ECRONICON OPEN ACCESS

PSYCHOLOGY AND PSYCHIATRY
Research Article

# Preliminary Development of Maladaptive Variants of Five-Factor Model Personality Dimensions Derived from the Item Pool of the ASEBA Youth-Self Report

**Robert A Semel***

*Licensed Psychologist, Brooklyn, New York, USA*

***Corresponding Author:** Robert A Semel, Licensed Psychologist, Brooklyn, New York, USA.

**Received:** August 04, 2020; **Published:** August 31, 2020

## Abstract

This study represents the first known attempt to operationalize maladaptive variants of the Five-Factor Model of personality (FFM) by indexing items from the ASEBA Youth Self-Report (YSR) using the normative YSR data set (N = 1057). Proxy scales of personality pathology from the YSR might serve to complement the traditional syndrome scales for research purposes. The existence of large archival databases of Youth Self-Reports could provide opportunities for both retrospective as well as concurrent and prospective study of personality in youth through use of a valid personality measure derived from the Youth Self-Report. Four trait domain scales (excluding Openness) were developed in this study in accordance with the constructs of the "Big Five" model or FFM, as well as the DSM-5 Alternative Model of Personality Disorders (AMPD). The fifth trait scale, labeled "Atypical Experiences", was intended to be aligned with the Psychoticism trait domain of the AMPD. The five preliminary scales that were developed ranged from good to acceptable internal consistency reliability. Exploratory Factor Analyses (subset sample n = 377) were followed by Confirmatory Factor Analyses (subset sample n = 700). Four of five CFA models required re-specification to achieve good model fit. Four models demonstrated full metric invariance by gender; the fifth CFA model demonstrated partial metric invariance. Only one CFA model demonstrated full scalar invariance. The proxy FFM scales were analyzed in an independent small sample data set from adjudicated juvenile delinquents (n = 77) who were undergoing psychological evaluation for court. Preliminary convergent and discriminant validity were suggested for several FFM scales. In comparison to the normative sample, the juvenile delinquent sample obtained significantly higher scores on Antagonism and Detachment and a significantly lower score on Neuroticism. Methodological limitations are discussed, along with suggestions for further research in this area.

***Keywords:*** *Five Factor Model of Personality; Maladaptive Personality; Youth Self-Report; DSM-5*

## Introduction

The study of childhood personality and of personality disorders (PDs) in youth has greatly expanded over the past two decades [1-3]. Research in the area of child personality has been augmented in part through findings that early-emerging child temperament traits can be interlinked with early-emerging personality traits consistent with adult models [4,5], especially regarding differences between children in their experiences and manifestations of positive and negative emotions and their abilities to regulate their emotions and behavior [6]. Normal and maladaptive personality traits and PDs can be identified during childhood and adolescence [3,6-8]. Prevalence of PDs in adolescents is sobering and demands serious attention from both a clinical and research perspective. PDs in community samples of

adolescents range from 6% to 17% [9], with estimates ranging between 41% to 64% in adolescent psychiatric samples [10,11] and with similar and even higher ranges in juvenile justice samples [3,12]. Peak prevalence for PDs occurs during early and middle adolescence [6,13]. PDs in adolescents are associated with increased risk of significant maladaptive outcomes, such as self-harm, substance abuse, harm to others, high-risk sexual behaviors, conflictual interpersonal relationships with family, friends, or romantic partners [7]. Normal and maladaptive personality traits reflect both change and continuity across the lifespan, and functional impairments associated with PDs have substantial duration over lifetime [3]. Therefore, early identification of maladaptive personality traits may facilitate interventions in earlier stages of life [14]. The current study aims to conduct a preliminary and exploratory examination whether a well-validated and standardized multi-informant measure of emotional and behavioral problems in youth, i.e. the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach and Rescorla, 2001) [15] might be utilized for research purposes (and possibly to screen for and) to identify potential personality problems in youth based on the Five-Factor Model of personality (FFM). Preliminary scales to assess maladaptive variants of the FFM were operationalized in this study using items from the ASEBA Youth Self-Report (YSR).

## Five-factor model of personality and its application to children and adolescents

The Five-Factor Model (FFM) of personality, sometimes referred to as the "Big Five" model, represents the current predominant conceptual and measurement model of personality in adults. This model conceptualizes five higher-order domains or dimensions of personality, i.e. Extraversion (E), Neuroticism (N), Conscientiousness (C), Agreeableness (A) and Openness to Experience (O), which are further delineated by lower-order facets within each of these domains. Ample empirical evidence exists pointing to the interrelationships between ("Axis I") clinical disorders, and ("Axis II") personality disorders [16,17], with earlier internalizing (e.g. depressive and anxiety disorders) and externalizing (disruptive behavior) disorders predicting elevated risk for later emergence and continuation of PDs into adulthood; conversely, earlier onset PDs predict greater risk for other adult psychiatric disorders [16]. There is also evidence of associations between symptoms of clinical disorders and personality traits as operationalized by the FFM [18] and studies have yielded empirical support for the position that children's and adolescent's personality traits can be reliably assessed using FFM measures, and that their personality traits are meaningfully associated with their behavior. In a study by Tackett., *et al*. [8] of children from five countries and four age groups, the FFM was found to be salient beginning in early childhood, with more similarities, overall, than differences across country and age. A striking finding in that study was that very low (A), or antagonism, and N (Neuroticism) were linked throughout the age groups, reflected by facets such as easily angry, irritable, quick tempered, mean, rude, aggressive.

## DSM-5 alternative dimensional model of personality disorders

A dimensional model of personality disorders is associated with the perspective that PDs represent maladaptive variants of normal personality traits as opposed to discrete categories of pathological personality traits [19-21]. The dimensional model assumes that extreme levels of traits may be present across different PDs and that diagnostic heterogeneity likely reflects a mixture of pathological traits in persons within a PD category [6]. The Diagnostic and Statistical Manual of Mental Disorders, 5$^{th}$ edition (DSM-5; APA, 2013) [22] includes an Alternative Model for Personality Disorders (AMPD) (in Section III - "Emerging Measures and Models") in which personality disorders are characterized by both impairments in personality functioning and pathological personality traits. Impairment in personality functioning is assessed by ratings of (levels of) impairment in self- (identity, self-direction) and interpersonal (empathy, intimacy) functioning. Pathological personality traits are organized into five broad domains of Negative Affectivity, Detachment, Antagonism, Disinhibition, and Psychoticism. The five domains map onto pathological variants of the Big Five trait domains, i.e. Emotional Stability, Extraversion, Agreeableness, Conscientiousness, and Lucidity, respectively. Within the broad PD domains are 25 trait facets. These traits and facets have been operationalized in a self-report measure, i.e. the Personality Inventory for DSM-5 (PID-5) [23]. Studies using the PID-5 in adolescents has thus far been sparse and further research of the PID-5 in younger age groups is needed.

## Use of the ASEBA to study personality disorder traits in youth

Given the interest in studying personality disorders or PD traits or pathology, and their precursors in children and youth, it is not surprising that researchers have studied the usefulness of the ASEBA for assessment of personality dysfunction. The ASEBA is one of the most widely researched and validated multi-informant behavioral rating scales that assesses a broad spectrum of problems as well as positive functioning in children and adolescents from the perspective of parents (Child Behavior Checklist; CBCL), teachers (Teacher Report Form; TRF), and from youth ages 11 to 18 (Youth Self-Report; YSR). Many studies have found significant associations between personality disorders or traits in adolescents and psychopathology as measured by the YSR scales. For example, Penney., *et al*. [24] found strong correlations between the Psychopathy Content Scale [25] and the YSR Delinquency (Rule-Breaking) scale. In a study of a large sample of incarcerated male and female adolescents, Kaszynski., *et al*. [12] found that almost 92% of the youths met criteria (based on structured diagnostic interviews) for at least one PD, and almost 25% met criteria for more than one PD. YSR Aggression and Delinquency mean scores were significantly higher in youths diagnosed with Antisocial, Borderline, Narcissistic and Schizotypal personality disorders.

Vrouva's [26] comprehensive review of the extant literature found that most studies concerning the usefulness of the ASEBA for assessing PD traits in youth focused on Antisocial Personality Disorder (ASPD)/psychopathy and Borderline Personality Disorder. Vrouva found that ASEBA scales demonstrated evidence of criterion validity in many studies and that extreme scores on ASEBA syndrome scales might alert clinicians to further assess for possible PD traits; however, ASEBA scales were not found to have adequate validity to assess for PD traits or emerging PD's.

Given the contemporary prominence of the FFM and the perspective that PDs represent maladaptive variants of the FFM, it might be fruitful to examine whether ASEBA items might be aggregated to form scales that are representative of the FFM and particularly maladaptive variants of the FFM. The remainder of this paper focuses on preliminary development of proxy scales from the ASEBA YSR that can represent valid measures of pathological variants of the Big Five or FFM trait domains. The current study is the first known study to explore the potential development of maladaptive variants of FFM dimensions based on the Youth Self-Report. The existence of large archival databases of ASEBA Youth Self-Reports, both in the U.S. and globally, could provide opportunities for both retrospective as well as concurrent and prospective study of personality in youth through use of a valid personality measure derived from the Youth Self-Report. Development of FFM scales derived from the YSR item pool would not be intended to substitute for the official and validated use of the ASEBA Youth Self-Report (in conjunction with parent and/or teacher forms) to assess for clinical problems, nor to substitute for other measures or interview-based assessments of personality problems in youth, and it is not at all clear that such FFM scales could be used for individual decision-making. Rather, FFM scales might serve to complement the traditional syndrome scales for research purposes and the potential of these scales to assess for screening or individual decision-making purposes would need to be determined through greater study. The preliminary ASEBA personality scales developed here from the normative sample of youth obtained by Achenbach and Rescorla [15] were examined for internal reliability, internal structure of the domain scales that were developed, similarities and differences in comparison with a smaller sample of adjudicated delinquent youth and preliminary convergent validity between the YSR-FFM scales and a separate personality measure administered concurrently in the sample of adjudicated delinquent youth.

## Method

### Participants

The normative sample for the YSR[1] included in this study consisted of 1057 non-referred youths (551 males, 506 females) recruited in a 1999 National Survey. Mean age was 14.2 (SD = 2.2). Regarding ethnicity, 60% of the youths were white, 20.2% were black, 8.4% were Hispanic, 8.8% mixed and other, 1.7% Asian, and 0.9% American Indian (see test manual for further details).

[1]US National Survey Data Set (YSR) was acquired from ASEBA/Research Center for Children, Youth and Families.

A juvenile court sample consisted of 83 youths (66 males, 17 females), adjudicated delinquent, who underwent court-ordered mental health evaluation for disposition purposes in a family court clinic located in a Northeastern jurisdiction in the U.S. The same sample was used by this author in a separate study [27]. Youths ranged in age from 12 to 17 (mean age = 15.2, SD = 1.1). Complete item data were available for 77 of the 83 youths.

### Procedure

Records-based research was conducted in the juvenile court sample, consistent with policy for the protection of human subjects, ethical principles and guidelines for research involving human subjects. Data were compiled without individually identifiable information so that subjects could not be identified, directly or through identifiers linked to the subjects.

### Measures

#### ASEBA youth self-report

The ASEBA Youth Self-Report (YSR) contains 112 items that yield scores on broad-band Internalizing, Externalizing, and Total Score scales and eight empirically based clinical syndrome scales. The test scales are supported by a solid research base and are technically sound, with strong reliability and validity data [28,29].

#### Convergent measure

The Jesness inventory-revised [30] was administered concurrently with the YSR in 36 out of 70 cases in which the JI-R was administered in the juvenile court sample. The JI-R is a self-report multi-scale measure of personality and psychopathology that is applicable to youth with severe behavioral problems, including youth involved in the juvenile justice system. The JI-R includes 11 personality scales, 9 personality subtype scales, and 2 DSM-IV scales (Conduct Disorder, Oppositional Defiant Disorder).

#### Scale development approach for FFM scales derived from the ASEBA youth self-report

Development of ASEBA YSR-FFM scales was guided by reference to the negative pole of FFM domains and facets and pathological or maladaptive personality trait facets in order to cover a broad range of negative or problematic personality traits that might be captured by YSR items. Thus, I considered the description of FFM domains and facets as presented in the Five Factor Model Rating Form (FFMRF) [31] which is composed of the five domains and six facets within each domain as per the NEO-PI-R [32]; the International Personality Item Pool (IPIP) [33,34], a non-proprietary measure assessing constructs similar to those in the NEO PI-R; the Big Five Aspect Scales [35]; and the Personality Inventory for DSM-5 (PID-5) [23]. The NEO-PI-R and IPIP are comprised of the same five domains, which include Neuroticism versus Emotional Stability; Extraversion versus Introversion; Openness versus Closedness to One's Own Experience; Agreeableness versus Antagonism; Conscientiousness versus Undependability. Both measures contain facet scales with similar labels and are comprised of similar items. The PID-5 incorporates maladaptive variants of the FFM trait dimensional model, and there is a clear degree of overlap between the PID-5 and FFM measures. However, there are also differences between the FFM and the AMPD trait domains for the DSM-5 [36-38]. For example, there is much research on the FFM that supports placement of suspiciousness as being on the low, or negative pole of trust within the trait domain of Agreeableness versus Antagonism. In contrast, suspiciousness is placed primarily within the domain of Detachment, and secondarily within the domain of Negative Affectivity in the AMPD. The AMPD does not include maladaptive characteristics associated with the opposite pole of Introversion/Detachment, that is, potential problems associated with excessive levels of extraversion (e.g. intrusiveness, high risk-taking behaviors).

Given both overlaps and discrepancies between the FFM (including IPIP and Big Five measures) and AMPD models, I attempted to integrate the models in judging whether each of the 112 items on the YSR were relevant to any of the FFM domains at either a high or low

level. In some instances, I followed placement of items consistent with the FFM, e.g. depressivity within Neuroticism rather than within Detachment as in the AMPD. One of the challenges in selecting and excluding items was to discriminate between items on the YSR that might confound personality characteristics with specific behaviors that may be associated with such characteristics, and which, in association with contextual conditions, might predict such behavior. This is not always a straightforward distinction given that personality traits are considered to be relatively stable and enduring patterns of thinking, feeling and behaving [39]. Personality traits and personality functioning overlap [40] and can be examined with respect to their affective, behavioral, and cognitive content components (ABC model); see for example [41]. Pytlik Zillig., *et al*. [41] examined the affective, behavioral, and cognitive content components of multiple big five measures using multiple raters. Conscientiousness was found to be dominated by the behavioral component whereas Neuroticism was dominated by the affective component. Extraversion was found to have both strong behavioral and affective content components. Thus, it follows that some items on personality measures would emphasize more of a behavioral component or aspect than other items.

An additional consideration in the selection of items was the principle expounded by Paunonen [42], that good Big Five personality measures should predict complex socially and culturally significant criterion behaviors that are multifaceted, have multiple determinants including genetic and environmental, and the behaviors do not simply represent alternative measures of the Big Five factors. Paunonen included alcohol consumption, dating behavior, tobacco consumption, participation in sports, donating blood, driving automobiles at excessive speeds, as examples of such complex behaviors. Paunonen found that alcohol consumption was positively associated with extraversion and negatively associated with conscientiousness. The YSR includes one item referencing alcohol use, and another item referencing drug use. Thus, it was judged that some YSR items, e.g. items pertaining to drug and alcohol use, stealing, fire setting, would better be considered as behavioral outcomes potentially associated with personality characteristics, rather than personality characteristics, per se. Judgments were difficult to make in some instances. For example, the YSR includes an item referencing self-harmful or suicidal behavior and another item referencing suicidal ideation. Whereas the FFM measures do not include such items, the PID-5 Depressivity trait scale includes an item implicitly referring to thoughts of suicide and an item that explicitly references talking about suicide. A decision was made here not to include the YSR items referencing suicidal ideation or self-harmful behavior.

Notwithstanding the general guiding principle of not including explicit behaviors that might be outcomes of, or associated with, personality traits, certain items referencing particular behaviors were included if there was a basis for considering such behavior as an indicator or manifestation of an FFM trait domain facet, or as emphasizing more of a behavioral component of a trait facet. For example, an item referencing rule breaking was included, as it was judged to be not behavior specific and to be associated with the Dutifulness facet of Conscientiousness, and is in fact included on the IPIP NEO-300 [34], which includes items such as "break rules", "break my promises". A YSR item referencing fighting was included as a specific manifestation of Antagonism, and indeed similar items are included in the IPIP (e.g. "love a good fight"). Admittedly, others might argue that fighting would better be considered as a criterion variable. Another challenge was to judge whether certain items belonged best to one rather than another domain scale. It is noted that the FFM, as assessed by various measures, does not obtain a perfect factorial simple structure, with some traits, such as angry hostility, occupying interstitial space, loading significantly on more than one domain, and that some of the domains and facets of the FFM correlate with other domains and facets [43,44].

On the basis of the preceding guidelines, as well as analyses of inter-item and item-total correlations, the following preliminary scales were formed: Neuroticism, Antagonism, Disinhibition, Detachment, Atypical Experiences. Table 1 lists the preliminary proposed YSR items assigned to each FFM scale, as well as the syndrome scale in which each item is included on the YSR. Due to proprietary restrictions, the complete items cannot be listed.

| YSR-FFM Scale | YSR Item # | YSR Syndrome Scale |
|---|---|---|
| Neuroticism | 35, 45, 50, 52, 71, 112 | Anxious/Depressed |
| | 75, 103 | Withdrawn/Depressed |
| | 87, 95[a] | Aggressive Behavior |
| | 27 | Social Problems |
| Antagonism | 3, 16, 37, 57, 86[b], 94, 95[a], 97 | Aggressive Behavior |
| | 43, 90[b] | Rule-Breaking Behavior |
| Detachment | 5, 42, 69, 102, 111 | Withdrawn/Depressed |
| | 34, 48 | Social Problems |
| | 89 | Aggressive |
| Disinhibition | 4, 8, 10, 41, 61, 78 | Attention Problem |
| | 28, 101[b] | Rule-Breaking Behavior |
| Atypical Experiences | 9, 40, 66, 70, 84, 85 | Thought Problems |
| | 13[b] | Attention Problems |

***Table 1:*** *Preliminary proposed list of items comprising the youth self-report FFM scales.*

*[a]: Item was subsequently removed from Neuroticism and added to Antagonism. [b]: Item was subsequently omitted.*

## Data analysis

Psychometric properties of the FFM YSR scales were evaluated primarily in the normative sample, and secondary analyses were conducted with the juvenile court sample. Analyses were conducted using IBM SPSS Version 26.0 [45]. Exploratory Factor Analyses were conducted to identify the factor structure or model for each of the five YSR-FFM sets of variables, followed by Confirmatory Factor Analyses to assess the accuracy of the hypothesized construct models. Given that confirmatory factor analyses (see below) resulted in minor modifications of several YSR-FFM domain scales, descriptive statistics are presented following presentation of the CFA's. Internal consistency reliability estimates (coefficient alpha) and bivariate scale intercorrelations were calculated. The availability of test scores from a fairly small sample of juvenile court youths allowed for mean score comparisons between the normative and juvenile court samples on the YSR-FFM scales. Additionally, a smaller sample of juvenile court youths who completed both a YSR and JI-R allowed for a very preliminary examination of convergent and discriminant validity between the YSR-FFM scales and a separate personality measure.

## Exploratory factor analyses

The normative sample of youths (n = 1057) was randomly sorted into a group of 357 youths for the purpose of EFAs, and 700 youths for CFAs. The 11 items of the Neuroticism construct variable were subjected to a Principal Factor Analysis (PAF), with Oblique (Oblimin) rotation. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .88. Bartlett's Test of Sphericity was significant ($p < .001$). PAF revealed two factors with eigenvalues exceeding 1 (3.259, 1.085) explaining 29.14% and 4.49% of the variance, respectively. However, Parallel Analysis supported retention of only the first factor. The EFA Pattern Matrix indicated that item #s 87, 95, both of which

suggest intense emotionality, and item # 112, formed a second factor. This scale was subsequently tested by CFA to test the hypothesized one-factor model.

Results of Parallel analysis of the nine items of the Antagonism construct variable as well as the scree plot and eigenvalues suggested a two-factor model. KMO verified sampling adequacy (.79) and Bartlett's Test of Sphericity was significant ($p < .001$). As indicated in the Pattern Matrix, the second factor was composed of item #s 37, 57, and 97, each of which involve aggression, and were negatively loaded with the construct, indicating very low scores on these items for some youths. As such, CFA was used to test whether a one-factor model or a two-factor correlated model provided a better fit.

Results of the EFA of the eight items of the Detachment variable indicated a one-factor model, with only one eigenvalue which exceeded 1 and a scree plot indicating one factor. KMO verified sampling adequacy (.80), and Bartlett's Test of Sphericity was significant ($p < .001$). Subsequently, a one-factor CFA model was tested.

Results of parallel analysis and scree plot of the 8-item Disinhibition variable suggested a one-factor solution, although the initial two factors had eigenvalues that exceeded 1. KMO verified sampling adequacy (.78), and Bartlett's Test of Sphericity was significant ($p < .001$). As indicated in the Pattern Matrix, the first factor included six items, i.e., items 4, 8, 10, 41, 61, and 78, which reference problems with attention and distractibility, impulsive behavior and poor school work. The second factor included Items 28 and 101, one which references generalized rule-breaking (Item # 28), whereas Item # 101 references a specific type of rule-breaking; the loading for Item #101 was weak (.325 in the pattern matrix) as was its communality after extraction, suggesting consideration of removal of this item.

Results of parallel analysis of the 7-item Atypical Experiences variable suggested a one-factor solution, although the first and second eigenvalue exceeded 1. KMO verified sampling adequacy (.72) and Bartlett's Test of Sphericity was significant ($p < .001$). As indicated in the Pattern Matrix, Items #13 and 40, which reference mental confusion and unusual auditory experiences, were included in a second factor, which was indicated also in the Structure Matrix, although Item #40 cross-loaded.

### Confirmatory factor analyses

Although the YSR-FFM scales are newly developed, extensive theoretical and empirical support for the Five-Factor Model of personality, upon which these scales are modeled, and results of the preceding EFAs, suggested that CFA was an appropriate means for assessing the latent structure of each of the FFM scales, particularly to test the extent to which the five hypothesized latent variable models corresponding with the FFM personality trait constructs are consistent (goodness-of-fit) with the sample data. Five single-factor CFA's were conducted using AMOS Version 23.0 [46]. Model fit was assessed using goodness-of-fit statistics including the chi-square statistic and the chi-square/degrees of freedom ratio, for which a value in the range of 2 - 5 indicates a reasonable fit [47]. Models that fit well are indicated by Chi-square non-significance. However, it is known that the chi-square statistic tends to increase along with the sample size [48]. This limitation of the chi-square Likelihood Ratio Test statistic has resulted in development of pragmatic approaches to goodness-of-fit indices [49]. Included in the current study are the robust comparative fit index (CFI) [50]; Root Mean Square Error of Approximation (RMSEA) [51] with 90% confidence intervals; the standardized Root Mean Square Residual (SRMR) [52]. CFI values of .95 or above suggest good model fit; values of .90 - .94 may suggest adequate model fit. RMSEA values less than 0.05 are good, values between 0.05 and 0.08 are acceptable, with 90% confidence intervals between .00 and .10 [53]. SRMR values below .08 are recommended [53]. Additionally, each of the single-factor models was tested for multigroup invariance based on gender by fitting a series of nested multigroup CFA models. Particularly, the model was tested for configural and measurement invariance, and, if metric invariance was met, scalar invariance would be explored. Metric invariance (factor loadings are constrained to be equal across groups and the model is compared to the configural model) tests whether the items that serve as indicators of the latent construct operate equivalently across the two groups [54]. Non-invariance is indicated if the chi-square difference between the two models is statistically significant. Given the excessive stringency of the chi-square

difference test, it has been proposed that invariance may be indicated by a difference in CFI value having a probability of < 0.01 [55]. Additionally, differences in RMSEA values < .015, and differences in SRMR values of < .03 (comparing configural and metric models) and < .01 (when comparing metric and scalar models) have been recommended by Chen [56]. Scalar invariance is tested by constraining factor loadings and intercepts to equality. Scalar invariance implies that the meaning of the construct and the levels of the underlying items are equal in both groups and that the groups can be compared on their scores on the latent variable [57]. Results of the CFA's of the YSR-FFM scales are presented in table 2 and measurement invariance of the YSR-FFM scales is presented in table 3.

## Results

### CFA models

### Neuroticism

The latent structure of the Neuroticism scale was assessed in the randomly selected group (n = 700). Given that the assumption of multivariate normality was not met, maximum-likelihood (ML) estimation was used and the data were bootstrapped, including 90% bias-corrected confidence intervals. As seen in table 2, the specified model for Neuroticism appeared to have a marginally acceptable fit. Inspection of Modification Indices (MI) indicated by the statistical software and the standardized residual covariance values suggested misspecification involving the observed variables # 87 and # 95. In reviewing these items, there appeared to be content overlap in that both suggest a sudden shift or surge of affect (in mood or temper). It was noted in the EFA that these two items, along with Item #112, formed a separate factor in the Pattern Matrix. It is recognized that re-specification and re-estimation of a model is, in its truest sense, conducted within an exploratory, rather than a confirmatory mode [52]. There were further indications that only 14.5% of variance in the indicator of Item # 95 was explained by the latent variable. Therefore, given that the analysis was now being conducted within an exploratory mode, Item # 95 was omitted from the analysis. This revised model, without any covariance specified, resulted in an adequate fit (CFI = .958; RMSEA = .053). Next, I tested for multigroup invariance of this model based on gender. The baseline model fit was assessed for each gender group separately. The model had acceptable fit for boys (CFI = .930, RMSEA = .051; and adequate fit girls (CFI = .957; RMSEA = .061). Configural invariance was supported in that the re-specified model for Neuroticism had an adequate fit across both genders. All factor loadings were significant (ps < .001) and ranged from .39 to .55 (males) and .49 to .72 (females). Next, metric invariance was tested. Full metric invariance was supported by the non-significant p value associated with the difference in chi-square values and the difference in CFIs. Scalar invariance (equivalence of loadings and item intercepts) was not supported, ($\chi^2$ test was significant at $p < .001$; $\Delta CFI = -.041$; albeit, $\Delta RMSEA = .01$ and $\Delta SRMR = .008$ were supportive of invariance).

### Antagonism

Based on the EFA which suggested a two-factor model, and further examination of inter-item correlations, the hypothesized model for Antagonism was revised to include Item # 95 and omit Item # 90 which was found to correlate with several items on the Disinhibition scale. One-factor and two-factor models were compared. Given that the assumption of multivariate normality was not met, maximum-likelihood (ML) estimation was used and the data were bootstrapped, including 90% bias-corrected confidence intervals. An initial nine-item, one-factor model evidenced a poor fit. There appeared to be misspecification associated with the pairing of error terms associated with Items 57 and 97. A review of these items suggested content overlap in that both items involve aggression (i.e. physical harm and threat of physical harm). Re-specification of the model was introduced by including an error covariance associated with items 57 and 97. This resulted in improved model fit. However, misspecification was indicated by a large standardized residual covariance between Items 3 and 86. There appeared to be conceptual overlap between these items which reference being obstinate and contentious. However, further re-specification was indicated given that Item # 86 had a squared multiple correlation of .09, indicating that only 9% of the variance in this indicator variable was accounted for by the latent construct. This item may be less specific to the Antagonism construct than Item

3 which references arguing frequently. Therefore, the model was re-specified with removal of Item #86. This resulted in a chi-square of 45.121 (DF = 19); CFI = .961; RMSEA = .044.

A nine-item, correlated two-factor model was examined, with Items 3, 16, 43, 86, 94 and 95 representing contentious, quick-tempered, mean qualities, and with Items 37, 57, and 97 composing a second factor representing tendencies toward physical aggression (including threat of physical aggression). This resulted in a marginal fit (CFI of .917). Removal of Item # 86 was suggested due to a large standardized residual covariance with Item # 3, and a low standardized estimate. Model fit improved; however, CFI was modestly lower and RMSEA was modestly higher compared to the one-factor model, standardized residuals were high for three pairs of items, and standard errors were high for four items. Therefore, the one-factor model was deemed to best fit the sample data. The model was examined separately by gender with adequate model fit for both males and females. A multigroup model was estimated for configural invariance, which indicated that the model had adequate fit across genders. All factor loadings were significant (ps < .001) and ranged from .38 to .53 (males) and .31 to .58 (females). Full metric invariance was indicated by a nonsignificant difference in chi-square values between metric and configural models as well as differences in CFI and RMSEA values that met criteria for metric invariance. Scalar invariance was not supported (chi-square difference p < .001; ΔCFI = -.036; however, ΔRMSEA = 0.006; ΔSRMR = < .001 suggested scalar invariance).

### Disinhibition

The hypothesized one-factor model of Disinhibition was tested. Given that the assumption of multivariate normality was not met, maximum-likelihood (ML) estimation was used and the data were bootstrapped, including 90% bias-corrected confidence intervals. Results yielded a marginal fit (CFI = .910, RMSEA = .067; PCMIN/DF = 5.802). An examination of MIs and standardized residual covariances indicated model misspecification, particularly given a very large standardized residual between Items 101 and 28, as well as MI between error terms for these items. These items were noted to load on a separate factor in the EFA, whereas the loading for Item 101 was low. The CFA analysis indicated a very low squared multiple correlation (coefficient of determination) for Item #101, indicating that only 6.4% of the variance in this item was attributed to the latent construct. Therefore, Item # 101 was omitted from the model. This resulted in improved model fit (Chi-square = 39.709; df = 14; CFI = .958; RMSEA = .051). Next, a multigroup model was estimated for configural invariance. This resulted in a chi-square value indicating that the model had adequate fit across genders. All factor loadings were significant (ps < .001) and ranged from .44 to .58 (males) and .39 to .68 (females). Next, I tested for invariance of factorial measurement across both gender groups. The chi-square difference between metric and configural models supported full metric invariance between the two groups. Scalar invariance was not supported (p < .001; ΔCFI = 0.042; however, ΔRMSEA (0.01) and ΔSRMR = 0.004, were consistent with scalar invariance). Using a backward sequential approach, intercept parameters were freed one at a time which resulted in 4 of 7 parameters invariant based on chi-square difference.

### Detachment

As seen in table 2, the hypothesized model for Detachment demonstrated adequate fit between the sample covariance matrix and the population covariance implied by the model, as noted by a CFI of .946 and RMSEA of .042. Given that the assumption of multivariate normality was not met, maximum-likelihood (ML) estimation was used and the data were bootstrapped, including 90% bias-corrected confidence intervals. A multigroup model was estimated to test simultaneously for configural (baseline) invariance across gender groups. The model had adequate fit across genders. All factor loadings were significant (ps < .001) and ranged from .32 to .48 (males) and .30 to .61 (females). Subsequently, the model was tested for metric invariance. Results of this model were compared to the configural model and testing for invariance of factor loadings was examined for differences between chi-square and CFI values. This yielded $\Delta\chi^2_{(7)} = 89.468$, p < .01 and ΔCFI = 0.027; ΔRMSEA = 0.003, ΔSRMR = 0.0100. Thus, full metric invariance could not be concluded. Using a backward sequential approach, I systematically freed parameter constraints one at a time to determine which factor loading(s) were noninvariant. Seven of eight factor loading parameters were found to be invariant across gender, with cumulatively maintained equality constraints, suggesting

the model to be partially invariant [58,59]. Item #102, which references lack of vigor or pep, was not invariant, indicating that this item operates somewhat differently for boys and girls in its measurement of the intended latent variable. Scalar invariance was not supported (p = .015; ΔCFI = .024; however, ΔRMSEA = .001; ΔSRMR = 0). The intercept associated with Item #102 was not invariant. However, with all other item intercepts constrained equal, there was no significant difference between the metric and scalar models, suggesting partial scalar invariance.

### Atypical experiences

The model for Atypical Experiences demonstrated inadequate fit, as seen in table 2. The model was noted for an error covariance associated with observed variables for Items # 40 and # 70. A large standardized residual covariance for these items were noted as well. Both items suggest unusual beliefs and experiences, one involving unusual auditory experiences, the other involving unusual visual experiences, and thus seem to have substantive similarity. Therefore, an error covariance parameter was included, which resulted in significant improvement to model fit, however, items 13 and 40 had a large standardized residual covariance indicating misspecification and there did not appear to be a substantive rationale for including these correlated errors, although it was noted that in the EFA these two items were identified as a second factor in the Pattern Matrix. Additionally, further examination of item correlations indicated that Item #13 had low correlations with other items. Therefore, it was decided to re-specify the Atypical Experiences model with removal of Item 13. This resulted in significant improvement in the model ($\chi^2_{(8)} = 12.389$; $\Delta\chi^2_{(5)} = 34.69$, $p < .001$). A multigroup model was estimated for configural invariance; the model had adequate fit across genders. All factor loadings were significant (ps < .001) and ranged from .38 to .58 (males) and .41 to .63 (females). Subsequently, the model was tested for metric invariance. Full metric invariance was supported. Next, Scalar invariance was tested. Scalar invariance was supported $\Delta\chi^2_{(6)} = 11.28$, p = 0.08 (NS), which implies that the meaning of the construct and the levels of the underlying items are equal in both groups.

| Model | $\chi^2$ | df | CFI | SRMR | RMSEA (90% CI) |
|---|---|---|---|---|---|
| Neuroticism | 138.83 | 44 | .914 | .05 | .080 (.07, .10) |
| Neur-R[a] | 103.34 | 35 | .958 | .03 | .053 (.04, .07) |
| Antagonism | 130.633 | 27 | .864 | .05 | .077 (.06, .09) |
| Antag-R[b] | 45.121 | 19 | .961 | .05 | .044 (.028, .061) |
| Disinhibition | 116.038 | 20 | .910 | .05 | .067 (.056, .080) |
| Disinhib-R[c] | 39.709 | 14 | .958 | .03 | .051 (.03, .07) |
| Detachment | 44.260 | 20 | .946 | .03 | .042 (.025, .058) |
| Atypical Experiences | 78.598 | 14 | .896 | .05 | .08 (.06, .10) |
| Atypical-R[d] | 12.389 | 8 | .991 | .02 | .028 (.000, .057) |

***Table 2:*** *Goodness-of-fit indicators of YSR-FFM models (n = 700).*

*[a]: Neur-R = Neuroticism-Revised; Model omitted Item #95. [b]: Antag-R = Antagonism-Revised; Model includes covariances between items 57 and 97, and omission of item # 86. [c]: Disinhib-R = Disinhibition-Revised; Model omits item #101.*

*[d]: Atypical-R = Atypical Experiences-Revised; Model includes covariance between items 40 and 70 and deletion of Item # 13.*

Comparisons of ML standard error estimates with bootstrap standard errors for each of the preceding final CFA models found that for Neuroticism, and Disinhibition, for both males and females, the bootstrap standard errors suggested that the distribution of the parameter estimates was not much wider than would be expected under normal theory assumptions. For Antagonism, for males most items had less than a 50% increase in the bootstrap standard error; for females, bootstrap standard errors displayed between 25% to

70% increases, suggesting that the distribution of the parameter estimates appeared to be wider than would be expected under normal theory assumptions. For Detachment, increases in bootstrap standard errors ranged from less than 20% to 80% for males; for females, increases in bootstrap standard errors ranged from 45% to 67%, suggesting that the distribution of the parameter estimates for both males and females appeared to be wider than would be expected under normal theory assumptions. For Atypical Experiences, increases in bootstrap standard errors ranged from less than 6% to 115% for males; for females, increases in bootstrap standard errors ranged from 39% to 73%, suggesting that the distribution of the parameter estimates for both males and females appeared to be wider than would be expected under normal theory assumptions. However, for all of the models, p-values for all items for both genders indicated that the confidence interval would have to be above the 99% level before the lower bound value would be zero.

## Supplementary CFA analyses

As noted earlier, the AMPD does not include maladaptive characteristics associated with the opposite pole of Introversion/Detachment, that is, potential problems associated with excessive levels of extraversion (e.g. intrusiveness, high risk-taking behaviors). The AMPD domain of Antagonism incorporates Attention Seeking and Grandiosity facets. Yet, attention seeking and more broadly, histrionic features, have been associated with FFM Extraversion and weakly associated with Agreeableness [20]; albeit, Gore., *et al*. [60] obtained empirical support for placing histrionism primarily within antagonism. The YSR includes several items suggestive of possible maladaptive or excessive extraversion quality (e.g. items referring to boisterous, garrulous behavior and antics). Indeed, the items of this type that I identified are included in a 6-item "Intrusive" scale in the ASEBA Adult Self-Report (ASR) [61], which is one of three ASR Externalizing scales. This scale does not appear on the YSR. One of the items on the Adult Intrusive scale referring to excessive teasing was included here in the YSR Antagonism scale. I modeled a 5-item scale, tentatively titled "Exhibitionism/Intrusiveness", which includes Items 7, 74 and 93, identified on the YSR as "Other Problems", and Items 19 and 104, from the YSR Aggressive Behavior scale. However, an error covariance parameter between two items that are similar in content (Items #93 and # 104) was necessary to achieve good fit (Chi-square = 10.808; p = .03; CFI = .992; RMSEA = .040 (.012, .070); SRMR = .02). Invariance was not tested given that this scale was not originally planned.

Given the availability of items that appear to capture extraversion, I additionally modeled a 6-item Extraversion scale (Items 59, 60, 80, 88, 92, 107, all from the YSR "Positive Qualities" scale). Given a large MI and standardized residual covariance value between Items 59 and 60, an error covariance parameter was introduced between those two items which resulted in an adequate fit (Chi-square = 29.251; df = 8; p < .01; CFI = .976; RMSEA = .050; SRMR = .03). Invariance was not tested given that this scale was not originally planned. It may be useful for research purposes to have such a scale available.

| Scale | Model | $\chi^2$ | df | $\Delta^{x2}$ | Δdf | p | ΔCFI | ΔSRMR | ΔRMSEA |
|---|---|---|---|---|---|---|---|---|---|
| Neuroticism | Configural | 146.754 | 70 | | | | | | |
| | Metric | 160.628 | 79 | 13.88 | 9 | .13 | .003 | .008 | .002 |
| | Scalar | 230.845 | 89 | 70.22 | 10 | <.001 | .041 | .008 | .010 |
| Antagonism | Configural | 72.096 | 38 | | | | | | |
| | Metric | 84.626 | 45 | 12.53 | 7 | .08 | .008 | .007 | .012 |
| | Scalar | 117.126 | 53 | 32.5 | 8 | <.001 | .036 | < .001 | .006 |
| Disinhibition | Configural | 51.896 | 28 | | | | | | |
| | Metric | 58.416 | 34 | 6.52 | 6 | .368 | <.001 | .006 | .003 |
| | Scalar | 91.796 | 41 | 33.38 | 7 | <.001 | .043 | <.001 | .0 |
| Detachment | Configural | 69.953 | 40 | | | | | | |
| | Metric | 89.438 | 47 | 19.49 | 7 | <.01 | .027 | .010 | .003 |
| | Partial Metric | 81.895 | 46 | 12.24 | 6 | .063 | .013 | .007 | 0 |
| | Scalar | 108.434 | 55 | 18.96 | 8 | .015 | .024 | 0 | .001 |
| | Partial Scalar | 89.159 | 53 | 7.26 | 6 | .117 | .014 | 0 | .005 |
| Atypical Experiences | Configural | 32.214 | 16 | | | | | | |
| | Metric | 34.899 | 21 | 2.69 | 5 | .75 | .004 | .000 | .007 |
| | Scalar | 46.170 | 27 | 11.28 | 6 | .08 | .011 | .000 | .001 |

***Table 3:*** *Multigroup CFA measurement invariance across sexes (366 boys vs. 344 girls).*

## Descriptive statistics

Descriptive statistics are reported in table 4. High coefficient of variation for each scale reflected high levels of dispersion around the mean. The Shapiro-Wilk test for normality indicated that all five scales had significance values below .001, indicating non-normal distributions. All scales were positively skewed which may have resulted in decrease in average inter-item correlation and decreased alpha [62]. Internal consistency (Cronbach's alpha) was good for Neuroticism, acceptable for Antagonism and Disinhibition, and questionable to acceptable for Detachment and Atypical Experiences, although alpha coefficients in this range are often considered acceptable in exploratory research. Comparisons of mean scores by gender indicated that females obtained significantly higher scores on Neuroticism, lower scores on Disinhibition, and higher scores on Atypical Experiences. Cohen's d indicated low effect sizes, below 0.2 for Disinhibition and Atypical Experiences, and a small to medium effect size for Neuroticism.

| YSR-FFM | Mean | SEM | SD | Median | Range | Skew | Kurtosis | Alpha | MIIC |
|---|---|---|---|---|---|---|---|---|---|
| Neuroticism | 4.01 | .10 | 3.39 | 3.00 | 19.00 | 1.04 | 1.01 | .80 | .29 |
| Antagonism | 2.58 | .07 | 2.25 | 2.00 | 14.00 | 1.15 | 1.67 | .69 | .25 |
| Disinhibition | 4.56 | .10 | 3.20 | 4.00 | 16.00 | .480 | -.310 | .71 | .26 |
| Detachment | 2.86 | .07 | 2.32 | 2.00 | 14.00 | 1.01 | 1.35 | .63 | .18 |
| Atypical | 1.55 | .06 | 1.79 | 1.00 | 11.00 | 1.59 | 2.91 | .62 | .24 |

***Table 4:*** *Descriptive statistics for the normative youth self-report FFM scales (N = 1057).*

Table 5 displays Spearman's rank order correlations among the YSR-FFM scales developed in the normative sample. This table displays 95% confidence intervals based on 1000 bootstrap samples. Correlation coefficient estimates were moderate to large, ranging from .37 (Antagonism and Atypical Experiences) to .61 (Neuroticism and Detachment) with a mean correlation estimate coefficient of .47. For comparison purposes, Pearson's correlations were conducted, with very similar results.

| Scale | Neurot | Antag | Disconst | Detach | Atypical |
|---|---|---|---|---|---|
| Neurot | _ | | | | |
| Antag | .42**<br>[.37, .47] | _ | | | |
| Disinhib | .44**<br>[.38, .49] | .58**<br>[.54, .62] | _ | | |
| Detach | .61**<br>[.57, .65] | .44**<br>[.40, .49] | .44**<br>[.39, .49] | _ | |
| Atypical | .50**<br>[.45, .54] | .37**<br>[.32, .42] | .40**<br>[.35, .45] | .48**<br>[.43, .53] | _ |

***Table 5:*** *Inter-correlations (Spearman's rho) among normative youth self-report FFM scales.*

*Note. Sample size (N) for normative youth = 1057.*

*Neurot: Neuroticism; Antag: Antagonism; Disinhib: Disinhibition; Detach: Detachment; Atypical: Atypical Experiences.*

***Correlation is significant at the 0.01 level (2-tailed); bootstrap 95% CIs reported in brackets.*

### YSR-FFM mean comparisons between normative and juvenile delinquent samples

The availability of a relatively small sample of Youth Self-Reports in adjudicated juvenile delinquent youths who were evaluated for disposition purposes, permitted a comparison of mean scores on the YSR-FFM scales between the normative and juvenile delinquent samples. No a priori hypotheses were made. Data were pooled for males and females in the juvenile delinquent sample (no gender differences were found on the FFM proxy scales that were developed in this study).

As seen in table 6, the juvenile delinquent group obtained significantly higher scores on the YSR Antagonism and Detachment scales, and significantly lower scores on the Neuroticism scale - all effect sizes (Cohen's d) were small. Given the very unequal sample sizes between the two groups, I alternatively selected a sample of exactly 77 cases from the normative sample. Means and SDs were similar to those in the total sample size and all group differences were consistent regardless of comparison methods. The difference between groups on Neuroticism is mitigated by the finding that there was no significant difference between the normative male sample and the juvenile delinquent sample (t (626) = 0.62, p = .53).

| | Normative Sample | | Juvenile Delinquent | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | M1-M2 95% CI | t | p | Cohen's d 95% CI |
| **YSR-FFM** | | | | | | | | |
| Neurot | 4.01 | 3.39 | 3.05 | 3.90 | 0.96 [0.17, 1.75] | 2.37 | 0.02 | 0.28 [0.05, 0.51] |
| Antag | 2.58 | 2.25 | 3.35 | 2.86 | -0.77 [-1.30, - 2.38] | 2.31 | 0.02 | 0.34 [0.10, 0.57] |
| Disinhib | 4.56 | 3.20 | 4.12 | 2.97 | 0.44 [-0.30, 1.18] | 1.17 | 0.24 | 0.14 [-0.09, 0.37] |
| Detach | 2.86 | 2.32 | 3.49 | 2.91 | -0.63 [-0.08, -1.78] | 2.26 | 0.02 | -0.27 [-0.03, -0.50] |
| Atypical | 1.55 | 1.79 | 1.36 | 1.53 | 0.19 [-0.22, 0.60] | 0.91 | 0.36 | 0.11 [-0.12, 0.34] |

***Table 6:*** *Mean Score Comparisons Between Normative (N = 1057) and Juvenile Delinquent Sample (N = 77) on YSR-FFM Scales. Note. Neurot: Neuroticism; Antag: Antagonism; Disinhib: Disinhibition; Detach: Detachment; Atypical: Aberrant Ideas; CI: Confidence Interval.*

### Convergent validity of YSR-FFM scales

A small sub-sample (n 36-38) was available in which juvenile delinquent youths completed both YSR and JI-R inventories (See table 7). With respect to the associations between the YSR-FFM scales and the JI-R scales, for the most part, the associations were moderate to strong and were non-discriminant with respect to strength of associations. It is noted that many of the JI-R scales are highly intercorrelated due to item overlap; item overlap results in less clear, less discriminating associations with other measures [63]. The overall moderate inter-correlations between the YSR-FFM scales may also have accounted in part for limited discriminant associations between YSR-FFM and JI-R scales. However, there were some convergent and discriminant associations which were conceptually consistent. For example, the correlation between YSR Antagonism and JI-R Conduct Disorder (r = .64) was significantly stronger than the correlations between Antagonism and JI-R Withdrawal/Depression (Fisher's Z = 2.07, p = 0.02) and JI-R Social Anxiety (Fisher's Z = 2.25, p = .01). Neuroticism was more strongly associated with Withdrawal/Depression (Fisher's Z = 1.81, p = .04), Social Anxiety (Fisher's Z = 1.92, p = .03), and Manifest Aggression (Fisher's Z = 1.70, p = .04) than with Oppositional Defiant Disorder.

| Scale | Neurot | Antag | Disinhib | Detach | Atypical |
|---|---|---|---|---|---|
| Neurot | _ | | | | |
| Antag | .60** | _ | | | |
| Disinhib | .59** | .57** | _ | | |
| Detach | .73* | .54 | .47* | _ | |
| Atypical | .67** | .62 | .54* | .61* | _ |
| Asocial | .02 | .33* | .09 | -.12 | .15 |
| SM | .35** | .51** | .41* | .28 | .38* |
| VO | .40* | .54** | .44** | .36* | .38* |
| IMM | .46** | .45** | .49** | 52** | 43** |
| Autism | .35* | .41** | .25 | .36* | .33* |
| Alienation | .27 | .45** | .40* | .41* | .30 |
| Scale | Neurot | Antag | Disinhib | Detach | Atypical |
| Manifest Aggression | .48** | .53** | .47** | .45** | .40* |
| Withdrawal/Depression | .50** | .24 | .53** | .53** | .42* |
| Social Anxiety | .52** | .21 | .56** | .46** | .41* |

| | | | | | |
|---|---|---|---|---|---|
| Repression | -.21 | -.30 | -.42** | -.35* | -.29 |
| Denial | -.40* | -.34* | -.51** | -.54** | -.46** |
| Conduct Disorder | .25 | .64** | .17 | .21 | .20 |
| Oppositional Defiant Disorder | .11 | .32 | .35** | .23 | .29 |

***Table 7:*** *Correlations (Spearman's rho) between youth self-report FFM scales and JI-R Scales in juvenile delinquent sample.*

*Note. Sample size (N) for youth who completed both the YSR and the JI-R = 36.*

*Neurot: Neuroticism; Antag: Antagonism; Disinhib: Disinhibition; Detach: Detachment; Atypical: Atypical Experiences; SM: Social Maladjustment; VO: Value Orientation; IMM: Immaturity.*

***Correlation is significant at the 0.01 level (2-tailed), *Correlation is significant at the 0.05 level (2-tailed).*

## Discussion

The current study represents the first known attempt to operationalize maladaptive variants of the Five-Factor Model of personality by indexing items from the ASEBA Youth Self-Report. Four trait domain scales (excluding Openness) were developed in this study in accordance with the constructs of the "Big Five" or Five Factor Model as well as the DSM-5 Alternative Model of Personality Disorders (AMPD). AMPD domains are intended to represent maladaptive variants of the Big Five domains. YSR-FFM Detachment was more aligned with the AMPD than with the FFM. The fifth trait scale, labeled here "Atypical Experiences", was clearly intended to be aligned with the Psychoticism trait domain of the AMPD rather than the opposite pole of FFM Openness. The five preliminary scales that were developed ranged from good to acceptable internal consistency reliability. Not surprisingly, all YSR-FFM scales were positively skewed which may have resulted in decrease in average inter-item correlation and decreased alpha [62,64]. The lowest reliability coefficient was on the shortest scale (i.e., Atypical Experiences). Additionally, reliability of the Atypical Experiences scale likely was attenuated by the restricted range of scores [65]. Most of the YSR-FFM scales were moderately intercorrelated, indicating that each trait scale domain comprises shared and unique variance in accounting for relationships with other FFM scales.

Given the shared variance between ASEBA Internalizing and Externalizing domains as broad dimensions of psychopathology [15], (the mean correlation between Internalizing and Externalizing scores for the parent, teacher, and youth-self report forms was .53), and that YSR-FFM Neuroticism and Antagonism represent fair proxies of internalizing and externalizing problems, respectively, it is not surprising that the YSR item pool appears to have a greater selection of items that may be representative of Neuroticism and Antagonism scales than the other YSR-FFM scales. The other FFM scales were somewhat more difficult to model from the YSR item pool. Each scale will be further discussed below.

The YSR Neuroticism scale appears to capture core features of the construct such as anxiety, fear, nervousness, worry, self-consciousness, timidity, and moodiness. A 10-item CFA model demonstrated good fit and demonstrated metric invariance, but not scalar invariance, across genders. The initial hypothesized model included Item # 95 which references being quick-tempered; however, this item did not fit well with the population covariance implied by the model. Again, given overlap between the Neuroticism and Antagonism personality domains, particularly in aspects pertaining to being quick to anger, it was not surprising that Item # 95 fit better with the Antagonism model. Due to small sample size in the juvenile delinquent sample, only preliminary support can be asserted for convergent validity of Neuroticism as indicated by the finding that Neuroticism was more strongly associated with Withdrawal/Depression, Social Anxiety, and Manifest Aggression than with Oppositional Defiant Disorder. Interestingly, the mean score of Neuroticism was significantly lower in the juvenile delinquent group compared to the normative group. Although externalizing behavior problems have greater frequencies than internalizing problems in juvenile delinquent samples, internalizing problems are not uncommon particularly in detained and incarcer-

ated youth offenders [66]. However, the mean score difference on Neuroticism between normative and juvenile delinquent groups became non-significant when only normative males were compared with the juvenile delinquent group, which was composed primarily of male youths. Nevertheless, it is possible that the current juvenile delinquent sample, overall, was more defensive with regard to reporting internalizing/emotional problems compared to the normative sample.

The YSR Antagonism scale appears to capture features such as being mean, deceptive, hostile, verbally and physically aggressive. The initial hypothesized 9-item model did not have adequate fit and there appeared to be reasonable justification for re-specification given content overlap between two items that involve aggression. However, further re-specification was required before the model had good fit. Although each model developed in this study was intended to be theory or concept-driven rather than data-driven, this study nevertheless represents an initial and preliminary attempt to model personality scales from the YSR. Item # 86, which references being obstinate, was omitted in order to obtain a well-fitting model. Incidentally, however, in a very recent study of the Aggressive Behavior syndrome scale of the YSR [67], Item # 86 was included in a lower-order factor of symptoms related more peripherally to Aggressive Behavior - such symptoms were noted to be more often common to multiple psychiatric conditions - rather than a targeted aggression group factor which items were relatively more specific to aggression. Item # 86 also did not have a strong loading on either the related symptoms lower-order factor or the general factor of the bifactor model in the above-referenced study. The current study did not examine a bifactor model of Antagonism, however, the single-factor model appeared to have better fit than a correlated two-factor model. Preliminary convergent and discriminant validity of the Antagonism scale was indicated by the finding that the correlation between YSR Antagonism and JI-R Conduct Disorder was significantly stronger than the correlations between Antagonism and JI-R Withdrawal/ Depression and JI-R Social Anxiety. Additionally, while also a preliminary finding, the mean score for Antagonism was higher in the juvenile delinquent group as compared to the normative group which also supports construct validity of this scale.

The Disinhibition construct is intended to align with the negative pole of Conscientiousness and includes facets in the FFM such as lax, negligent, disorganized, careless, and facets in the AMPD such as irresponsibility, impulsivity, risk taking, distractibility. However, given the intention in scale development to minimize selection of YSR items that might confound personality characteristics with specific behaviors that may be associated with such characteristics, the Disinhibition scale in its present form might fall short of capturing characteristics referred to above, whereas it likely better assesses distractibility, and to a lesser extent impulsiveness and rule-breaking. Item # 101, which references truancy, did not fit well with the hypothesized model and was subsequently omitted. The re-specified model which omitted Item #101 had good model fit. Still, it is noted that whereas the juvenile delinquent group had a higher score on Antagonism compared to the normative sample, no difference in mean scores on the Disinhibition scale was found between the two groups. Given that low conscientiousness/disinhibition has been rated as high on the behavioral component of personality [40,41], the Disinhibition scale might have incorporated YSR items referencing defiance of parents, defiance in school, stealing, etc., which, in association with contextual conditions, might predict impulsive, irresponsible, reckless behavior. Future study might find a better balance in item selection between minimizing explicit behavior and capturing a broader range of the Disinhibition construct.

With respect to the development of a scale to measure introversion or detachment, or the negative pole of extraversion, the YSR includes a 14-item scale labeled "Positive Qualities", (also referred to as socially desirable items), which appear to reflect facets or aspects of both Extraversion and Agreeableness, e.g. gregarious, assertive, energetic qualities and also altruistic and straightforward qualities. Thus, it would appear that the YSR item pool contains items, which, reverse scored, would be consistent with a measure of introversion. However, when six reverse scored items were included in a preliminary scale with items that tap withdrawal, anhedonia, and lethargy, four of the six reverse scored items had item-total correlations below .25, and the mean inter-item correlation was .113. Thus, given the low item-total correlations for reverse scored items, the Detachment scale was modeled in accordance with the AMPD Detachment construct which incorporates suspiciousness along with facets such as withdrawal, anhedonia, restricted affectivity, depressivity. The depressivity facet was subsumed here within the YSR-FFM Neuroticism scale. Still, given the overlap between the AMPD domains of Negative Affectivity and

Detachment (both include facets of depressivity and suspiciousness) it is not surprising that the YSR-FFM Neuroticism and Detachment scales were strongly correlated. Preliminary support for convergent and discriminant validity of the Detachment scale was suggested by the significantly higher correlation between Detachment and JI-R Withdrawal/Depression than with the JI-R Social Maladjustment, Conduct Disorder, and Oppositional Defiant Disorder scales. Additionally, a significantly higher score on Detachment was found in the juvenile delinquent sample compared with the normative sample. This finding stands in notable contrast to the finding that the mean score on Neuroticism was significantly lower in the juvenile delinquent sample compared with the normative sample. One possible explanation for this finding is that the Detachment scale might be more susceptible to "state", or situation-specific characteristics, e.g. withdrawal, given that most of the juvenile delinquents were in detention at the time of evaluation, whereas the Neuroticism scale might better assess more stable and enduring "trait" characteristics. The AMPD Detachment facet of Intimacy Avoidance was not captured by the YSR item pool. However, as suggested by others [2], certain areas of personality development and pathology might not be relevant or exhibit presence in youth in the same manner as with adults, and vice versa. In particular, De Clercq., *et al*. [2] noted that Intimacy Avoidance does not emerge as an important lower-order personality trait in youth as in adults but might be covered by or manifested in Withdrawal.

The YSR item pool does not include many items that capture the AMPD domain of Psychoticism. Two of the six items on the Atypical Experiences scale of the YSR-FFM are similar to items within the Eccentricity facet of Psychoticism as measured by the PID-5; two items are similar to the Unusual Beliefs and Experiences facet as measured by the PID-5. The Atypical Experiences scale was robustly correlated with the YSR-FFM Neuroticism/Negative Affect scale and also moderately correlated with the other YSR-FFM scales. In a study of the hierarchical structure and construct validity of the PID-5 in a sample of adolescents, De Clercq., *et al*. [2] found that the Psychoticism factor was robustly correlated with the four remaining factors and also with separate measures of Disagreeableness, Emotional Instability, and Introversion. This led De Clercq., *et al*. [2] to suggest that personality dysfunction assessed in content areas of eccentricity, unusual beliefs and experiences, and perceptual dysregulation might be more closely related to general psychological dysfunction in younger persons as compared to adults. Still, due to the small number of items, the YSR-FFM Atypical Experiences scale might have limited construct coverage. However, the CFA model proposed here had good fit and demonstrated both metric and scalar invariance, albeit, good fit was obtained after including an error covariance parameter between two items which had substantive similarity (one involving unusual auditory experiences, the other involving unusual visual experiences), and after omitting Item # 13 which had low correlations with the other items and which inclusion resulted in the model being misspecified.

## Limitations and Future Directions

Notwithstanding that this study represents the first known published attempt to index maladaptive variants of the Five Factor Model of personality in youth through item content of the ASEBA Youth Self-Report, there are notable limitations which will need to be addressed going forward. These limitations were principally within the area of methodology. Typically, in scale construction of new scales from existing item content, researchers conduct at least two independent ratings to obtain a consensus among expert raters in assigning items to new, preliminary scales based on definitions of the constructs to be measured. The preliminary scales developed here were based on assignment of items to scales conducted only by this author which results in greater potential for the judgment of this one rater to be biased or go astray of the constructs to be measured. Further development and refinement of the YSR-FFM scales should utilize multiple expert raters. Additionally, the YSR-FFM scales might be further refined through application of Item Response Theory (IRT) methods which can improve measurement accuracy and reliability. In this connection, and notwithstanding possible limitations of short scales for individually applied purposes, Lambert., *et al*. [68] found that only 27 YSR items had good discrimination, the remaining items having poor or mediocre discrimination estimates. They concluded that psychometrically sound information could be obtained from much fewer items on the YSR. Additionally, multiple methods ideally should be utilized since reliance on self-reports alone increases the potential for common method bias.

Among the strengths of this study is that primary analyses were based on the normative YSR data set which supports generalizability of the current results. Four of the five CFA models developed here, while requiring re-specification to achieve good model fit, demonstrated full metric invariance for gender, and the fifth model demonstrated partial metric invariance. However, only one model (Atypical Experiences) demonstrated full scalar invariance and another model (Detachment) demonstrated partial scalar invariance. Scalar invariance is required in order to compare the means of the construct across gender groups and implies that individuals with the same latent construct score would have the same observed scores regardless of the gender group [57]. The nonequivalence of the intercepts may be due to actual differences in the mean levels between genders on the other YSR-FFM scales, which currently limits comparisons of mean scores between males and female youth on those scales. Also, some of the item factor loadings were not strong indicators of their designated factor; however, such items would not be used as single measures.

While preliminary findings in the development of FFM domains from the YSR item pool are encouraging, the trait domains do not include facet scales at this time. Particularly for the shortest scales, it might be difficult to identify coherent facet scales. However, it is possible that either facet scales or "nuance" scales might be developed for all scales. Oltmanns and Widiger [69] developed a Five-Factor Personality Inventory for ICD-11; the ICD-11 contains only domain scales and not facet scales. The FFiCD facet-level measure for the ICD-11 trait model developed by Oltmanns and Widiger includes both facets and shorter "nuance" scales to provide an even more specific and homogeneous assessment of personality. The nuance scales, which are based on convergence of items, are composed of either two or three items. Hence, upon further study, facet and/or nuance scales might be developed for the YSR-FFM measure.

Future study to develop proxy FFM scales from the ASEBA forms might include the Child Behavior Checklist and the Teacher Report Form both of whose item pools mostly overlap with the YSR item pool. As the ASEBA is intended to include youth, parent, and teacher forms, given that each observer/rater may provide unique information about youths, it would be advantageous to have separate personality scales that would similarly permit comparisons between different sources of information concerning youths.

## Conclusion

In conclusion, this study represented a preliminary effort to explore the feasibility of developing FFM scales based on the item content of the ASEBA Youth Self-Report. A potentially significant benefit of having such scales available is that the existence of large archival databases of ASEBA Youth Self-Reports could provide opportunities for both retrospective as well as concurrent and prospective study of personality in youth through use of a valid personality measure derived from the Youth Self-Report. Notwithstanding serious methodological limitations that will need to be addressed going forward, some preliminary findings are encouraging concerning development of meaningful personality trait scales from this frequently used measure of emotional and behavioral problems in youth. Although the focus of this study concerns personality trait measurement, ultimately the purposes of personality measures are prevention, early intervention, and treatment. Shiner and Allen [7] opined that the AMPD model can facilitate identification of treatment targets, e.g., negative affectivity, disinhibition, as well as impairment in areas of identity, self-direction, empathy, or intimacy, so that treatment modalities can be matched to those targets, such as cognitive treatments designed to improve mood regulation. Shiner and Allen [7] also noted that adolescent's strengths may be identified for example by features of low Antagonism and low Disinhibition, which can promote resilience, a sense of encouragement and optimism for youths, their caregivers, and treatment providers.

## Bibliography

1. Caspi A., *et al*. "Personality development: Stability and change". *Annual Review of Psychology* 56 (2005): 453-484.

2. De Clercq B., *et al*. "The Hierarchical structure and construct validity of the PID-5 trait measure in adolescence". *Journal of Personality* 82 (2014): 158-169.

3. Kongerslev MT., *et al*. "Personality disorder in childhood and adolescence comes of age: a review of the current evidence and prospects for future research". *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology* 3 (2015): 31-48.

4. De Pauw SSW., *et al*. "How are traits related to problem behavior in preschool children?: Similarities and contrasts between temperament and personality". *Journal of Abnormal Child Psychology* 37 (2009): 309-325.

5. Shiner RL. "Mapping the landscape of personality in childhood and adolescence". *Social and Personality Psychology Compass* 4 (2010): 1084-1097.

6. Shiner RL and Tackett JL. "Personality disorders in children and adolescents". In Marsh EJ and Barkley RA. (Editions.), Child Psychopathology (3rd edition.). New York, NY: Guilford Press. (2014): 848-896.

7. Shiner RL and Allen TA. "Assessing personality disorders in adolescents: Seven guiding principles". *Clinical Psychology Science and Practice* 20 (2013): 361-377.

8. Tackett JL., *et al*. "The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence". *Journal of Personality* 80 (2012): 847-879.

9. Johnson JG., *et al*. "Personality disorders". In Wolfe DA and Mash EJ (Editions.), Behavioral and emotional disorders in adolescents: Nature, assessment, and treatment. New York, NY: Guilford Publications. (2006): 463-484.

10. Feenstra DJ., *et al*. "Prevalence and comorbidity of Axis I and Axis II disorders among treatment refractory adolescents admitted for specialized psychotherapy". *Journal of Personality Disorders* 25 (2011): 842-850.

11. Grillo CM., *et al*. "Frequency of personality disorders in two age cohorts of psychiatric inpatients". *American Journal of Psychiatry* 155 (1998): 140-142.

12. Kaszynski K., *et al*. "Incarcerated youth with personality disorders: Prevalence, comorbidity and convergent validity". *Personality and Mental Health* 8 (2013): 42-51.

13. Johnson JG., *et al*. "Age-related change in PD trait levels between early adolescence and adulthood: A community-based longitudinal investigation". *Acta Psychiatrica Scandinavica* 102 (2000): 265-275.

14. Newton-Howes G., *et al*. "Personality disorder across the life course". *The Lancet* 385 (2015): 727-734.

15. Achenbach TM and Rescorla LA. "Manual for the ASEBA School-Age Forms and Profiles". Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families (2001).

16. Cohen P., *et al*. "The children in the community study of developmental course of personality disorder". *Journal of Personality Disorders* 19 (2005): 466-486.

17. Links PS., *et al*. "The relationship of personality disorders and Axis I clinical disorders". In Widiger TA. (Edition.), Oxford handbook of personality disorders. New York: Oxford University Press (2012): 237-259.

18. Malouff JM., *et al*. "The relationship between the five-factor model of personality and symptoms of clinical disorders: A meta-analysis". *Journal of Psychopathology and Behavioral Assessment* 27 (2005): 101-114.

19. Widiger TA. "The DSM-5 dimensional model of personality disorder: Rationale and empirical support". *Journal of Personality Disorders* 25 (2011): 222-234.

20. Lynam DR and Widiger TA. "Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach". *Journal of Abnormal Psychology* 110 (2001): 401-412.

21. Trull TJ. "The five-factor model of personality disorder and DSM-5". *Journal of Personality* 80 (2012): 1697-1720.

22. American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (5th edition., DSM-5). Washington, DC: The American Psychiatric Association (2013).

23. Krueger RF., *et al*. "The Personality Inventory for DSM-5 (PID-5)-Adult". Washington, DC: American Psychiatric Association (2013).

24. Penney SR., *et al*. "Structural validity of the MACI psychopathy and narcissism scales: Evidence of multidimensionality and implications for use in research and screening". *Journal of Clinical Child and Adolescent Psychology* 37 (2008): 422-433.

25. Murrie DC and Cornell DG. "Psychopathy screening of incarcerated juveniles: A comparison of measures". *Psychological Assessment* 14 (2002): 390-396.

26. Vrouva I. "The Achenbach System of Empirically Based Assessment as a measure of personality disorder traits in children and adolescents". (Doctoral Thesis), University College: London, UK (2012).

27. Semel RA. "Mean score comparisons on the ASEBA Youth-Self Report between a sample of adjudicated delinquent youths evaluated for legal disposition and the normative sample of non-clinically referred male adolescents". *EC Psychology and Psychiatry* 9.2 (2020): 01-09.

28. Flanagan R. "Review of the Achenbach System of Empirically Based Assessment". In: Spies RA and Plake BS (Editions), The sixteenth mental measurements yearbook. Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from Mental Measurements Yearbook with Tests in Print database (2005).

29. Watson TS. "Review of the Achenbach System of Empirically Based Assessment". In: Spies RA and Plake BS (Editions.), The sixteenth mental measurements yearbook. Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from Mental Measurements Yearbook with Tests in Print database (2005).

30. Jesness CF. "Jesness Inventory Revised". Technical Manual. North Tonawanda, NY: MHS (2003).

31. Mullins-Sweatt SN., *et al*. "Psychometric properties of an abbreviated instrument of the five-factor model". *Assessment* 13 (2006): 119-137.

32. Costa PT and McCrae RR. "Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual". Odessa, FL: Psychological Assessment Resources (1992).

33. Goldberg LR. "A broad-bandwidth, public domain, personality inventory measuring the lower-order facets of several five-factor models". In Mervielde I., *et al*. (Editions.), Personality Psychology in Europe (Volume 7), Tilburg, The Netherlands: Tilburg University Press (1999): 7-28.

34. Goldberg LR., *et al*. "The International Personality Item Pool and the future of public-domain personality measures". *Journal of Research in Personality* 40 (2006): 84-96.

35. DeYoung CG., *et al*. "Between facets and domains: 10 aspects of the big five". *Journal of Personality and Social Psychology* 93 (2007): 880-896.

36. Trull TJ. "The Five-Factor Model of personality and DSM-5". *Journal of Personality* 80 (2012): 1697-1720.

37. Widiger TA. "The DSM-5 dimensional model of personality disorder: Rationale and empirical support". *Journal of Personality Disorders* 25 (2011): 222-234.

38. Crego C and Widiger TA. "The convergent, discriminant, and structural relationship of the DAPP-BQ and SNAP with the ICD-11, DSM-5, and FFM trait models". *Psychological Assessment* 32 (2020): 18-28.

39. McCrae RR and Costa PT. "Personality in adulthood: A five-factor theory perspective (2nd edition)". New York: Guilford Press (2003).

40. Nuzum H., *et al*. "Affect, behavior, and cognition in personality and functioning: An item-content approach to clarifying empirical overlap". *Psychological Assessment* 31 (2019): 905-912.

41. Pytlik Zillig LM., *et al*. "What do we assess when we assess a Big 5 trait? A content analysis of the affective, behavioral, and cognitive processes represented in Big 5 personality inventories". *Personality and Social Psychology Bulletin* 28 (2002): 847-858.

42. Paunonen SV. "Big five factors of personality and replicated predictions of behavior". *Journal of Personality and Social Psychology* 84 (2003): 411-424.

43. Mullins-Sweatt SN and Widiger TA. "The Five-Factor Model of personality disorder: A translation across science and practice". In Krueger RF and Tackett JL (Editions.), Personality and Psychopathology. New York: Guilford Press. (2006): 39-70.

44. Crego C and Widiger TA. "The convergent, discriminant, and structural relationship of the DAPP-BQ and SNAP with the ICD-11, DSM-5, and FFM trait models". *Psychological Assessment* 32 (2020): 18-28.

45. IBM Corp. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp (2019).

46. Arbuckle JL. Amos 23.0 User's guide. Chicago: IBM SPSS. (2014).

47. Marsh HW and Hocevar D. "Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups". *Psychological Bulletin* 97 (1985): 562-582.

48. Kline RB. "Principles and practice of structural equation modeling (3rd edition)". New York: Guilford Press (2011).

49. Byrne BM. "Structural equation modeling with Amos: Basic concepts, applications, and programming". New York: Routledge (2016).

50. Bentler PM. "Comparative fix indexes in structural models". *Psychological Bulletin* 107 (1990): 238-246.

51. Steiger JH and Lind JC. "Statistically based tests for the number of common factors". Paper presented at the Psychometric Society Annual Meeting, Iowa City, IA (1980).

52. Joreskog KG and Sorbom D. "LISREL 8: Structural equation modeling with the SIMPLIS command language". Chicago, IL: Scientific Software International (1993).

53. Hu L-T and Bentler PM. "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives". *Structural Equation Modeling* 6 (1999): 1-55.

54. Meredith W. "Measurement invariance, factor analysis, and factorial invariance". *Psychometrika* 58 (1993): 525-543.

55. Cheung GW and Rensvold RB. "Evaluating goodness-of-fit indexes for testing measurement invariance". *Structural Equation Modeling* 9 (2002): 233-255.

56. Chen FF. "Sensitivity of goodness of fit indexes to lack of measurement invariance". *Structural Equation Modeling* 14 (2007): 464-504.

57. Brown TA. "Confirmatory factor analysis for applied research" (2nd edition.). New York, NY: Guilford Press.

58. Donahue BH. "The effect of partial measurement invariance on prediction". (PhD Thesis). University of Georgia, Athens (2006).

59. Vandenberg RJ and Lance CE. "A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research". *Organizational Research Methods* 3 (2000): 4-70.

60. Gore WL., *et al*. "The home for histrionism". *Personality and Mental Health* 5 (2011): 57-72.

61. Achenbach TM and Rescorla LA. "Manual for the ASEBA Adult Forms and Profiles". Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families (2003).

62. Piovesana A and Senior G. "How small is big: Sample size and skewness". *Assessment* 25 (2018): 793-800.

63. Weiner IB and Greene RL. "Handbook of personality assessment". Hoboken, NJ: Wiley and Sons (2008).

64. Sheng Y and Sheng Z. "Is coefficient alpha robust to nonnormal data?". *Frontiers in Psychology* 3 (2012): 1-13.

65. Haladyna TM and Rodriguez MC. "Developing and validating test items". New York, NY: Routledge. (2013).

66. Teplin LA., *et al*. "Psychiatric disorders in youth in juvenile detention". *Archives of General Psychiatry* 59 (2002): 1133-1143.

67. Kendall AD., *et al*. "Psychometric properties of the Aggressive Behaviors scale from the Youth Self-Report in juvenile offenders". *Journal of Personality Assessment* 102 (2020): 480-487.

68. Lambert MC., *et al*. "Is it prudent to administer all items for each Child Behavior Checklist cross-informant syndrome? Evaluating the psychometric properties of the Youth Self-Report dimensions with confirmatory factor analysis and item response theory". *Psychological Assessment* 15 (2003): 550-568.

69. Oltmanns JR and Widiger TA. "The Five-Factor Personality Inventory for ICD-11: A facet-level assessment of the ICD-11 trait model". *Psychological Assessment* 32 (2020): 60-71.

**Volume 9 Issue 9 September 2020**