Journal of
Attention Disorders

Age and Gender Measurement Noninvariance of the Adult ADHD Self-Report Scale Screener

Journal:	*Journal of Attention Disorders*
Manuscript ID	JAD-17-07-184.R1
Manuscript Type:	Article
Keywords:	ADHD, Adult, Age, Gender, Measurement Invariance

Abstract:	Objective: The nature and form of demographics-related differences in ADHD self-reported symptoms across adulthood is currently poorly understood. This study explored the psychometrics of the Adult ADHD Self-Report Scale Screener (ASRS-6), including its age- and gender-related measurement invariance. Methods: Structural equation models were used to analyze adult data—aged 16 to 95—from the 2007 British Adult Psychiatric Morbidity Survey. Results: The 3-factor model (Disorganization, Inattention, Hyperactivity) outperformed one- and two-factor models. Self-reported attentional deficits may be more pathognomonic of overall self-reported ADHD in adults than other symptoms. All items exhibited age measurement noninvariance while only a subset exhibited gender measurement noninvariance. Individuals who are male and younger tend to have lower thresholds for endorsement for ASRS-6 items than individuals who are female and older. Conclusions: The ASRS-6 does not appear to be unidimensional, and self-reported ASRS-6 symptomatology changes in meaning with age.

Abstract

Objective: The nature and form of demographics-related differences in ADHD self-reported symptoms across adulthood is currently poorly understood. This study explored the psychometrics of the Adult ADHD Self-Report Scale Screener (ASRS-6), including its age- and gender-related measurement invariance. Methods: Structural equation models were used to analyze adult data—aged 16 to 95—from the 2007 British Adult Psychiatric Morbidity Survey. Results: The 3-factor model (Disorganization, Inattention, Hyperactivity) outperformed one- and two-factor models. Self-reported attentional deficits may be more pathognomonic of overall self-reported ADHD in adults than other symptoms. All items exhibited age measurement noninvariance while only a subset exhibited gender measurement noninvariance. Individuals who are male and younger tend to have lower thresholds for endorsement for ASRS-6 items than individuals who are female and older. Conclusions: The ASRS-6 does not appear to be unidimensional, and self-reported ASRS-6 symptomatology changes in meaning with age.

Keywords: adult, ADHD, age, gender, measurement invariance

Age and Gender Measurement Noninvariance in the Adult ADHD Self-Report Scale Screener

Prospective studies have established attention-deficit/hyperactivity disorder (ADHD) symptoms may continue at clinically significant levels as children transition into adulthood (Mannuzza, Klein, Bessler, Malloy, & LaPadula, 1998; Rasmussen & Gillberg, 2000). Epidemiological studies found between 1 and 16.4% of adults meet criteria for ADHD (de Graaf et al., 2008; Faraone & Biederman, 2005; Fayyad et al., 2007; Kessler et al., 2006; Kooij et al., 2005; Montejano, Sasane, Hodgkins, Russo, & Huse, 2011; Polanczyk & Rohde, 2007; Simon, Czobor, Bálint, Mészáros, & Bitter, 2009). With 40 to 70% of children meeting criteria for ADHD showing clinical levels in adulthood (Barkley, Murphy, & Fischer, 2008; Faraone, Biederman, & Mick, 2006; Kessler et al., 2010; Weiss & Hechtman, 1993), there is clear need to examine how ADHD indicators vary in their measurement properties across age.

Factor Structure of ADHD Symptoms

Investigations into ADHD symptom structure often compared two- (i.e., inattention and hyperactivity/ impulsivity) and three-factor (i.e., inattention, hyperactivity, impulsivity) models; resulting in support more often for the two-factor model (Manor et al., 2012; Reuter, Kirsch, & Hennig, 2006; Smith & Johnson, 1998; Span, Earleywine, & Strybel, 2002; Spencer et al., 2010) than the three-factor model (Caterino, Gómez-Benito, Balluerka, Amador-Campos, & Stock, 2009; Glöckner-Rist, Pedersen, & Rist, 2013; Proctor & Prevatt, 2009). However, when studies included higher-order models, structural models for children and adolescents (Martel, von Eye, & Nigg, 2010; Normand, Flora, Toplak, & Tannock, 2012; Toplak et al., 2009, 2012) as well as adults (Gibbins, Toplak, Flora, Weiss, & Tannock, 2012) have provided support for a bifactor structure in ADHD symptomatology. Bifactor models consist of a general ADHD factor and specific factors for inattention and hyperactivity-impulsivity, supporting the notion that general ADHD risk may be separate from specific risk for either inattention or hyperactivity-impulsivity.

Age-Related Differences in ADHD Symptoms

Longitudinal studies in children and adolescents show that age is clearly relevant to the presentation of ADHD symptoms (Hart, Lahey, Loeber, Applegate, & Frick, 1995; Lahey, Pelham, Loney, Lee, & Willcutt, 2005; Larsson, Dilshad, Lichtenstein, & Barker, 2011). Inattention tends to persist into adulthood more than hyperactivity or impulsivity (Biederman et al., 2006); however, this may only be true for self-report (Bramham et al., 2012). Others have posited the same presentation represents increasing ADHD severity as individuals age (Barkley et al., 2008; Halperin, Trampush, Miller, Marks, & Newcorn, 2008). Formal statistical modeling (e.g., measurement invariance analyses) of age-related differences in adult ADHD symptom endorsements are needed to distinguish whether apparent differences in ADHD symptom severity may be due to measurement bias (Meredith, 1993; Meredith & Teresi, 2006).

While several studies approximate this goal and have unique methodological strengths, more research is needed. For example, Toplak and colleagues (2012) tested measurement invariance in children. Polanczyk and colleagues (2010) used Rasch analyses to test the fit of item response models in multiple age groups, though Rasch models are likely overly restrictive in their assumptions and cannot distinguish between poor fit due to measurement variance and poor fit due to inadequacies of the model. We are aware of no study using comprehensive structural analyses to disentangle possible reasons for age-related differences in adult self-reported ADHD symptoms across a broad range of ages.

Gender Differences in ADHD Symptoms

Gender differences exist in mean-level symptom endorsements with boys exhibiting more

ADHD symptoms than girls (Gaub & Carlson, 1997; Gershon, 2002; Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007; Wiesner, Windle, Kanouse, Elliott, & Schuster, 2015). Gender-by-age interactions have also been found, with ADHD prevalence appearing to decrease with age in male-dominant samples but increase in female-dominant samples (Simon et al., 2009). Consistent with gender differences across the life span, and in contrast to gender differences found in childhood, an adult sample assessed with the ASRS-6 showed no gender differences in mean-level ADHD symptoms (Das, Cherbuin, Butterworth, Anstey, & Easteal, 2012). The magnitude, direction, and nature of gender differences tend to vary widely across studies, making findings difficult to reconcile and interpret.

Given possible gender differences and difficulty in interpreting them, it is important to determine to what extent other sources of variation are being confounded with "true" gender differences in ADHD symptoms (Meredith, 1993; Meredith & Teresi, 2006). Indeed, a common inference made across many studies (Arcia & Conners, 1998; Gaub & Carlson, 1997; Gershon, 2002; Simon et al., 2009) is the notion that gender bias—arising from referral or sample source, type of informant, or methods of ADHD identification—may influence prevalence estimates and mean-level severities. Clearly, these concerns point to the importance of understanding the instruments used in assessing ADHD symptoms. We must distinguish between gender bias at the level of the informant, referral process, or clinician, as compared to gender bias in the symptom descriptions. The former is more likely to influence measurement of childhood ADHD symptoms when others are reporting symptoms, while the latter has the potential to continue to influence measurement of adult ADHD symptoms even when individuals self-report symptoms.

The Psychometric Properties of the Adult ADHD Self-Report Scale Screener

Taylor, Deb, and Unwin (2011) reviewed existing diagnostic measures of adult ADHD, following study quality criteria set by the Quality Assessment for Diagnostic Accuracy Studies (Whiting, Rutjes, Reitsma, Bossuyt, & Kleijnen, 2003). The development, accurate and detailed reporting, and validation of adult ADHD scales are important since symptom profiles have been shown to vary between childhood and adulthood (Biederman et al., 2006; Mannuzza, Klein, & Moulton, 2003; Weiss, Milroy, & Perlman, 1985; Zwi & York, 2004). The authors concluded that additional high-quality studies with detailed reporting were needed in order to explicate the psychometric properties of adult ADHD scales. Specifically, Taylor and colleagues (2011) found that the Adult ADHD Self-Report Scale Screener (ASRS-6) has desirable psychometric properties but needed independent validation, having reviewed two studies by the developers of the ASRS-6 (Kessler et al., 2005, 2007). Taylor and colleagues (2011) did not review factor structure findings. Additional research is needed into the factor structure of the ASRS-6, including whether it is unidimensional and whether a total score is warranted.

The Adult ADHD Self-Report Scale (ASRS-18) was adapted from the clinician-rated ADHD Rating Scale (ADHD-RS; DuPaul, Power, Anastopoulos, & Reid, 1998) into a self-administered scale for screening at-risk individuals in primary care settings by professionals without lengthy training in ADHD diagnosis (Adler & Alperin, 2015). The ASRS-18 was designed to assess the frequency of current symptoms that map onto the DSM-IV Criterion “A” for ADHD (Adler et al., 2006). Items were carefully worded for self-administration, adult-relevant situations, and assessment of current symptoms. The ASRS-18 and ASRS-6 are not intended for diagnosis. High scorers need further assessment, including an in-depth history and differential diagnosis (Adler & Alperin, 2015). The full version contained 18 items (ASRS-18), and the shorter ASRS-6 (Kessler et al., 2005) was created for general screening purposes. The authors found the ASRS-6 has better psychometric properties over the ASRS-18, concluding the shorter screener should be preferred to the full version in community and clinical samples. Validation studies have found high internal consistency, concurrent validity with the clinician-rated ADHD symptoms, and test-retest reliability (Kessler et al., 2007).

We know of no studies that have factor analyzed English versions of the ASRS-6 in a large community sample. Two-factor models fit best to studies using smaller samples, non-community samples, or the ASRS-6 in other languages (Carlucci, Ivanova, Bissada, & Tasca, 2017; Daigre et al., 2009; Doğan, Õncü, Varol-Saraçoğlu, & Küçükgöncü, 2009; Hesse, 2013). These studies did not attempt to fit higher-order models such as bifactor models, which have been found to fit the ASRS-18 (Morin, Tran, & Caci, 2013; Stanton, Forbes, & Zimmerman, 2018). We also know of no age or gender measurement invariance modeling of the ASRS-6. Finally, the ASRS-6 is remarkable because its content represents the overlap across a variety of adult ADHD instruments. Therefore, findings of multidimensionality and measurement noninvariance across age or gender in the ASRS-6 may lead to hypotheses about other adult ADHD instruments requiring independent validation.

The Present Study

The present study aimed to examine the factor structure of the ASRS-6 and evaluate the age- and gender-related measurement invariance of the ASRS-6 in a general population sample. The large sample size, broad age range, and oversampling of older adults in the present sample enhances the generalizability of findings compared with previous factor analyses. Importantly, we are aware of no previous study that formally modeled age-related and gender differences in the measurement properties of ADHD symptoms in adults.

Method

Procedure and Participants

The Adult Psychiatric Morbidity Survey 2007 is a nationally-representative survey of British adults in private households using stratified random probability sampling. McManus Meltzer, Brugha, Bebbington, and Jenkins (2009) outlined the surveying procedures and response. Interviews of individuals 16 years or older were conducted by lay interviewers using computer-assisted interviewing and assessment methods. The final sample of 7403 participants had a mean age of 51.1 (*SD* = 19.6) ranging from 16 to 95 years, with 56.8% of the sample being male. The sample consisted of 91.9% participants who identified as Caucasian, 2.5% Black, 2.7% South Asian, 2.2% Mixed or Other, and 0.7% participants who did not identify their race.

Measures

Adult Self-Report ADHD Scale Screener. The ASRS-18 (Adler et al., 2006) is a self-report measure designed to identify at-risk community adults who may need further evaluation for ADHD. Part A of the ASRS-18 consists of the six items (ASRS-6; Kessler et al., 2005) found to have the most desirable psychometric properties for detecting possible ADHD. Part B was not administered to the present sample. Items ask respondents to rate an ADHD symptom over the past six months on a Likert scale with 0 (“never”), 1 (“rarely”), 2 (“sometimes”), 3 (“often”), or 4 (“very often”). For reliability and validity data, see Kessler and colleagues (2005, 2007) for the ASRS-6 and Taylor and colleagues (2011) for ASRS-18.

Analyses

Confirmatory factor analyses. Multiple confirmatory factor analyses (CFAs) were fit to the data (Figure 1). Two- and three-factor models were fit once allowing factors to correlate and fit again fixing the correlations to zero. Factor correlations were restricted to zero in hierarchical models but allowed to vary freely in higher-order models. Models were analyzed using maximum likelihood estimation with robust standard errors in Mplus version 6.1 (Muthén & Muthén, 1998-2011). Analyses used sampling weights to account for selection probabilities within different sized households, non-response bias, and demographic representativeness based on national distribution (McManus et al., 2009).

Measurement invariance multiple-group modeling. Age-related models were fit using seven age groups (Jöreskog, 1971). Participants 16 to 75 years of age were separated into 10-year-span groups while participants 76 to 95 years of age were placed into a single group due to sparse endorsement among elderly participants. Models with different degrees of measurement invariance were compared (Vandenberg & Lance, 2000): configural invariance (Config), metric invariance (Metric), strict factorial invariance (Strict), and partial invariance models (Partial; Schmitt, Golubovich, & Leong, 2011). Item intercepts are estimated freely across groups in Config and Metric models, while factor loadings are estimated freely in Config but held constant in Metric models. Significant decreases in fit from Config to Metric models indicate age or gender noninvariance in factor loadings. Strict models were restrictive modifications on Metric models, with item intercepts held invariant across groups. Significant decreases in fit from Metric to Strict models would indicate age or gender bias in parts or all of the ASRS-6. Factor variances, covariances, and residual variances are freely estimated across groups.

Models were compared using likelihood ratio tests (LRTs) when possible, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Draper's BIC (Draper, 1995). LRTs and AICs are known for sometimes being biased in favor of over-parameterized models, while BICs are known for being biased in favor of overly simplistic models (Weakliem, 1999). Therefore, whenever LRTs and AICs disagree with BICs, Draper's BICs—which is a value between AICs and BICs—was used to select the best-fitting model.

Results

Confirmatory Factor Analyses

CFA results are shown in Table 1. The one-factor model (Model A) showed poor fit. Fit indices (AIC = 111455.043, BIC = 111683.034) identified the correlated three-factor model as best-fitting (Model F). In this model, Factor 1 consisted of Items 1 (λ_1 = 0.73, *SD* = 0.01) and 2 (λ_2 = 0.85, *SD* = 0.01), reflecting Disorganization. Factor 2 consisted of Items 3 (λ_3 = 0.57, *SD* = 0.02) and 4 (λ_4 = 0.68, *SD* = 0.01), reflecting Inattention. Factor 3 consisted of Items 5 (λ_5 = 0.85, *SD* = 0.02) and 6 (λ_6 = 0.53, *SD* = 0.02), reflecting Hyperactivity. Factors 1 and 2 were highly correlated (r_{12} = 0.91, *SD* = 0.02), while Factors 1 and 3 (r_{13} = 0.45, *SD* = 0.02) and Factors 2 and 3 (r_{23} = 0.64, *SD* = 0.02) were moderately correlated. Though r_{12} was high in Model F, Model D's relatively poor fit suggests Items 1 to 4 should not be combined into a single factor.

Of the complex models, bifactor models fit better than second-order models while models with two subordinate or specific factors fit better than those with three (Table 1). The best-fitting complex model was bifactor Model J, in which Items 1 and 2 loaded onto a specific factor, Items 5 and 6 loaded onto another specific factor, and all items loaded onto a general factor (AIC = 111506.194, BIC = 111727.276). No second-order or bifactor model fit better than the best-fitting three-factor Model F. Notably, Model G is a reparameterization of the best-fitting simple model but resulted in a warning of possible nonconvergence, with the second subordinate factor's loading onto the superordinate factor approaching one and residual variance approaching zero. In supplemental analyses using Bayesian estimation, Model F was the only model to have good fit to the data with a posterior positive predictive value (*PPPV*) of 0.06 (95% CI of χ^2= [-3.91, 42.44]). Models F and J's comparative fit index, Tucker Lewis index, and root mean square error of approximation are 0.996 and 0.990, 0.991 and 0.979, and 0.026 and 0.038, respectively; indicating good absolute fit according to supplemental weighted least-squares analyses.

Measurement Invariance Multiple-Group Modeling

Multiple-group models were performed on the best-fitting model and comparison results are shown in Table 2. Metric invariance was achieved with respect to age and gender, indicating factor loadings were invariant across groups. The age- and gender-invariant factor loadings of Items 1 to 6 onto their corresponding factors are broadly consistent with the pattern of loadings from the best-fitting model. Disorganization and Inattention are highly correlated across all age and gender groups ($0.83 < r_{12} < 0.97$), while Hyperactivity is more modestly correlated with Disorganization ($0.39 < r_{13} < 0.51$) and Inattention ($0.55 < r_{23} < 0.68$) across all groups.

For age and gender, pairwise comparisons between Metric and Strict models indicated bias in all or parts of the ASRS-6. Metric's intercepts are presented for age (Figure 2A) and gender (Figure 2B) to illustrate the potential degree of bias. To investigate partial invariance — some items are biased while others are not—intercepts were systematically constrained across groups for each item (Partial1 to Partial6). With respect to age, Partial models fit significantly worse than Metric. Taken together, these results suggest age bias in all items of the ASRS-6.

Gender-related Partial models suggest the existence of bias in Items 1 and 3 and the absence of bias in Items 2, 5, and 6. Metric fits better than the Partial1 or Partial3 models, suggesting gender bias in Items 1 and 3. Partial5 fits better than the Metric model, indicating Item 5 is not gender biased. The preponderance of fit indices favor Partial2, Partial4, and Partial6 over the Metric model, suggesting corresponding items are not biased. To account for possible overlap between these items in the amount of general ADHD variance they may explain, we restricted the intercepts of Items 2, 4, 5, and 6 simultaneously (Partial2456). The difference in

Draper's BIC—a relative fit index without an absolute threshold on significant differences for model comparisons—between Metric and Partial2456 models is subtle (i.e., 121708.14 – 121704.53 = 3.61). Therefore, we went a step further and relaxed the thresholds of Item 4 (Partial256)—the item with the greatest potential for gender bias, as evidenced in the fit indices of Partial4 versus Partial2, Partial5, and Partial6. Indeed, Partial256 fits better than Metric. Taking all of these model comparisons into consideration, evidence exists that Items 2, 5, and 6 are not gender biased, while the differential functioning of Item 4 across genders is inconclusive.

To illustrate possible inferences should item bias be ignored, age-related differences in factor means are presented using the Strict model (Figure 3). While Hyperactivity appears to decrease dramatically with age, differences in Disorganization and Inattention are much smaller. In the gender-related Strict model, females and males do not appear to differ in mean levels of Inattention ($p = 0.89$) and Hyperactivity ($p = 0.88$). Contrastingly, females appear to exhibit lower mean levels of Disorganization compared to males ($p < 0.001$). To be clear, Strict models are not the best-fitting multiple-group models because they assume that all ASRS-6 items are unbiased with respect to age or gender. Our analyses suggest this assumption is unwarranted.

Discussion

We sought to examine the factor structure of the ASRS-6 and evaluate the age- and gender-related measurement invariance of the ASRS-6 in a general population sample.

Factor Structure of the ASRS-6

The preponderance of structural evidence indicates the three-factor (i.e., Disorganization, Inattention, Hyperactivity) model best fit the present sample. However, parameter estimates from simple models raise an important question that cannot be answered without the complex models. If the Disorganization and Inattention factors correlate as highly as 0.91, what distinguishes between these items such that they do not simply form one factor? Complex models suggest this may because Inattention items covary with general ADHD while Disorganization items contain specific variance partially independent of general variance. Consistent with previously stated notions that inattention symptoms may be most closely related to poor executive functioning (Barkley et al., 2008), the comparatively better fit of complex models lacking a lower-order or subordinate factor for Items 3 and 4 raise the hypothesis that inattention items may be more pathognomonic of general ADHD vulnerability than other items. Note that "general ADHD" in this context refers to a statistically-derived construct capturing variation that is common to all items of the ASRS-6, used with clinical utility without claims about etiological validity (Barkley, 2006; Barkley et al., 2002; Faraone, 2005; Hinshaw, 2002; Lahey & Willcutt, 2002).

Complex model results suggest a two-factor bifactor model, with some complications to straightforward interpretation. Consistent with previous findings that bifactor models best characterize the structure of ADHD symptoms, our two-factor bifactor model fit the data well according to weighted least-squares analyses and was the best-fitting complex model. Because the three-factor second-order model reparameterized from the three-factor simple model did not converge and cannot be directly compared to other complex models, we cannot conclude that a bifactor structure fit better than the best-fitting simple model. Thus, replication efforts should consider complex models such as bifactor models alongside simple two- or three-factor structures. Complex models are also difficult to test with six items; the structure of adult ADHD symptoms may be best investigated with lengthier measures.

Regardless of whether the best-fitting structure is three-factor or bifactor, caution needs to be exercised when interpreting the ASRS-6. The present analyses strongly suggest the ASRS-6 may not be unidimensional and global scores may not be optimally informative. Distinguishing between multiple symptom dimensions may be important when assessing adult ADHD with the ASRS-6 in clinical and research contexts. Further, the present results suggest that Items 1 and 2, 3 and 4, and 5 and 6 covary in paired clusters. While Items 1 to 4 were conceptualized as inattention symptoms, our analyses found that the Disorganization and Inattention clusters do not function identically in relation to other items. Possibly, ASRS-6 items that were originally conceptualized as inattention symptoms do not comprise a unidimensional construct either.

Age-Related Noninvariance of the ASRS-6

Configural invariance was achieved, such that the relationship (i.e., loading) of each item to its latent factor is invariant across age. A lack of such invariance may mean changes in the underlying construct an item represents with age. For example, if item loadings are lower in older age groups than younger age groups, it may mean that the influence of latent vulnerabilities decrease as individuals age. The finding of configural invariance assures us that regardless of age, each item represents latent vulnerabilities to the same extent across the life span.

Measurement invariance analyses also examined whether items functioned similarly across ages (i.e., same intercepts for younger versus older individuals). Each intercept represents the level of the latent trait that an individual must possess in order to endorse that response level. Taken together, measurement noninvariance in intercepts would mean that an item or a set of items are biased with respect to age. Further, measurement noninvariance (i.e., the “difficulty” of the ASRS-6 increases with age) would be consistent with previous hypotheses that the same symptom presentation may represent increasing levels of ADHD severity as individuals age (Barkley et al., 2008; Halperin et al., 2008) as well as the recent implementation of a five-symptom cutoff for older adolescents and adults within the DSM-5.

Results showed all items may contain age bias. Graphical representations (Figure 2A) are consistent with quantitative modeling. For all items, intercepts appear to vary substantially between age groups for at least some response levels. This suggests some items are too "easy" to appropriately distinguish between different levels of ADHD severity in younger populations, whereas some are too "difficult" to appropriately screen for ADHD in more elderly populations. If we ignore measurement bias, it would be easy to conclude that hyperactive symptoms decline with age while disorganized and inattentive symptoms remain more stable (Figure 3).

Gender-Related Noninvariance of the ASRS-6

Similar to age-invariance analyses, configural invariance was achieved, such that each item represents latent vulnerabilities to the same extent regardless of gender. Measurement invariance analyses also examined whether the items functioned similarly across genders. Quantitative and qualitative (Figure 2B) results showed that two items of the ASRS-6 (i.e., "trouble wrapping up projects", "problems remembering appointments or obligations") are biased with regards to gender. Males have lower thresholds for endorsement for these items than females. These findings suggest that the same self-reported difficulty with organization or attention may represent higher levels of the latent trait in females compared to males. If we ignore measurement bias, it would be easy to conclude that females exhibit less disorganized symptoms compared to males.

Concerns about gender bias have appropriately focused on issues regarding differences between how females and males with ADHD are perceived by others and the types of symptoms that are noticed by others. It is also important to investigate the instruments themselves as a source of gender bias. This is especially relevant to self-report measures such as the ASRS-6, where item-level gender bias may obfuscate clinical or research findings even if the reporter (i.e., adult experiencing symptoms) accurately reports their experiences. The potential existence of gender bias in Items 1 and 3 of the ASRS-6 is also notable because ASRS-6 items represent the content overlap across other measures of adult ADHD symptoms. Therefore, we hope that these findings might spur future research into the gender-invariance of these other measures.

Strengths, Limitations, and Future Directions

The current study has certain methodological strengths worth mentioning. First, it is the first study that we are aware of to formally model age- and gender-related differences in the psychometric properties of a measure of adult ADHD symptoms. Further, it has an exceptionally wide age range and the oversampling of elderly populations may boost power to detect measurement noninvariance in such populations. Second, the current findings may hold implications for existing hypotheses about the changing meaning of symptoms across age and gender. While the results are most immediately applicable to the ASRS-6, that the ASRS-6 items represents the content overlap across other measures of adult ADHD symptoms should raise hypotheses about the measurement invariance of other measures. If future investigations uncover evidence of differential item functioning in other measures, more robust inferences may be drawn about the symptom criteria of adult ADHD in general. Clinical diagnoses, scoring and interpretation of assessment instruments, and research design decisions may potentially be impacted if the same symptom presentation represents increasing levels of severity of ADHD psychopathology in elderly or female populations. Third, psychometric validation and further improvement of the ASRS-6 may benefit from consideration of our findings.

Even though the current study yields many noteworthy findings, there are limitations. This study is limited by the sole use of the ASRS-6, to the extent that it obtains incomplete information about current symptoms. ADHD psychopathology was not independently verified against measures other than the ASRS-6. Because the ASRS-6 is designed for use as a screener for the possible existence of ADHD, a separate measure of ADHD severity would have strengthened the inferential ability of the current study. For example, whenever levels of latent traits are estimated, we cannot infer that these latent traits are specific aspects of ADHD psychopathology without such external validation. Even though self-report has been found to be most accurate with regard to adult ADHD symptoms, individuals with ADHD may underreport symptoms (Kooij et al., 2008; Young & Gudjonsson, 2005). Informant reports should be used whenever possible to supplement self-reports in future studies. These results also should be replicated using longer measures as well as longitudinal data. Additionally, performance measures of executive functioning, inattention, memory, and other forms of neurocognitive functioning may yield much needed information about covert difficulties of adults with ADHD psychopathology.

References

Adler, L. A., & Alperin, S. (2015). Diagnosing ADHD in adults in the primary care setting. In R. A. Barkley (Ed.), *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment, 4th ed.* (pp. 501-509). New York, NY: Guilford Press.

Adler, L. A., Spencer, T., Faraone, S. V., Kessler, R. C., Howes, M. J., Biederman, J., & Secnik, K. (2006). Validity of pilot Adult ADHD Self- Report Scale (ASRS) to rate adult ADHD symptoms. *Annals of Clinical Psychiatry, 18*(3), 145-148.

Arcia, E., & Conners, C. K. (1998). Gender differences in ADHD? *Journal of Developmental and Behavioral Pediatrics, 19*(2), 77-83.

Barkley, R. A. (2006). *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment, 3rd ed.* New York, NY: Guilford Press.

Barkley, R. A., Cook, E. H., Jr., Diamond, A., Zametkin, A., Thapar, A., & Teeter, A. (2002). International consensus statement on ADHD. January 2002. *Clinical Child and Family Psychology Review, 5*(2), 89-111.

Barkley, R. A., Murphy, K. R., & Fischer, M. (2008). *ADHD in Adults: What the Science Says*. New York, NY: Guilford Press.

Biederman, J., Faraone, S. V., Spencer, T. J., Mick, E., Monuteaux, M. C., & Aleardi, M. (2006). Functional impairments in adults with self-reports of diagnosed ADHD: A controlled study of 1001 adults in the community. *Journal of Clinical Psychiatry, 67*(4), 524-540.

Bramham, J., Murphy, D. G., Xenitidis, K., Asherson, P., Hopkin, G., & Young, S. (2012). Adults with attention deficit hyperactivity disorder: An investigation of age-related differences in behavioural symptoms, neuropsychological function and co-morbidity. *Psychological Medicine, 42*(10), 2225-2234.

Carlucci, S., Ivanova, I., Bissada, H., & Tasca, G. A. (2017). Validity and reliability of the attention deficit hyperactivity disorder self-report scale (ASRS-v1.1) in a clinical sample with eating disorders. *Eating Behaviors, 26*, 148-154.

Caterino, L. C., Gómez-Benito, J., Balluerka, N., Amador-Campos, J. A., & Stock, W. A. (2009). Development and validation of a scale to assess the symptoms of attention-deficit /hyperactivity disorder in young adults. *Psychological Assessment, 21*(2), 152-161.

Daigre, C., Ramos-Quiroga, J. A., Valero, S., Bosch, R., Roncero, C., Gonzalvo, B., ... Casas, M. (2009). Adult ADHD Self-Report Scale (ASRS-v1.1) symptom checklist in patients with substance use disorders. *Actas Españolas de Psiquiatría, 37*(6), 299-305.

Das, D., Cherbuin, N., Butterworth, P., Anstey, K. J., & Easteal, S. (2012). A population-based study of attention deficit/hyperactivity disorder symptoms and associated impairment in middle-aged adults. *PLoS ONE, 7*(2), e31500.

de Graaf, R., Kessler, R. C., Fayyad, J., ten Have, M., Alonso, J., Angermeyer, M., ... Posada-Villa, J. (2008). The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: Results from the WHO World Mental Health Survey Initiative. *Occupational and Environmental Medicine, 65*(12), 835-842.

Doğan, S., Õncü, B., Varol-Saraçoğlu, G., & Küçükgöncü, S. (2009). Validity and reliability of the Turkish version of the Adult ADHD Self-Report Scale (ASRS-v1.1). *Anadolu Psikiyatri Dergisi, 10*(2), 77-87.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 45-97.

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.

Faraone, S. V. (2005). The scientific foundation for understanding attention-deficit/hyperactivity disorder as a valid psychiatric disorder. *Euro J of Child and Adol Psychiatry, 14*(1), 1-10.

Faraone, S. V., & Biederman, J. (2005). What is the prevalence of adult ADHD? Results of a population screen of 966 adults. *Journal of Attention Disorders, 9*(2), 384-391.

Faraone, S. V., Biederman, J., & Mick, E. (2006). The age-dependent decline of attention deficit hyperactivity disorder: A meta-analysis of follow-up studies. *Psychol Med, 36*(2), 159-165.

Fayyad, J., de Graaf, R., Kessler, R., Alonso, J., Angermeyer, M., Demyttenaere, K., ... Jin, R. (2007). Cross-national prevalence and correlates of adult attention-deficit hyperactivity disorder. *British Journal of Psychiatry, 190*, 402-409.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *J of the American Academy of Child & Adolescent Psychiatry, 36*(8), 1036-1045.

Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders, 5*(3), 143-154.

Gibbins, C., Toplak, M. E., Flora, D. B., Weiss, M. D., & Tannock, R. (2012). Evidence for a general factor model of ADHD in adults. *Journal of Attention Disorders, 16*, 635–644.

Glöckner-Rist, A., Pedersen, A., & Rist, F. (2013). Conceptual structure of the symptoms of adult ADHD according to the DSM-IV and retrospective Wender-Utah criteria. *Journal of Attention Disorders, 17*(2), 114-127.

Halperin, J. M., Trampush, J. W., Miller, C. J., Marks, D. J., & Newcorn, J. H. (2008). Neuropsychological outcome in adolescents/young adults with childhood ADHD: Profiles of persisters, remitters and controls. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 49*(9), 958-966.

Hart, E. L., Lahey, B. B., Loeber, R., Applegate, B., & Frick, P. J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology 23*(6), 729-749.

Hesse, M. (2013). The ASRS-6 has two latent factors: Attention deficit and hyperactivity. *Journal of Attention Disorders, 17*(3), 203-207.

Hinshaw, S. P. (2002). Is ADHD an impairing condition in childhood and adolescence? In P. S. Jensen & J. R. Cooper (Eds.), *Attention-Deficit Hyperactivity Disorder: State of the Science, Best Practices* (pp. 5-1-5-22). Kingston, NJ: Civic Research Institute.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*(4), pp. 409-426.

Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S. V., Hiripi, E., ... Walters, E. E. (2005). The World Health Organization Adult ADHD Self-Report Scale (ASRS): A short screening scale for use in the general population. *Psychological Medicine, 35*(2), 245-256.

Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O., ... Zaslavsky, A. M. (2006). The prevalence and correlates of adult ADHD in the United States: Results from the National Comorbidity Survey Replication. *Am J of Psychiatry, 163*(4), 716-723.

Kessler, R. C., Adler, L. A., Gruber, M. J., Sarawate, C. A., Spencer, T., & Van Brunt, D. L. (2007). Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) Screener in a representative sample of health plan members. *International Journal of Methods in Psychiatric Research, 16*(2), 52-65.

Kessler, R. C., Green, J. G., Adler, L. A., Barkley, R. A., Chatterji, S., Faraone, S. V., ... Van Brunt, D. L. (2010). Structure and diagnosis of adult attention-deficit/hyperactivity disorder: Analysis of expanded symptom criteria from the Adult ADHD Clinical Diagnostic Scale. *Archives of Gen Psychiatry, 67*(11), 1168-1178.

Kooij, J. J., Boonstra, A. M., Swinkles, S. H. N., Bekker, E. M., de Noord, I., & Buitelaar, J. K. (2008). Reliability, validity, and utility of instruments for self-report and informant report concerning symptoms of ADHD in adult patients. *J of Attention Disorders, 11*, 445-458.

Kooij, J. J., Buitelaar, J. K., van den Oord, E. J., Furer, J. W., Rijnders, C. A., & Hodiamont, P. P. (2005). Internal and external validity of attention-deficit hyperactivity disorder in a population-based sample of adults. *Psychological Medicine, 35*(6), 817-827.

Lahey, B. B., Pelham, W. E., Loney, J., Lee, S. S., & Willcutt, E. (2005). Instability of the DSM-IV Subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry, 62*(8), 896-902.

Lahey, B. B., & Willcutt, E. (2002). Validity of the diagnosis and dimensions of attention deficit hyperactivity disorder. In P. S. Jensen & J. R. Cooper (Eds.), *Attention-Deficit Hyperactivity Disorder: State of the Science, Best Practices* (pp. 1-1-1-23). Kingston, NJ: Civic Research Institute.

Larsson, H., Dilshad, R., Lichtenstein, P., & Barker, E. D. (2011). Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: Genetic effects, family risk and associated psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 52*(9), 954-963.

Mannuzza, S., Klein, R. G., Bessler, A., Malloy, P., & LaPadula, M. (1998). Adult psychiatric status of hyperactive boys grown up. *American Journal of Psychiatry, 155*(4), 493-498.

Mannuzza, S., Klein, R. G., & Moulton, J. L., 3rd. (2003). Persistence of Attention-Deficit/Hyperactivity Disorder into adulthood: What have we learned from the prospective follow-up studies? *Journal of Attention Disorders, 7*(2), 93-100.

Manor, I., Vurembrandt, N., Rozen, S., Gevah, D., Weizman, A., & Zalsman, G. (2012). Low self-awareness of ADHD in adults using a self-report screening questionnaire. *European Psychiatry, 27*(5), 314-320.

Martel, M. M., von Eye, A., & Nigg, J. T. (2010). Revisiting the latent structure of ADHD: Is there a 'g' factor? *J of Child Psychol and Psychiatry and Allied Disciplines, 51*(8), 905-914.

McManus, S., Meltzer, H., Brugha, T., Bebbington, P., & Jenkins, R. (2009). *Adult Psychiatric Morbidity in England, 2007: Results of a Household Survey*. London: National Centre for Social Research.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), pp. 525-543.

Meredith, W. & Teresi, J. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*(Suppl 3), pp. S69-S77.

Montejano, L., Sasane, R., Hodgkins, P., Russo, L., & Huse, D. (2011). Adult ADHD: Prevalence of diagnosis in a US population with employer health insurance. *Current Medical Research and Opinion, 27*(Suppl 2), 5-11.

Morin, A. J., Tran, A., & Caci, H. (2013). Factorial validity of the ADHD Adult Symptom Rating Scale in a French community sample: Results from the ChiP-ARD study. *Journal of Attention Disorders, 20*(6), 530-541.

Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus User's Guide, Sixth Edition*. Los Angeles, CA: Muthén & Muthén.

Normand, S., Flora, D. B., Toplak, M. E., & Tannock, R. (2012). Evidence for a general ADHD factor from a longitudinal general school population study. *Journal of Abnormal Child Psychology, 40*, 555–567.

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry, 164*(6), 942-948.

Polanczyk, G., Laranjeira, R., Zaleski, M., Pinsky, I., Caetano, R., & Rohde, L. A. (2010). ADHD in a representative sample of the Brazilian population: Estimated prevalence and comparative adequacy of criteria between adolescents and adults according to the item response theory. *International Journal of Methods in Psychiatric Research, 19*(3), 177-184.

Polanczyk, G., & Rohde, L. A. (2007). Epidemiology of attention-deficit/hyperactivity disorder across the lifespan. *Current Opinion in Psychiatry, 20*(4), 386-392.

Proctor, B. E., & Prevatt, F. (2009). Confirming the factor structure of attention-deficit/hyperactivity disorder symptoms in college students using student and parent data. *Journal of Learning Disabilities, 42*(3), 250-259.

Rasmussen, P., & Gillberg, C. (2000). Natural outcome of ADHD with developmental coordination disorder at age 22 years: A controlled, longitudinal, community-based study. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(11), 1424-1431.

Reuter, M., Kirsch, P., & Hennig, J. (2006). Inferring candidate genes for attention deficit hyperactivity disorder (ADHD) assessed by the World Health Organization Adult ADHD Self-Report Scale (ASRS). *Journal of Neural Transmission, 113*(7), 929-938.

Schmitt, N., Golubovich, J., & Leong, F. T. (2011). Impact of measurement invariance on construct correlations, mean differences, and relations with external correlates: An illustrative example using Big Five and RIASEC measures. *Assessment, 18*(4), pp. 412-427.

Simon, V., Czobor, P., Bálint, S., Mészáros, Á., & Bitter, I. (2009). Prevalence and correlates of adult attention-deficit hyperactivity disorder: Meta-analysis. *The British Journal of Psychiatry, 194*, 204-211.

Smith, E. V., Jr., & Johnson, B. D. (1998). Factor structure of the DSM-IV criteria for college students using the Adult Behavior Checklist. *Measurement & Evaluation in Counseling & Development, 31*(3), 164-183.

Span, S. A., Earleywine, M., & Strybel, T. Z. (2002). Confirming the factor structure of attention deficit hyperactivity disorder symptoms in adult, nonclinical samples. *Journal of Psychopathology and Behavioral Assessment, 24*(2), 129-136.

Spencer, T. J., Adler, L. A., Meihua, Q., Saylor, K. E., Brown, T. E., Holdnack, J. A., ... Kelsey, D. K. (2010). Validation of the adult ADHD investigator symptom rating scale (AISRS). *Journal of Attention Disorders, 14*(1), 57-68.

Stanton, K., Forbes, M. K., & Zimmerman, M. (2018). Distinct dimensions defining the Adult ADHD Self-Report Scale: Implications for assessing inattentive and hyperactive/impulsive symptoms. *Psychological Assessment*. Advanced online publication. doi: 10.1037/pas0000604

Taylor, A., Deb, S., & Unwin, G. (2011). Scales for the identification of adults with attention deficit hyperactivity disorder (ADHD): A systematic review. *Research in Developmental Disabilities, 32*(3), 924-938.

Toplak, M. E., Pitch, A., Flora, D. B., Iwenofu, L., Ghelani, K., Jain, U., & Tannock, R. (2009). The unity and diversity of inattention and hyperactivity/impulsivity in ADHD: Evidence for a general factor with separable dimensions. *J of Abnormal Child Psych, 37*(8), 1137-1150.

Toplak, M. E., Sorge, G. B., Flora, D. B., Chen, W., Banaschewski, T., Buitelaar, J., ... Faraone, S. V. (2012). The hierarchical factor model of ADHD: Invariant across age and national groupings? *J of Child Psychology and Psychiatry and Allied Disciplines, 53*(3), 292-303.

Vandenberg, R. J., & Lance, C. E. (2000). Review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), pp. 4-70.

Weakliem, D. L. (1999). A critique of the Bayesian Information Criterion for model selection. *Sociological Methods & Research, 27*, 359-397.

Weiss, G., & Hechtman, L. (1993). *Hyperactive Children Grown Up*. New York, NY: Guilford Press.

Weiss, G., Hechtman, L., Milroy, T., & Perlman, T. (1985). Psychiatric status of hyperactives as adults: A controlled prospective 15-year follow-up of 63 hyperactive children. *Journal of the American Academy of Child Psychiatry, 24*(2), 211-220.

Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., & Kleijnen, J. (2003). The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology, 3*, 25.

Wiesner, M., Windle, M., Kanouse, D. E., Elliott, M. N., & Schuster, M. A. (2015). DISC Predictive Scales (DPS): Factor structure and uniform differential item functioning across gender and three racial/ethnic groups for ADHD, conduct disorder, and oppositional defiant disorder symptoms. *Psychological Assessment, 27(4)*, 1324-1336.

Young, S. (2004). The YAQ-S and YAQ-I: The development of self and informant questionnaires reporting on current adult ADHD symptomatology, comorbid and associated problems. *Personality and Individual Differences, 36*, 1211-1223.

Young, S., & Gudjonsson, G. H. (2005). Neuropsychological correlates of the YAQ-S and YAQ-I self- and informant-reported ADHD symptomatology, emotional and social problems and delinquent behaviour. *British Journal of Clinical Psychology, 44*(Pt 1), 47-57.

Zwi, M., & York, A. (2004). Attention-deficit hyperactivity disorder in adults: Validity unknown. *Advances in Psychiatric Treatment, 10*, 248-256.

Table 1. Comparative fits of the confirmatory factor analysis models.

Model	Correlated Factors?	# of Factors	# of Parameters	AIC	BIC
A	n/a [a]	1	30	112418.416	112625.681
B	Yes	2	31	112286.529	112500.703
B	No	2	30	114194.937	114402.202
C	Yes	2	31	112044.633	112258.807
C	No	2	30	113492.629	113699.894
D	Yes	2	31	111647.392	111861.566
D	No	2	29	112717.341	112917.697
E	Yes	2	31	112172.545	112386.719
E	No	2	29	114474.842	114675.198
F	Yes	3	33	**111455.043**	**111683.034**
F	No	3	27	115283.943	115470.481
G [b]	Yes	4	36	111513.673	111762.391
H	Yes	3	33	111572.856	111800.847
I	No	4	33	111508.216	111736.208
J	No	3	32	**111506.194**	**111727.276**

[a] Only one factor in the model. No multiple factors to correlate.
[b] Problem with model convergence using maximum likelihood estimation with robust standard errors. The second subordinate factor's loading onto the superordinate factor approaches 1 and residual variance approaches 0.

Table 2. Model comparisons used in age- and gender-invariance testing of the three-correlated-factors model.

	Null Model	No. of free parameters of null model	AIC	BIC	Draper's BIC	Alternative Model	Chi-square statistic	Inference based on likelihood ratio testing, AIC, BIC, Draper's BIC
Age	Config	255	138859.70	140621.45	139132.86			
	Metric	219	**138814.65**	140327.68	**139049.25**	Config	35.75	Factor loadings invariant
	Strict	93	139002.24	**139644.76**	139101.86	Metric	377.35 ***	Item intercepts vary wrt age
	Part1	195	138949.90	140297.12	139158.78	Metric	148.54 ***	Item 1 intercepts vary wrt age
	Part2	195	138934.52	140281.74	139143.41	Metric	145.67 ***	Item 2 intercepts vary wrt age
	Part3	195	139023.19	140370.41	139232.08	Metric	230.76 ***	Item 3 intercepts vary wrt age
	Part4	195	138900.00	140247.22	139108.89	Metric	114.91 ***	Item 4 intercepts vary wrt age
	Part5	195	139253.73	140600.95	139462.62	Metric	413.92 ***	Item 5 intercepts vary wrt age
	Part6	195	139309.54	140656.76	139518.43	Metric	448.95 ***	Item 6 intercepts vary wrt age
Gender	Config	70	121644.28	122127.90	121719.26			
	Metric	64	121637.58	122081.75	121708.14	Config	10.23	Factor loadings invariant
	Strict	43	121687.83	**121984.91**	121733.89	Metric	70.98 ***	Item intercepts vary wrt gender
	Part1	60	121661.59	122076.12	121725.87	Metric	22.85 **	Item 1 intercepts vary wrt gender
	Part2	60	121641.19	122055.72	121705.46	Metric	7.40	Item 2 intercepts invariant
	Part3	60	121663.08	122077.61	121727.36	Metric	25.44 ***	Item 3 intercepts vary wrt gender
	Part4	60	121642.11	122056.64	121706.38	Metric	8.02	Item 4 intercepts invariant
	Part5	60	**121634.36**	122048.89	121698.64	Metric	2.09	Item 5 intercepts invariant
	Part6	60	121637.31	122051.84	121701.58	Metric	4.46	Item 6 intercepts invariant
	Part2456	48	121653.11	121984.73	121704.53	Metric	35.15 *	Items 2, 4, 5, and/or 6 intercepts may vary wrt gender
	Part256	52	121635.49	121994.75	**121691.20**	Metric	15.39	Items 2, 5, and 6 intercepts invariant

Note: Config = configural invariance. Metric = metric invariance. Strict = strict factorial invariance. Part = partial invariance; the subsequent numerals indicate the item(s) whose intercepts are restricted to being invariant across groups. wrt = with respect to. Significant likelihood ratio tests indicate a significant reduction in model fit from the null model to the more constrained alternative model. *** $p < .0001$, ** $p < .001$, * $p < .01$, bolded values are the best-fitting models according to each fit index. In calculating the Draper's BIC, $n_{\text{DBIC}} = n / (2\pi)$.

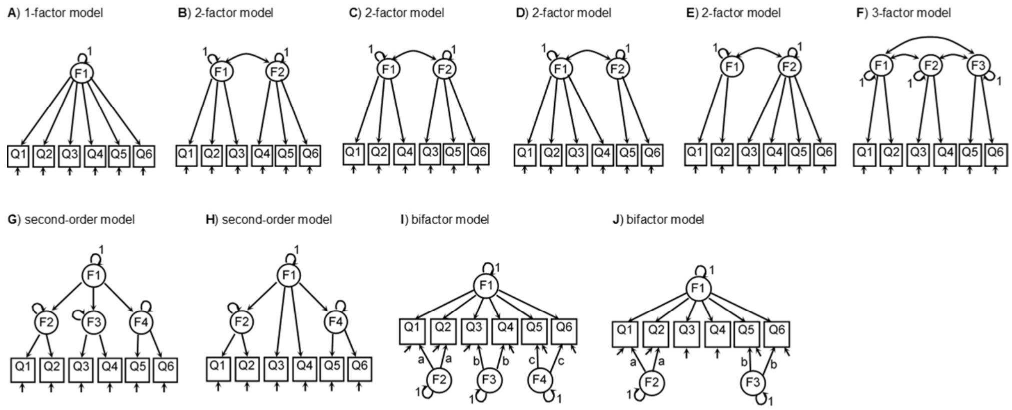

Figure 1. Confirmatory factor analysis models.

165x67mm (300 x 300 DPI)

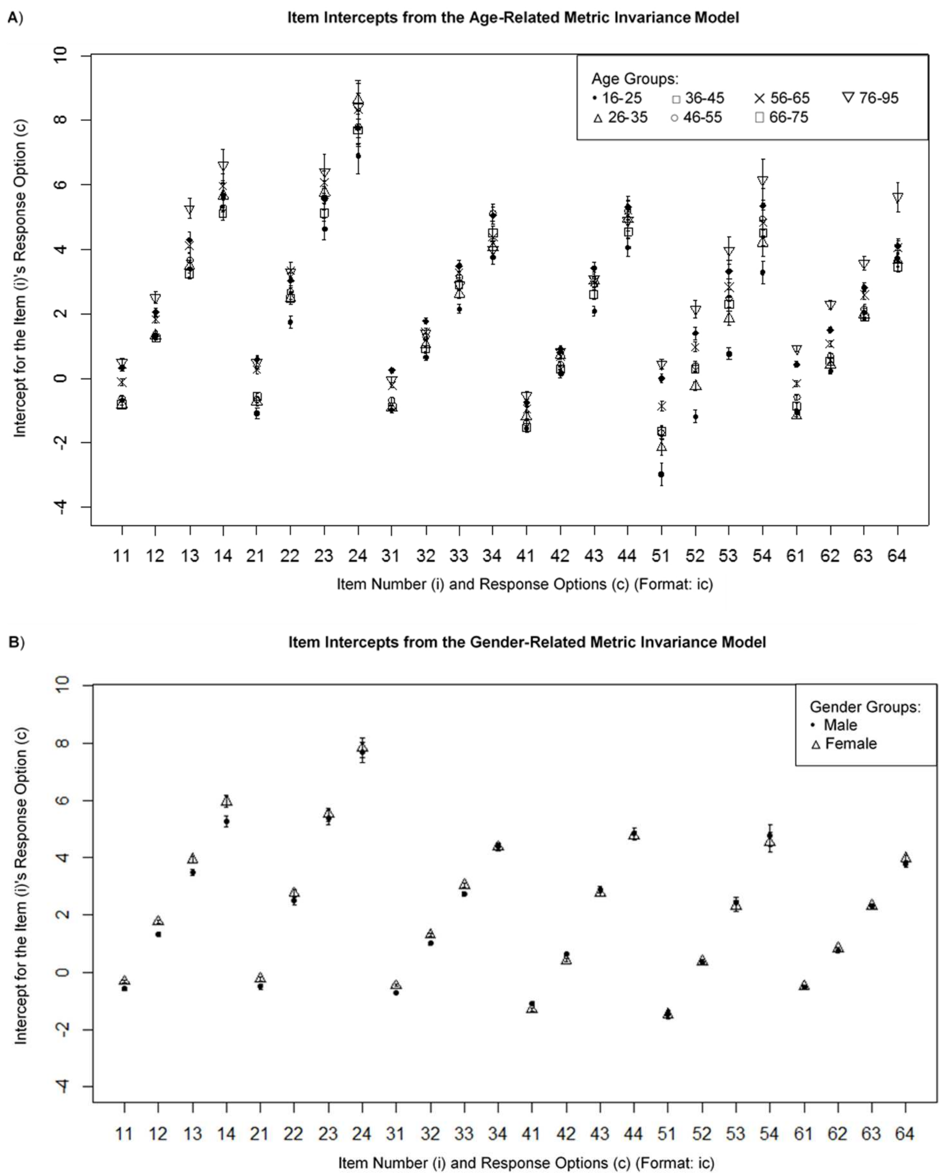

Figure 2. Item intercepts from the best-fitting multiple-groups model for both age and gender, in which factor loadings are invariant across groups while item intercepts vary across groups.

165x210mm (300 x 300 DPI)

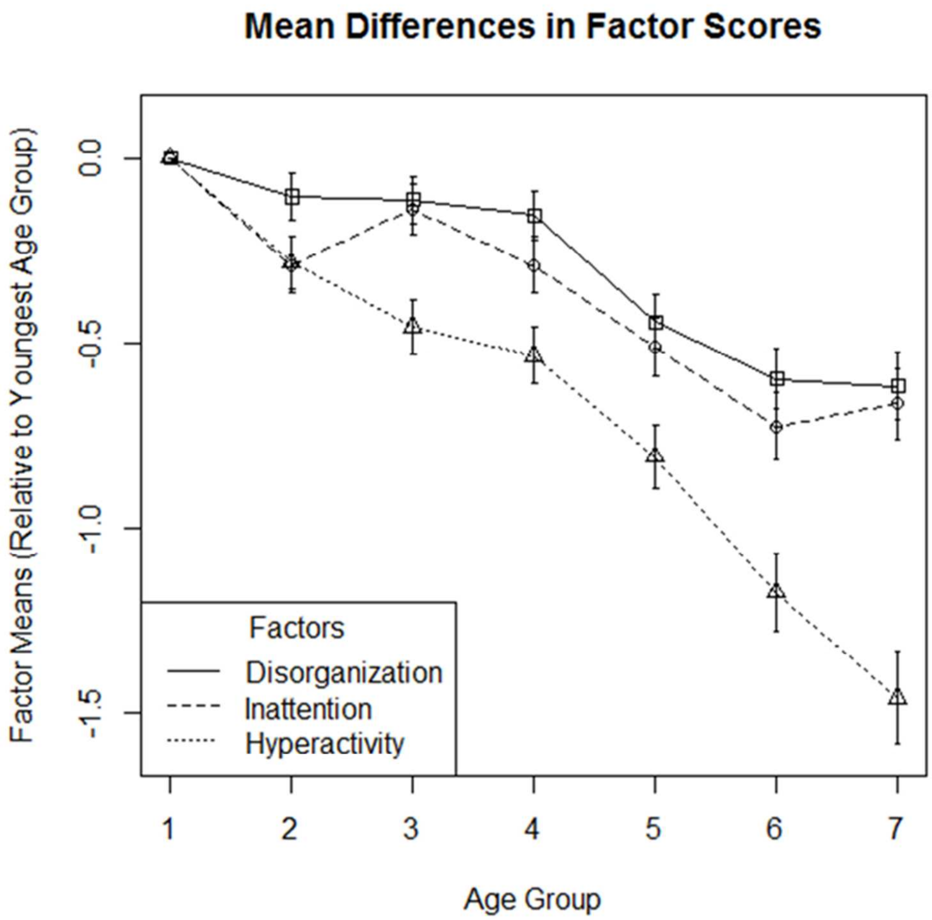

Figure 3. Age differences in latent ADHD factor means, when measurement variance in item intercepts are ignored.

76x76mm (300 x 300 DPI)