stats

MDPI

Article

Psychometric Properties of the Adult Self-Report: Data from over 11,000 American Adults

Michelle Guerrero [1,*], Matt Hoffmann [2] and Laura Pulkki-Råback [3,4]

1 Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON K1H 8L1, Canada
2 Department of Kinesiology, California State University, Fullerton, CA 92831, USA; matthoffmann@fullerton.edu
3 Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, 00100 Helsinki, Finland; laura.pulkki-raback@helsinki.fi
4 Department of Clinical Medicine, Division of Child Psychiatry, University of Turku, 20014 Turku, Finland
* Correspondence: mguerrero@cheo.on.ca

Received: 8 September 2020; Accepted: 25 October 2020; Published: 29 October 2020

Abstract: The first purpose of this study was to examine the factor structure of the Adult Self-Report (ASR) via traditional confirmatory factor analysis (CFA) and contemporary exploratory structural equation modeling (ESEM). The second purpose was to examine the measurement invariance of the ASR subscales across age groups. We used baseline data from the Adolescent Brain Cognitive Development study. ASR data from 11,773 participants were used to conduct the CFA and ESEM analyses and data from 11,678 participants were used to conduct measurement invariance testing. Fit indices supported both the CFA and ESEM solutions, with the ESEM solution yielding better fit indices. However, several items in the ESEM solution did not sufficiently load on their intended factors and/or cross-loaded on unintended factors. Results from the measurement invariance analysis suggested that the ASR subscales are robust and fully invariant across subgroups of adults formed on the basis of age (18–35 years vs. 36–59 years). Future research should seek to both CFA and ESEM to provide a more comprehensive assessment of the ASR.

Keywords: confirmatory structural equation modeling; exploratory structural equation modeling; measurement invariance; mental health

1. Introduction

The Achenbach System of Empirically Based Assessment (ASEBA) is a comprehensive evidence-based assessment system designed to measure adaptive and maladaptive functioning across individuals from age 1½ to over 90 years [1]. Quantitative data from ASEBA-inventories can be conceptualized into functioning scales, syndrome scales, substance use scales, *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*)-oriented scales, and critical item scales (items used primarily by clinicians). ASEBA-inventories have been translated in over 110 languages and referenced and used in more than 8000 publications. ASEBA inventories are widely used in various settings, including education, research, clinical, and healthcare (mental health and medical).

The Adult Self-Report (ASR) is one component of the ASEBA [2]. This self-report inventory comprises eight syndrome scales: *anxious/depressed*, *withdrawn*, *somatic complaints*, *thought problems*, *attention problems*, *aggressive behavior*, *rule-breaking behavior*, and *intrusiveness*. The syndrome scales can be examined independently, or aggregated into two broad-band scales, internalizing (i.e., *anxious/depressed*, *withdrawn*, and *somatic complaints*) and externalizing (i.e., *aggressive behavior*, *rule-breaking behavior*, and *intrusiveness*).

One gap within the ASR literature is that few studies have examined its factor structure [3,4]. The eight syndrome scales were initially supported via confirmatory factor analysis (CFA) during the development of the ASR [2]. Ivanova et al. [5] more recently used CFA to test the generalizability of the eight-syndrome model across 29 societies, such as Brazil, China, Russia, and the UK. Results showed that the correlated eight syndrome model converged for all 29 samples. Specifically, the Root Mean Square Error of Approximation (RMSEA) values demonstrated good model fit for all societies (<0.06). The Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) values indicated inadequate model fit for 13 of the 29 societies, with values falling below the recommended 0.90 threshold. Ivanova et al.'s work provides an up-to-date overview of the ASR's generalizability. However, their study did not include American adults. More research using data from the United States is needed to support and extend our knowledge of the ASR's psychometric properties. This is especially true as newer, more contemporary analytical techniques are advanced to test the psychometric properties of instruments.

While CFA has traditionally been used to assess the factorial validity of an instrument, there are criticisms of this method. Mainly, cross-loadings on non-targeted factors are assumed to be zero in CFA, which is considered overly restrictive in social sciences and idealistic [6,7]. When cross-loadings are constrained to zero, CFA will result in inflated factor correlations, which can ultimately undermine validity evidence (e.g., discriminant, predictive) of a multidimensional instrument [8]. Fortunately, more contemporary methods have been proposed in recent years to remedy the limitations of established analytic techniques such as CFA.

Exploratory structural equation modeling (ESEM) is one alternative to CFA [8]. ESEM is a flexible technique that integrates exploratory factor analysis, CFA, and structural equation modeling. The primary strength of ESEM is that it allows for minor cross-loadings, while simultaneously preserving the advantages of CFA (e.g., invariance testing) [8]. Allowing cross-loadings does not undermine constructs by adding 'noise' but rather allows them to be estimated using all of the relevant information [9]. ESEM studies comparing CFA and ESEM models across various domains have often reported more favorable results of the ESEM model [7,10,11].

Measurement invariance of the ASR subscales is another disparity in the literature. Studies that have examined whether ASR subscales function equivalently across groups have been primarily conducted with dyadic data (i.e., romantic couples) [4,12]. This research has generally shown that the ASR subscales are conceptualized differently (i.e., noninvariant) across self- and partner-reports as well as across males and females. Only one study has examined age measurement invariance, testing equivalency on one of the eight ASR syndrome scales (i.e., *thought problems*) [13]. Results showed that the *thought problems* subscale was invariant across three age groups (12–18 years, 19–27 years, and 28–59 years). Taken together, these findings can have significant implications for research and applied work.

Analyzing the factorial structure and measurement invariance of the ASR may extend the instrument's validity and provide further support for its psychometric properties in the American population. Therefore, the first purpose of the current study was to compare CFA and ESEM approaches. The second purpose was to investigate the invariance of the syndrome scales of the ASR across age groups. To the best of our knowledge, this is the first study to use ESEM with the ASR and one of the few to test for measurement invariance across age groups.

2. Materials and Methods

2.1. Data Source

We used data from the Adolescent Brain Cognitive Development (ABCD) study, a broadly representative ongoing 10-year longitudinal study examining psychological and neurobiological changes from preadolescence to early adulthood. Baseline assessments (2016 and 2018) included recruiting 11,875 children (9–10 years) and their parents/guardians from 21 sites across the United States. Baseline data revealed that the sample was relatively diverse on several socio-demographic factors: 59% of parents had a bachelor's or postgraduate degree; household income was <$50,000 for 30% of the

families and between $50,000–99,999 for 26% of the families; 52% of families identified as White, 15% as Black, 20% as Hispanic, 2% as Asian, and 11% as other (e.g., biracial). Complete recruitment details can be found elsewhere [14]. Study procedures received ethics board approval. Prior to participation, parents/guardians provided written informed consent and children provided assent.

2.2. ASR

Participants answered the 99 items corresponding to the ASR's eight syndrome scales [2]: *anxious/depressed* (18 items), *withdrawn* (9 items), *somatic complaints* (12 items), *thought problems* (10 items), *attention problems* (15 items), *aggressive behavior* (15 items), *rule-breaking behavior* (14 items), and *intrusiveness* (6 items). Items are rated on a 3-point scale: 0 = *not true*, 1 = *somewhat or sometimes true*, or 2 = *very true or often true*. Participants were instructed to answer each item based on the preceding 6 months.

2.3. Statistical Analysis

For the main analysis, we conducted CFA and ESEM using the latent variable modeling program Mplus Version 8.3. Data for these analyses were based on responses from 11,773 parents. A total of 102 participants were removed from the analyses because of incomplete data on ASR items (n = 5) or they exceeded the age threshold (i.e., >60 years; n = 97). We used the weighted least squares mean and variance-adjusted (WLSMV) estimator, which is suitable for ordered categorical (ordinal) variables wherein the assumption of normality is not assumed [15]. We used target rotation for the ESEM solution and set all unintended factor loadings (i.e., cross-loadings) to be close to zero. We used the following fit indices to evaluate the model fit of both the CFA and ESEM solutions: CFI, TLI, and RMSEA. Model fit was deemed acceptable if CFI and TLI values were ≥0.90 and RMSEA was ≤0.08 [16,17]. Model fit was considered good if CFI and TLI values were ≥0.95 and RMSEA was ≤0.06 [18]. We also reported the Chi-Square statistic (χ^2), but because of its sensitivity to large sample sizes [19], we did not use it to gauge model fit. Factor loadings for both the CFA and ESEM solutions were required to be ≥0.30 for items to be considered salient and used in defining constructs. Loadings >0.71 were considered excellent, >0.63 very good, >0.55 good, >0.45 fair, and ≥0.30 poor [20,21]. Whereas cross-loadings were automatically zero for the CFA solution, they were required to be <0.30 for the ESEM solution. Had both the CFA and ESEM solutions met all criteria for an adequate fitting model, our intention was to compare the two solutions by examining the model fit statistics for both models. However, as described in detail in the Results section, the ESEM solution demonstrated issues with some of the factor loadings. Consequently, we deemed the CFA solution to be the more favorable model.

Because we deemed the CFA solution as the more favorable model, we tested for measurement invariance by examining changes in model fit statistics using CFA. Data for the invariance testing were based on responses from 11,678 participants (M_{age} = 39.76, SD = 6.48). A total of 95 participants did not self-report their age and were therefore excluded from the analysis. We tested for configural, metric, and scalar invariance to determine whether the underlying factor structure of the ASR was the same across age groups (i.e., 18–35 years [n = 3094] vs. 36–59 years [n = 8584]). These age categories were selected in order to align with the normed scales for syndromes outlined in the ASEBA manual [2]. Configural invariance assesses the fit of the factor structure when there are no invariance constraints imposed across groups. Metric invariance assesses the invariance of factor loadings across groups. Scalar invariance assesses the invariance of both factor loadings and item intercepts across groups. We assessed metric invariance after configural invariance was established, and scalar invariance after metric invariance was established. Thus, this process involved testing a series of increasingly constrained models. There is support for a more constrained model when the CFI decreases by less than 0.010 and the RMSEA increases by less than 0.015 [22]. There is strong measurement invariance when all three tests demonstrate invariance across groups [23].

Last, we used standardized factor loadings to calculate composite reliability scores for the latent factors [24]. We assessed reliability using a composite reliability index. Given the persistent use of

alpha coefficients as a measure of reliability, alpha coefficients were also calculated, though it should be noted that alpha has been increasingly criticized as a measure of reliability [25,26].

3. Results

The results of the CFA indicated that the 8-factor model fit acceptably well based on CFI (0.919) and TLI (0.917) values and very well based on the RMSEA (0.026; 90% CI [0.026–0.026]), χ^2 (4724) = 42,010.26, $p < 0.001$. All standardized factor loadings for the CFA were significant ($p < 0.001$), ranging from 0.36 to 0.90 (see Table 1).

Table 1. Standardized factor loadings and standard errors for CFA and ESEM models.

	ESEM										
Items	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**CFA**	**ESEM S.E.**	**CFA S.E.**
F1. Anxious/depressed											
12. Lonely	**0.56**	0.19	0.11	−0.04	0.03	0.01	0.13	−0.01	**0.75**	0.02	0.01
13. Confused	0.29	0.06	0.23	0.09	0.33	0.03	−0.03	0.01	**0.78**	0.02	0.01
14. Cries	**0.46**	0.00	0.19	−0.07	−0.09	0.17	0.19	0.01	**0.69**	0.02	0.01
22. Worry future	**0.51**	−0.05	0.05	−0.07	0.00	0.07	0.21	−0.03	**0.58**	0.02	0.01
31. Fears doing bad	**0.42**	0.14	0.02	0.12	0.01	0.05	0.13	0.10	**0.67**	0.02	0.02
33. Unloved	**0.57**	**0.35**	0.16	−0.10	−0.11	0.04	0.14	0.08	**0.79**	0.02	0.01
34. Out to get him/her	**0.37**	0.26	0.14	0.08	−0.09	0.13	0.08	0.16	**0.74**	0.03	0.02
35. Worthless	**0.68**	0.19	0.04	−0.00	0.11	−0.02	0.03	0.03	**0.83**	0.02	0.01
45. Nervous	**0.66**	−0.12	0.06	0.16	0.08	0.16	−0.22	0.07	**0.76**	0.02	0.01
47. Lacks self-conf.	**0.64**	0.15	−0.05	−0.02	0.29	−0.02	−0.13	−0.06	**0.74**	0.02	0.01
50. Fearful	**0.75**	−0.16	0.05	0.17	0.07	0.11	−0.27	0.04	**0.75**	0.02	0.01
52. Too guilty	**0.63**	−0.09	0.01	0.06	0.13	0.08	−0.08	0.10	**0.70**	0.02	0.01
71. Self-conscious	**0.54**	0.09	−0.06	0.06	0.16	0.11	−0.17	−0.10	**0.63**	0.02	0.01
91. Thinks suicide	**0.56**	0.10	0.08	0.21	−0.10	−0.04	0.22	−0.02	**0.75**	0.04	0.02
103. Sad	**0.51**	0.07	0.15	0.05	0.06	0.16	0.12	−0.19	**0.85**	0.01	0.01
107. Can't succeed	**0.43**	0.14	0.02	−0.08	0.27	−0.01	0.14	−0.04	**0.70**	0.02	0.01
112. Worries	**0.68**	−0.18	0.07	0.01	−0.02	0.24	−0.01	−0.05	**0.73**	0.01	0.01
113. Worries opp. sex	**0.41**	0.10	0.01	0.01	0.00	0.03	0.32	0.03	**0.65**	0.02	0.01
F2. Withdrawn											
25. Not getting along	0.00	**0.50**	0.07	0.01	0.05	0.28	−0.04	0.11	**0.66**	0.02	0.01
30. Poor opp. sex	**0.30**	0.28	0.04	−0.02	−0.02	0.05	0.33	−0.02	**0.71**	0.02	0.01
42. Rather be alone	0.07	**0.40**	−0.01	0.28	0.12	0.15	−0.11	−0.24	**0.60**	0.02	0.01
48. Not liked	0.14	**0.55**	0.06	0.02	0.12	0.17	−0.06	0.12	**0.77**	0.02	0.01
60. Enjoys little	0.28	**0.34**	0.11	0.06	0.17	0.14	0.05	−0.16	**0.86**	0.02	0.01
65. Won't talk	0.07	**0.42**	0.04	0.18	0.09	0.20	0.09	−0.14	**0.76**	0.02	0.02
67. No friends	0.19	**0.57**	−0.05	0.10	0.22	0.10	−0.09	−0.07	**0.76**	0.02	0.01
69. Secretive	0.08	**0.31**	0.05	0.25	0.01	0.16	0.13	−0.16	**0.65**	0.02	0.01
111. Withdrawn	0.07	**0.43**	−0.03	0.20	0.13	0.12	0.00	−0.20	**0.61**	0.02	0.01
F3. Somatic complaints											
51. Dizzy	0.16	−0.02	**0.54**	0.11	0.10	0.00	−0.13	0.03	**0.70**	0.02	0.01
54. Feels tired	0.15	−0.02	**0.32**	0.15	0.31	0.11	−0.03	−0.23	**0.84**	0.02	0.01
56a. Aches	−0.06	−0.07	**0.66**	0.10	0.09	0.03	0.00	−0.02	**0.64**	0.01	0.01
56b. Headaches	0.04	−0.09	**0.65**	−0.01	0.01	0.07	−0.07	−0.06	**0.56**	0.01	0.01
56c. Nausea	0.04	−0.01	**0.87**	−0.03	0.01	−0.08	0.01	0.08	**0.75**	0.01	0.01
56d. Eye problems	−0.06	0.05	**0.48**	0.03	0.01	0.03	0.08	−0.00	**0.50**	0.02	0.02
56e. Skin problems	−0.02	0.02	**0.41**	0.04	0.09	0.02	−0.06	0.03	**0.45**	0.02	0.02
56f. Stomach aches	0.08	−0.04	**0.76**	−0.04	0.01	−0.04	−0.02	0.08	**0.67**	0.02	0.01
56g. Vomits	−0.06	0.13	**0.87**	−0.05	−0.09	−0.06	0.05	0.10	**0.69**	0.02	0.02
56h. Heart pounds	0.25	−0.07	**0.50**	0.14	0.03	0.01	−0.08	0.04	**0.73**	0.02	0.01
56i. Numbness	−0.02	−0.04	**0.60**	0.10	0.08	0.06	0.01	−0.02	**0.68**	0.02	0.01
100. Sleep problems	0.16	−0.02	0.29	0.12	0.09	0.15	−0.04	−0.11	**0.66**	0.02	0.01
F4. Thought problems											
9. Can't get mind off	**0.40**	−0.06	0.07	0.18	0.10	0.17	−0.01	0.05	**0.77**	0.02	0.01
18. Harms self	**0.52**	0.07	0.15	0.15	−0.06	−0.09	**0.32**	0.06	**0.83**	0.06	0.04
36. Gets hurts	−0.01	−0.01	0.28	0.19	0.27	0.02	−0.10	0.10	**0.55**	0.02	0.02
40. Hears things	0.04	0.26	**0.32**	**0.44**	−0.08	−0.01	0.18	−0.06	**0.79**	0.05	0.03
46. Twitches	0.16	0.01	0.28	0.29	0.11	0.06	−0.02	0.05	**0.72**	0.02	0.01
63. Pref. older people	−0.05	0.24	0.17	0.13	0.03	0.16	0.13	−0.08	**0.55**	0.02	0.01
66. Repeats acts	0.07	0.11	0.11	0.27	0.04	0.18	0.14	0.05	**0.70**	0.02	0.01
70. Sees things	0.00	0.25	**0.30**	**0.31**	−0.09	0.09	0.17	0.02	**0.73**	0.04	0.03
84. Strange behavior	−0.05	0.10	−0.03	**0.74**	0.06	0.00	0.15	0.15	**0.71**	0.02	0.01
85. Strange ideas	0.05	0.07	−0.04	**0.74**	0.07	−0.03	0.12	0.11	**0.74**	0.02	0.01

Table 1. *Cont.*

	ESEM										
Items	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**CFA**	**ESEM S.E.**	**CFA S.E.**
F5. Attention problems											
1. Forgetful	−0.07	0.01	0.19	−0.03	**0.63**	0.05	−0.04	0.03	**0.63**	0.01	0.01
8. Can't concentrate	0.05	−0.02	0.11	0.11	**0.55**	0.06	−0.06	0.06	**0.68**	0.01	0.01
11. Dependent	**0.39**	0.06	0.02	−0.06	**0.31**	−0.03	0.04	0.16	**0.66**	0.02	0.01
17. Daydreams	0.17	−0.01	0.05	0.24	0.22	−0.06	0.07	0.11	**0.54**	0.02	0.01
53. Can't plan	0.26	0.02	0.03	0.02	**0.46**	−0.02	0.27	−0.05	**0.81**	0.01	0.01
59. Fails to finish	0.02	0.10	−0.01	0.03	**0.71**	−0.03	0.09	0.02	**0.74**	0.01	0.01
61. Poor work	0.11	0.23	0.03	−0.02	**0.55**	−0.01	0.12	0.00	**0.77**	0.02	0.02
64. Can't prioritize	0.07	0.06	−0.07	−0.03	**0.74**	−0.03	0.17	0.03	**0.75**	0.01	0.01
78. Decisions	0.33	0.04	−0.06	−0.01	**0.53**	0.00	−0.07	−0.01	**0.69**	0.01	0.01
101. Avoids work	−0.06	0.09	0.03	0.11	0.24	0.07	**0.30**	−0.06	**0.53**	0.03	0.02
102. Lacks energy	0.11	0.00	**0.30**	0.09	**0.38**	0.13	−0.04	−0.28	**0.75**	0.01	0.01
105. Disorganized	−0.11	0.01	0.03	0.04	**0.76**	−0.07	0.14	0.14	**0.67**	0.02	0.01
108. Loses things	−0.13	0.01	0.17	0.00	**0.66**	0.08	0.03	0.06	**0.68**	0.01	0.01
119. Poor at details	−0.05	0.08	−0.04	−0.02	**0.59**	−0.01	0.13	0.07	**0.55**	0.02	0.01
121. Late	−0.08	−0.02	0.04	−0.08	**0.48**	0.02	0.14	0.10	**0.43**	0.02	0.01
F6. Aggressive behavior											
3. Argues	0.07	0.06	−0.05	−0.19	0.01	**0.68**	−0.02	0.18	**0.58**	0.02	0.01
5. Blames	0.28	0.21	−0.08	−0.16	0.18	0.18	−0.07	0.28	**0.53**	0.03	0.02
16. Mean	−0.09	0.28	0.01	−0.04	0.00	**0.47**	0.05	0.21	**0.58**	0.02	0.02
28. Badly with family	0.15	**0.34**	0.06	−0.07	0.03	0.27	0.08	0.04	**0.64**	0.02	0.01
37. Fights	−0.04	0.22	0.12	0.00	−0.03	**0.48**	0.10	0.22	**0.70**	0.04	0.03
55. Elation-depression	**0.30**	−0.01	0.18	0.16	−0.02	**0.36**	0.20	−0.14	**0.90**	0.02	0.01
57. Attacks	−0.12	0.16	0.18	0.10	−0.05	**0.55**	0.12	0.12	**0.74**	0.07	0.04
68. Screams	0.02	0.03	0.05	−0.22	−0.02	**0.77**	0.07	0.04	**0.66**	0.02	0.01
81. Behavior changes	0.03	0.03	−0.02	0.24	0.03	0.29	0.15	0.01	**0.56**	0.02	0.01
86. Stubborn	0.09	0.04	−0.01	0.17	0.01	**0.64**	−0.07	−0.08	**0.72**	0.02	0.01
87. Mood changes	0.15	−0.04	0.06	0.20	−0.05	**0.59**	0.17	−0.16	**0.83**	0.02	0.01
95. Temper	−0.03	−0.00	−0.03	−0.01	−0.03	**0.84**	0.00	0.05	**0.69**	0.02	0.01
97. Threatens	−0.03	0.17	0.12	0.18	−0.07	**0.41**	0.19	0.22	**0.78**	0.05	0.03
116. Upset	0.26	−0.04	−0.02	−0.06	0.03	**0.73**	−0.01	−0.04	**0.82**	0.02	0.01
118. Impatient	0.07	−0.10	−0.10	0.03	0.19	**0.64**	−0.01	0.03	**0.68**	0.02	0.01
F7. Rule-breaking behav.											
6. Uses drugs	0.10	−0.01	−0.06	0.23	−0.07	−0.05	**0.34**	0.17	**0.38**	0.03	0.03
20. Damages things	0.04	0.21	0.11	0.09	−0.06	**0.34**	0.13	0.14	**0.73**	0.05	0.03
23. Breaks rules	−0.02	0.16	−0.05	0.17	0.16	0.05	0.24	0.29	**0.59**	0.02	0.02
26. Lacks guilt	−0.10	0.17	−0.01	0.05	0.01	0.14	0.24	0.09	**0.37**	0.02	0.02
39. Bad companions	−0.08	0.11	0.11	0.16	0.07	0.15	**0.32**	0.19	**0.67**	0.04	0.03
41. Impulsive	0.02	0.03	0.03	0.14	0.23	0.34	0.13	0.13	**0.82**	0.02	0.01
43. Lies, cheats	0.04	0.17	−0.01	0.08	0.10	0.11	0.28	0.22	**0.63**	0.03	0.02
76. Irresponsible	0.13	0.04	−0.01	0.21	0.25	0.06	**0.32**	0.19	**0.84**	0.02	0.02
82. Steals	0.12	0.06	−0.04	0.25	0.04	0.03	**0.32**	0.22	**0.64**	0.04	0.03
90. Gets drunk	0.08	−0.09	−0.12	0.21	0.06	0.04	0.21	0.14	**0.36**	0.03	0.02
92. Trouble with law	0.03	0.05	−0.07	**0.32**	0.02	0.08	**0.43**	0.18	**0.69**	0.03	0.02
114. Doesn't pay debts	0.03	−0.18	0.08	−0.09	0.23	0.11	**0.77**	−0.22	**0.71**	0.02	0.01
117. Money troubles	0.02	−0.22	0.02	−0.07	**0.35**	0.14	**0.70**	−0.16	**0.74**	0.02	0.01
122. Can't keep job	0.04	0.20	0.03	0.01	0.20	0.14	**0.33**	−0.04	**0.72**	0.03	0.02
F8. Intrusiveness											
7. Brags	0.04	0.05	0.01	0.08	0.03	0.05	0.04	**0.51**	**0.57**	0.02	0.02
19. Demands attention	0.18	−0.06	0.03	0.00	0.07	−0.01	0.16	**0.59**	**0.70**	0.02	0.02
74. Shows off	−0.05	−0.07	0.05	0.20	0.08	0.01	0.12	**0.54**	**0.60**	0.02	0.02
93. Talks too much	0.03	**−0.31**	0.11	0.00	0.13	0.25	0.07	**0.48**	**0.69**	0.02	0.01
94. Teases	−0.06	−0.01	−0.05	0.15	0.16	0.23	−0.05	**0.47**	**0.70**	0.02	0.02
104. Loud	−0.07	−0.24	0.07	0.02	0.06	0.41	0.12	**0.44**	**0.76**	0.02	0.02

Note. S.E. = standard error. Factor loadings that are larger than 0.30 are shown in boldface. All CFA factor loadings are significant ($p < 0.001$). All ESEM values loading on their intended factors are significant ($p < 0.01$). Item numbers correspond with those in the ASEBA manual [2].

The 8-factor model using ESEM had very good model fit based on the CFI (0.970), TLI (0.965), and RMSEA (0.017; 90% CI [0.017–0.017]), χ^2 (4087) = 17,706.41, $p < 0.001$. However, not all items loaded cleanly on their intended factors. The *anxious/depressed* factor had one intended item that loaded <0.30, and three intended items that cross-loaded ≥0.30 on unintended factors. The *withdrawn* factor had two intended items that loaded <0.30, and one intended item that cross-loaded >0.30 on an unintended

factor. The *somatic complaints* factor had one intended item that loaded <0.30, yet no intended items that cross-loaded ≥0.30 on unintended factors. The *thought problems* factor was particularly problematic; seven of ten intended items loaded <0.30, and three intended items cross-loaded ≥0.30 on unintended factors. The *attention problems* factor had three intended items that loaded <0.30, and two intended items that cross-loaded >0.30 on an unintended factor. The *aggressive behavior* factor had three intended items that loaded <0.30, and one intended item that cross-loaded >0.30 on an unintended factor. The *rule-breaking behavior* factor was quite problematic; six of fourteen intended items loaded <0.30, and four intended items cross-loaded ≥0.30 on unintended factors. All intended items for the *intrusiveness* factor loaded >0.30, and one intended item cross-loaded >0.30 on an unintended factor. Important to note is that in some cases, secondary loadings were larger than the primary loadings. Factor loadings (including cross-loadings) for the ESEM solution are reported in Table 1.

The results of measurement invariance testing using CFA demonstrated evidence of acceptable model fit. All changes in fit indices pertaining to configural, metric, and scalar invariance were well within acceptable ranges [22]. Consequently, we found support for strong measurement invariance across age groups (i.e., 18–35 years vs. 36–59 years). These results are summarized in Table 2.

Table 2. Measurement invariance testing across age groups (18–35 years vs. 36–59 years).

	CFI	TLI	RMSEA	CM	ΔCFI	ΔTLI	ΔRMSEA
1a. Configural	0.926	0.924	0.024	-	-	-	-
1b. Metric	0.929	0.927	0.023	1a	+0.003	+0.003	−0.001
1c. Scalar	0.931	0.931	0.023	1b	+0.002	+0.004	0.000

Note. CM = Comparison model.

Descriptive statistics, as well as latent factor correlations (polychoric) and internal consistencies for the CFA and ESEM solutions, are reported in Table 3. Reliability scores for the CFA solution were good (0.83 to 0.95 for composite reliability and 0.82 to 0.95 for alpha coefficient) and scores for the ESEM solution ranged from poor to good (0.59 to 0.88).

Table 3. Descriptive statistics, and latent factor correlations and composite reliability for the CFA and ESEM models.

Variables	F1	F2	F3	F4	F5	F6	F7	F8
F1. Anxious/depressed	-	0.35	0.47	0.37	0.54	0.51	0.32	0.01 *
F2. Withdrawn	0.80	-	0.30	0.29	0.20	0.37	0.38	0.17
F3. Somatic complaints	0.72	0.64	-	0.39	0.38	0.49	0.32	0.04
F4. Thought problems	0.76	0.78	0.70	-	0.34	0.43	0.35	0.20
F5. Attention problems	0.79	0.69	0.68	0.72	-	0.48	0.30	0.13
F6. Aggressive behavior	0.78	0.77	0.65	0.75	0.69	-	0.39	0.27
F7. Rule-breaking behavior	0.66	0.70	0.56	0.73	0.75	0.75	-	0.23
F8. Intrusiveness	0.37	0.35	0.35	0.54	0.48	0.59	0.56	-
M	5.04	1.56	2.94	1.44	4.64	3.36	1.19	1.00
SE	0.05	0.02	0.03	0.02	0.04	0.03	0.02	0.01
Composite reliability (CFA)	0.95	0.90	0.90	0.91	0.92	0.93	0.91	0.83
Alpha coefficient (CFA)	0.95	0.89	0.90	0.89	0.92	0.93	0.90	0.82
Composite reliability (ESEM)	0.88	0.66	0.87	0.59	0.85	0.86	0.66	0.68

Note. CFA and ESEM correlations are presented below and above the diagonal, respectively. * Correlation is non-significant ($p > 0.05$); all remaining ESEM correlations are significant ($p < 0.01$). All CFA correlations are significant ($p < 0.001$).

4. Discussion

Scholars have argued that the classic CFA models may not be appropriate for many well-established psychological measures. ESEM is a new and evolving methodological approach that has proven to overcome the typical shortcomings of CFA. The purpose of the current study was to compare the factor

structure of the ASR using both CFA and ESEM and to test for age measurement invariance. Results of the current study showed that both measurement models met conventional model fit criteria, with the ESEM solution showing an *improvement* in fit over CFA but demonstrating poor factor loadings on intended factors and substantial cross-loading on unintended factors. Furthermore, invariance testing revealed that the ASR is robust and invariant across two age groups.

The ESEM solution showed an improvement in model fit statistics over the CFA solution. This finding is not surprising given that ESEM yields better fit in the presence of nontrivial cross-loadings [27,28]. However, several (n = 26) items in the ESEM solution did not sufficiently load on their intended factors (<0.30) and/or cross-loaded on unintended factors (≥0.30). There were nine instances wherein the cross-loading on the unintended factor was greater than the factor loading on the intended factor. Two factors that were particularly problematic were *thought problems* and *rule-breaking behavior*, with more than half of the items (13 of 24) being problematic. Moreover, factor loadings on intended factors were substantially lower in the ESEM solution compared to the CFA solution. The strength of factor loadings on intended factors in the ESEM model generally ranged from 'poor' to 'fair' and very few were considered 'excellent', whereas the strength of factor loadings in the CFA solution typically hovered around 'good' and 'excellent'. Overall, results of the ESEM solution revealed that: (a) factor loadings on intended factors were generally weak; (b) factor loadings did not sufficiently load on intended factors; (c) there were substantial cross-loadings on unintended factors; (d) in some cases, secondary loadings were larger than the primary loadings.

Many studies examining the factorial validity of psychological measures seek to compare results from traditional CFA to the newer method of ESEM. In these studies, the conclusion that the ESEM solution was superior to the CFA solution is more common than not [11]. However, in some cases [29], and certainly in the current study, ESEM may not always provide the best overall solution—a notion acknowledged within the statistical literature [6]. Booth and Hughes [29] used ESEM as an alternative to CFA and found little evidence supporting the application of ESEM over CFA. They acknowledged that there are several nuances that must be considered when selecting between the two analytic strategies. For example, they discuss how in ESEM the inclusion of all possible parameters ensures that no substantive parameters are omitted, but that this approach may also be interpreted as overly inclusive and contradictory to the principle of parsimony. In their study, additional parameters included in the ESEM solution added very little to the definition of latent constructs. Highlighting the advantages and disadvantages of both CFA and ESEM, Booth and Hughes illustrate how CFA and ESEM can be used together to improve measurement of constructs.

The results of measurement invariance analysis suggested that the ASR is robust and fully invariant across subgroups of adults formed on the basis of age (18–35 years vs. 36–59 years). That is, the factor structure was the same regardless of the adults' age (configural invariance), the strength of the factor loadings were similar across age (metric invariance), and no significant age differences were noted in the intercepts (full scalar invariance). This finding is consistent with previous research that showed the *thought problems* subscale was invariant across three age cohorts (12–18 years, 19–27 years, and 28–59 years) [13]. Results of the current study are encouraging and confirm that the ASR's psychometric properties are robust across adulthood.

The results of this study should be considered in light of some limitations. Results provide evidence of the ASR's factorial validity only. Subsequent research should consider examining other types of construct validity, such as criterion or concurrent. Furthermore, we were unable to test for gender invariance because data on the participants' gender were unavailable. Existing research with dyadic data (romantic couples) has shown that the ASR subscales are noninvariant across females and males [12]. Specifically, five subscales (*withdrawn*, *attention problems*, *aggressive behavior*, *rule-breaking behavior*, and *intrusiveness*) met full metric invariance, two subscales (*anxious/depressed* and *somatic complaints*) met partial metric invariance, and all subscales met partial scalar invariance [12]. Future research should seek to explore the measurement invariance of the ASR across other factors (e.g., socio-demographic status) as well as across multiple time points as more ABCD data is released.

The results of the current study should be carefully considered in future studies using the ASR. Results of the ESEM solution uncovered important findings regarding factor loadings on intended and unintended factors, especially for the *thought problems* and *rule-breaking behavior* syndromes. The ESEM results therefore may raise concern for some regarding the sole reliance on CFA to examine the factor structure of the ASR. It is hoped that future researchers use CFA and ESEM in harmony rather than in isolation. Previous works examining the factor structure of psychological measures have successfully illustrated how ESEM and CFA can be used to complement each other [28,29]. Future research should also explore the psychometric properties of the ASR using other modeling approaches, such as partial least squares SEM, bifactor models that incorporate a general factor encompassing all syndromes [30] and Bayesian structural equation modeling (BSEM) [31]. The latter approach is similar to ESEM in that it allows cross-loadings. A key difference between ESEM and BSEM is that prior knowledge (referred to as priors) can inform the statistical model via statements regarding model parameters (e.g., mean, path coefficient). Thus, the model is tested against a set of known parameters rather than against a null hypothesis. An obvious challenge for researchers seeking to use BSEM is the choice of the priors, since using improper priors can have a negative impact on estimates [31]. Another advantage of BSEM is that it performs well with small sample sizes.

5. Conclusions

Findings showed that both the eight-syndrome CFA and ESEM solutions met model fit criteria.

As anticipated and in line with previous research, the ESEM solution revealed an improvement in fit statistics and lower factor correlations than the CFA solution; however, examination of the item factor loadings revealed that several items had poor factor loadings on intended factors and/or significant cross-loadings on unintended factors. Future research should seek to use both CFA and ESEM to provide a more comprehensive assessment of the ASR. Measurement invariance testing revealed that the ASR subscales are equivalent for younger and older adults. To the best of our knowledge, this study was the first to explore the factor structure of the ASR using ESEM and the first to test measurement invariance of the individual items from all ASR subscales across age groups.

Author Contributions: Conceptualization, M.G. and L.P.-R.; methodology, M.G. and L.P.-R.; software, M.H.; validation, M.H.; formal analysis, M.G. and M.H.; writing—original draft preparation, M.G. and M.H.; writing—review and editing, M.H. and L.P.-R.; visualization, M.H.; supervision, L.P.-R. All authors have read and agreed to the published version of the manuscript.

Funding: The Jenny and Antti Wihuri Foundation and The Yrjo Jahnsson Foundation, Finland to L.P.-R.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Achenbach, T.M. As Others See Us Clinical and Research Implications of Cross-Informant Correlations for Psychopathology. *Curr. Dir. Psychol. Sci.* **2006**, *15*, 94–98. [CrossRef]
2. Achenbach, T.M.; Rescorla, L.A. *Manual for the ASEBA Adult Forms & Profiles*; University of Vermont Research Center for Children, Youth, & Families: Burlington, VT, USA, 2003.
3. Michelini, G.; Barch, D.M.; Tian, Y.; Watson, D.; Klein, D.N.; Kotov, R. Delineating and Validating Higher-Order Dimensions of Psychopathology in the Adolescent Brain Cognitive Development (ABCD) Study. *Transl. Psychiatry* **2019**, *9*, 21–25. [CrossRef] [PubMed]
4. Claxton, S.E.; Deluca, H.K.; Van Dulmen, M.H.M. Testing Psychometric Properties in Dyadic Data Using Confirmatory Factor Analysis: Current Practices and Recommendations. *Test. Psychom. Methodol. Appl. Psychol.* **2015**, *22*, 181–198.
5. Ivanova, M.Y.; Achenbach, T.M.; Rescorla, L.A.; Turner, L.V.; Ahmeti-Pronaj, A.; Au, A.; Maese, C.A.; Bellina, M.; Caldas, J.C.; Chen, Y.C.; et al. Syndromes of Self-Reported Psychopathology for Ages 18–59 in 29 Societies. *J. Psychopathol. Behav. Assess.* **2015**, *37*, 171–183. [CrossRef] [PubMed]

6. Marsh, H.W.; Morin, A.J.S.; Parker, P.D.; Kaur, G. Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. *Annu. Rev. Clin. Psychol.* **2014**, *10*, 85–110. [CrossRef] [PubMed]
7. Marsh, H.W.; Liem, G.A.D.; Martin, A.J.; Morin, A.J.S.; Nagengast, B. Methodological Measurement Fruitfulness of Exploratory Structural Equation Modeling (ESEM): New Approaches to Key Substantive Issues in Motivation and Engagement. *J. Psychoeduc. Assess.* **2011**, *29*, 322–346. [CrossRef]
8. Asparouhov, T.; Muthén, B. Exploratory Structural Equation Modeling. *Struct. Equ. Model.* **2009**, *16*, 397–438. [CrossRef]
9. Asparouhov, T.; Muthén, B.; Morin, A.J.S. Bayesian Structural Equation Modeling with Cross-Loadings and Residual Covariances. *J. Manag.* **2015**, *41*, 1561–1577. [CrossRef]
10. Hoffmann, M.D.; Loughead, T.M. Preliminary Development of a Questionnaire to Assess Peer Athlete Mentoring Functions: The Athlete Mentoring Questionnaire (AMQ). *Meas. Phys. Educ. Exerc. Sci.* **2019**, *23*, 10–25. [CrossRef]
11. Joshanloo, M.; Jose, P.E.; Kielpikowski, M. The Value of Exploratory Structural Equation Modeling in Identifying Factor Overlap in the Mental Health Continuum-Short Form (MHC-SF): A Study with a New Zealand Sample. *J. Happiness Stud.* **2017**, *18*, 1061–1074. [CrossRef]
12. DeLuca, H.K.; Sorgente, A.; Van Dulmen, M.H.M. Dyadic Invariance of the Adult Self-Report and Adult Behavior Checklist: Evidence from Young Adult Romantic Couples. *Psychol. Assess.* **2019**, *31*, 192–209. [CrossRef]
13. Abdellaoui, A.; De Moor, M.H.M.; Geels, L.M.; Van Beek, J.H.D.A.; Willemsen, G.; Boomsma, D.I. Thought Problems from Adolescence to Adulthood: Measurement Invariance and Longitudinal Heritability. *Behav. Genet.* **2012**, *42*, 19–29. [CrossRef] [PubMed]
14. Garavan, H.; Bartsch, H.; Conway, K.; Decastro, A.; Goldstein, R.Z.; Heeringa, S.; Jernigan, T.; Potter, A.; Thompson, W.; Zahs, D. Recruiting the ABCD Sample: Design Considerations and Procedures. *Dev. Cogn. Neurosci.* **2018**, *32*, 16–22. [CrossRef]
15. Li, C.H. Confirmatory Factor Analysis with Ordinal Data: Comparing Robust Maximum Likelihood and Diagonally Weighted Least Squares. *Behav. Res. Methods* **2016**, *48*, 936–949. [CrossRef]
16. Brown, T. *Confirmatory Factor Analysis for Applied Research*, 2nd ed.; Guilford Press: New York, NY, USA, 2015.
17. Browne, M.W.; Cudeck, R. Alternative Ways of Assessing Model Fit. *Sociol. Methods Res.* **1992**, *21*, 230–258. [CrossRef]
18. Hu, L.T.; Bentler, P.M. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Struct. Equ. Model.* **1999**, *6*, 1–55. [CrossRef]
19. Kline, R.B. *Principles and Practice of Structural Equation Modeling*, 3rd ed.; Guilford Press: New York, NY, USA, 2011.
20. Comrey, A.L.; Lee, H.B. *A First Course in Factor Analysis*, 2nd ed.; Erlbaum: Hillsdale, NJ, USA, 1992.
21. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics*, 6th ed.; Pearson: Boston, MA, USA, 2012.
22. Chen, F.F. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Struct. Equ. Model.* **2007**, *14*, 464–504. [CrossRef]
23. Marsh, H.W.; Parker, P.D.; Morin, A.J.S. Invariance Testing across Samples and Time: Cohort-Sequence Analysis of Perceived Body Composition. In *An Introduction to Intermediate and Advanced Statistical Analyses for Sport and Exercise Scientists*; Ntoumanis, N., Myers, N., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2016; pp. 101–130.
24. Raykov, T. Estimation of Composite Reliability for Congeneric Measures. *Appl. Psychol. Meas.* **1997**, *21*, 173–184. [CrossRef]
25. McNeish, D. Thanks Coefficient Alpha, We'll Take It From Here. *Psychol. Methods* **2017**. [CrossRef]
26. Martin, J.J.; Beasley, V.L.; Guerrero, M.D. Sport Psychology Research: Proper Standards and Limitations. In *APA Handbooks in Psychology Series: APA Handbook of Sport and Exercise Psychology*; Ansel, M., Petrie, T., Steinfeldt, J., Eds.; American Psychological Association: Washington, DC, USA, 2019; Volume 1, pp. 17–40.
27. Marsh, H.W.; Nagengast, B.; Morin, A.J.S.; Parada, R.H.; Craven, R.G.; Hamilton, L.R. Construct Validity of the Multidimensional Structure of Bullying and Victimization: An Application of Exploratory Structural Equation Modeling. *J. Educ. Psychol.* **2011**, *103*, 701–732. [CrossRef]
28. Joshanloo, M. Factorial/Discriminant Validity and Longitudinal Measurement Invariance of MHC-SF in Korean Young Adults. *Curr. Psychol.* **2020**, *39*, 51–57. [CrossRef]

29. Booth, T.; Hughes, D.J. Exploratory Structural Equation Modeling of Personality Data. *Assessment* **2014**, *21*, 260–271. [CrossRef]
30. Chen, F.F.; Hayes, A.; Carver, C.S.; Laurenceau, J.P.; Zhang, Z. Modeling General and Specific Variance in Multifaceted Constructs: A Comparison of the Bifactor Model to Other Approaches. *J. Pers.* **2012**, *80*, 219–251. [CrossRef] [PubMed]
31. Muthén, B.; Asparouhov, T. Bayesian Structural Equation Modeling: A More Flexible Representation of Substantive Theory. *Psychol. Methods* **2012**, *17*, 313–335. [CrossRef] [PubMed]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).