Tilburg University

Longitudinal validation of the Levenson Self-Report Psychopathy (LSRP) Scale in a high-risk Dutch community sample

Wissenburg, Sophie; Garofalo, Carlo; Blokland, Arjan; Palmen, Hanneke; Sellbom, Martin

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Wissenburg, S., Garofalo, C., Blokland, A., Palmen, H., & Sellbom, M. (2022). Longitudinal validation of the Levenson Self-Report Psychopathy (LSRP) Scale in a high-risk Dutch community sample. *Assessment*, *29*(3), 367-384. https://doi.org/10.1177/1073191120975130

Longitudinal Validation of the Levenson Self-Report Psychopathy (LSRP) Scale in a

High-Risk Dutch Community Sample

Abstract

The Levenson Self-Report Psychopathy (LSRP) scale is a self-report measure that can be used to assess psychopathic traits in community samples, and recent research suggested that its three-factor model (Egocentricity, Callousness, and Antisocial) has promising psychometric properties. However, no study to date has validated the LSRP in a longitudinal framework. The present study sought to validate the LSRP scale in a longitudinal design using a sample of Dutch emerging adults (Ns = 970 and 693 at time point 1 and 2, respectively). We assessed longitudinal measurement invariance and the stability of psychopathic traits over an 18-month time period, from age 20 to age 21.6. Furthermore, we replicated and extended findings on the factor structure, reliability, and construct validity of the Dutch LSRP scale. Confirmatory factor analysis revealed that the three-factor model fit the data well. Evidence of partial longitudinal measurement invariance was observed, which means that the Dutch translation of the LSRP scale is measuring an equivalent construct (and overall latent factor structure) over time. Psychopathic traits were relatively stable over time. The three LSRP subscales showed largely acceptable levels of internal consistency at both time points and showed conceptually expected patterns of construct validity and predictive validity, with a few notable exceptions.

Keywords: psychopathy, Levenson Self-Report Psychopathy scale, confirmatory factor analysis (CFA), longitudinal measurement invariance, temporal stability

Psychopathy is a personality disorder characterized by severe socially deviant traits and behavior, such as egocentricity, manipulativeness, irresponsibility, impulsivity and a lack of empathy, remorse and guilt (Skeem et al., 2011). Individuals with high levels of psychopathy are responsible for a disproportionate amount of crime and violence in society. Among offenders, psychopathic traits are associated with elevated recidivism rates, especially violent and sexual recidivism (Hawes et al., 2013; Leistico et al., 2008). At the same time, there seems to be a significant role of psychopathy outside the criminal justice system. For example, psychopathic traits share similar associations with harassment, aggression, low relational satisfaction and exploitation in the general population (Benning et al., 2005; Muñoz et al., 2011; Weiss et al., 2018). In order to improve upon our understanding of psychopathy and to design developmentally appropriate interventions, it is critical to develop reliable and valid measures to assess this construct (Neumann & Hare, 2008). The current study aims to longitudinally validate the Dutch translation of the Levenson Self-Report Psychopathy (LSRP) scale, an instrument designed to assess psychopathic traits outside of prison and other institutional settings. Research on the Dutch translation of this questionnaire is scarce, with only one prior study documenting its translation and initial validation in convenience samples (Garofalo et al., 2019). Additionally, the present study is the first to validate the LSRP scale in a longitudinal framework.

The Self-Report Assessment of Psychopathy Using the LSRP

Psychopathy research has flourished with the development of the Psychopathic Checklist-Revised (PCL-R; Hare, 1991/2003), a clinician-rated instrument based on a semi-structured interview and file information. Although it is widely used in both legal and clinical practices (Sellbom et al., 2018), PCL-R assessment is highly cumbersome and labor-intensive, and the need to access institutional files makes it less suitable for some research settings (Sellbom et al., 2018).

Evidence indicates that psychopathy is dimensional rather than categorical in nature (e.g., Sellbom & Drislane, 2020) and psychopathic traits are also continuously distributed among community samples, wherein they present a comparable nomological network (Colins et al., 2017; Lilienfeld et al., 2014). Assessment of community samples is important, as it might provide insights into why some individuals with psychopathic traits become involved in severe forms of criminal and other forms of antisocial activities that makes them end up in contact with the criminal justice system, whereas others manage to achieve relatively successful careers or otherwise non-criminal enterprises in society.

Several self-report measures have been developed that can be used to assess psychopathic traits outside of the prison walls (see Sellbom et al., 2018, for a review). The LSRP (Levenson et al., 1995) is one such frequently used self-report questionnaire. Compared to other instruments, this questionnaire is particularly promising for application in research settings because it is freely available and of short length, thereby saving both time and costs. The LSRP consists of 26 items scored on a 4-point Likert-scale. The LSRP was initially developed to capture psychopathic traits as was done in the two factors of Hare's (1991) Psychopathy Checklist, resembling Karpman's (1948) distinction between primary and secondary psychopathy with its two subscales (Levenson et al., 1995). Specifically, primary psychopathy was characterized by deficits in affective (e.g., callousness) and interpersonal (e.g., manipulation) traits. Secondary psychopathy, on the other hand, was characterized by anger dysregulation and disinhibited behavior. The LSRP total score showed adequate construct validity, in light of positive associations with measures of antisocial behavior, impulsivity, alcohol and drug abuse, low agreeableness, and low conscientiousness (Brinkley et al., 2001; Colins et al., 2017; Levenson et al., 1995; Lynam et al., 1999; McHoskey et al., 1998; Salekin et al., 2014; Seibert et al., 2011). Additionally, research has shown separate correlates for the two LSRP subscales. For instance, primary psychopathy was negatively

related to thrill and adventure seeking (Levenson et al., 1995). On the other hand, secondary psychopathy was positively related to aggression and trait anxiety (Levenson et al., 1995; Lynam et al., 1999; McHoskey et al., 1998; Salekin et al., 2014; Seibert et al., 2011).

Although these findings are in line with theoretical expectations, some studies also raise some concerns about the correlation of the LSRP primary and secondary scales with external measures. For example, according to Levenson et al. (1995), trait anxiety should be negatively related to primary psychopathy and positively related to secondary psychopathy. However, studies have failed to find a negative association between trait anxiety and primary psychopathy, thereby questioning the construct validity of this scale (Levenson et al., 1995; Lynam et al., 1999; McHoskey et al., 1998; Poythress et al., 2010; Salekin et al., 2014; Seibert et al., 2011). Besides this issue, several other psychometric concerns regarding the LSRP two-factor structure have been reported in the literature as well, such as low internal consistency of the secondary scale and poor model fitting of two-factor model using confirmatory factor analysis (CFA) (Brinkley et al., 2008; Sellbom, 2011). In addition, a conceptual problem with the two-factor structure of the LSRP is that it conflates what was intended as a person-centered distinction (i.e., primary and secondary psychopathy as two *subtypes* of psychopathy) with a variable-center approach to psychopathy measures (i.e., primary and secondary psychopathy as two *subscales* in psychopathy measures). Given these controversies, the two-factor model does not seem optimal in describing the underlying structure of psychopathy from the LSRP method of operationalization.

In recent years, an alternative three-factor model of the LSRP has received increased attention in the literature (see Sellbom et al., 2018, for a review). This model was first described by Brinkley and colleagues (2008) using exploratory factor analysis (EFA). They reported that 19 of the 26 items could best be modeled with a three-factor structure, thereby excluding seven items of the original LSRP scale (Brinkley et al., 2008). In contrast to the

two-factor model, the three-factor structure makes a distinction between interpersonal and affective traits (previously conflated in the LSRP primary psychopathy subscale), which increases the alignment with other contemporary measures of psychopathy (e.g., Self-Report Psychopathy Scale, Triarchic Psychopathy Measure, Youth Psychopathic traits Inventory). The first factor – Egocentricity – captures features such as self-centeredness, interpersonal manipulation and antagonism. The second factor – Callousness – captures low empathy and remorselessness. The third factor – Antisocial – captures traits like anger, poor frustration tolerance, boredom susceptibility, and impulsivity. Research over the past several years has repeatedly observed that the three-factor model better fits observed data in comparison to the two-factor model across a number of countries on four separate continents (Christian & Sellbom, 2016; Garofalo et al., 2019; Salekin et al., 2014; Sellbom, 2011; Shou et al., 2017; Somma et al., 2014; Wang et al., 2018).

To date, the construct validity of these three LSRP factors has received promising yet mixed support. In support of the scales' construct validity, the Egocentricity subscale has been related to external measures of narcissism, Machiavellianism, antagonism, psychopathic meanness, and diminished perception of social responsibility (Few et al., 2013; Salekin et al., 2014; Sellbom, 2011). Moreover, the Callousness subscale tends to be associated with measures of guiltlessness and low empathy (Salekin et al., 2014; Sellbom, 2011; White, 2014). Further, the Antisocial subscale has been found to be positively associated with indices of disinhibition, externalizing and a history of antisocial behavior (Brinkley et al., 2008; Salekin et al., 2014; Sellbom, 2011).

Although these findings provide evidence for the construct validity of the LSRP scale scores, other findings were contrary to theoretical expectations, especially concerning the Callousness subscale. For example, some studies reported weak correlations with measures of empathy (Salekin et al., 2014; Sellbom, 2011), positive correlations with neuroticism, and

negative associations with fearlessness and stress immunity (Few et al., 2013; Salekin et al., 2014; Sellbom, 2011). Another issue concerning construct validity regards the discriminant validity of the LSRP Antisocial subscale. While the Antisocial subscale has been positively related with measures of negative emotions, like distress, anxiety, and anger, which is in line with expectations from the literature, the magnitudes of these correlations are quite large. In some cases, these correlations are larger than those reported between the Antisocial subscale with externalizing behavior (Brinkley et al., 2008; Garofalo et al., 2019; Salekin et al., 2014; Sellbom, 2011). Taken these conflicting findings into consideration, the construct validity of the three-factor model warrants further investigation.

Stability of Psychopathic Traits

Psychopathic traits are often conceptualized as stable components of human behavior, but research into the influences of critical periods of development on the longitudinal trajectories of these traits is limited. For instance, little is known about the expression of psychopathic traits in emerging adulthood, a developmental period from late adolescence to early adulthood (Hawes et al., 2014). Due to termination of control and legal responsibility by caregivers, emerging adulthood is known for an increase in freedom and independence, especially in Western countries (Arnett, 2000). At the same time, the continuous development of the prefrontal cortex provides the neural substrate responsible for complex behaviors such as planning and decision making (Crone & Dahl, 2012). Considering this period of significant change, research on psychopathic traits in this developmental window may yield theoretical insight into the etiology of psychopathy.

In order to investigate the development of psychopathic traits, two characteristics are particularly important, namely rank-order stability and mean-level stability. Rank order stability reflects the stability of an individual's trait to higher or lower values relative to other individuals in the population. Mean level stability reflects the consistency of the level of a

certain trait in the population over time. To date, only few studies have investigated rank-order and mean-level stability of psychopathic traits among emerging adults. For example, a study among a community sample of male and female twins in the United States showed both rank-order and mean-level stability of the interpersonal/affective dimension of psychopathy from age 17 to 24, using the Multidimensional Personality Questionnaire (Blonigen et al., 2006). In contrast, mean levels of traits related to the behavioral dimension of psychopathy were found to decline over time, in line with the development of the prefrontal cortex. Another study found a more general pattern of decline of psychopathic traits over time in a large sample of male offenders using the Youth Psychopathic Trait Inventory from age 17 to 23 (Hawes et al., 2014). However, effect sizes of these declines were relatively small. Taking these findings together, it seems likely that psychopathic traits are relatively stable throughout emerging adulthood, though more research is needed to elaborate on these findings. In addition, rank-order stability and mean-level stability have not been examined previously using the LSRP method of operationalization.

A preliminary consideration is worth noting and has often be neglected in previous studies. In order to make valid conclusions about the stability of traits over time, a particular instrument must measure an equivalent construct (and overall latent factor structure) at different time points. This means that scores at baseline represent the same theoretical construct as scores at follow up time, which is generally referred to as longitudinal measurement invariance (Horn & Mcardle, 1992; Meredith, 1993). For example, change scores over time may reflect changes in subjects' perception of items over time (de Beurs et al., 2015), rather than an actual improvement or worsening of symptoms over time. Thus, it is of great importance to longitudinally validate an instrument before using it to assess subjects over time. Yet, in spite of a large body of prospective studies that used psychopathy measures to predict later outcomes, relatively scant attention has been paid to the longitudinal validation

of self-report measures of psychopathy (i.e., assessing it at different time points to evaluate its longitudinal measurement invariance). Indeed, most research to date on the measurement invariance of self-report measures of psychopathy (including research with the LSRP scale) has been cross-sectional (see Sellbom et al., 2018, for a review).

The Present Study

The present study was designed to expand upon the construct validity research base of the LSRP in general, and the Dutch translation (Garofalo et al., 2019) of this instrument in particular. Specifically, our aims were addressed using a large archival longitudinal sample of Dutch emerging adults, with an equal proportion of men and women. To our knowledge, the present study represented the first *longitudinal* validation of the LSRP as well as any other self-report inventory of psychopathy. First, the present study examined longitudinal measurement invariance of the LSRP. We hypothesized that the three-factor model would be equivalent over time, as this model has repeatedly been found optimal in describing the underlying structure of psychopathy from the LSRP method of operationalization (Christian & Sellbom, 2016; Garofalo et al., 2019; Marion & Sellbom, 2011; Salekin et al., 2014; Wang et al., 2018). Next, in line with findings discussed earlier (Blonigen et al., 2006; Hawes et al., 2014), we hypothesized that the LSRP would exhibit temporal stability of psychopathic traits from late adolescence into early twenties. Further, this study aimed to replicate prior findings on the LSRP construct validity (see Sellbom et al., 2018, for an overview), by testing its association with constructs that have conceptually been related to psychopathy, such as delinquency, aggression, empathy, agreeableness, conscientiousness and impulsivity, both concurrently and over time. The following hypotheses were formulated in terms of preferential patterns of associations (i.e., significant concurrent correlations and significant prediction over time). We expected LSRP Egocentricity to be negatively associated with agreeableness, positive attitudes towards authority, prosocial behavior and sensation seeking

(Few et al., 2013; Garofalo et al., 2019; Salekin et al., 2014). Next, we expected that

Callousness would be negatively associated with empathy and agreeableness (Sellbom, 2011).

Finally, we expected that Antisocial would be positively associated with likelihood of future

criminal behavior, frequency of committed offences, criminal attitudes, alcohol and drug use,

low self-control, sensation seeking, impulsivity and aggression; and negatively associated

with agreeableness, conscientiousness, pro-social behavior and positive attitudes toward

authority (Brinkley et al., 2008; Marion & Sellbom, 2011; Wang et al., 2018).

Method

Participants

The present study used data from the first and fourth waves of the 'Transition in

Amsterdam' (TransAM) project, a longitudinal study that was carried out by a Dutch research

institute between 2010 and 2014 (Blokland, 2014). The TransAM project used both official

and self-reported data on delinquent behavior and markers of emerging adulthood in a

multiethnic sample of emerging adults. The sample consist of an equal proportion of men and

women from Dutch, Moroccan and Dutch-Caribbean descents, defined by parental birth

country. Potential participants resided in Amsterdam and were randomly selected from the

municipal registry. However, in order to have sufficient variation in the sample regarding

delinquent behavior those with registered police contact prior to age 17 years were

oversampled. Moroccan and Dutch-Caribbean ethnicities were chosen because they are most

overrepresented in crime rates in The Netherlands (Blokland, 2014). Of the 3,408 potential

participants selected, 28% gave their informed consent and completed the first interview

wave. Sampled participants came from all Amsterdam's city districts and from 84 of 89

neighborhoods; very small neighborhoods (i.e., with population < 2,000) were excluded.

Participants were interviewed 4 times at 6-month intervals. The present study used data from

the first and fourth interview waves, with an 18-month time interval, when the LSRP was

administered. Of the 970 participants who completed the first interview wave (henceforth 'time point 1'), 54% was female. Of the 693 (70% of time point 1) participants who completed the fourth interview wave (henceforth 'time point 2'), 58% was female. On average, participants were aged 20.0 ($SD = 1.4$) at time point 1 and 21.6 ($SD = 1.3$) at time point 2. At both time points, most participants were from Dutch descent (43% at time point 1 and, 51% at time point 2), followed by Moroccan descent (38% at time point 1, and 30% at time point 2) and Caribbean descent (19% at both time points). Most participant were attending education at time point 1 (79%) and time point 2 (73%). Additionally, most participants had a paid job at both time points (65% at time point 1, and 70% at time point 2). Half of our sample (46% at time point 1, and 54% at time point 2) had a high SES (based on parental education), followed by low parental SES (29% at time point 1, and 24% at time point 2) and medium parental SES (21% at time point 1, and 19% at time point 2). Almost half of our sample was in a relationship at time point 1 (40%) and time point 2 (47%). Participants were more likely to complete assessment at time point 2 if they were females, if they had Dutch ethnicity and if they had higher socioeconomic status. Participation at time point 2 was not predicted by self-reported delinquency at time point 1, or having a police record prior to age 17 years (Hill et al., 2016). Moreover, attrition at time point 2 was not predicted by higher LSRP scores at time point 1.

Measures[1]

      Levenson Self-Report Psychopathy scale. The LSRP (Levenson et al., 1995; Dutch translation: Garofalo, 2018) is a 26-item self-report questionnaire to measure psychopathic traits. The LSRP items were scored on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree). Higher scores indicate a higher level of psychopathic traits. The LSRP was described in detail earlier. The current study used the Dutch translation of the

---

[1] Descriptive statistics and internal consistency coefficients for the criterion measures are reported in Supplemental Table S1.

19-item LSRP version, which has received promising construct validity support in a sample of

nonclinical Dutch participants (Garofalo et al., 2019). All items on the Callousness scale are

reversed. Although this might indicate a potential method factor, the scale has received

adequate validity support. Internal consistency α coefficients in the present study are

presented in the Results section and in Table 5.

Delinquent Attitudes and Behavior.

Future likelihood[2]. Participants were asked to indicate the likelihood that they would engage

in future criminal behaviors. This scale consists of 8 items. Items were responded to on a 6-

point Likert scale ranging from 1 (not at all likely/definitely not) to 6 (very likely/definitely).

Higher scores indicate a greater likelihood of future crimes. This scale was designed by

TransAM (Blokland, 2014). Items were derived from the Hope/Self-Efficacy Scale, which

has received promising support for its reliability among a sample of male offenders with

multiple prior offences (LeBel et al., 2008).

Frequency. Participants had to indicate from a list of 48 offenses how many times in

the previous 6-months, or since their previous interview, they had committed each offense.

Some participants reported very high numbers of offenses. In order to control the skewedness

of the data distribution, this variable was adjusted to allow for a maximum of 10 times for

each offense. The list of offenses used was adapted from the Self-Report Delinquency Study

(Junger-Tas et al., 1994) and the South Holland Study (Hofstra et al., 2001), which both

provide construct validity support for these ratings among samples of young adults.

Criminal Attitudes and Associates. The Measure of Criminal Attitudes and Associates

(MCAA; Mills, Kroner, & Forth, 2002) is composed of two parts, that both show reasonable

construct validity among a population of incarcerated adult males sentenced to 2 years or

---

[2] These questions were only asked to people that had reported on past committed offenses

more, and criterion validity as evidenced by positive associations with criminal history (Mills et al., 2002).

Part A is a measure of criminal associates and part B measures four attitudes domains, namely Violence, Entitlement, Antisocial Intent, and Associates. The current study used a shortened version of part B, that omitted the Violence subscale, consisting of 25 items. These items are divided into the domains Entitlement (4 items), Antisocial Intent (11 items) and Associates (10 items). Items were responded to on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores indicate greater degrees of criminal attitudes.

Attitudes Towards Authority. The General Attitude Towards Institutional Authority Scale (GAIAS; Rigby, 1982) originally consisted of 32 items, divided into four subscales named Army, Police, Law and Teachers. The current study used a shortened version of the GAIAS consisting of the subscale 'Law'. This subscale consists of 8 items that are responded to on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores indicate more positive attitudes towards authority. The GAIAS has received promising support for its construct validity among a sample of Australian and British tertiary students. This was evidenced in the scale's positive relationship with a measure of general conservatism (Rigby, 1982).

Substance Use. The broader category substance use was measured by assessing the sub-categories alcohol and drug use. Items belonging to both sub-categories were derived from Composite International Diagnostic Interview (CIDI; World Health Organization, 1994) and Addiction Severity Index (ASI; (McLellan et al., 1980). Both instruments have validity support in substance abuse samples (Andrews & Peters, 1998; McLellan et al., 1980).

Alcohol[3]. In order to assess alcohol use, participants were asked how often in the last six months they had drunk between 1-4 drinks and 5 or more drinks at one occasion. Both items were responded to on a 5-point Likert scale from 1 (nearly every day) to 5 (less than once month). These scale items reflect alcohol frequency (Blokland, 2014). Although Blokland (2014) assessed excessive alcohol use (which is defined as drinking more than 5 drinks at one occasion), we reversed coded the response options and calculated a sum scored based on both items. This aggregate score indicates overall alcohol frequency.

Drugs. A dichotomous variable (i.e., "yes" or "no") was used to indicate whether participants had used drugs in either the previous 6-months, or since their previous interview.

Psychological Constructs.

Empathy. The Interpersonal Reactivity Index (IRI; Davis, 1983) was used to assess empathy. This questionnaire consists of 28-items that are responded to on a 5-point Likert scale ranging from 1 (not at all true of myself) to 5 (true of myself). Items aggregate onto four subscales labeled Perspective Taking, Fantasy, Empathic Concern and Personal Distress. Each scale consists of 7 items and some items were reversed in order to control for response sets. Higher scores indicate higher degrees of empathy. The current study used the Dutch translation of the IRI, which has demonstrated good psychometric properties among Dutch adults and adolescents and good construct validity as evidenced by positive associations with helping behavior, negative associations with aggression, and null associations with conceptually unrelated constructs such as extraversion (De Corte et al., 2007; Hawk, Keijsers, Branje, Van Der Graaff, et al., 2013).

Agreeableness and Conscientiousness. The Dutch Big Five Inventory (BFI; Denissen, Geenen, Van Aken, Gosling, & Potter, 2008) was used to assess agreeableness and conscientiousness, which are two domains from the Five Factor Model of personality

---

[3] These questions were not asked to non-drinkers

(McCrae & Costa, 1987). Both scales consist of 9 items that are responded to on a 5-point Likert Scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores indicate higher levels of agreeableness and conscientiousness. The Dutch BFI has received promising validity support among a diverse sample of the Dutch population (Denissen et al., 2008). For example, Conscientiousness was positively related to educational attainment (Denissen et al. 2008)

       Pro-Social Behavior. Prosocial behavior was examined using the 11-item subscale prosocial behavior from the Self-Report of Aggression and Social behavior Measure (Morales & Crick, 1998, 1999). The current study used the Dutch translation of this scale (Van Lier et al., 2008). Previous research provided good validity support for this instrument among early and late adolescents (Clark et al., 2015; Hawk, Keijsers, Branje, Van Der Graaff, et al., 2013). Items were scored to on a 7-point Likert scale ranging from 1 (never true) to 7 (definitely true). Higher scores reflect greater degrees voluntary behavior intended to benefit others.

       Self-control. The Self-Control Scale (Grasmick et al., 1993) consists of 24 items that aggregate onto six subscales labeled Impulsivity (4 items), Simple tasks (4 items), Risk Taking (4 items), Physical Activities (4 items), Self-Centeredness (4 items), and Temper (4 items). Items are scored to on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree). Higher scores reflect lower self-control. Scale scores have shown acceptable reliability and good support for validity across offender samples and the general population (Piquero & Rosay, 1998).

       Sensation seeking. The Brief Sensation Seeking Scale (BSSS) was used in order to assess sensation seeking tendencies (Hoyle et al. 2002). This questionnaire consists of 8 items that are scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores reflect higher degrees of sensation seeking. Hoyle et al. (2002) supported the validity of this scale across a large sample of adolescents. Positive correlations

were found between BSSS scores and measures of drug-related attitudes and behaviors, and

deviance. Negative correlations were found between BSSS scores and law abidance (Hoyle et

al. 2002).

Impulsivity. A short version of the Barratt Impulsiveness Scale (BIS) was used to

assess impulsivity (Spinella, 2007). This questionnaire consists of 15 items that are scored on

a 4-point Likert scale ranging from 1 (rarely or never) to 4 (nearly always). Higher scores

reflect greater degrees of impulsivity. This short version of the BIS maintained good

reliability and validity in a community sample (Spinella, 2007).

Aggression. The Aggression Questionnaire was used to assess aggressive behavior

(Buss & Perry, 1992) This questionnaire consist of 29 items, divided into four subscales

labeled Physical Aggression (9 items), Verbal Aggression (5 items), Anger (7 items), and

Hostility (8 items). Items were scored to on a 5-point Likert scale ranging from 1 (extremely

unlike me) to 5 (extremely like me). Higher scores reflect greater degrees of aggression. The

current study used the Dutch version of the Aggression Questionnaire, which has received

promising validity support among a sample of Dutch university students (Meesters et al.,

1996).

Procedures

Potential participants were contacted through mail by researchers of the TransAM

project, which was followed by a home visit. Of the 3,408 reached, 970 (28%) agreed to

participate and gave informed consent. Collected data included information on romantic, peer

and parental relationships, education and employment status, delinquent behavior, and a range

of psychological constructs. For those measures for which no Dutch translation was available,

items were first translated to Dutch, then reviewed until agreement on a satisfactory version

was ensured. The translated version of the questionnaire was then back-translated into English

in order to control for conceptual and cultural equivalence. Discrepancies were discussed until

a satisfactory version was reached. Data were collected using computer assisted personal interviews (CAPI). The data are archival and their use for the present investigation was approved by the local university Ethics Committee.

Data Analysis

Two different confirmatory factor analyses (CFAs) were conducted to test the fit of the three-factor model; one for time point 1 and one for time point 2.[4] At both time points, items were constrained to load on the latent factor (Antisocial, Egocentricity, Callousness) to which they belong based on Brinkley's (2008) original findings. A mean and variance adjusted Weighted Least Squares (WLSMV) estimator was used in M*plus* 8, and this estimator was chosen because it is considered more appropriate in order to model ordered categorical data (Beauducel & Herzberg, 2006). Model fit was assessed using global fit indices, including the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), and the root mean square error of approximation (RMSEA). For CFI and TLI, the conventional cutoff values of .90 or greater for acceptable fit were used. RMSEA values equal to or smaller than .08 were used for acceptable fit (Brown, 2015). Initial CFA models rarely achieve acceptable fit, and re-specification of the three-factor model was considered to the extent conceptually defensible. Based on Sellbom's (2011) work, two pairs of residuals were selected as candidates (e.g. 'I have been in a lot of shouting matches with other people' and 'when frustrated, I often let off steam by blowing my top'; 'For me, what's right is whatever I can get away with' and 'In today's world, I feel justified in doing anything I can get away with to succeed').

In order to assess longitudinal measurement invariance, three models were specified and examined across time (see Little, 2013, for an overview). First, a configural invariance

---

[4] Because there were no *a priori* theoretical expectations of structural variance across gender, and because previous research has corroborated empirically that the LSRP was invariant across men and women for multiple samples (Sellbom, 2011) as well as showing the same predictive utility across genders (Marion & Sellbom, 2011), the present study did not examine measurement invariance across gender to avoid unnecessarily increases in length and complexity.

model was estimated across the two time points. In this configural invariance model, items were constrained to load on the latent factors to which they belong at both time points, and the two pairs of residuals were correlated at both time points. Second, a weak invariance model was estimated in which the factor loadings were constrained to be equal across time intervals. Third, a strong invariance model was estimated, in which the loadings as well as the item thresholds were constrained to be equal across time. The relative model fit of two nested models was compared using a chi-square difference test specifically designed for WLSMV estimation. The chi-square values across models are not on the same scale, so the DIFFTEST function in M*plus* was used to rescale the two models for direct comparison. A statistically nonsignificant DIFFTEST was indicative of invariance. In instances of either weak or strong non-invariance, modification indices were inspected first, to determine the possibility for partial weak and/or partial strong invariance by determining which loading or threshold constraint release might lead to the largest improvement in model fit. If the release of these loadings or thresholds did not result in partial weak and/or partial strong invariance, or when no modification indices were available to inform us, freely estimated loadings in the configural model, and freely estimated thresholds in the partial weak invariance model were inspected. Loadings and thresholds that showed large differences over time qualified for release. Temporal stability was assessed by examining rank-order stability and mean-level stability. To assess rank order stability, the weak or partial weak invariance model was used and latent factor correlations of the three LSRP subscales over time were inspected. In a meta-analysis on the stability of personality traits, Roberts and DelVecchio (2000) found an average test-retest correlation coefficient of .55 over a 1-year period. Because our study was conducted over an 18-month time period, we considered stability values of 0.50 acceptable. In order to assess mean level stability, the strong or partial strong invariance model was considered, and the latent means between time 1 and time 2 were compared statistically and

with Cohen's *d* effect sizes. In line with Cohen's guidelines, effect sizes were classified as small ($d = 0.20$-0.49), medium ($d = 0.50$-0.79) and large ($d \geq 0.80$) (Cohen, 1988).

Construct validity was assessed by estimating the associations between each external criterion measure and the three latent factor scores and a latent total score concurrently (correlations at both time points) and over time (i.e., a latent regression model whereby LSRP factors and latent total scores at time point 1 predicted external criterion variable scores at time point 2), using the weak (or partial weak) invariance model. A separate analysis was performed for each external measure (including subscales when applicable). Concurrent and predictive validities were assessed in separate models to avoid statistical anomalies due to excessive partialing. Regarding predictive validity, we simultaneously regressed external criterion scores onto the three LSRP factor scores over time, thus to represent the unique contribution of each LSRP factor. Latent "total" LSRP scores were estimated in separate analyses by specifying a second-order model where the three first-order factors (Egocentricity, Callousness, Antisocial) loaded on a second order latent factor.

Results

Factor Structure, Inter-Correlations, and Internal Consistency of the LSRP

Table 1 shows the model fit indices for the three-factor model at time point 1 and 2, both before and after modification based on Sellbom (2011). The results indicate that the three-factor model by Brinkley et al. (2008) exhibited marginal model fit at best at both time points. Based on Sellbom (2011), two pairs of residuals were correlated, which elevated the three-factor model to acceptable fit across all indices at both time points. Table 2 shows the factor loadings for individual items at both time points. All factor loadings were acceptable (i.e., > 0.40) at both time points. The correlations between latent factors were 0.472 (Egocentricity and Callousness), 0.522 (Egocentricity and Antisocial) and 0.157 (Callousness and Antisocial) at time point 1 and 0.480 (Egocentricity and Callousness), 0.581 (Egocentricity

and Antisocial), 0.299 (Callousness and Antisocial) at time point 2. Furthermore, the internal

consistencies were calculated for Egocentricity, Callousness and Antisocial subscales as well

as for the total LSRP scores at both time points. The internal consistency of the total LSRP

scale ($\alpha = 0.82$ at time point 1, and $\alpha = 0.85$ at time point 2) and subscale Egocentricity ($\alpha =$

0.82 at time point 1, and $\alpha = 0.86$ at time point 2) were good at both time points. The internal

consistency of subscale Callousness was acceptable at both time points ($\alpha = 0.71$ at time point

1, and $\alpha = 0.70$ at time point 2) and; the internal consistency of the Antisocial subscale was

questionable at both time points ($\alpha = 0.65$ at time point 1, and $\alpha = 0.69$ at time point 2),

although these values were consistent with the LSRP literature (Sellbom, 2011). Also,

observed mean scores on the three factors were calculated at both time points, by averaging

scores on individual items. Mean scores were 1.99 (SD = 0.74), 1.99 (SD = 0.68), 2.06 (SD =

0.72) at time point 1, and 1.91 (SD = 0.70), 1.98 (SD = 0.64), 1.99 (0.66) at time point 2 for

Egocentricity, Callousness, and Antisocial, respectively.

Longitudinal Measurement Invariance

In order to assess whether the LSRP measured the same construct at both time points, the

fit of three different models (configural, weak, strong) was examined. As can be seen in Table

3, all three models fit the data well (CFIs > 0.9 and RMSEAs < 0.05). Next, the fit of the three

models was compared in order to examine longitudinal invariance. First, the weak model was

compared to the configural model, hypothesizing that the weak model would not have worse

model fit. The DIFFTEST was significant, which means that the weak invariance model

worsened model fit relative to the configural model. Therefore, the weak model was re-

specified to evaluate for potential partial weak invariance. Modification indices could not

guide us here, so absolute differences of factor loadings in the non-constrained configural

model were inspected. Upon inspection of the factor loadings at both time points in the

configural model, the factor loading constraint of item 1 was released, because the factor

loadings of this item showed the biggest difference over time ($\lambda = 0.521$, at time point 1; and λ = 0.642, at time point 2) compared to the factor loadings of other items. After releasing the constraint of item 1, the partial weak invariance model's fit was improved and the DIFFTEST was nonsignificant when compared to the configural invariance model.

Next, the partial weak invariance model was compared to the strong invariance model, hypothesizing that the strong invariance model would not show a decrement in model fit after constraining the item thresholds to be equal across time points. Again, the DIFFTEST was significant, which means that the strong invariance model had a decrement in model fit relative to the partial weak invariance model. First, modification indices were inspected, and based on these indices, thresholds of two items were released. Since this did not lead to a nonsignificant DIFFTEST, freely estimated thresholds in the partial weak invariance model were inspected, and a partial strong model was specified, releasing one constraint at a time. Threshold constraints that qualified for release showed an absolute difference over time of 0.1 or bigger in the partial weak invariance model. Once the constraints of 18 thresholds (32% of all thresholds[5]) had been released, the partial strong invariance model was associated with a nonsignificant DIFFTEST. To determine the degree to which these differences were meaningful, we performed two sets of calculations. First, effect sizes were calculated with Cohen's *d* effect sizes. Table S2 in the online supplemental materials shows the differences in thresholds over time of the released thresholds, and effect sizes of these differences. In line with Cohen's guidelines, the magnitudes of these differences in thresholds over time did not appear substantial, since all effect sizes were below 0.30. In addition, inspection of those thresholds with small effects revealed that most of them had very little psychologically

---

[5] The 19 LSRP items are scored to on a 4-point Likert-Scale. This means that there are three thresholds per item. In the strong invariance model, item threshold of both time points are set to be equal over time. This makes 57 threshold constraints in total (19x3). In the partial strong invariance model, 18 threshold constraints were released, which is 32% of the total number of possible threshold constraints.

intuitive reason for being variant. Second, to evaluate the practical significance of the non-invariance at the threshold level, we compared factor scores (across time) for the three factors derived from the strong invariance model (i.e., thresholds constrained to be equal) and the final partial strong invariance model (i.e., in which 18 of 57 constraints were released). Figures S1a-c in the online supplemental materials shows the average factor scores for the two time points across the two invariance models. We conducted repeated measures ANOVAs for each LSRP factor score and examined whether time*invariance model interactions were significant, which would reveal a potentially meaningful practical difference (i.e., a different pattern of mean stability over time). For Callous and Antisocial, there were no factor mean differences across invariance models at any time point (all ps > .05) and time*invariance interactions were non-significant (F[1, 1936) = 0.39, p =.843 [Callous] and F[1, 1936] = 2.25, p =.143 [Antisocial]). For Egocentricity, there was a significant time*invariance interaction (F[1, 1936] = 12.65, p <.001), but the effect size (η_p^2 = .006) associated with this difference was trivial. As a result of these two sets of analyses, it was deemed acceptable to free the constraints of these thresholds without substantially violating measurement invariance in any practical sense, and the partial strong invariance model was thus retained.

Rank-order and mean-level stability

In terms of rank-order stability, the correlations between the three latent factors over time were significant and of large magnitude, specifically, 0.708 (Egocentricity), 0.488 (Antisocial) and 0.572 (Callousness). Regarding mean-level stability, the latent means of the Egocentricity (M_{diff} = -0.082, p < .05) and Antisocial (M_{diff} = -0.176, p < .001) factors were significantly different over time, whereas the latent means for Callousness were not (M_{diff} < -0.0001, p = 0.995). However, since the current study consisted of a large sample, the likelihood of observing significant results was quite high, so Cohen's d effect sizes were

calculated. Effect sizes were -0.08, -0.17, and < -0.001 for Egocentricity, Antisocial, and Callousness, respectively. Thus, the latent mean level differences over time were trivial per Cohen's guidelines. These small effect sizes point to evidence of substantial mean level stability of psychopathic traits across these two time points.

Construct Validity and Predictive Validity

Table 4 shows the zero-order correlations between the three LSRP latent factors (Egocentricity, Callousness, Antisocial) and external criterion measures at both time points, correlations between total LSRP scores and external criterion measures at both time points, the regression coefficients of the LRSP latent factor scores at time point 1 predicting external criterion measure scores at time point 2, and the regression coefficients of total LSRP scores at time point 1 predicting external criterion measures at time point 2. Given the large number of models estimated, we only focused on correlations associated with $p < .01$. Hypothesized correlations appear in boldface in the table.

Concurrent Validity. The overall pattern of correlations indicates that each of three factors showed largely expected patterns of correlations with external measures at both time points. As hypothesized, the Egocentricity subscale was negatively correlated with agreeableness, positive attitudes towards authority and pro-social behavior at both time points. Further, Callousness was found to correlate negatively with empathy and agreeableness at both time points. Finally, at both time points, the Antisocial subscale correlated positively with likelihood of future criminal behavior, frequency of committed offenses, criminal attitudes, low self-control, impulsivity, and aggression; and, negatively, with pro-social behavior, conscientiousness, agreeableness, drug use and positive attitudes towards authority. While these findings support our hypotheses, other correlations of the Antisocial subscale with external measures were partly inconsistent with our predictions. For example, positive correlations with sensation seeking, were significant only at time point 1. Additionally, the

latent total LSRP scores showed evidence for good construct validity, in light of its meaningful associations with almost all measures, except alcohol and drug use.

Predictive Validity. Inspection of the predictive correlations revealed several unique patterns of correlations that confirmed our hypotheses. First, as expected, the Callousness subscale negatively predicted empathy. Second, in line with our hypotheses, the Antisocial subscale significantly predicted frequency of committed offenses, criminal attitudes, drug use, low self-control, impulsivity, and aggression (all positively), as well as positive attitudes towards authority, conscientiousness, agreeableness and pro-social behavior negatively. However, in contrast with our expectations, the Antisocial subscale did not predict future likelihood of criminal behavior, alcohol use and sensation seeking. Moreover, in contrast to our expectations, no associations over time were found between Egocentricity and agreeableness, positive attitudes towards authority and sensation seeking, and a positive association was found with conscientiousness. With the exception of measures for drug and alcohol use, latent total LSRP scores were correlated with all measures over time.

Discriminant Validity. Our results provided mixed support for the discriminant validity of the LSRP scales. At the bivariate level, the three LSRP latent factors correlated concurrently with external criterion scores with similar magnitudes at both time points. For example, although we expected both the latent factor Egocentricity and Antisocial to be correlated with negative attitudes towards authority, all three LSRP latent factors seem to predict this external measure with more or less similar magnitudes at both time points. Yet, when examining the predictive validity over time while controlling for the shared variance among LSRP latent factors, the Antisocial factor was the strongest predictor of negative attitudes towards authority. Moreover, we expected that both Callousness and Egocentricity were correlated with low empathy, though Callousness should be the best predictor. Indeed, our findings revealed that both latent subscales were correlated with low empathy. However,

the pattern of correlations was not so clearly favoring Callousness over Egocentricity. Furthermore, we expected that the Antisocial factor was correlated with both future likelihood of criminal behavior and frequency of criminal offenses. While both scales were correlated with the latent subscale Antisocial at both time points, the correlations were not so clearly favoring Antisocial over Egocentricity and Callousness. In fact, when looking at the predictive validity it became clear that Callousness and Antisocial were the best predictors of crime frequency but not of future likelihood of criminal behavior.

Discussion

The aim of the current study was to validate the LSRP scale in a longitudinal design with a community sample of Dutch emerging adults over-sampled for youth at risk of criminal justice system involvement. More specifically, the present study aimed at replicating and extending findings on the factor structure, reliability, stability, and construct validity of the LSRP scale in its Dutch translation. In addition, the present study was the first to examine longitudinal measurement invariance of the LSRP scales and, to our knowledge, of any self-report measure of psychopathic traits. First, this study provided evidence that the 19-item three-factor model fit the data well, and that the three LSRP factors showed largely acceptable levels of internal consistency at both time points. Second, this three-factor model was partially invariant over time, and psychopathic traits have been found relatively stable from late adolescent into early twenties. Lastly, the three-factor model showed conceptually expected patterns of construct validity and predictive validity, although some exceptions were noted. Below, the results of this study will be discussed in the context of the broader literature.

Factor Structure, Longitudinal Measurement Invariance, and Temporal Stability

The present findings provided evidence that the 19-item three-factor model properly captures the internal structure of psychopathy as operationalized in the LSRP scale. Acceptable model fit was achieved when two pairs of residuals were correlated, replicating

findings from previous studies on different independent samples (Christian & Sellbom, 2016; Garofalo et al., 2019; Sellbom, 2011; Shou et al., 2017). Moreover, the model fit results of the current study are in line with those of a recent study by Garofalo et al. (2019), who also demonstrated evidence of acceptable model fit of the 19-item three-factor model in a non-clinical sample of Dutch adults. Also, the strength of the latent inter-correlations were consistent with those reported in Garofalo et al. (2019). Specifically, the inter-correlations between Egocentricity and Callousness and between Egocentricity and Antisocial were relatively stronger than the correlation between Callousness and Antisocial. However, Garofalo et al. (2019) observed lower levels of internal consistency reliability of the Callousness and Antisocial subscales scores, which was likely due to the range restriction in their sample. Due to oversampling of at-risk youth in the current study, the current sample had, in all likelihood, higher levels and variation of psychopathic traits. Further, observed mean scores on the three factors were slightly higher compared to those of Garofalo et al. (2019), which is likely also due to the higher proportion of at-risk youth in our sample compared to Garofalo et al.'s (2019) sample. The current findings revealed that the three subscales achieved acceptable levels of internal consistency reliability at both time points, with the partial exception of the Antisocial subscale. However, when inspecting the factor loadings of the Antisocial subscale, none of the factors loadings was remarkably low. Taken these findings together, the 19-item three-factor model seemed acceptable in describing the internal structure of psychopathy from the LSRP method of operationalization in a Dutch sample of emerging adults.

Of note, our findings indicated that the longitudinal measurement invariance of the LSRP over an 18-month period was partly met. Specifically, the release of some constraints was needed before evidence of invariance emerged: one factor loading and 18 thresholds (32% of total) were sufficiently different over time to significantly violate the invariance

assumption. It should be noted, however, that the difference between the weak invariance model and the partial weak invariance model (in which one factor loading was released) was minimal, and likely of trivial practical significance. With regard to the released thresholds, it should be emphasized that the magnitude of differences in the thresholds that were allowed to differ over time were trivial and thus unlikely to be meaningful. There were also no practical differences across the strong and partial strong invariance models with respect to estimating latent means over time. Taken together, the overall trend that emerged was that the magnitudes of the differences in threshold values and estimates factor means across invariance models over time were not indicative of meaningful differences, and likely an artifact of a very large sample size. Hence, the re-specified partial invariance model can confidently be interpreted as substantially invariant, which means that the Dutch translation of the LSRP scale is measuring an equivalent construct (and overall latent factor structure) over time. As such, the LSRP scale – at least in its Dutch translation – holds promise for use in longitudinal designs that can inform the developmental literature on psychopathy.

Once evidence supported longitudinal measurement invariance of the LSRP scale, the present study examined rank-order and mean-level stability of psychopathic traits over time. First, we reported large rank-order stability of the Egocentricity, Callousness, and Antisocial subscales over an 18-month period. These findings imply that our sample generally maintained the same rank-ordering of psychopathic traits over time, in line with Hawes et al.'s (2014) findings, which observed large rank-order stability of the affective, interpersonal and behavioral dimensions of psychopathy over a one-year period, from age 20 to age 21. It should be noted that rank-order stability is affected by test-retest reliability. A time period of 1 to 2 weeks between assessments is often considered appropriate, as longer time intervals generally result in test-retest variability (Bisgaard et al., 2007; Roberts & DelVecchio, 2000). These measurement errors might lower the rank-order stability levels. Given that our study

was conducted over an 18-month time interval, the observed rank-order stability here is thus likely an underestimation of the real effects. When inspecting the test-retest correlations, it can be concluded that the magnitudes of these correlations are mostly far above the expected trait correlation levels reported for these time intervals. These findings underscore the relatively high trait consistency levels that have been observed in the current study.

Second, we found evidence for mean-level stability of psychopathic traits, as the Egocentricity, Callousness, and Antisocial subscale remained stable over an 18-month period in participants' early twenties. These findings indicated that our sample, as a whole, showed no changes in psychopathic features over time, suggesting that changes related to emerging adulthood do not influence the expression of psychopathic features. Our findings are partly consistent with Blonigen et al. (2006), who found mean-level stability of the interpersonal and affective dimension of psychopathy, akin to the LSRP Egocentricity and Callousness subscales, respectively. However, Blonigen et al. (2006) found a decline in mean-levels of the antisocial dimension. This finding is not surprising giving that externalizing behaviors reduce in severity by early adulthood (Roberts, Caspi, & Moffitt, 2001). Similarly, the age-crime curve perspective posits the increase in criminal offending during teenage years, followed by a sharp drop when individuals enter their early twenties (Blumstein, Cohen, & Farrington, 1988). Additionally, these behavioral changes parallel maturation of the prefrontal cortex (Blonigen, 2010). Considering these findings, the question arises why our study found a relatively stable pattern of expression of antisocial tendencies throughout emerging adulthood. One possible answer is that most longitudinal investigations are conducted over a longer period of time and starting at an earlier age, for example focusing on the period from age 18 to about age 24 (Blonigen et al., 2006, 2008; Hawes et al., 2014). In contrast, our study was conducted over an 18-month time period, from age 20 to about age 21.5. Although this is a critical developmental stage marking early adulthood, the relatively short time-interval, with

most participants being already 20 at the beginning of the study, might have made it difficult

to capture the whole set of changes that appear to occur during emerging adulthood.

Therefore, in order to be more conclusive on potential mean-level changes of psychopathic

traits during the transition into adulthood, future longitudinal investigations should be

conducted over a longer time period.

Construct Validity of the LSRP

A latent estimation of an LSRP total score showed promising construct validity across

almost all external measures, which is not surprising since we selected a priori criteria we

assumed would be related to psychopathy. Also, associations of the three LSRP subscales

with external measures were mostly in line with theoretical expectations. For example, the

Antisocial subscale was the strongest predictor of low self-control, aggression and

impulsivity. In addition, this subscale was associated to antisocial attitudes and associates,

low conscientiousness, drug use and was very strongly related to the temper and anger

components of self-control and aggression, respectively. These findings clearly confirm that

the Antisocial subscale captures the externalizing, disinhibited traits of psychopathy.

The Egocentricity subscale broadly captures a selfish, self-centered and generally

antagonistic style (Shou et al., 2017). Our findings supported this conceptualization, as

negative correlations were observed with pro-social behavior and positive attitudes towards

authority. However, Egocentricity was not uniquely correlated to these external measures, as

even stronger correlations were observed with the Antisocial subscale. These findings seem to

mirror some previous work indicating that the Antisocial subscale captures more than solely

antisocial behaviors and impulsivity. Indeed, this subscale was more broadly related to

problems in self- and interpersonal functioning (Garofalo et al., 2019).

Further, we hypothesized that both Egocentricity and Callousness were associated

with low empathy, though Callousness should emerge as the best predictor, as it captures a

more general remorseless attitude. However, the results of the current study were not so clearly favoring Callousness over Egocentricity in this regard. These findings are in line with other studies that found inconsistent and low associations between Callousness and diminished empathy (Salekin et al., 2014; Sellbom, 2011). An often mentioned reason for these construct validity issues is the short length of the Callousness subscale. A recent attempt of Christian and Sellbom (2016) to improve construct coverage by including additional items proved to be successful, as stronger associations with low empathy and clearer discriminant validity were observed. Thus, revisions of the Callousness subscale are important in order to improve coverage of the psychopathy construct as measured via the LSRP.

In terms of discriminant validity, however, additional mixed results warrant some discussion. As expected, the Personal Distress scale of the IRI had differential associations with LSRP Antisocial (positively related to distress) as opposed to LSRP Egocentricity and Callousness (negatively related or unrelated to distress). In addition, Conscientiousness was largely unrelated, or positively related, to the LSRP Egocentricity and Callousness scales. Yet, a clear pattern of discriminant validity did not emerge for the self-control subscales, and, to a lesser extent, with impulsivity. Thus, the discriminant validity with regard to disinhibition-related traits received mixed support. Similarly, the Antisocial scale did not show the expected preferential association with sensation seeking and likelihood of antisocial behavior. A previous investigation of the LSRP found a shared correlation of the Callousness subscale and the Antisocial subscale with low social desirability (Garofalo et al., 2019). These findings together may suggest that the LSRP has better construct validity in terms of total scores than discriminant validity at the subscale level. A conceptually plausible alternative is that it is not only the behavioral component of psychopathy, but also the callous features of the disorder that predict a certain carelessness to act in accordance with social rules, although they may derive from different motivations (e.g., lack of concern for norms or impulsivity).

Further, some associations of the latent factors with external variables seem to be explained by statistical suppressor effects. For example, the association of Egocentricity with frequency of committed offenses seems inflated by the strong inter-correlations that exist between Egocentricity and Antisocial, and between Egocentricity and Callousness. Indeed, at the predictive level, the Egocentricity coefficient changed direction, and only unique correlations of Antisocial and Callousness with frequency of committed offenses could be observed. Also, a similar pattern of inflation can be observed for the association of Egocentricity with aggression, and to a lesser extent, with self-control.

Notably, some hypothesized correlations were not supported by our results. First, we did not find a correlation between LSRP total score and substance misuse, and the Antisocial subscale and alcohol use, which mirrors some previous work by Shou et al. (2017). Shou and colleagues (2017) did not find a significant relationship between the LSRP subscales and problematic drinking, due to a possible range restriction associated with low levels of problematic drinkers in their sample. Our sample might also be affected by a range restriction associated with low levels of substance misuse, which could partly be explained by cultural factors. A relatively large proportion of our sample is namely from Moroccan descent and likely to be raised according to Muslim tradition. In Muslim culture, alcohol use is much less tolerated compared to either the Dutch or the Caribbean culture. Another possible explanation for the lacking correlation between the latent LSRP total score and alcohol frequency, and Antisocial subscale and alcohol frequency is that we generated a (possibly invalid) aggregated score for alcohol use. This variable is likely not specific enough to distinguish alcohol use from alcohol misuse. Second, the findings of the current study did not support all of the expected associations between psychopathic traits and basic personality dimensions. More specifically, the Egocentricity subscale has often been negatively related to agreeableness. Surprisingly, we did not find a unique contribution of Egocentricity on the prediction of

agreeableness. Future studies should continue to examine the construct validity of this subscale in order to be more conclusive on this finding.

Limitations

The results of this study must be interpreted in light of several limitations. First, this study solely relied on self-report questionnaires, which could unduly bias validity analyses due to shared method variance. As a result, magnitudes of observed correlations might be inflated to an unknown degree. Second, some constructs were assessed with relatively brief measures consisting of very few items. Although brief measures increase the efficacy of longitudinal investigations, they are likely to have less predictive validity than scores on well-constructed longer inventories (Credé et al., 2012). Third, some of the scales showed rather low internal consistency levels, such as the LSRP Antisocial subscale, the Big Five agreeableness subscale and the likelihood of future crime scale. It should be noted that these scales measure quite broad constructs, which could explain the poor interrelatedness of items, especially when considering that these measures consist of relatively few items. In turn, low internal consistency coefficients may have resulted in attenuated correlations for these scales. The subscales that showed alpha levels that were deemed too low (e.g., the MCAA entitlement subscale and the verbal aggression subscale), were not used in separate analyses, but only the full scale scores were used, of which internal consistency levels were adequate. Fourth, the degree to which certain variables predict attrition might lead to some problems with generalizability, as participants who dropped out might differ from those who stayed in the study, although attrition rates were not particularly high and associations with key variables were trivial or non-significant. Further, the study response rate might seem low and therefore affect generalizability of findings, though our response rate is not entirely inconsistent with other work on youth delinquency in the same geographical area (Junger-Tas et al., 1994). Also, some associations might be inflated due to content overlap between

constructs measured by items on the LSRP scale and external measures. Lastly, this investigation was limited to only two time points that were relatively close to each other. Thereby, our conclusion was restricted to a relatively small developmental window and thus only partly captures changes related to emerging adulthood.

Conclusions

Despite these limitations, the present study relied on a large sample of emerging adults over-sampled for participants at risk of deviant behavior, and reported incremental promising evidence for the structural properties, longitudinal measurement invariance, and construct validity of the LSRP, at least for its Dutch translation. Psychopathic traits, as operationalized by the LSRP, may be better captured disentangling at least three separate components, and have shown high degrees of temporal stability in emerging adulthood. By and large, the LSRP subscales displayed an adequate pattern of convergent and discriminant validity both within and across time, although few notable exceptions were found. Taken together, these findings reinforce the appropriateness of the LSRP as a brief, effective instrument to investigate psychopathic traits in developmental studies based on longitudinal designs with community samples, although revisions of some subscales (e.g., Christian & Sellbom, 2016) may further enhance its psychometric properties.

References

Andrews, G., & Peters, L. (1998). The psychometric properties of the Composite International

Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology*, *33*(2), 80–88.

https://doi.org/10.1007/s001270050026

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through

the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037//0003-

066X.55.5.469

Beauducel, A., & Herzberg, P. Y. (2006). On the Performance of Maximum Likelihood

Versus Means and Variance Adjusted Weighted Least Squares Estimation in CFA.

Structural Equation Modeling: A Multidisciplinary Journal, *13*(2), 186–203.

https://doi.org/10.1207/s15328007sem1302_2

Benning, S. D., Patrick, C. J., Blonigen, D. M., Hicks, B. M., & Iacono, W. G. (2005).

Estimating Facets of Psychopathy From Normal Personality Traits. *Assessment*, *12*(1),

3–18. https://doi.org/10.1177/1073191104271223

Bisgaard, T., Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L.,

Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for

measurement properties of health status questionnaires. *Journal of Clinical

Epidemiology*. https://doi.org/10.1016/j.jclinepi.2006.03.012

Blokland, A. A. J. (2014). "School, Intensive work, Excessive alcohol use and Delinquency

During Emerging Adulthood." In *Criminal Behaviour From School to the Workplace:

Untangling the Complex Relations Between Employment, Education and Crime* (pp. 87–

107). https://doi.org/10.4324/9780203670446

Blonigen, D. M. (2010). Explaining the relationship between age and crime: Contributions

from the developmental literature on personality. In *Clinical Psychology Review*.

https://doi.org/10.1016/j.cpr.2009.10.001

Blonigen, D. M., Carlson, M. D., Hicks, B. M., Krueger, R. F., & Iacono, W. G. (2008).

Stability and Change in Personality Traits From Late Adolescence to Early Adulthood: A

Longitudinal Twin Study. *Journal of Personality*, *76*(2), 229–266.

https://doi.org/10.1111/j.1467-6494.2007.00485.x

Blonigen, D. M., Hicks, B. M., Krueger, R. F., Patrick, C. J., & Iacono, W. G. (2006).

Continuity and change in psychopathic traits as measured via normal-range personality:

A longitudinal-biometric study. *Journal of Abnormal Psychology*, *115*(1), 85–95.

https://doi.org/10.1037/0021-843X.115.1.85

Brinkley, C. A., Diamond, P. M., Magaletta, P. R., & Heigel, C. P. (2008). Cross-Validation

of Levenson's Psychopathy Scale in a Sample of Federal Female Inmates. *Assessment*,

15(4), 464–482. https://doi.org/10.1177/1073191108319043

Brinkley, C. A., Schmitt, W. A., Smith, S. S., & Newman, J. P. (2001). Construct validation

of a self-report psychopathy scale: Does Levenson's self-report psychopathy scale

measure the same constructs as Hare's psychopathy checklist-revised? *Personality and

Individual Differences*, *31*(7), 1021–1038. https://doi.org/10.1016/S0191-

8869(00)00178-1

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research (2nd ed.). New

York, NY: Guilford.*

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and

Social Psychology*, *63*(3), 452–459. https://doi.org/10.1037/0022-3514.63.3.452

Chin, A., Markey, A., Bhargava, S., Kassam, K. S., & Loewenstein, G. (2017). Bored in the

USA: Experience sampling and boredom in everyday life. *Emotion*, *17*(2), 359–368.

https://doi.org/10.1037/emo0000232

Christian, E., & Sellbom, M. (2016). Development and Validation of an Expanded Version of

the Three-Factor Levenson Self-Report Psychopathy Scale. *Journal of Personality*

Assessment, *98*(2), 155–168. https://doi.org/10.1080/00223891.2015.1068176

Clark, C. M., Dahlen, E. R., & Nicholson, B. C. (2015). The Role of Parenting in Relational

Aggression and Prosocial Behavior among Emerging Adults. *Journal of Aggression,*

Maltreatment & Trauma, *24*(2), 185–202.

https://doi.org/10.1080/10926771.2015.1002653

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale,

NJ: Lawrence Earlbaum Associates. *Lawrence Earlbaum Associates*.

https://doi.org/http://dx.doi.org/10.1016/B978-0-12-179060-8.50008-6

Colins, O. F., Fanti, K. A., Salekin, R. T., & Andershed, H. (2017). Psychopathic Personality

in the General Population: Differences and Similarities Across Gender. *Journal of*

Personality Disorders, *31*(1), 49–74. https://doi.org/10.1521/pedi_2016_30_237

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the

consequences of using short measures of the Big Five personality traits. *Journal of*

Personality and Social Psychology. https://doi.org/10.1037/a0027403

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective

engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650.

https://doi.org/10.1038/nrn3313

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a

multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–

126. https://doi.org/10.1037/0022-3514.44.1.113

de Beurs, D. P., Fokkema, M., de Groot, M. H., de Keijser, J., & Kerkhof, A. J. F. M. (2015).

Longitudinal measurement invariance of the Beck Scale for Suicide Ideation. *Psychiatry*

Research, *225*(3), 368–373. https://doi.org/10.1016/J.PSYCHRES.2014.11.075

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007).

Measuring Empathic Tendencies: Reliability And Validity of the Dutch Version of the

Interpersonal Reactivity Index. *Psychologica Belgica*, *47*(4), 235.

https://doi.org/10.5334/pb-47-4-235

Denissen, J. J. A., Geenen, R., Van Aken, M. A. G., Gosling, S. D., & Potter, J. (2008).

Development and validation of a Dutch translation of the Big Five Inventory (BFI).

Journal of Personality Assessment, *90*(2), 152–157.

https://doi.org/10.1080/00223890701845229

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G. (2006). Psychopathic, not

psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal

of Abnormal Psychology*. https://doi.org/10.1037/0021-843X.115.1.131

Few, L. R., Miller, J. D., & Lynam, D. R. (2013). An examination of the factor structure of

the Elemental Psychopathy Assessment. *Personality Disorders: Theory, Research, and

Treatment*, *4*(3), 247–253. https://doi.org/10.1037/per0000016

Garofalo, C., Noteborn, M. G. C., Sellbom, M., & Bogaerts, S. (2019). Factor Structure and

Construct Validity of the Levenson Self-Report Psychopathy Scale (LSRP): A

Replication and Extension in Dutch Nonclinical Participants. *Journal of Personality

Assessment*, *101*(5), 481–492. https://doi.org/10.1080/00223891.2018.1519830

Grasmick, H. G., Tittle, C. R., Bursik, R. J., & Arneklev, B. J. (1993). Testing the core

empirical implications of gottfredson and hirschi's general theory of crime. *Journal of

Research in Crime and Delinquency*, *30*(1), 5–29.

https://doi.org/10.1177/0022427893030001002

Hare, R. (1991).The Hare Psychopathy Checklist—Revised.Toronto,Ontario, Canada: Multi-

Health Systems.

Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination

of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic

findings using the Psychopathy Checklist—Revised. *Psychological Assessment*, *25*(1),

233–243. https://doi.org/10.1037/a0030391

Hawes, S. W., Mulvey, E. P., Schubert, C. A., & Pardini, D. A. (2014). Structural coherence and temporal stability of psychopathic personality features during emerging adulthood. *Journal of Abnormal Psychology*, *123*(3), 623–633. https://doi.org/10.1037/a0037078

Hawk, S. T., Keijsers, L., Branje, S. J. T., Graaff, J. Van der, Wied, M. de, & Meeus, W. (2013). Examining the Interpersonal Reactivity Index (IRI) Among Early and Late Adolescents and Their Mothers. *Journal of Personality Assessment*, *95*(1), 96–106. https://doi.org/10.1080/00223891.2012.696080

Hill, J. M., Blokland, A. A. J., & van der Geest, V. R. (2016). Desisting from Crime in Emerging Adulthood. *Journal of Research in Crime and Delinquency*, *53*(4), 506–535. https://doi.org/10.1177/0022427816628586

Hofstra, M. B., Van Der Ende, J., & Verhulst, F. C. (2001). Adolescents' self-reported problems as predictors of psychopathology in adulthood: 10-Year follow-up study. *British Journal of Psychiatry*. https://doi.org/10.1192/bjp.179.3.203

Horn, J. L., & Mcardle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3), 117–144. https://doi.org/10.1080/03610739208253916

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personal Individ Differ*, *32*, 401–414. https://doi.org/10.1016/S0191-8869(01)00032-0

Junger-Tas, J., Terlouw, G.-J., & Klein, M. W. (1994). Delinquent Behavior Among Young People in the Western World: First Results of the International Self-Report Delinquency Study. In *Studies on Crime and Justice: A Series from the Dutch Research and Documentation Centre*.

Karpman, B. (1948). The Myth Of The Psychopathic Personality. *American Journal of*

Psychiatry, *104*(9), 523–534. https://doi.org/10.1176/ajp.104.9.523

LeBel, T. P., Burnett, R., Maruna, S., & Bushway, S. (2008). The 'chicken and egg' of

   subjective and social factors in desistance from crime. *European Journal of*

   Criminology, *5*(2), 131-159.

Leistico, A.-M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-

   analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and*

   Human Behavior, *32*(1), 28–45. https://doi.org/10.1007/s10979-007-9096-6

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing Psychopathic

   Attributes in a Noninstitutionalized Population. *Journal of Personality and Social*

   Psychology, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Lilienfeld, S. O., Latzman, R. D., Watts, A. L., Smith, S. F., & Dutton, K. (2014). Correlates

   of psychopathic personality traits in everyday life: results from a large community

   survey. *Frontiers in Psychology*, *5*, 740. https://doi.org/10.3389/fpsyg.2014.00740

Lynam, D. R., Whiteside, S., & Jones, S. (1999). Self-Reported Psychopathy: A Validation

   Study. *Journal of Personality Assessment*, *73*(1), 110–132.

   https://doi.org/10.1207/S15327752JPA730108

Marion, B. E., & Sellbom, M. (2011). An Examination of Gender-Moderated Test Bias on the

   Levenson Self-Report Psychopathy Scale. *Journal of Personality Assessment*, *93*(3),

   235–243. https://doi.org/10.1080/00223891.2011.558873

McCrae, R. R., & Costa, P. T. (1987). Validation of the Five-Factor Model of Personality

   Across Instruments and Observers. *Journal of Personality and Social Psychology*, *52*(1),

   81–90. https://doi.org/10.1037/0022-3514.52.1.81

McHoskey, J. W., Worzel, W., & Szyarto, C. (1998). Machiavellianism and Psychopathy.

   Journal of Personality and Social Psychology. https://doi.org/10.1037/0022-

   3514.74.1.192

McLellan, A. T., Luborsky, L., Woody, G. E., & O'Brien, C. P. (1980). An improved

diagnostic evaluation instrument for substance abuse patients. The Addiction Severity

Index. *The Journal of Nervous and Mental Disease*, *168*(1), 26–33.

Meesters, C., Muris, P., Bosma, H., Schouten, E., & Beuving, S. (1996). Psychometric

evaluation of the Dutch version of the aggression questionnaire. *Behaviour Research and*

Therapy. https://doi.org/10.1016/0005-7967(96)00065-4

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance.

Psychometrika, *58*(4), 525–543. https://doi.org/10.1007/BF02294825

Mills, J. F., Kroner, D. G., & Forth, A. E. (2002). Measures of Criminal Attitudes and

Associates (MCAA). *Assessment*, *9*(3), 240–253.

https://doi.org/10.1177/1073191102009003003

Morales, J. R., & Crick, N. R. (1998). Self-report measure of aggression and victimization.

Unpublished Manuscript; University of Minnesota; Minneapolis. https://doi.org/10.1037/

t29632-000.

Morales, J. R., & Crick, N. R. (1999). Hostile attribution and aggression in adolescent peer

and romantic relationships. In *Poster presented at the Biennial Meetings of the Society*

for Research in Child Development.

Muñoz, L. C., Khan, R., & Cordwell, L. (2011). Sexually coercive tactics used by university

students: A clear role for primary psychopathy. *Journal of Personality Disorders*.

https://doi.org/10.1521/pedi.2011.25.1.28

Neumann, C. S., & Hare, R. D. (2008). Psychopathic Traits in a Large Community Sample:

links to voilence, alcohol use, and intelligence. *Journal of Consulting and Clinical*

Psychology, *76*, 893–899.

Piquero, A., & Rosay, A. (1998). The reliability and validity of Grasmick et al.'s Self-control

scale: comment Longshore et al. *Criminology*, *36*(1), 157–174.

https://doi.org/10.1111/j.1745-9125.1998.tb01244.x

Poythress, N. G., Lilienfeld, S. O., Skeem, J. L., Douglas, K. S., Edens, J. F., Epstein, M., &

Patrick, C. J. (2010). Using the PCL-R to Help Estimate the Validity of Two Self-Report

Measures of Psychopathy With Offenders. *Assessment*, *17*(2), 206–219.

https://doi.org/10.1177/1073191109351715

Rigby, K. (1982). A concise scale for the assessment of attitudes towards institutional

authority. *Australian Journal of Psychology*, *34*(2), 195–204.

https://doi.org/10.1080/00049538208257811

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits

from childhood to old age: A quantitative review of longitudinal studies. *Psychological

Bulletin*. https://doi.org/10.1037/0033-2909.126.1.3

Salekin, R., Chen, D., & Sellbom, M. (2014). Examining the factor structure and convergent

and discriminant validity of the Levenson Self-Report Psychopathy Scale: Is the two-

factor model the best fitting model? *Personality Disorders*.

https://doi.org/10.1037/per0000073

Seibert, L. A., Miller, J. D., Few, L. R., Zeichner, A., & Lynam, D. R. (2011). An

Examination of the structure of self-report psychopathy measures and their relations with

general traits and externalizing behaviors. In *Personality Disorders: Theory, Research,

and Treatment*. https://doi.org/10.1037/a0019232

Sellbom, M. (2011). Elaborating on the construct validity of the Levenson Self-Report

Psychopathy Scale in incarcerated and non-incarcerated samples. *Law and Human

Behavior*, *35*(6), 440–451. https://doi.org/10.1007/s10979-010-9249-x

Sellbom, M., Lilienfeld, S., Fowler, K., & McCrary, K. (2018). The self-report assessment of

psychopathy: Challenges, pitfalls, and promises. *Handbook of Psychopathy*, 211–258.

Sellbom, M., & Drislane, L. E. (2020). The classification of psychopathy. *Aggression and*

Violent Behavior. Advanced online publication.

https://doi.org/10.1016/j.avb.2020.101473

Shou, Y., Sellbom, M., & Han, J. (2017). Evaluating the Construct Validity of the Levenson

Self-Report Psychopathy Scale in China. *Assessment*, *24*(8), 1008–1023.

https://doi.org/10.1177/1073191116637421

Skeem, J. L., Polaschek, D. L. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic

Personality: Bridging the Gap Between Scientific Evidence and Public Policy.

Psychological Science in the Public Interest, *12*(3), 95–162.

https://doi.org/10.1177/1529100611426706

Somma, A., Fossati, A., Patrick, C., Maffei, C., & Borroni, S. (2014). The three-factor

structure of the Levenson Self-Report Psychopathy Scale: Fool's gold or true gold? A

study in a sample of Italian adult non-clinical participants. *Personality and Mental*

Health, *8*(4), 337–347. https://doi.org/10.1002/pmh.1267

Spinella, M. (2007). Normative data and a short form of the Barrat Impulsiveness Scale.

International Journal of Neuroscience, *117*(3), 359–368.

https://doi.org/10.1080/00207450600588881

Wang, M.-C., Shou, Y., Deng, Q., Sellbom, M., Salekin, R. T., & Gao, Y. (2018). Factor

structure and construct validity of the Levenson Self-Report Psychopathy Scale (LSRP)

in a sample of Chinese male inmates. *Psychological Assessment*.

https://doi.org/10.1037/pas0000537

Weiss, B., Lavner, J. A., & Miller, J. D. (2018). Self- and partner-reported psychopathic

traits' relations with couples' communication, marital satisfaction trajectories, and

divorce in a longitudinal sample. *Personality Disorders: Theory, Research, and*

Treatment. https://doi.org/10.1037/per0000233

White, B. A. (2014). Who cares when nobody is watching? Psychopathic traits and empathy

in prosocial behaviors. *Personality and Individual Differences*, *56*, 116–121.

https://doi.org/10.1016/J.PAID.2013.08.033

Table 1

Confirmatory factor analysis fit indices for tested models at time point 1 (N=965) and 2 (N=691).

		χ^2	df	P	RMSEA	90% CI LL	90%CI UL	Close fit p	CFI	TLI
Time 1	Model A	1189.53	149	<.0001	0.085	0.081	0.090	<.001	.892	.877
	Model B	992.17	147	<.0001	0.077	0.073	0.082	<.001	.913	.898
Time 2	Model A	1017.21	149	<.0001	0.092	0.087	0.097	<.001	.919	.906
	Model B	706.52	147	<.0001	0.074	0.069	0.080	<.001	.947	.939

Note. Model A = Brinkley et al.'s (2008) three-factor model. Model B = Brinkley et al.'s (2008) three-factor model, as re-specified in Sellbom (2011). χ^2 = goodness of fit test; df = degrees of freedom; RMSEA = root mean square error of approximation; CI = confidence interval; LL = lower level; UL = upper level; CFI = comparative fit index; TLI = Tucker-Lewis Index.

Table 2

Standardized factor loadings (λ) and standard error (SE) for the Levenson Self Report Psychopathy (LSRP) scales at time point 1 and 2.

Factor	Item	Factor loading time point 1		Factor loading time point 2	
		λ	SE	λ	SE
Egocentricity	Success is based on survival of the fittest; I am not concerned about losers.	0.525	0.025	0.626	0.025
	For me, what's right is whatever I can get away with.[a]	0.607	0.025	0.658	0.023
	In today's world, I feel justified in doing anything I can get away with to succeed.[a]	0.714	0.020	0.723	0.024
	My main purpose in life is getting as many goodies as I can.	0.712	0.021	0.808	0.017
	Making a lot of money is my most important goal.	0.655	0.023	0.749	0.019
	I let others worry about higher values; my main concern is with the bottom line.	0.665	0.023	0.772	0.022
	People who are stupid enough to get ripped off usually deserve it.	0.656	0.022	0.757	0.021
	I tell other people what they want to hear so that they will do what I want them to do.	0.737	0.019	0.719	0.022
	I often admire a really clever scam.	0.456	0.027	0.516	0.030
	I enjoy manipulating other people's feelings.	0.554	0.027	0.623	0.027
Callousness	I make a point of trying not to hurt others in pursuit of my goals. [R]	0.460	0.029	0.417	0.035
	I feel bad if my words or actions cause someone else to feel emotional pain. [R]	0.768	0.022	0.724	0.025
	Even if I were trying very hard to sell something, I wouldn't lie about it. [R]	0.822	0.020	0.886	0.022
	Cheating is not justified because it is unfair to others. [R]	0.774	0.022	0.787	0.027
Antisocial	I find myself in the same kinds of trouble time after time.	0.570	0.034	0.665	0.036
	I am often bored.	0.501	0.033	0.500	0.038
	I quickly lose interest in tasks I start.	0.505	0.034	0.448	0.040
	I have been in a lot of shouting matches with other people.[b]	0.737	0.041	0.808	0.040
	When frustrated, I often let off steam by blowing my top.[b]	0.505	0.041	0.587	0.041

Note. All factor loadings are significant at p < .001. [R] indicates reverse-keyed items. Items whose residuals were allowed to correlate are flagged with the same superscript ([a] and [b]), and were derived from Sellbom (2011).

Table 3

Model fit indices and difference tests for measurement invariance models.

Model	Chi-Square Test of Model Fit			RMSEA	RMSEA 90%-CI	CFI
	Value	d.f.	p-value			
A. Configural invariance[1]	2105.824	627	<.001	0.049	0.047 - 0.052	0.930
B. Weak invariance[2]	1983.843	643	<.001	0.046	0.044 - 0.049	0.936
C. Partial Weak Invariance[3]	1982.195	642	<.001	0.046	0.044 - 0.049	0.936
D. Strong Invariance[4]	2090.496	697	<.001	0.045	0.043 - 0.048	0.934
E. Partial Strong Invariance[5]	2024.284	679	<.001	0.045	0.043 - 0.048	0.936
Model comparison	Chi-Square for Difference Testing					
	Value	d.f.	p-value			
B vs. A	27.968	16	<.05			
C vs. A	23.356	15	0.077			
D vs. C	1003.43	55	<.001			
E vs. C	76.527	36	0.173			

Note. [1] items constrained to load on the latent factors they belong, [2] factor loadings constrained to be equal across time intervals, [3] constraint of item 1 released, [4] factor loadings and item thresholds constrained to be equal across time, [5] 18 of 57 constraints released. All fit indices based on robust standard errors and WLSMV estimator. d.f. = degrees of freedom; RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index.

Table 4

Overview of correlations and regression analysis of relationship between latent LSRP subscales and external measures.

Measure	Subscales	Time point 1 [a]				Time point 2 [a]				Predictive [b]			
		Total	Ego	Callous	Anti	Total	Ego	Callous	Anti	Total [c]	Ego	Callous	Anti
Likelihood of ASB	n/a	0.38***	0.32***	0.23***	**0.20***	0.28***	0.19**	0.24***	**0.24***	0.22**	-0.01	0.18	**0.22**
Frequency of ASB	n/a	0.24***	0.18***	0.10**	**0.25***	0.23***	0.15***	0.08***	**0.32***	0.18***	-0.09	0.22***	**0.21***
MCAA	Total	0.60***	0.46***	0.35***	**0.47***	0.61***	0.46***	0.46***	**0.42***	0.51***	0.12	0.32***	**0.36***
	Entitlement	0.28***	**0.23***	0.07	0.29***	0.23***	**0.17**	0.17***	0.17***	0.16**	**0.02**	0.15**	0.04
	Antisocial intent	0.62***	0.48***	0.38***	**0.45***	0.60***	0.46***	0.46***	**0.40***	0.50***	0.27***	0.17**	**0.11**
	Antisocial associates	0.55***	0.41***	0.35***	**0.45***	0.57***	0.43***	0.42***	**0.40***	0.49***	0.08	0.22***	**0.33***
GAIAS	Law	-0.49***	**-0.38***	-0.29***	**-0.40***	-0.57***	**-0.43***	-0.38***	**-0.48***	-0.48***	**-0.15**	-0.17**	**-0.41***
Alcohol	n/a	0.07	0.04	0.24	**0.25**	0.01	-0.04	0.07	**0.07**	-0.08	-0.28***	0.24***	**0.11**
Drugs	n/a	0.02	-0.02	0.03	**0.14***	0.02	-0.02	0.06	**0.12**	0.03	-0.17**	0.14**	**0.16**
IRI	Total	-0.49***	-0.42***	**-0.36**	-0.14**	-0.46***	-0.36***	**-0.45***	-0.13**	-0.50***	-0.40***	**-0.26***	0.10
	Fantasy	-0.36***	-0.31***	**-0.27***	-0.08	-0.32***	-0.28***	**-0.27***	-0.05	-0.37***	-0.37***	**-0.13**	0.14
	Empathic concern	-0.54***	-0.46***	**-0.40***	-0.18***	-0.54***	-0.42***	**-0.48***	-0.18***	-0.52***	-0.42***	**-0.26***	0.07
	Perspective taking	-0.41***	-0.32***	**-0.29***	-0.28***	-0.45***	-0.32***	**-0.40***	-0.29***	-0.42***	-0.15**	**-0.27***	-0.18**
	Personal distress	-0.06	-0.09**	**0.05**	0.13***	-0.02	-0.01	**-0.14***	0.17***	-0.11**	-0.18**	-0.08	0.24***
Big Five	Agreeableness	-0.39***	**-0.26***	-0.30***	**-0.31***	-0.42***	**-0.29***	-0.31***	**-0.35***	-0.39***	**-0.01**	-0.28***	**-0.31***
	Conscientiousness	-0.10**	0.01	-0.07	**-0.37***	-0.16***	-0.06	-0.06	**-0.41***	-0.16***	0.40***	-0.16**	**-0.66***
Pro-social behavior	n/a	-0.39***	**-0.26***	-0.35***	**-0.23***	-0.40***	**-0.28***	-0.34***	**-0.28***	-0.36***	**-0.13**	-0.20***	**-0.20***
Self-control	Total	0.66***	0.52***	0.33***	**0.57***	0.69***	0.55***	0.37***	**0.57***	0.61***	0.16**	0.16**	**0.56***
	Impulsivity	0.37***	0.28***	0.18***	**0.32***	0.44***	0.38***	0.18***	**0.34***	0.47***	0.03	0.08	**0.44***
	Simple task	0.37***	0.29***	0.12***	**0.37***	0.37***	0.28**	0.10**	**0.43***	0.39***	-0.09	0.03	**0.54***
	Risk taking	0.29***	0.21***	0.17***	**0.26***	0.39***	0.31***	0.23***	**0.28***	0.37***	0.02	0.16**	**0.29***
	Physical activities	0.41***	0.34***	0.17***	**0.30***	0.34***	0.28***	0.18***	**0.21***	0.43***	0.19**	0.02	**0.20**
	Self-centeredness	0.59***	**0.49***	0.33***	0.34***	0.66***	**0.52***	0.48***	0.39***	0.62***	**0.17**	0.29***	0.37***
	Temper	0.53***	0.36***	0.26***	**0.55***	0.58***	0.42***	0.28***	**0.61***	0.46***	-0.35***	0.29***	**0.83***
Sensation seeking	n/a	0.15***	**0.11**	0.07	**0.16***	0.14**	**0.12**	0.08	**0.09**	0.16***	**0.04**	0.07	**0.11**
Impulsivity	n/a	0.37***	0.23***	0.21***	**0.49***	0.41***	0.27***	0.23***	**0.47***	0.38***	-0.17***	0.18**	**0.63***
Aggression	Total	0.59***	0.43***	0.23***	**0.68***	0.46***	0.34***	0.14***	**0.58***	0.34***	-0.25**	0.16**	**0.74***
	Physical	0.63***	0.46***	0.29***	**0.58***	0.45***	0.33***	0.22***	**0.46***	0.37***	-0.27**	0.22***	**0.66***
	Verbal	0.35***	0.25***	0.11**	**0.40***	0.22***	0.18***	0.00	**0.30***	0.17***	-0.26**	0.12*	**0.45***
	Anger	0.44***	0.29***	0.13***	**0.60***	0.35***	0.22***	0.08*	**0.52***	0.21***	-0.58***	0.23**	**0.84***
	Hostility	0.41***	0.28***	0.12**	**0.51***	0.41***	0.31***	0.12**	**0.51***	0.32***	-0.41***	0.19**	**0.79***

Note. [a] Coefficients of the within time zero order correlations at time 1 and time 2, [b] Regression models in which criterion was regressed onto LSRP factors simultaneously, [c] Total Psychopathy was examined in a separate regression model. ASB = Antisocial Behavior. MCAA = Measure of Criminal Attitudes and Associates. GAIAS = The General Attitude Towards Institutional Authority Scale. IRI = Interpersonal Reactivity Index. Hypothesized correlations are shown in boldface. [**] $p < .01$. [***] $p < .001$.