Published in final edited form as:
J Res Pers. 2018 April ; 73: 180–188. doi:10.1016/j.jrp.2018.02.003.

Psychopathic features across development: Assessing longitudinal invariance among Caucasian and African American youths

Samuel W. Hawes[a,*], **Amy L. Byrd**[b], **Shannon E. Kelley**[c], **Raul Gonzalez**[a], **John F. Edens**[c], **Dustin A. Pardini**[d]

[a]Florida International University, Department of Psychology, Center for Children and Families, United States

[b]University of Pittsburgh, Department of Psychiatry, United States

[c]Texas A&M University, Department of Psychology, United States

[d]Arizona State University, School of Criminology and Criminal Justice, United States

Abstract

Objectives—Psychopathy is associated with severe forms of antisocial and violent behavior in adults. There is also a rapidly growing body of research focused on extending features of adult psychopathy downward to youth. To date however, the degree to which these features can be consistently and comparatively assessed at these younger ages, remains unclear. This study addresses this issue by investigating measurement invariance of underlying features of psychopathy across childhood and adolescence in a racially diverse sample of youth.

Methods—Three cohorts of youth (n =1517) were assessed annually from childhood to adolescence (ages ~7–16). Underlying features of psychopathy commonly assessed in youth (e.g. lack of guilt, impulsivity) were examined within a longitudinal bi-factor framework using multi-dimensional item-response theory (IRT) techniques. Differential item functioning was used to assess invariance across development and participant's race (African-American and Caucasian), using two distinct approaches: (1) traditional item-response theory (IRT) methods; and (2) a recently developed Bayesian structural equation modeling (BSEM) approach.

Results—Psychopathy features assessed in this study exhibited measurement consistency across development (~ages 7–16) and were found to tap into the same underlying construct as intended across measurement occasions, and equivalently for African-American and Caucasian youth. Results were similar when assessed using traditional IRT procedures for longitudinal invariance testing and when implementing the more recent BSEM methodology.

Conclusions—Findings provide the first evidence that features of psychopathy can be assessed consistently in youth and improve our understanding of important developmental and sociocultural factors associated these features during earlier periods of development.

[*]Corresponding author at: 11200 SW 8th Street, Office #434, Miami, FL 33139, United States. Shawes@fiu.edu (S.W. Hawes).
Appendix A. Supplementary material
Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jrp.2018.02.003.

Keywords

Psychopathy; Longitudinal; Measurement invariance; Development; Bayesian

1. Introduction

Adult features of psychopathy, which are characterized by callousness, shallow affect, a lack of remorse, and irresponsibility, have been linked to chronic and severe forms of violent behavior and criminal recidivism (Kahn, Byrd, & Pardini, 2012; Neumann & Hare, 2008). More recently, a rapidly growing body of literature has begun to focus on delineating similar features in youth. While this area of research has made substantial inroads, several key issues are still in need of being addressed. One notable limitation of existing research is the paucity of studies to have examined whether there is continuity in the assessment of these features across distinct phases of development. In addition, there remains a significant gap in our knowledge of potential differences in how these features are measured across different races and ethnicities, particularly during early periods of development. Advancing this area of research is essential, as currently, this gap in literature constrains the general conclusions that can be inferred from currently published studies. Indeed, establishing measurement consistency is considered a critical priority across a number of scientific fields, due to the potential of measurement inconsistency to obscure study results and lead to spurious conclusions (Borsboom, 2006).

1.1. Establishing measurement invariance

Measurement invariance is used by researchers to determine whether or not a measure indexes a given latent construct equivalently over time and between groups. A first step in testing measurement invariance requires establishing a basic structural model. It is worth noting here that unlike most other areas of psychopathology research, significant (and at times contentious) debate remains regarding fundamental conceptual and structural models of psychopathy (Lilienfeld, Watts, Francis Smith, Berg, & Latzman, 2015). Indeed, many major researchers in the field continue to be sharply divided as to what precisely constitutes the core underlying features of psychopathy, and whether these features should be modeled as a unidimensional, multidimensional, or higher-order construct. To this point, a range of hypothesized models have been proposed, consisting of anywhere from two to eight separate factors (e.g., Cooke & Michie, 2001; Neumann, Hare, & Newman, 2007; Vitacco, Neumann, & Jackson, 2005). Highlighting but a few differences across these models shows that some conceptualizations emphasize the assessment of distinct personality and behavioral dimensions (e.g., Hare, 2003; Neumann et al., 2007), some include content that taps into potentially adaptive features (e.g., Patrick, Fowles, & Krueger, 2009), while others focus on personality factors and exclude items intended to assess overt antisocial behaviors (e.g., Cooke & Michie, 2001; Lynam, 2002). In particular, while some researchers suggest that antisocial behaviors are an integral component of psychopathy (Hare & Neumann, 2010), others have expressed concerns that these behaviors represent a sequelae, rather than core component of the syndrome (Skeem & Cooke, 2010a). In addition, some research suggests that personality based conceptualizations may provide greater insight into psychopathy's developmental precursors, noting that some behavioral manifestations are less likely to emerge during earlier periods of development (Dotterer et al., 2017; Frick, Ray, Thornton, & Kahn, 2014) (for a more thorough tx of these issues, see Lilienfeld, Watts, Francis Smith, Berg, & Latzman, 2014).

Lack of consensus regarding psychopathy's underlying features notwithstanding, there does appear to be *relative* agreement that commonly assessed sub-dimensions (e.g., callousness, impulsive/irresponsible, manipulative/deceitful) tap into an overarching psychopathy factor. Considerable research shows that it is this overarching factor that confers the greatest risks, as youth with the highest overall psychopathy scores tend exhibit the most severe and violent forms of delinquent behavior (Lynam, Miller, Vachon, Loeber, & Stouthamer-Loeber, 2009), the worst treatment outcomes (Spain, Douglas, Poythress, & Epstein, 2004), and are at increased risk for displaying psychopathic personality features into adulthood (Lynam et al., 2009). Recently, several studies have provided evidence indicating that bifactor models, which posit that items on a measure are simultaneously influenced by general (e.g., psychopathy) and specific factors (e.g., callousness, impulsivity), can be useful for integrating distinct features of psychopathy into a basic structural model (e.g., Dotterer et al., 2017; Hawes, Mulvey, Schubert, & Pardini, 2014; Patrick, Hicks, Nichol, & Krueger, 2007; Waller, McCabe, Dotterer, Neumann, & Hyde, 2017). In particular, this approach can help to reconcile findings that show dimensions of psychopathy to be strongly influenced by genetic factors, while also exhibiting unique associations with theoretically meaningful constructs, such as negative emotionality (Bezdjian, Raine, Baker, & Lynam, 2011; Forsman, Lichtenstein, Andershed, & Larsson, 2008).

1.2. Longitudinal measurement invariance

A conceptual and methodological concern when assessing features of psychopathy across development is whether certain features may become more or less indicative of the overarching psychopathy construct over time (Obradović, Pardini, Long, & Loeber, 2007; Salekin & Frick, 2005; Seagrave & Grisso, 2002). When scale scores designed to index a construct change in a systematic manner over time, this is commonly interpreted as capturing "true" developmental change. However, in the absence of having established measurement invariance, change may merely reflect temporal inconsistencies among indicators used to assess the construct, rather than reflecting actual "developmental" change. For example, bifactor models produce item factor loadings and thresholds (when using categorical data) that correspond to item difficulty and discrimination parameters from multidimensional item response theory (MIRT) models (Muthén & Asparouhov, 2013b). Establishing longitudinal measurement invariance (referred to as differential item functioning [DIF] in IRT parlance) for these models requires ascertaining that item loading and threshold parameters remain consistent across development. To date, existing research that has conducted longitudinal invariance testing with items used to assess features of psychopathy is extremely limited, especially during early periods of development.

1.3. Invariance across race

The degree to which measures used to render diagnoses and assess psychological constructs generalize across different subgroups of individuals can have important clinical, legal, and policy implications (Skeem, Edens, Camp, & Colwell, 2004). This is certainly true for the construct of psychopathy given its widespread use for informing legal decisions (Edens, Colwell, Desforges, & Fernandez, 2005; Jones, 2008) and it holds even more true among racial and ethnic minorities, given their disproportionate involvement with the criminal justice system. Additionally, public perceptions that psychopathy is more prevalent among minority group members (particularly African American men) may further exacerbate these disparities (Skeem et al., 2004), making it critical that instruments used to assess this construct do so in a precise and consistent fashion across these groups.

To date, a limited but accumulating body of evidence suggests that, in general, psychopathy and its underlying features can be assessed equivalently across African American and Caucasian adults. In a meta-analysis of race differences on the PCL family of measures, Skeem et al. (2004) found that levels of psychopathy were equivalent for African Americans and Caucasians, although there was considerable heterogeneity in reported effect sizes among some features, most notably among factor 2 items (e.g., social deviance/antisocial lifestyle). In addition, research also provides evidence of structural invariance across African Americans and Caucasians for the PCL family of measures (Jackson, Neumann, & Vitacco, 2007; Jones, Cauffman, Miller, & Mulvey, 2006; Neumann & Hare, 2008; Skeem, Mulvey, & Grisso, 2003; Vitacco et al., 2005) and that they discriminate similarly across these groups (Cooke & Michie, 2001).

Yet, not all findings have provided support for measurement equivalence across these groups. In particular, research comparing African American and Caucasian samples has shown modest to substantial differences in item slope and threshold estimates (Cooke & Michie, 2001; Tsang, Piquero, & Cauffman, 2014) and differences in underlying factor structures of psychopathy have been reported as well (Kosson, Smith, & Newman, 1990; Sullivan & Kosson, 2006). For example, Tsang et al. (2014) found differential item functioning of estimated thresholds in 10 of the 20 PCL:YV items among Caucasian and African American boys, and further noted four items that were more sensitive to change (i.e., had larger estimated slopes) for Caucasian compared to African American participants. In addition, it is important to point out that the overwhelming majority of this work has focused on adult correctional and forensic psychiatric samples, using PCL based measures. This has led researchers to highlight the need to expand this area of investigation to community-based samples (Skeem et al., 2004), and across a diversity of assessment tools (Skeem & Cooke, 2010b). Finally, and perhaps most relevant to the current study, extant research to examine these issues among youth is exceedingly sparse, and none have used longitudinal assessments in order to situate findings in a developmentally informative context.

1.4. Current study

This study aims to identify whether features of psychopathy commonly assessed in youth can be measured consistently across childhood and adolescence, and equivalently between African American and Caucasian youth. To investigate these issues, three grade-based cohorts of boys (n = 1517) were prospectively followed and assessed annually from childhood to adolescence (spanning ~ages 7–16). Longitudinal and multi-group invariance testing was carried out using two distinct approaches: (1) a traditional IRT procedure for investigating differential item functioning; and (2) a more recently developed BSEM approach for assessing measurement invariance.

2. Method

2.1. Design and participants

Data were collected as part of the Pittsburgh Youth Study (PYS), a longitudinal investigation aimed at understanding the development of delinquency, substance use, and mental health problems among a community sample of boys (Loeber, Farrington, Stouthamer-Loeber, & Van Kammen, 1998).[1] The PYS consists of three cohorts recruited from the Pittsburgh Public Schools in the 1st, 4th, and 7th grades. A screening assessment was conducted on a random sample of boys selected from a list of enrolled students in each grade. The screening used parent-, teacher-, and youth-report measures of conduct problems to create a risk index for each grade. Boys rated in the top 30% of the grade-based risk index, as well as a randomly selected equal number of boys from the remaining 70%, were selected for longitudinal follow-up. The resulting sample is 1517 boys across three cohorts (youngest, $n = 503$; middle, $n = 508$; oldest, $n = 506$). All procedures were reviewed and approved by the Institutional Review Board at the University of Pittsburgh. At each assessment, the participant's parent or guardian signed an informed consent, and the child was provided with an opportunity to assent or decline participation. For more details about the sample selection procedures, see Loeber et al. (1998).

Mean ages at screening for the youngest, middle, and oldest cohorts were 6.96 ($SD = 0.55$), 10.25 ($SD = 0.79$), and 13.38 ($SD = 0.79$), respectively. The racial/ethnic composition of each cohort was primarily Caucasian (40.6–42.7%) and African American (52.4–55.7%), with a small percent of Hispanic (0.2%), Asian (0.4–1.0%), or mixed race/ethnicity (2.4–3.8%) participants. Boys in the screening sample did not differ from the follow-up sample in terms of race, California Achievement Test reading scores, single parent household, or parent education (for further details see Loeber et al., 1998).

2.2. Procedures

The current study used data collected at 1-year intervals beginning at the screening assessment, which included 8 assessments for the youngest cohort (~ages 7–14), 4 assessments for the middle cohort (~ages 10–13) and 4 assessments for the oldest cohort (~ages 13–16).

2.3. Measures

2.3.1. Youth psychopathic features—Annual assessments of psychopathic features in youth were collected using a short-form of the teacher-reported Childhood Psychopathy Scale (CPS-SF; Lynam, 1997; Lynam et al., 2009). The CPS, which was originally constructed using items from the Childhood Behavior Checklist (CBCL; Achenbach, 1991) and the Common-Language Q-sort (CCQ; Caspi et al., 1992), was developed to serve as a relatively pure measure of personality and therefore does not include items that tap into

[1]As per requirements by the Journal of Research in Personality, we note that this study was not pre-registered.

more overt antisocial behaviors (Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007). Across a number of studies, the CPS has demonstrated associations with other measures commonly used to assess psychopathy, including the Antisocial Process Screening Device (ASPD; Frick & Hare, 2001) (*r*'*s* = 0.57–0.61; Bijttebier & Decoene, 2009), the Triarchic Psychopathy Measure (TriPM; Patrick et al., 2009) ($r = 0.38$; Drislane, Patrick, & Arsal, 2014), the Inventory of Callous- Unemotional Traits (ICU; Frick, 2004) ($r = 0.49$; Roose, Bijttebier, Claes, Decoene, & Frick, 2009) and the PCL-R across an 11-year span ($r = 0.31$; Lynam et al., 2007). In addition, the CPS has been shown to be predictive of theoretically relevant constructs including oppositional defiant disorder and conduct disorder (*r*'s = 0.53–0.70; Bijttebier & Decoene, 2009), criminal history over a 13-year span (Lynam et al., 2009), recidivism and poor treatment outcomes in adolescence (Falkenbach, Poythress, & Heide, 2003), and patterns of electrodermal hyporesponsivity in line with other psychopathy measures (Fung et al., 2005).

A shortened, 18-item version of the CPS was developed using items from the extended version of the CBCL (Lynam et al., 2009). This shortened version has demonstrated evidence of construct validity and reliability across childhood and adolescence, strong correlations with the original CPS ($r = 0.91$), and moderate to strong rank-order stability across childhood and adolescence (*r*'s - 0.49–0.80) (Lynam et al., 2009). Use of teacher-report data has been used to assess psychopathy in a number of studies and exhibits correlates similar to those seen with self- and parent-report (for review see Frick et al., 2014). As teachers in the current study completed the standard CBCL Teacher Report Form (TRF; Achenbach & Edelbrock, 1986), rather than the extended version, 4 items from the original were unavailable.[2] Of the 14 CPS items used in this study, 9 map onto the interpersonal/affective dimension of psychopathy (e.g., 'doesn't feel guilty', 'manipulates') and 5 map onto an impulsive/behavioral dimension (e.g., 'behaves irresponsibly', 'impulsive'; see Table 2 for a complete listing). All items were rated on a 3-point scale from 0 (not true) to 2 (very true). Internal consistency in the current study remained high across each follow-up assessment for the CPS total (α range = 0.93–0.96) and factor scores (α range = interpersonal/affective factor − 0.90–0.94; impulsive/behavioral factor − 0.87–0.89).

2.4. Data analysis plan

Two distinct approaches for examining measurement invariance were implemented using the Mplus 7.2 software package (Muthén & Muthén, 1998–2012). First, a traditional invariance testing approach was conducted using a mean and variance adjusted weighted least squares estimator (WLSMV) for ordinal items. The use of WLSMV for these models is statistically equivalent to an IRT model (Reise, Horan, & Blanchard, 2011; Wirth & Edwards, 2007). However, we report results using factor analytic terminology (thresholds and loadings) rather than IRT parameterization (discrimination and difficulty), as this framework is considered to be more familiar to readers (Reise et al., 2011). Formulas for converting between the two parametrizations can be found in Muthén and Asparouhov (2013a).

[2]The four unavailable items from the extended CBCL teacher report form were 'takes credit from others accomplishments', 'rarely or never saves money, 'borrows and doesn't pay back', 'blames others excessively').

Configural invariance requires the same underlying factor structure to adequately fit the measure's items across time/groups. We examined the fit of a bifactor model, with the general factor and both underlying dimension specified to be orthogonal to one another. This model was compared to a unidimensional and correlated 2-factor model. Several statistical indices (i.e., omega coefficient, omega hierarchical, factor determinacy, explained common variance, and percentage of uncontaminated correlations) were also calculated to allow for further evaluation of the bifactor model, in addition to information provided by indices of model fit (for a detailed review, see Bonifay, Lane, & Reise, 2017; Rodriguez, Reise, & Haviland, 2016). Subsequent to establishing configural invariance of the bifactor model, we investigated scalar invariance by allowing the loadings and thresholds of the same item to be freely estimated across assessments and contrasting this with a more constrained model in which loadings and thresholds were held equal across assessments (Horn & McArdle, 1992). These models were assessed using several global fit indices including the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). Cutoff values of 0.90 or greater were used to indicate acceptable fit and 0.95 or greater to indicate good fit for both CFI and TLI (Hu & Bentler, 1999; McDonald & Ho, 2002a). RMSEA values between 0.05 and 0.10 were considered to represent an acceptable fit, while values less than 0.05 were considered to indicate good fit (Browne & Cudeck, 1993; McDonald & Ho, 2002b).

Consistent with traditional invariance testing procedures, competing models were compared using a difference testing approach (DIFFTEST procedure in MPlus 7.2; Muthén & Muthén, 1998–2012), which provides a corrected chi-square difference test for nested models. However, as this method has been shown to be sensitive to sample size and violations of normality (Brannick, 1995), we also compared differences in absolute fit indices among these nested models. Prior research suggests that changes in CFI equal to or less than 0.01, and changes in RMSEA of equal to or less than 0.015 provides evidence in support of invariance (Chen, 2007; Cheung & Rensvold, 2002), although some research suggests using a more stringent criteria (i.e., change in CFI equal to or less than 0.002; Meade, Johnson, & Braddy, 2008).

Some research has noted that this traditional invariance testing suffers from two primary limitations: (1) Chi-square difference testing is overly sensitive to relatively minor differences in model fit when examined using large samples; and (2) empirically-derived standards for examining substantive changes in overall model fit when testing for measurement invariance of ordinal indicators have not been developed (see Meade, Johnson, & Braddy, 2006). To address such issues, Muthén and colleagues have recently developed a Bayesian structural equation modeling (BSEM) approach for testing measurement invariance (Muthén & Asparouhov, 2013a, 2013b). Implementation of this approach is particularly suited for applications that involve high-dimensional latent variable models, such as the present case, where maximum-likelihood estimation proves to be too computationally demanding and has difficulties obtaining precise likelihood values (Muthén & Asparouhov, 2012).

BSEM analysis specifies approximate measurement invariance of all loadings and thresholds across time/groups by specifying small-variance priors for these parameters. That is, the zero-fixed equality constraints placed on loadings and thresholds across time for the traditional approach to invariance testing are replaced by approximate zeros (i.e., 0.01 small-variance priors). In this sense, approximate invariance may be thought of as non-invariance that is too small in magnitude to be of practical significance. The difference between each measurement parameter and that parameters average across time/groups is then estimated.

The methods used to test for invariance using the BSEM as based on procedures described in Muthén and Asparouhov (2012) (also see Muthén and Asparouhov, 2013a, 2013b). Specifically, the posterior distribution of Bayesian estimation was achieved through the Markov chain Monte Carlo (MCMC) algorithm with 50,000 iterations specified. We used thinning to estimate the posterior distribution on the basis of every 10th iteration, rather than every iteration in order deal with potential autocorrelation in the chain (Muthén & Asparouhov, 2012). For BSEM, model fit was assessed via the posterior predictive p value (PPP), which is a measure of model misspecification. Unlike more common indices of model fit used in traditional invariance testing (e.g., RMSEA), there is no clear guideline for PPP values that indicate whether or not a model provides acceptable fit. However, according to simulation research, small positive values (e.g., 0.005) are suggested to be indicative of a poorly fitting model, whereas PPP values closer to 0.5 are considered to demonstrate excellent fit (Muthén & Asparouhov, 2012). As BSEM difference testing in MPlus has not yet been implemented for thresholds of polytomous items, it was necessary to collapse item categories for the BSEM analyses in the current study. This was done by collapsing the lowest base rate category (category 3, 'very true') with category 2 ('somewhat true'), which had the next lowest base rate of endorsement.

Subsequently, we examined multi-group invariance to investigate potential differences between African American and Caucasian participants at each assessment wave. Assessment of multi-group measurement invariance included the specification of an initial model allowing item loadings and thresholds to differ across African American and Caucasian participants. Next, item loadings and thresholds were constrained to equality across these groups. These models were compared using the same guidelines as previously specified for traditional invariance testing and this procedure was conducted at each assessment point, separately for each cohort. Primary study data can be accessed via the Journal of Research in Development's data repository.

3. Results

3.1. Measurement invariance across development

Separate confirmatory factor analyses (CFA) were conducted at each assessment wave for each cohort. Although our primary focus was on establishing fit of an initial bifactor structural model, we also examined and compared the relative fit of a unidimensional and correlated two-factor model at each assessment wave. Results from these analyses indicated that the bifactor model provided adequate to good fit across each assessment for all three cohorts (16 total CFAs; CFI's > 0.98; TLI's > 0.98; RMSEA's < 0.08). Further, while attenuated but adequate fit was found for a unidimensional (CFI's > 0.97; TLI's > 0.97; RMSEA's < 0.10) and correlated two-factor model (CFI's > 0.97; TLI's > 0.96; RMSEA's < 0.09), chi-square difference testing showed the bifactor model to provide significantly better fit than either of these models when compared at each assessment point across all 3 cohorts (*ps* < 0.001). An examination of item loadings across models revealed that after accounting for the variance attributed to the general factor, the item "sudden change in mood or feelings" exhibited a significant negative association with the Impulsive/Behavioral sub-domain. This is not uncommon among such models (Muthén & Asparouhov, 2013b), and in turn this item was specified to load only onto the general factor. Subsequently, a single longitudinal configural invariance model was specified for each cohort by including the bifactor model across each assessment wave in a single model overall model (for each cohort). Within each longitudinal configural invariance model, item loadings and thresholds were allowed to vary across time. This configural invariance model provided a good fit to the data for each cohort (see Table 1).

The unconstrained configural model for each cohort was compared to a more rigorous, longitudinal scalar invariance model (i.e., item loadings and thresholds constrained to equality across assessment waves). Results from chi-square difference testing (Table 1) revealed that for each cohort, the configural model provided better fit than the scalar invariance model (*ps* < 0.001). However, further inspection of absolute fit indices (CFI, RMSEA) revealed almost identical fit for these models. Item loadings and thresholds for the scalar invariant bifactor model's general and subfactors for each cohort are provided in Table 2. Strong loadings onto the general factor indicates that the model did a good job of tapping into the overarching psychopathy factor. In addition, although the sub-factor loadings were of less magnitude than for the general factor (as expected), examination of these loadings shows that useful information is still provided toward these factors.

To further investigate evidence of longitudinal measurement invariance across each of the cohorts, we also examined invariance using a BSEM framework. For the BSEM analyses, PPP values were obtained for the youngest (PPP = 0.11), middle (PPP = 0.12) and oldest (PPP = 0.26) cohorts. Using this framework, a difference between each measurement parameter and the average of that parameter across assessments was estimated. By specifying a small variance prior (0.01) under the assumption that the model parameters at least approximate longitudinal invariance, results from these analyses point to any parameters exhibiting non-invariance. For the youngest cohort, of 304 possible instances of non-invariance (24 item loadings,[3] 14 item thresholds, 8 assessment waves), while for both, the middle and oldest cohorts, there were 152 possible instances of non-invariance (24 item loadings, 14 item thresholds, 4 assessment waves).

BSEM results showed that across cohorts, less than 5% (30 out of the 604) of the model parameters tested exhibited any evidence of potential non-invariance (see Table 3). For the youngest cohort, only 6% (19 out of 304) of the parameters (all item thresholds) were identified as potentially non-variant. Of these, no item was found to display non-invariance at greater than 4 assessment waves and at no assessment wave were there greater than 5 items exhibiting potential non-invariance. For the middle cohort, only 3 parameters were detected as exhibiting evidence of non-invariance (3 item thresholds at the final assessment),

[3]For BSEM analyses, the first item loading for the general and domain specific factors is fixed to 1, resulting in no item loading being estimated for these items.

while the oldest cohort demonstrated 8 total invariant thresholds (3 items at the first assessment, 1 item at the second assessment, 1 item at the third assessment, and 3 items at the final assessment). All potentially non-invariant parameters resulted from item thresholds (opposed to factor loadings), which may reflect differences in the probability of endorsing an item when levels of the underlying construct are equivalent. However, these findings appear to provide little evidence of any meaningful non-invariance, as only a very small number of parameters (<5%) exhibited any non-invariant properties and as outlined above, the ones that did appeared to do so in a non-systematic manner (also see Table 3).

3.2. Measurement invariance across racial groups

Multi-group measurement invariance testing first examined a model with item loadings and thresholds free to differ across African American and Caucasian participants (Model 1). Results indicated that this model fit well at each assessment (see Table 4). Next, item loadings and thresholds were constrained to equality across these groups (Model 2). This model continued to demonstrate good fit and very little difference from the unconstrained model, across cohorts. An inspection of Table 4 shows that none of the models differed according to absolute fit indices and even the more stringent chi-square difference testing results indicated there were no significant differences between the overwhelming majority of these models. These findings were corroborated when using the BSEM approach to investigate measurement invariance across these groups. For these analyses, of the 604 potential instances of non-invariance across all time points for each cohort, results indicated that there were only 3 potential instances of non-invariance (< 1%).

4. Discussion

To gain insight into individual differences in psychological phenomena it is essential that these phenomena be measured comparatively over time or across groups. The expeditious growth of research focused on extending features of adult psychopathy downward to youth, makes it imperative to establish that these features can be assessed with precision and consistency at younger ages. The present study provides the first empirical evidence to show that features of psychopathy can be assessed consistently across childhood and adolescence, and comparably between African American and Caucasian youth. Notably, these results were commensurate when examined using a more traditional ‘item response theory’ approach to invariance testing, as well as when using a ‘Bayesian structural equation modeling’ framework. Findings from this study further our understanding of important developmental and sociocultural factors associated with psychopathy’s underlying features in youth.

Demonstrating longitudinal invariance is particularly important for developmental research, as it helps to ensure that findings of intra-individual change reflect a true change a psychological phenomenon of interest. Results from this study provide support for the idea that psychopathy’s underlying features can be comparatively measured across childhood and adolescence. This helps to address concerns that behavioral manifestations of psychopathy may differ substantially from childhood to adolescence; for example, as a result of poor abstract reasoning skills related to impediments in cognitive development. Although this still may be true in the case of very young children, these findings suggest that features of psychopathy as assessed here remained equally indicative of the construct from ages 7 to 16.

As a psycholegal construct, psychopathy is used to inform a wide-range of important legal decisions, which range from placement into treatment programs to assessing future risk and determining sentence length. Establishing invariance of psychopathy's underlying features across ethnic and racial groups is imperative considering the well-publicized discrepancies in criminal justice involvement among minority groups, particularly African American men (Skeem et al., 2004). These results provide initial evidence supporting the measurement equivalency of psychopathy features between these groups across childhood and adolescence. Interestingly, these results differ somewhat from recent findings presented by Tsang et al. (2014), who demonstrated evidence of DIF on several items used to assess psychopathy between African-American and Caucasian youth. It is worth noting that one of the major differences between that study and the findings presented here is that Tsang et al. (2014) used the PCL:YV to assess psychopathy. Unlike the measure used in the current study, the PCL:YV includes content intended to assess antisocial and delinquent behaviors, and notably, evidence of DIF between African-American and Caucasian youth reported by Tsang et al. (2014) was largely relegated to that subset of items. Of course, it may also be the case that these divergent findings are due, at least in part, to informant differences, as the PCL: YV is scored by raters as opposed to teacher report used in the current study.

4.1. Limitations

The current study was characterized by a number of strengths including a prospective longitudinal design spanning a large window (i.e., 9 years) during an important developmental transition, as well as a racially diverse sample. However, study findings also need to be considered within the context of several limitations. Specifically, the study was comprised of an all-male sample recruited from an urban school district. As such, additional research focused on other populations, such as female and juvenile offender populations, is needed. In addition, features of psychopathy were assessed using a relatively small number of items, which unlike some existing measures, was intentionally developed to omit items tapping conduct problems and criminal behavior (e.g., repeat offending, juvenile delinquency). It will also be important for future studies to further examine associations convergent and divergent associations between the bifactor model used in the current study with other measures of psychopathy and relevant external correlates. Although multiple indices were used to evaluate the bifactor model, it should be noted that these models can be challenging to interpret and can be susceptible to data overfitting. Additionally, although strengths and weaknesses have been associated with the use of distinct types of informant report measures (e.g., teacher, parent, youth), the use of teacher-report in the current study provided a multiple-informant assessment of this construct across time. Nonetheless, the current results should be replicated using alternative or combined informant methods.

4.2. Conclusions

The extension of adult psychopathy to youth has led to a rapid acceleration in prospectively designed research investigating this facet of psychopathology during childhood and adolescence. In turn, it is essential to ascertain that underlying features of psychopathy can be assessed with consistency across earlier development and in a manner that accurately reflects the overarching construct as intended. Findings from this study offer support to the notion that early assessments of psychopathy can produce consistent measurements that allow for comparative observations across development and between youth of different racial backgrounds. Further, and in addition to noted developmental and sociocultural implications, results from this study stress the need for future research aimed at identifying early risk markers of psychopathy in order inform prevention and treatment efforts during earlier periods of development.[4]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/support

Grant support provided by the National Institute of Mental Health (MH51091-01A1, MH094467, MH48890, MH50778, MH078039-01A1); National Institute on Drug Abuse (DA411018); and the Office of Juvenile Justice and Delinquency Prevention (96-MU-FX-0012).

References

Achenbach TM (1991). Manual for the child behavior checklist/4–18 and 1991 profiles. Burlington: University of Vermont, Department of Psychiatry.

Achenbach TM, & Edelbrock C (1986). Manual for the teacher's report form and teacher version of the child behavior profile. Burlington: University of Vermont, Department of Psychiatry.

Bezdjian S, Raine A, Baker L, & Lynam D (2011). Psychopathic personality in children: genetic and environmental contributions. Psychological Medicine, 41 (03), 589–600. [PubMed: 20482945]

Bijttebier P, & Decoene S (2009). Assessment of psychopathic traits in children and adolescents further validation of the antisocial process screening device and the childhood psychopathy scale. European Journal of Psychological Assessment, 25(3), 157–163. 10.1027/1015-5759.25.3.157.

Bonifay W, Lane SP, & Reise SP (2017). Three concerns with applying a bifactor model as a structure of psychopathology. Clinical Psychological Science, 5(1), 184–186.

Borsboom D (2006). When does measurement invariance matter? Medical Care, 44 (11), S176–S181. [PubMed: 17060825]

Brannick MT (1995). Critical comments on applying covariance structure modeling. Journal of Organizational Behavior, 16(3), 201–213.

Browne MW, & Cudeck R (1993). Alternative ways of assessing model fit In Bollen KA & Long JS (Eds.), Testing structural equation models (pp. 136–162). Newbury Park, CA: Sage.

Caspi A, Block J, Block JH, Klopp B, Lyman D, Moffitt TE, & Stouthamer-Loeber M (1992). A "common language" version of the California child Q-set for personality assessment. Psychological Assessment, 4(4), 512–523.

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling, 14(3), 464–504.

Cheung GW, & Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233–255.

Cooke DJ, & Michie C (2001). Refining the construct of psychopathy: Towards a hierarchical model. Psychological Assessment, 13(2), 171–188. [PubMed: 11433793]

[4]Dr. XXXX and Dr. XXXX each contributed to the study design, analysis, and writing. Dr. XXXX and Dr. XXXX each contributed to manuscript writing. Dr. XXXX contributed to the study design and manuscript writing.

Dotterer HL, Waller R, Cope LM, Hicks BM, Nigg JT, Zucker RA, & Hyde LW (2017). Concurrent and developmental correlates of psychopathic traits using a triarchic psychopathy model approach. Journal of Abnormal Psychology, 126(7), 859. [PubMed: 29106272]

Drislane LE, Patrick CJ, & Arsal G (2014). Clarifying the content coverage of differing psychopathy inventories through reference to the triarchic psychopathy measure. Psychological Assessment, 26(2), 350. [PubMed: 24320762]

Edens JF, Colwell LH, Desforges DM, & Fernandez K (2005). The impact of mental health evidence on support for capital punishment: are defendants labeled psychopathic considered more deserving of death? Behavioral Sciences and the Law, 23(5), 603–625. [PubMed: 16170787]

Falkenbach DM, Poythress NG, & Heide KM (2003). Psychopathic features in a juvenile diversion population: Reliability and predictive validity of two selfreport measures. Behavioral Sciences and the Law, 21(6), 787–805. [PubMed: 14696031]

Forsman M, Lichtenstein P, Andershed H, & Larsson H (2008). Genetic effects explain the stability of psychopathic personality from mid- to late adolescence. Journal of Abnormal Psychology, 117(3), 606–617. 10.1037/0021-843x.117.3.606. [PubMed: 18729612]

Frick PJ (2004). The inventory of callous-unemotional traits. Unpublished rating scale.

Frick PJ, & Hare RD (2001). Antisocial process screening device technical manual. Toronto, Canada: Multi-Health Systems.

Frick PJ, Ray JV, Thornton LC, & Kahn RE (2014). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. Journal of Child Psychology and Psychiatry, 55(6), 532–548. [PubMed: 24117854]

Fung MT, Raine A, Loeber R, Lynam D, Steinhauer SR, Venables PH, & Stouthamer-Loeber M (2005). Reduced electrodermal activity in psychopathyprone adolescents. Journal of Abnormal Psychology, 114(2), 187–196. [PubMed: 15869350]

Hare RD (2003). Hare psychopathy checklist-revised (PCL-R) (second ed.). North Toawanda, NY: Multi-Health Systems Inc.

Hare RD, & Neumann CS (2010). The role of antisociality in the psychopathy construct: Comment on Skeem and Cooke (2010). Psychological Assessment, 22 (2), 446–454. 10.1037/A0013635. [PubMed: 20528070]

Hawes SW, Mulvey EP, Schubert CA, & Pardini DA (2014). Structural coherence and temporal stability of psychopathic personality features during emerging adulthood. Journal of Abnormal Psychology, 123(3), 623. [PubMed: 24978692]

Horn JL, & McArdle JJ (1992). A practical and theoretical guide to measurement invariance in aging research. Experimental Aging Research, 18(3–4), 117–144. [PubMed: 1459160]

Hu L, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1–55.

Jackson RL, Neumann CS, & Vitacco MJ (2007). Impulsivity, anger, and psychopathy: The moderating effect of ethnicity. Journal of Personality Disorders, 21(3), 289–304. [PubMed: 17536941]

Jones S (2008). Juvenile psychopathy and judicial decision making: An empirical analysis of an ethical dilemma. Behavioral Sciences and the Law, 26(2),151–165. [PubMed: 18344169]

Jones s., Cauffman E, Miller JD, & Mulvey E (2006). Investigating different factor structures of the psychopathy checklist: Youth version: Confirmatory factor analytic findings. Psychological Assessment, 18(1), 33–48. [PubMed: 16594810]

Kahn RE, Byrd AL, & Pardini DA (2012). Callous-unemotional traits robustly predict future criminal offending in young men. Law and Human Behavior. 10.1037/b0000003.

Kosson DS, Smith SS, & Newman JP (1990). Evaluating the construct validity of psychopathy in black and white male inmates: Three preliminary studies. Journal of Abnormal Psychology, 99(3), 250–259. [PubMed: 2212275]

Lilienfeld SO, Watts AL, Francis Smith S, Berg JM, & Latzman RD (2014). Psychopathy deconstructed and reconstructed: Identifying and assembling the Personality building blocks of Cleckley's chimera. Journal of Personality.

Lilienfeld SO, Watts AL, Francis Smith S, Berg JM, & Latzman RD (2015). Psychopathy deconstructed and reconstructed: Identifying and assembling the personality building blocks of Cleckley's chimera. Journal of Personality, 83(6), 593–610. [PubMed: 25091380]

Loeber R, Farrington DP, Stouthamer-Loeber M, & Van Kammen WB (1998). Antisocial behavior and mental health problems: Explanatory factors in childhood and adolescence. New York, NY: Routledge.

Lynam DR (1997). Pursuing the psychopath: Capturing the fledgling psychopath in a nomological net. Journal of Abnormal Psychology, 106(3), 425–438. [PubMed: 9241944]

Lynam DR (2002). Fledgling psychopathy: A view from personality theory. Law and Human Behavior, 26(2), 255–259. [PubMed: 11985302]

Lynam DR, Caspi A, Moffitt TE, Loeber R, & Stouthamer-Loeber M (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. Journal of Abnormal Psychology, 116(1), 155–165. [PubMed: 17324026]

Lynam DR, Charnigo R, Moffitt TE, Raine A, Loeber R, & Stouthamer-Loeber M (2009). The stability of psychopathy across adolescence. Development and Psychopathology, 21(4), 1133–1153. 10.1017/S0954579409990083. [PubMed: 19825261]

Lynam DR, Miller DJ, Vachon DD, Loeber R, & Stouthamer-Loeber M (2009). Psychopathy in adolescence predicts official reports of offending in adulthood. Youth Violence and Juvenile Justice, 7(3), 189–207. [PubMed: 22661910]

McDonald RP, & Ho MHR (2002a). Principles and practice in reporting structural equation analyses. Psychological Methods, 7(1), 64–82. 10.1037//1082-989x.7.1.64. [PubMed: 11928891]

McDonald RP, & Ho MR (2002b). Principles and practice in reporting structural equation analyses. Psychological Methods, 7(1), 64–82. 10.1037/1082-989x.7.1.64. [PubMed: 11928891]

Meade AW, Johnson EC, & Braddy PW (2006). The utility of alternative fit indices in tests of measurement invariance. Paper presented at the Academy of management proceedings.

Meade AW, Johnson EC, & Braddy PW (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. Journal of Applied Psychology, 93(3), 568. [PubMed: 18457487]

Muthén BO, & Asparouhov T (2013a). BSEM measurement invariance analysis. Mplus web notes: No. 17 Available online at: <www.statmodel.com>.

Muthén BO, & Asparouhov T (2013b). Item response modeling in Mplus: A multidimensional, multi-level, and multi-timepoint example. Handbook of item response theory: models, statistical tools, and applications, pp. 1–29.

Muthén LK, & Muthén BO (1998–2012). Mplus user's guide, seventh ed. Muthén & Muthén, Los Angeles, CA.

Muthén BO, & Asparouhov T (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. Psychological Methods, 17(3), 313. [PubMed: 22962886]

Neumann CS, & Hare RD (2008). Psychopathic traits in a large community sample: Links to violence, alcohol use, and intelligence. Journal of Consulting and Clinical Psychology, 76(5), 893–899. 10.1037/0022-006x.76.5.893. [PubMed: 18837606]

Neumann CS, Hare RD, & Newman JP (2007). The super-ordinate nature of the psychopathy checklist-revised. Journal of Personality Disorders, 21(2), 102–117. [PubMed: 17492916]

Obradović J, Pardini DA, Long JD, & Loeber R (2007). Measuring interpersonal callousness in boys from childhood to adolescence: An examination of longitudinal invariance and temporal stability. Journal of Clinical Child and Adolescent Psychology, 36(3), 276–292. [PubMed: 17658974]

Patrick CJ, Fowles DC, & Krueger RF (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. Development and Psychopathology, 21(3), 913–938. 10.1017/S0954579409000492. [PubMed: 19583890]

Patrick CJ, Hicks BM, Nichol PE, & Krueger RF (2007). A bifactor approach to modeling the structure of the psychopathy checklist-revised. Journal of Personality Disorders, 21(2), 118–141. [PubMed: 17492917]

Reise SP, Horan WP, & Blanchard JJ (2011). The challenges of fitting an item response theory model to the Social Anhedonia Scale. Journal of Personality Assessment, 93(3), 213–224. [PubMed: 21516580]

Rodriguez A, Reise SP, & Haviland MG (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. Psychological Methods, 21(2), 137. [PubMed: 26523435]

Roose A, Bijttebier P, Claes L, Decoene S, & Frick PJ (2009). Assessing the affective features of psychopathy in adolescence: A further validation of the Inventory of Callous and Unemotional Traits. Assessment. doi: 1073191109344153 [pii]10.1177/1073191109344153.

Salekin RT, & Frick PJ (2005). Psychopathy in children and adolescents: The need for a developmental perspective. Journal of Abnormal Child Psychology, 33 (4), 403–409. [PubMed: 16118988]

Seagrave D, & Grisso T (2002). Adolescent development and the measurement of juvenile psychopathy. Law and Human Behavior, 26(2), 219–239. [PubMed: 11985299]

Skeem JL, & Cooke DJ (2010a). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. Psychological Assessment, 22(2), 433–445. 10.1037/A0008512. [PubMed: 20528069]

Skeem JL, & Cooke DJ (2010b). One measure does not a construct make: Directions toward reinvigorating psychopathy research-Reply to Hare and Neumann (2010). Psychological Assessment, 22(2), 455–459. 10.1037/A0014862. [PubMed: 20528071]

Skeem JL, Edens JF, Camp J, & Colwell LH (2004). Are there ethnic differences in levels of psychopathy? A meta-analysis. Law and Human Behavior, 28(5), 505–527. [PubMed: 15638207]

Skeem JL, Mulvey EP, & Grisso T (2003). Applicability of traditional and revised models of psychopathy to the Psychopathy Checklist: Screening version. Psychological Assessment, 15(1), 41–55. [PubMed: 12674723]

Spain SE, Douglas KS, Poythress NG, & Epstein M (2004). The relationship between psychopathic features, violence and treatment outcome: The comparison of three youth measures of psychopathic features. Behavioral Sciences and the Law, 22, 85–102. [PubMed: 14963882]

Sullivan EA, & Kosson DS (2006). Ethnic and cultural variations in psychopathy. Handbook of psychopathy, 437–458.

Tsang S, Piquero AR, & Cauffman E (2014). An examination of the Psychopathy Checklist: Youth Version (PCL: YV) among male adolescent offenders: An item response theory analysis. Psychological Assessment, 26(4), 1333. [PubMed: 25181395]

Vitacco MJ, Neumann CS, & Jackson RL (2005). Testing a four-factor model of psychopathy and its association with ethnicity, gender, intelligence, and violence. Journal of Consulting and Clinical Psychology, 73(3), 466–476. [PubMed: 15982144]

Waller R, McCabe HK, Dotterer HL, Neumann CS, & Hyde LW (2017). Unique and interactive associations between maltreatment and complex emotion recognition deficits and psychopathic traits in an undergraduate sample. Journal of Personality Disorders, 1–19.

Wirth R, & Edwards MC (2007). Item factor analysis: current approaches and future directions. Psychological Methods, 12(1), 58. [PubMed: 17402812]

Table 1

Fit statistics and model comparisons for traditional longitudinal invariance testing.

	χ^2	*df*	*p*	Absolute Fit Statistics			Difference Test of Relative Fit			
				CFI	TLI	RMSEA	Comp.	$^\wedge\chi^2$	*df*	*p*
Youngest Cohort										
Model 1 (configural)	6403.284	5916	<0.001	0.992	0.992	0.013				
Model 2 (scalar)	6695.867	6182	<0.001	0.992	0.992	0.013	1 vs. 2	346.813	266	<0.001
Middle Cohort										
Model 1 (configural)	1699.414	1414	<0.001	0.993	0.992	0.020				
Model 2 (scalar)	1849.772	1528	<0.001	0.992	0.992	0.020	1 vs. 2	181.456	114	<0.001
Oldest Cohort										
Model 1 (configural)	1534.000	1414	<0.05	0.997	0.996	0.013				
Model 2 (scalar)	1683.352	1527	<0.01	0.996	0.996	0.014	1 vs. 2	178.350	113	<0.001

Notes. $^\wedge\chi^2$= Chi-Square change from Model 1 (configural) to Model 2 (scalar).

Table 2

Item loadings and thresholds for final invariance models.

	Youngest			Middle			Oldest		
	λ	τ^1	τ^2	λ	τ^1	τ^2	λ	τ^1	τ^2
General									
Exaggerates	0.78	0.78	1.58	0.82	0.48	1.32	0.84	0.58	1.31
Lying or cheating	0.81	0.66	1.54	0.82	0.34	1.17	0.85	0.59	1.30
Manipulates people	0.77	0.78	1.52	0.84	0.51	1.28	0.87	0.63	1.33
Fast or smooth talker	0.70	0.73	1.44	0.80	0.41	1.15	0.80	0.48	1.21
Doesn’t seem to feel guilty after misbehaving	0.82	0.35	1.05	0.83	−0.02[a]	0.73	0.89	0.10[†]	0.78
Cruelty, bullying, or meanness to others	0.89	0.60	1.39	0.90	0.35	1.09	0.86	0.54	1.28
Teases a lot	0.81	0.52	1.30	0.84	0.17	0.96	0.80	0.24	0.97
Cannot always trust what he says	0.81	0.59	1.46	0.84	0.21	1.03	0.90	0.40	1.16
Sudden changes in mood or feelings	0.80	0.69	1.48	0.84	0.35	1.17	0.87	0.41	1.21
Impulsive or acts without thinking	0.89	0.12	1.06	0.91	−0.06[a]	0.88	0.88	0.06[a]	0.98
Wants to have things right away	0.82	0.54	1.34	0.87	0.31	1.17	0.85	0.40	1.22
Behaves irresponsibly	0.78	0.51	1.13	0.78	0.23	0.90	0.88	0.25	0.95
Easily frustrated, demands must be met immediately	0.84	0.53	1.22	0.87	0.29	1.12	0.87	0.46	1.22
Behaves explosively and unpredictably	0.85	0.74	1.38	0.93	0.43	1.16	0.92	0.52	1.22
Interpersonal									
Exaggerates	0.27			0.30			0.28		
Lying or cheating	0.42			0.41			0.34		
Manipulates people	0.33			0.33			0.33		
Fast or smooth talker	0.48			0.39			0.35		
Doesn’t seem to feel guilty after misbehaving	0.21			0.16			0.12		
Cruelty, bullying, or meanness to others	0.16			0.11			0.15		
Teases a lot	0.17			0.11			0.16		
Cannot always trust what he says	0.48			0.43			0.39		
Impulsive									
Impulsive or acts without thinking	0.01[a]			0.06[a]			0.08[*]		

	Youngest			Middle			Oldest		
	λ	τ^1	τ^2	λ	τ^1	τ^2	λ	τ^1	τ^2
Wants to have things right away	0.39			0.37			0.37		
Behaves irresponsibly	-0.07^a			0.02^a			0.03^a		
Easily frustrated, demands must be met immediately	0.43			0.40			0.29		
Behaves explosively and unpredictably	0.21			0.11			0.07^a		

Notes. λ = Item Loading; τ^1 = Item Threshold 1; τ^2 = Item Threshold 2; Unless otherwise specified, all loadings significant at $p < .01$.

* $p < .05$.

[a]Non-significant.

Table 3

Assessment age for item thresholds exhibiting significant differences relative to their average (prior variance = 0.01).

	Youngest Cohort Age	Middle Cohort Age	Oldest Cohort Age
Items			
Exaggerates	8,10,14		
Lying or cheating	11		
Manipulates people	7, 8, 12		
Fast or smooth talker	7, 8	13	
Doesn't seem to feel guilty after misbehaving			14, 16
Cruelty, bullying, or meanness to others			13, 16
Teases a lot	7, 8, 12, 14		13, 16
Cannot always trust what he says			
Sudden changes in mood or feelings			
Impulsive or acts without thinking	7, 11	13	
Wants to have things right away			
Behaves irresponsibly	7, 8, 13, 14	13	
Easily frustrated, demands must be met immediately			13, 15
Behaves explosively and unpredictably			

Notes. Only threshold parameters are displayed, as there was no evidence of item loading invariance for any cohort across assessments.

Table 4

Multi-group invariance of psychopathic features between African-American and Caucasian participants.

	Model 1 (configural)		Model 2 (scalar)		Difference test of absolute fit		Difference test of relative fit		
	CFI	**RMSEA**	**CFI**	**RMSEA**	**^CFI**	**^RMSEA**	**^χ^2**	**df**	**p**
Psychopathic Features									
Youngest									
Age 7	0.987	0.065	0.990	0.052	0.003	0.013	34.42	35	0.49
Age 8	0.993	0.053	0.995	0.042	0.002	0.011	36.37	35	0.40
Age 9	0.992	0.059	0.993	0.048	0.001	0.011	39.10	35	0.29
Age 10	0.991	0.067	0.991	0.059	<0.001	0.008	52.95	35	0.03
Age 11	0.992	0.068	0.992	0.061	<0.001	0.007	57.95	35	0.01
Age 12	0.994	0.058	0.996	0.044	0.002	0.014	30.66	35	0.67
Age 13	0.994	0.056	0.995	0.045	0.001	0.011	34.93	35	0.47
Age 14	0.994	0.060	0.996	0.046	0.002	0.014	29.44	35	0.73
Age 15	0.995	0.052	0.996	0.044	0.001	0.008	37.60	35	0.35
Middle									
Age 10	0.995	0.054	0.996	0.041	0.001	0.013	29.36	35	0.74
Age 11	0.992	0.060	0.992	0.052	<0.001	0.008	48.99	35	0.06
Age 12	0.989	0.080	0.989	0.069	<0.001	0.011	53.19	35	0.03
Age 13	0.994	0.056	0.993	0.055	0.001	0.001	65.54	35	0.01
Oldest									
Age 13	0.994	0.058	0.994	0.051	<0.001	0.007	47.85	35	0.07
Age 14	0.995	0.047	0.993	0.047	0.002	<0.001	60.11	35	0.01
Age 15	0.995	0.057	0.995	0.050	<0.001	0.007	44.80	35	0.12
Age 16	0.996	0.051	0.997	0.040	0.001	0.011	34.62	35	0.49

Notes. ^CFI = Change in CFI from Model 1 to Model 2 (all < 0.01 threshold); ^RMSEA = Change in RMSEA from Model 1 to Model 2 (all < 0.015 threshold); ^χ^2 = Chi-Square change from Model 1 to Model 2.