables of interest (See Yerys et al. 2013 for an example).

This study informs clinical practice because it suggests that providers should exercise caution when interpreting ratings of ADHD symptoms in youth with ASD. Because several of the ADHD-RS-IV items do not map to expected ADHD dimensions, it is important for providers to conduct more rigorous questioning around ADHD symptoms that have a high social demand. For example, providers will need to determine if the symptom (e.g., "Does not listen when spoken to directly") is related to a difficulty with understanding social expectations (i.e., does not know teacher was talking to them), or whether the child understands the expectations but violates them due to poor sustained attention abilities (i.e., child becomes bored/distracted). We propose a clinical approach that combines existing ADHD tools with additional clinical interviewing aimed at understanding the persistence of Inattention and Hyperactivity/Impulsivity across social and non-social settings for making a comorbid diagnosis. This approach should be the gold standard until ADHD screeners are refined to improve factor structure fit within an ASD sample or until specific items that are most sensitive *and* specific to Inattention and Hyperactivity/Impulsivity symptoms are identified in future research, thereby reducing potential false positives for individuals with ASD.

It is important to note limitations of this study. First, the present study did not use semi-structured interviews to provide a "best practices" diagnosis of ADHD, and therefore

our ability to establish diagnostic convergence is limited to comparing the present study with other samples. While the cross study comparison is encouraging, a better approach would be to utilize a rating scale and diagnostic interview in the same sample. Second, this study did not assess the level of impairment associated with ADHD symptoms, which may inflate the estimate of ADHD diagnoses (DuPaul et al. 2014). Third, we used a convenience sample of ASD rather than a community sample. This approach has the potential to add bias with respect to which families chose to participate in research with a high demand (i.e., travel time and testing time) versus those families that do not participate. In our future work it will be important to obtain ratings from a community sample to validate any future ADHD rating scale versions in youth with ASD. Fourth, the present study excluded youth with significant cognitive impairments (IQ < 60). It remains unclear whether the ADHD-RS-IV would show similar results with a more severely impaired population, or whether other measures would be recommended. Fifth, with the release of DSM-5 there are some additional examples associated with ADHD symptoms to improve the assessment of ADHD in older adolescents and adults, as well as an adaptation of one hyperactivity/impulsivity item to include restlessness, which are now reflected in the ADHD-RS-5 (DuPaul et al. 2016). While these changes are minor, future research examining the validity of using informant reports of ADHD among youth with ASD should use validated measures aligned with DSM-5 criteria.

In summary, the ADHD-RS-IV provides similar estimates of ADHD prevalence in ASD as semi-structured interviews, and demonstrates expected relationships with child characteristics and behaviors in real world settings. However, factor analyses demonstrate that the ADHD-RS-IV does not capture the latent constructs of Inattention and Hyperactivity/Impulsivity with sufficient precision. Secondary analyses suggest that Inattention items (#3, #5, #15) and a Hyperactivity/Impulsivity item (#2) may capture both ADHD and ASD symptoms. These items need refinement to improve the accuracy of our measurement of the latent constructs of Inattention and Hyperactivity/Impulsivity. This limitation is likely to be present for other well-used ADHD screeners that pull items directly from the DSM-IV-TR (e.g., the Vanderbilt ADHD Diagnostic Parent Rating Scale). The field needs to demonstrate that newer tools (e.g., ADHD-RS-5) have adequate psychometric properties in an ASD population. Until then, existing tools can be used for screening and diagnostic purposes only when combined with a follow-up interview to distinguish the reasons for elevated ratings, with a particular eye toward those items which were found to be problematic in ASD.

Acknowledgments We would like to thank the youth and family members who gave their time and energy to participate in this study. The study was sponsored by grants from the National Institute of Mental Health (K23MH086111; PI: B.E. Yerys, R21MH092615; PI: B.E. Yerys, RC1MH088791; R.T. Schultz), and a New Program Development Award to B.E. Yerys through the Intellectual and Developmental Disabilities Research Center funded by the Eunice Kennedy Shriver National Institute of Child and Human Development (P30HD026979; PI: M. Yudkoff), a grant from the Philadelphia Foundation, a grant from the Pennsylvania Department of Health (SAP #4100042728) to R.T. Schultz, a grant from the Pennsylvania Department of Health (SAP #4100047863) to R.T. Schultz, a grant from Pfizer to R.T. Schultz, and a grant from the Robert Wood Johnson Foundation, #6672 to R.T. Schultz. T.J. Power receives compensation for the ADHD-IV rating scale.

Author Contributions BEY, JNT, AdM, TJP, and RTS conceived the study's key aims and objectives. LA contributed to data collection, data cleaning, and preliminary analyses. MWW conducted factor analyses, interpretation, and drafting of factor analyses. BEY drafted the initial manuscript with significant revisions from all authors.

Compliance with Ethical Standards

Conflict of Interest Thomas J. Power receives financial compensation for the ADHD rating scale 4th edition. Benjamin E. Yerys declares that he has no conflict of interest. Jenelle Nissley-Tsiopinis declares that she has no conflict of interest. Ashley de Marchena declares that she has no conflict of interest. Marley W. Watkins declares that he has no conflict of interest. Ligia Antezana declares that she has no conflict of interest. Robert T. Schultz declares that he has no conflict of interest.

Ethical Approval All procedures performed in this study with human participants were in accordance with the ethical standards of the institutional research board committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This included obtaining informed parental consent and child assent for all participants.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition* (4th ed.). Arlington, VA: American Psychiatric Publishing, Inc.

Andersen, P. N., Hovik, K. T., Skogli, E. W., Egeland, J., & Oie, M. (2013). Symptoms of ADHD in children with high-functioning autism are related to impaired verbal working memory and verbal delayed recall. *PLoS One, 8*(5), e64842. doi:10.1371/journal.pone.0064842.

Antezana, L., Mosner, M. G., Troiani, V., & Yerys, B. E. (2016). Social-emotional inhibition of return in children with autism spectrum disorder versus typical development. *Journal of Autism and Developmental Disorders, 46*(4), 1236–1246.

Baumgaertel, A., Wolariach, M. L., & Dietrich, M. (1995). Comparison of diagnostic criteria for attention deficit disorders in a German elementary school sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*(5), 629–638.

Bentler, P. M., & Wu, E. J. C. (2002). *EQS for Windows user's guide*. Encino, CA: Multivariate Software, Inc.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.

Chevallier, C., Parish-Morris, J., Tonge, N., Le, L., Miller, J., & Schultz, R. T. (2014). Susceptibility to the audience effect explains performance gap between children with and without autism in a theory of mind task. *Journal of Experimental Psychology: General, 143*(3), 972–979.

Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research, 8*(5), 620–628.

Child, D. (2006). *The Essentials of factor analysis*. London: A&C Black.

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research, 166*(2–3), 210–222. doi:10.1016/j.psychres.2008.02.005.

Doyle, A. E., Faraone, S. V., Seidman, L. J., Willcutt, E. G., Nigg, J. T., Waldman, I. D., et al. (2005). Are endophenotypes based on measures of executive functions useful for molecular genetic studies of ADHD?. *Journal of Child Psychology and Psychiatry*, 46(7), 774–803. doi:10.1111/j.1469-7610.2005.01476.x.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD rating scale-IV: Checklists, norms, and clinical interpretation*. New York, NY: Guilford Press.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (2016). *ADHD rating scale-5 for children and adolescents: Checklists, norms, and clinical interpretation*. New York, NY: Guilford Press.

DuPaul, G. J., Reid, R., Anastopoulos, A. D., Lambert, M. C., Watkins, M. W., & Power, T. J. (2016). Parent and teacher ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Psychological Assessment, 28*(2), 214–225. doi:10.1037/pas0000166.

DuPaul, G. J., Reid, R., Anastopoulos, A. D., & Power, T. J. (2014). Assessing ADHD symptomatic behaviors and functional impairment in school settings: Impact of student and teacher characteristics. *School Psychology Quarterly, 29*(4), 409–421. doi:10.1037/spq0000095.

Elliott, C. D. (2007). *Differential ability scales-II (DAS-II)*. San Antonio, TX: Pearson Assessments.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*(4), 466–491. doi:10.1037/1082-989X.9.4.466.

Gioia, G. A., Isquith, P. K., Guy, S., & Kenworthy, L. (2000). *BRIEF: Behavior rating inventory of executive function*. Odessa, FL: Psychological Assessment Resources.

Goldberg, L. R., & Velicer, W. F. (2006). Principles of exploratory factor analysis. In S. Strack (Ed.), *Differentiating normal and abnormal personality* (2nd ed., pp. 209–237). New York, NY: Springer.

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, 68(3), 532–560. doi:10.1207/s15327752jpa6803_5.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37*(4), 613–627.

Granader, Y., Wallace, G. L., Hardy, K. K., Yerys, B. E., Lawson, R. A., Rosenthal, M., et al. (2014). Characterizing the factor structure of parent reported executive function in autism spectrum disorders: The Impact of cognitive Inflexibility. *Journal of Autism and Developmental Disorders, 44*(12), 3056–3062.

Herrington, J. D., Miller, J., Pandey, J., & Schultz, R. T. (2016). Anxiety and social deficits have distinct relationships with amygdala function in autism spectrum disorder. *Social Cognitive and Affective Neuroscience, 11*(6), 907–914.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. doi:10.1080/10705519909540118.

Humphreys, K. L., Mehta, N., & Lee, S. S. (2012). Association of parental ADHD and depression with externalizing and internalizing dimensions of child psychopathology. *Journal of Attention Disorders, 16*(4), 267–275. doi:10.1177/1087054710387264.

Johnston, K., Dittner, A., Bramham, J., Murphy, C., Knight, A., & Russell, A. (2013). Attention deficit hyperactivity disorder symptoms in adults with autism spectrum disorders. *Autism Research*. doi:10.1002/aur.1283.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., et al. (1994). DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *The American Journal of Psychiatry, 151*(11), 1673–1685.

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., et al. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders, 36*(7), 849–861. doi:10.1007/s10803-006-0123-0.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685.

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., et al. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology, 17*(7), 643–662.

McCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (7nd ed.). Los Angeles, CA

Parish-Morris, J., Chevallier, C., Tonge, N., Letzen, J., Pandey, J., & Schultz, R. T. (2013). Visual attention to dynamic faces and objects is linked to face processing skills: A combined study of children with autism and controls. *Frontiers in Psychology, 4*.

Pillow, D. R., Pelham, W. E. J., & Hoza, B. (1998). Confirmatory factor analyses examining attention deficit hyperactivity disorder symptoms and other childhood disruptive behaviors. *Journal of Abnormal Child Psychology, 26*(4), 293–309.

Pugliese, C. E., Kenworthy, L., Bal, V. H., Wallace, G. L., Yerys, B. E., Maddox, B. B., et al. (2015). Replication and comparison of the newly proposed ADOS-2, module 4 algorithm in ASD without ID: A multi-site study. *Journal of Autism and Developmental Disorders, 45*(12), 3919–3931.

Reiersen, A. M., & Todorov, A. A. (2013). Exploration of ADHD subtype definitions and co-occurring psychopathology in a Missouri population-based large sibship sample. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology, 1*(1), 3.

Revelle, W. (2015). *psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University. Retrieved from http://CRAN.R-project.org/package=psych.

Reynolds, C., & Kamphaus, R. (2004). *BASC-2: Behavior assessment systerm for children* (2nd edn). Odessa, FL: Psychological Assessment Resources.

Schwarz, N. (2010). Measurement as a cooperative communication: What research participants learn from questionnaires. In G. Walford, E. Tucker & M. Viswanathan (Eds.), *The SAGE handbook of measurement* (pp. 43–59). Thousand Oaks, CA: Sage Publications, Inc.

Sikora, D. M., Vora, P., Coury, D. L., & Rosenberg, D. (2012). Attention-deficit/hyperactivity disorder symptoms, adaptive functioning, and quality of life in children with autism

spectrum disorder. *Pediatrics, 130*(Suppl 2), S91–S97. doi:10.1542/peds.2012-0900G.

Sinzig, J., Walter, D., & Doepfner, M. (2009). Attention deficit/hyperactivity disorder in children and adolescents with autism spectrum disorder: Symptom or syndrome? *Journal of Attention Disorders, 13*(2), 117–126. doi:10.1177/1087054708326261.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment* (pp. 41–71). New York, NY: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4615-4397-8_3.

Wegener, D. T., & Fabrigar, L. R. (2000). Analysis and design for nonexperimental data. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 412–450). New York, NY: Cambridge University Press.

Willcutt, E. G., Chhabildas, N., Kinnear, M., DeFries, J. C., Olson, R. K., Leopold, D. R., et al. (2014). The internal and external validity of sluggish cognitive tempo and its relation with DSM–IV ADHD. *Journal of Abnormal Child Psychology, 42*(1), 21–35. doi:10.1007/s10802-013-9800-6.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry, 57*(11), 1336–1346. doi:10.1016/j.biopsych.2005.02.006.

Yerys, B. E., Kenworthy, L., Jankowski, K. F., Strang, J., & Wallace, G. L. (2013). Separate components of emotional go/no-go performance relate to autism versus attention symptoms in children with autism. *Neuropsychology, 27*(5), 537–545. doi:10.1037/a0033615.