AN INVESTIGATION OF THE STRUCTURE OF EXTERNALIZING PSYCHOPATHOLOGY

A THESIS

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF ARTS

BY

MCKENZIE R. CREMEANS

DR. TAYLA T. C. LEE – ADVISOR

BALL STATE UNIVERSITY

MUNCIE, INDIANA

DECEMBER 2019

Acknowledgements

I would like to acknowledge my sincere gratitude for the support of several individuals without whom the completion of this project would not have been possible.

First, I would like to express my deep appreciation to my thesis committee members, Drs. Johnathan Forbey and Thomas Holtgraves, for your time, support, and guidance during the process of completing this project. Thank you for your valuable feedback and unique expertise that encouraged me to expand my thinking about the concepts investigated in this thesis.

To my committee chair, Dr. Tayla Lee, thank you for first inspiring my interest in this realm of research during my time as an undergraduate in your lab. Thank you for investing your time and support into my work with you on this project and throughout my Master's program, and for always offering me honest and meaningful feedback when I needed it most. I will be forever grateful for your mentorship and guidance during the first steps of my academic career.

To the Personality and Psychopathology Lab at Ball State University, especially to its principle investigators, Drs. Tayla Lee and Johnathan Forbey, thank you for making this project possible by allowing me to conduct my analyses using this dataset.

To my loving family – my mother, father, brother, and sister – thank you for being the best support system that I could ask for. I am certain that it was not always easy being there for me while I completed my graduate work, but each of you always were. Thank you for instilling in me the determination and passion for education that were crucial to completing this project.

Finally, to Logan, thank you for your unwavering love and support during my time working on this thesis and throughout all of my academic endeavors. You have always kept me believing in myself, even when you were 700 miles away. There is no one else I would rather take on life with.

An Investigation of the Structure of Externalizing Psychopathology

Externalizing disorders are those characterized by a general tendency toward disinhibition and risky, impulsive behaviors that can elicit high costs to individuals and families affected by these problems, as well as to society as a whole (Krueger & Markon, 2006). Discrete difficulties associated with externalizing psychopathology include substance use disorders and antisocial behavior disorders. These problems have been demonstrated to share behavioral tendencies and genetically-based personality characteristics (Krueger et al., 2002). To better differentiate these dysfunctions, Krueger and colleagues (2007) developed a bifactor model describing both the shared disposition for tendencies toward externalizing behaviors, as well as the distinct factors that contribute uniquely to specific manifestations of the externalizing liability. While this model holds great promise as a framework of externalizing psychopathology that accounts for shared and unique variance among discrete disorders, it has only been replicated in one other study to date. As such, the current study aims to investigate whether this model can be replicated in another sample when using the measurement scale devised by Krueger and colleagues, as well as whether the model remains consistent when other measures of externalizing psychopathology are used.

Externalizing Disorders

The widely used *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)* defines substance use disorders and antisocial behavior disorders as discrete categories with distinct behavioral tendencies and diagnostic criteria. The essential diagnostic feature of a substance use disorder is a combination of cognitive, behavioral, and physiological symptoms suggesting that use of a substance is continued despite significant problems that result from use (American Psychiatric Association, 2013). Separately, antisocial behavior disorders, such as

antisocial personality disorder and conduct disorder, are assigned as a diagnosis for individuals who display a prevalent pattern of disregarding and violating societal regulations and the rights of others (American Psychiatric Association, 2013). These disorders are all generally associated with disinhibited personality and impulsive behaviors ranging from alcohol and illicit substance misuse to acts of violence and aggression (Krueger, Markon, Patrick, Benning, & Kramer, 2007).

Behaviors and outcomes associated with externalizing psychopathology are of high cost to individuals experiencing them, as well as to their family members and society. Individuals with antisocial behavior disorders and substance use difficulties frequently become involved with the legal system, often resulting in incarceration. This has long-lasting negative impact on their lives, as studies have shown that employment and income rates are much lower after incarceration, and these individuals frequently are arrested again after being released (Freudenberg, Daniels, Crum, Perkins, & Richie, 2008). In addition to legal, financial, and employment difficulties, these individuals experience numerous health, family, and social problems. These include psychological impairment and other mental health difficulties, higher risk for a variety of physical illnesses, and marital and family problems (Newcomb & Bentler, 1988). Societally, it is estimated that substance use, including the misuse of alcohol, tobacco, and illicit drugs, has an annual price of $510.8 billion in the US due to medical, legal, and other associated national costs (Miller & Hendrie, 2008). Yearly, alcohol misuse alone causes approximately 88,000 deaths, making it the third leading preventable cause of death in the US (NIAAA, 2010). The costs of antisocial personality disorder are also immense – the estimated US annual cost of criminal behavior associated with externalizing difficulties approaches $460 billion (Kiehl & Hoffman, 2011). Due to the magnitude of harmful consequences experienced by individuals with these disorders, as well as the large amount of money and societal damage

linked with them, it is crucial to develop an understanding of these difficulties in order to create appropriate treatments and interventions.

Difficulties associated with substance misuse and antisocial behavior have typically been studied and understood as discrete categories of psychopathology (Krueger, Markon, Patrick, Benning, & Kramer, 2007). The categorical approach to understanding mental disorders first came about in order to provide a common nomenclature that allowed individuals around the world communicate about and study mental health difficulties (Blashfield, Flanagan, & Raley, 2010). Traditional diagnostic nosologies, such as those described in the *Diagnostic and Statistical Manual of Mental Disorders (DSM),* are products of the categorical model, initially created to address all aspects of psychological disturbance and to classify mental health difficulties that emerge across inpatient and outpatient populations. This model of understanding psychopathology has remained popular, and the DSM is widely used across settings to classify symptomology and behavioral patterns.

Although helpful in expanding our understanding of psychopathology, the categorical model is limited. Arbitrary numeric symptom cutoffs between diagnostic categories create false separations between similar disorders that are not well validated as true differences (Decker, 2013). These diagnostic categories place many individuals into categories that they do not completely fit into or that do not describe all of their problems. This model places an emphasis on differences between disorders, which results in an inability to explain the variety of characteristics many disorders share (Decker, 2013). This is best exemplified by high levels of comorbidity among discrete disorders, which has been demonstrated statistically to be the rule rather than the exception in psychopathology (Widiger & Samuel, 2005). Comorbidity holds true among externalizing diagnoses, as substance use disorders and antisocial behavior disorders

are both diagnosed more often than expected by chance in children and adults (Hann, 2001; Swendsen et al., 2010; Krueger et al., 2002).

Given these limitations, recent psychological research has investigated alternative models that propose explanations for shared symptomology and behavioral patterns among several discrete disorders (Kotov et al., 2017). Researchers in this area have used statistical modeling to determine which disorders co-occur most often and to investigate whether there might be underlying liabilities that give rise to multiple disorders with similar symptomologies. These models propose that observed disorders are unique manifestations of underlying predispositions toward a broad category of difficulties, rather than distinct patterns of difficulty (Eaton, South, & Krueger, 2012; Krueger & Markon, 2006). Underlying dispositional liabilities toward internalizing and externalizing difficulties are the two most well-replicated dispositions toward frequently occurring mental disorders to come out of this research. A visual representation of the MCLM is displayed in Figure 1. This model suggests that disorders such as depressive disorders and anxiety disorders fall under the category of internalizing and are characterized by inward distress, while substance abuse disorders and antisocial behavior disorders lie on the externalizing dimension of psychopathology and are related based on tendencies toward disinhibition.

The externalizing liability is hypothesized to be a disinhibitory personality disposition characterized by disagreeableness and lack of conscientiousness (Krueger & Markon, 2006). This conceptualization can be thought of transdiagnostically (Nolen-Hoeksema & Watkins, 2011), suggesting a genetic-based distal liability toward disinhibition in combination with other environmental or individual difference factors that act as proximal risk to give rise to externalizing difficulties. Genetic risk for the externalizing liability has been supported in

several studies supporting the heritability of a reduced P300 component of brain event-related potentials (ERPs), which has been linked in twin studies to a variety of externalizing problems, including substance misuse and antisocial behavior (Kendler, Prescott, Myers, and Neale, 2003; Krueger et al., 2002, Hicks et al., 2007). Specifically, this component is thought of as a metric for cognitive functioning processes such as attention and working memory when ERPs are measured during decision-making processes (Linden, 2005). The reduced amplitude of the P300 component has therefore been observed in ERP data in response to deficits in attention and working memory during paradigms measuring decision-making. This disposition is thought to manifest as disinhibited personality, characterized by both behavioral- and cognitive-based difficulties with inhibition, such as tendencies toward impulsivity often characterized by aggression, increased discounting of delayed rewards, lower working memory capacity, and deficits in executive control. These characteristics are, in turn, thought to contribute risk for developing externalizing problems (Nigg, 2000; Krueger, Markon, Patrick, Benning, & Kramer, 2007; Bobova, Finn, Rickert, & Lucas, 2009; Patrick, Curtain, & Krueger, 2010).

Alternative Model of Externalizing Psychopathology

Although the genetically-based personality liability of externalizing psychopathology has been well-replicated, limitations of its ability to fully explain these disorders have been demonstrated. Though the model has the ability to describe what makes each of the externalizing disorders similar, it does not provide clarification of their differences. Specifically, the single liability model does not provide explanation for which difficulties and disorders are most likely to manifest in individuals high on the externalizing liability (Kendler, Prescott, Myers, & Neale, 2003). Hicks and colleagues (2007) demonstrated that, although the externalizing spectrum of difficulties share a genetic liability toward disinhibition, there are other

genetic and environmental risk factors that make them different as well. For example, the single liability model cannot explain why some individuals who are liable for externalizing psychopathology develop only substance-related difficulties, while others solely develop problems with antisocial behavior.

In response to these shortcomings, Krueger and colleagues (2007) conducted a study examining individual differences in externalizing behavior among participants from a sample of both college students and incarcerated prisoners. The goal of this work was to determine whether there was a model that better explained risk factors for externalizing psychopathology than the single liability model. To achieve this goal, the researchers surveyed for externalizing tendencies with items developed to measure a wide range of behaviors associated with disinhibition drawn from previously existing measures of externalizing psychopathology. Twelve construct domains were targeted, including aggression, lack of remorse, blame externalization, alcohol use, marijuana use, drug use, antisocial behavior, impulsivity, irresponsibility, rebelliousness, sensation seeking, and dependability (Krueger et al., 2007). After items assessing these domains were identified, data were collected in three waves of independent samples and structural analyses were used to compare the previously suggested single liability model with possible alternative models of externalizing psychopathology.

Results of analyses from these efforts indicated that the best fitting model for the externalizing items was a bifactor model. This model proposes that, not only is the externalizing spectrum based on a general underlying liability toward externalizing difficulties, but two additional, independent underlying liabilities also contribute risk for these disorders as well. As seen in Figure 2, the externalizing liability is the general factor that contributes risk for all observed difficulties with disinhibition. The model's other proposed factors specifically give

rise to aggressive- and substance-related behaviors and difficulties. The bifactor model suggests these two factors account for shared risk among behaviors not accounted for by the externalizing liability alone (Krueger, Markon, Patrick, Benning, & Kramer, 2007). For example, the observable behavior of marijuana misuse would be explained in this model as a result of the externalizing and substance-related liabilities, as this discrete difficulty is linked to both of these factors. Alternatively, the observable behavior of physical aggression would be suggested by this model to be explained by influence from both the externalizing and aggressive liabilities. There are also observed behaviors, such as attentional difficulties, which are explained by the externalizing liability alone in this model.

This derived bifactor model has potential to overcome the limitations of the previous single-factor model of externalizing psychopathology by furthering the conceptualization of the individual differences that make these difficulties related, but also distinct. Specifically, this model potentially provides a method of explaining other genetic, environmental, and individual factors that contribute alongside the externalizing liability to the specific manifestation of these disorders. However, although a full and brief-form measure were developed to capture the constructs identified by this model (Krueger, Markon, Patrick, Benning, & Kramer, 2007; Patrick, Kramer, Krueger, & Markon, 2013), the proposed structure has since been replicated for consistency across samples when using an alternative measure of externalizing psychopathology in only one study (Sellbom, 2016). This study used a variety of relevant Minnesota Multiphasic Personality Inventory – 2 – Restructured Form (MMPI-2-RF) items to compare proposed externalizing psychopathology models and found that Krueger and colleagues' (2007) expanded structure provided the best fit. Nonetheless, although MMPI-2-RF scales have been demonstrated to measure multiple facets of aggression relevant to the externalizing model, there

are few scales scored on the MMPI-2-RF that distinctly measure substance-related problems and impulsivity (Ben-Porath & Tellegen, 2008). Due to these limitations in item content, further support for this structure using scales from other measures of externalizing psychopathology would provide more holistic support for the bifactor model.

Current Study

Understanding the empirical model of externalizing psychopathology is important to the understanding of these costly disorders as well as the development of appropriate prevention and treatment methods. Given potential benefits of the proposed bifactor model in providing a better explanation of the risk factors giving rise to externalizing psychopathology, it is important to determine whether it can be replicated again with alternative measures that target the entire range of behaviors associated with these disorders. Successful replication of the model would provide support for the understanding of distinct risk factors that give rise to disorders and difficulties characterized by substance misuse and antisocial personality. As such, the current study aimed to investigate whether the structure of externalizing disorders suggested by Krueger and colleagues (2007) can replicate across samples when using the measure created in their study – the Externalizing Spectrum Inventory – Brief Form (ESI-bf). Additionally, the current study aimed to investigate whether the proposed model of externalizing psychopathology is supported when using alternative measures of externalizing behavior. To examine these questions, the current study contrasted the single liability model of externalizing psychopathology with the expanded bifactor model developed by Krueger and colleagues (2007). As depicted in Panel A of Figure 3, the first model suggests that comorbidity in externalizing disorders is best explained by a single predisposition toward externalizing dysfunction. Alternatively, the bifactor model displayed in Panel B of Figure 3 suggests a bifactor pathway of risk for externalizing difficulties

that includes the externalizing liability, as well as liabilities toward aggressive- and substance-related behaviors that contribute beyond the risk captured by the externalizing liability alone.

To determine which model best explains the covariance among externalizing behaviors and problems, the two structural models just described were estimated in a large sample of college students who completed a wide variety of measures assessing externalizing difficulties. After models were calculated, they were statistically contrasted to determine which of the models best fits the data. These models were specified in two ways. First, the models were specified using the ESI-bf, which was developed as a product of Krueger and colleagues' (2007) work on models of externalizing psychopathology. Second, the models were specified with other self-report measures assessing externalizing psychopathology in order to determine which model fits the data best when measured with scales other than the ESI-bf. It was determined that if the bifactor model fit best, both when specified using the ESI-bf and when using alternative models of externalizing psychopathology, this would support the theory that there is a bifactor pathway of risk for difficulties with disinhibition – one that includes both general and specific risk factors for aggressive- and substance-related behaviors. However, if the bifactor model fit best only when using the ESI-bf, this might suggest that a better model is necessary to explain externalizing behaviors, given that this model does not hold up across measures.

Method

The current study used data from a university sample collected as part of a previous study at Ball State University. The larger study's purpose was to investigate the assessment of the construct of impulsivity using scales from the MMPI-2-RF and other measures of externalizing behavior. Participants completed informed consent procedures in which they provided permission for their data to be archived and used in future studies. Participants were assessed in

accordance with procedures approved by the Institutional Review Board (IRB) at the institution where data collection occurred.

Participants

Participants in the current study were 608 undergraduate students at a Midwestern university who completed both of the two sessions that were a part of the larger study. To be included in the study, participants had to be between the ages of 18 and 40 years, understand and be able to respond to questions in English, be able to read at a Grade 6 level, have 20/20 corrected vision, and be able to physically manipulate a mouse and a keyboard. Participants could not have major cognitive impairments (i.e., can understand and respond to all screening questions), have any history of traumatic brain injury or serious head injury, be taking medications for cancer, AIDS treatment, or epilepsy, or be taking other medications that would affect behavior, such as major tranquilizers or antipsychotics.

Data from participants with 10% or more missing self-report items were excluded in order to reduce error variance in analyses. Data from participants with invalid MMPI-2-RF profiles were excluded as well in order to minimize error variance. The exclusions for invalid responding on the MMPI-2-RF were determined using standard criteria for its validity scales. These scales are well-established measures of content non-responsiveness, random and fixed responding, as well as over- and under-reporting of psychological symptoms (Ben-Porath & Tellegen, 2011). More information about the previously established utility of these validity scales for determining invalid responding can be found in the *Measures* section.

After excluding participants with either 10% or more missing self-report data or invalid MMPI-2-RF profiles, demographic characteristics were contrasted between participants excluded and participants not excluded from the sample. This was done in order to determine whether the

participants excluded from analyses displayed significant demographic differences from the final sample. These analyses did not display a significant difference in age between the two groups, $t(605) = -0.38$, $p = 0.71$, $d = .04$. However, there were significant differences observed in the sex and racial categories. Specifically, participants who identified as White were more likely than other racial groups to produce valid and complete profiles, $\chi 2$ (2, N = 608) = 7.37, $p = 0.03$, $\varphi = 0.11$. Individuals who identified as women or another gender were more likely than those who identified as men to produce valid and complete profiles, $\chi 2$ (2, N = 607) = 6.29, $p = 0.04$, $\varphi = 0.10$. Although these differences were observed, it was noted that these differences displayed a small effect size and the groups were relatively comparable. The final sample consisted of 513 participants (148 men, 364 women, 1 non-binary) who ranged in age from 18 to 34 years ($M = 18.82$, $SD = 1.38$). Most participants (431) identified as White, while 37 identified as Black, and 45 identified as another racial category.

Measures

For measure reliability values and descriptive statistics for ESI subscales, see *Table 1.* To see this information for other self-report measures used, see *Table 2.*

Study Criteria Questionnaire. This questionnaire's purpose was to determine whether participants met inclusionary and exclusionary criteria required to take part in the study. Participants responded to questions with either "true" or "false" regarding information about physical ability, major cognitive impairments, history of traumatic brain injuries, and medication use that would disqualify them from participating. If all questions were answered with "false," the participant was allowed to participate; if any questions were answered "true," the participant was unable to continue in the study.

Demographics Questionnaire. This set of questions probed for participants' age, gender, and race. Participants responded by choosing or typing in the response that best matched their personal demographic information.

Externalizing Spectrum Inventory – Brief Form (ESI-bf; Patrick, Kramer, Krueger, & Markon, 2013). The ESI-bf is a 160-item self-report measure created to assess externalizing behaviors including risk-taking, antisociality, substance use, and aggression. It is a shortened form of the original Externalizing Spectrum Inventory (ESI; Krueger et al., 2007). Participants responded to items on a four-point Likert scale labeled "True," "Somewhat True," "Somewhat False," and "False." There are 23 lower-order facet scales subsumed under the three higher-order factor scales of *General Disinhibition*, *Callous-Aggression*, and *Substance Abuse*. The 23 facet scales are derived from Krueger and colleagues' (2007) model of externalizing psychopathology that is displayed in *Figure 2.* For the current study, scores from the 23 facet scales were used to measure the 23 indicators in the models displayed in *Figures 4* and *5.* Scores on this measure's facet scales have been demonstrated to display high levels of internal consistency, α = .85-.98 (Patrick, Kramer, Krueger, & Markon, 2013). Additionally, scores on the three higher-order scales have been supported to have good criterion validity when compared with similar scales on the Multidimensional Personality Questionnaire (MPQ; Patrick, Kramer, Krueger, & Markon, 2013).

Minnesota Multiphasic Personality Inventory – 2 – Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011). The MMPI-2-RF is a 338-item broadband self-report inventory that assesses a wide range of personality and psychopathology traits. Participants responded to items by answering "true" or "false" in response to questions about their behavioral, social, and emotional functioning. The instrument includes nine validity scales

assessing response styles and 42 substantive scales assessing social, emotional, and behavioral tendencies. Three scales for indexing triarchic psychopathy constructs using MMPI-2-RF items were also developed by Sellbom and colleagues (2016). These three scales include *Boldness, Meanness,* and *Disinhibition,* and assess the three-factor model of psychopathy (Patrick, Fowles, & Krueger, 2009).

The current study utilized the MMPI-2-RF validity scales, the Restructured Clinical scale *RC3,* the externalizing dysfunction scales *Substance Abuse (SUB)* and *Aggression (AGG),* and the triarchic psychopathy scales *Meanness* and *Boldness.* Scores from the validity scales were used to exclude invalid participant profiles from analyses. Previous literature has supported the use of scores from the *VRIN-r* and *TRIN-r* scales in predicting random and fixed responding styles, respectively, in both clinical and non-clinical samples (Handel, Ben-Porath, Tellegen, & Archer, 2010). The over-reporting scales *F-r, Fs, Fp-r,* and *FBS-r* have been supported as valid measures of over-reporting of symptoms of psychopathology in undergraduate samples (Sellbom & Bagby, 2010). Additionally, the under-reporting scales *L-r* and *K-r* have been demonstrated to be useful measures of under-reporting of symptoms of psychopathology in undergraduate samples (Crighton, Marek, Dragon, & Ben-Porath, 2017).

Scores on *RC3* were used to measure the alienation facet of aggression in this study. *RC3* has 15 items and has been demonstrated to measure cynical views of human nature as well as feelings of social alienation (Ben-Porath & Tellegen, 2011), a concept which is conceptually relevant to the construct of alienation described by Krueger and colleagues (2007). The *RC3* scale has been displayed to have good internal consistency reliability in the MMPI-2 normative sample ($\alpha = .81$ for men and $\alpha = .80$ for women) and adequate test-retest reliability ($r = .76$ for men and $r = .87$ for women). Scores on this scale have also been demonstrated to be

conceptually related to other measures of alienation in undergraduate samples (Ingram, Kelso, & McCord, 2011).

Scores on *SUB* were used to measure the drug problems facet of substance use difficulties in this study. This seven-item scale was designed to measure problematic use of substances including alcohol, illicit drugs, and prescription medications (Ben-Porath & Tellegen, 2011). In the current study, *SUB* was used to measure the indicator *Drug Problems* of Krueger and colleagues' (2007) model. Scores on *SUB* have been demonstrated to show high test-retest reliability (r = .87) in the MMPI-2 normative sample. Additionally, scores on this scale have been demonstrated to be related to a history of problematic drug use and a history of difficulties resulting from drug use (Ben-Porath & Tellegen, 2011).

Scores on *AGG* were used to measure physical aggression in the current study. This nine-item scale was created to assess engagement in physically violent behavior toward others, violence in response to anger, and the enjoyment of thoughts about physical aggression toward others (Ben-Porath & Tellegen, 2011), which is conceptually relevant to the construct of physical aggression described by Krueger and colleagues (2007). Scores on *AGG* have been demonstrated to show decent levels of internal consistency (α = .66) in the MMPI-2 normative sample. Additionally, scores on this scale have been demonstrated to be related to a history of violent behavior toward others as well as abusive tendencies (Ben-Porath & Tellegen, 2011).

Scores on *Boldness* were used to measure relational aggression in this study. This scale, which represents a facet of the triarchic psychopathy scales, was created to measure social dominance. *Boldness* has been demonstrated to display high levels of internal consistency reliability, α = .89/.82 (Sellbom et al., 2016). It has been demonstrated to be conceptually related to other well-validated measures of boldness and social potency (Sellbom & Phillips, 2013),

supporting its use for the measurement of Krueger and colleagues' facet of relational aggression (2007).

Meanness was used to measure the empathy facet of aggression in the current study. This scale has been demonstrated to display high levels of internal consistency reliability, α = .90/.88 (Sellbom et al., 2016). Scores on this scale have been supported to measure tendencies toward disdain for others, exploitativeness, and antagonism in undergraduate samples (Sellbom et al., 2016), which is conceptually relevant to Krueger and colleagues' (2007) description of the empathy facet of aggression. For this study, items 26, 41, 84, 231, 316, and 329 were dropped from the *Meanness* scale because of overlap with the *AGG* scale that is also being used. Additionally, items 36, 55, 87, 99, 142, 185, and 213 were dropped from the *Meanness* scale due to overlap with the *RC3* scale that is being used as well. This assured that variance due to item overlap was accounted for.

Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). The AUDIT consists of 10 items that measure various aspects of problematic alcohol use and difficulties resulting from alcohol use. Participants responded to items with a five-item Likert scale ranging from zero to four, and a sum of the score on all items indicates severity of difficulties with alcohol use. Questions consist of content relating to frequency of alcohol use as well as frequency of problems with alcohol, primarily over the past year. In the current study, this measure was used to assess the substance use facet of alcohol problems described by Krueger and colleagues (2007). The AUDIT has been demonstrated to have adequate test-retest reliability (r = .60-.95) as well as good internal consistency (α = .80; de Meneses-Gaya et al., 2009). Scores on this measure have been previously supported to be related to other measures of alcohol use difficulties as well as measures of biological markers of

these problems, contributing to the convergent validity evidence for its use in detecting difficulties associated with alcohol use in both clinical and non-clinical samples (de Meneses-Gaya et al., 2009). Scores on the AUDIT have displayed a specificity of .71 and a sensitivity of .84 in their ability to correctly identify college students with alcohol use disorder, and this measure has displayed prominent advantages over other measures of alcohol use disorders in their ability to identify undergraduates for these difficulties (Fleming, Barry, & Macdonald, 1991).

Cannabis Use Disorders Identification Test (CUDIT; Adamson & Sellman, 2003). The CUDIT consists of 10 items adapted from the AUDIT to measure problematic cannabis use and difficulties due to the use of cannabis. Participants responded to the first eight items on a five-point Likert scale ranging from zero to four, and the last two items with either "no" or "yes." These items probe for frequency of cannabis use as well as frequency or experience with problems related to cannabis use, primarily across the previous six months. A sum of scores on all items produces a total score indicative of severity of problematic cannabis use. In the current study, this measure was used to assess the substance use facet of marijuana problems described by Krueger and colleagues (2007). Scores on the CUDIT have been demonstrated to have high levels of internal consistency reliability $\alpha = .72$ in non-clinical samples. Previous research has provided convergent validity evidence for scale scores in non-clinical young adult samples by supporting the association between scores on the CUDIT with other measures of problems or dysfunction as a result of cannabis use (Annaheim, Rehm, & Gmel, 2008). This measure has also been validated for its use in identifying young adults with cannabis use difficulties (Annaheim, Rehm, & Gmel, 2008), as well as for identifying individuals with a *DSM* diagnosis of Cannabis Use Disorder (Adamson & Sellman, 2003).

Drug Use Disorders Identification Test (DUDIT; Berman, Bergan, Palmstierna, & Schlyter, 2003). The DUDIT is an 11-item measure, also adapted from the AUDIT, consisting of items measuring substance-related problems and substance use frequency. Participants responded to items one through nine on a five-point Likert scale ranging from zero to four, and the last two items with either "no" or "yes." Scores for this measure are summed for a total score on the measure indicating severity of difficulties with substance use. In the current study, this measure was used to assess the substance use facet of drug use described by Krueger and colleagues (2007). A recent review of previous literature found that scores on the DUDIT have high internal consistency reliability in both non-clinical and clinical samples ($\alpha > .90$), as well as favorable sensitivity (.85-1.00) and specificity (.75-.92) in both non-clinical and clinical samples. Scores on the DUDIT have been previously supported as being associated with other measures of problematic substance use in both clinical and non-clinical samples, supporting the validity of this measure for assessing drug use problems within samples of participants with and without severe substance use difficulties (Voluse et al., 2012).

Premeditation and Negative Urgency Subscales of the UPPS-P Impulsive Behavior Scale (Lynam, Smith, Whiteside, & Cyders, 2006). The UPPS-P is a 59-item self-report measure created to assess impulsivity using the Five Factor Model of personality. Its facets include *Lack of Premeditation* (11 items), *Negative Urgency* (12 items), *Positive Urgency* (14 items), *Sensation Seeking* (12 items), and *Lack of Perseverance* (10 items; Cyders et al., 2007; Whiteside & Lynam, 2001). For the current study, *Premeditation* was used to measure planful control and *Negative Urgency* was used to measure impatient urgency as a part of Krueger and colleagues' model of the externalizing liability (2007). Participants responded to items on a four-item Likert scale ranging from one to four. Item content varies among scales but generally

relates to experiences with and tendencies toward impulsive behavior. Total score and scale scores are calculated by summing normally scored and reverse-scored items for scores that indicate levels of impulsivity across scales.

The subscales of the UPPS-P have been demonstrated to have high levels of internal consistency when used with undergraduate samples, $\alpha = .82\text{-}.91$ (Whiteside & Lynam, 2011; Cyders et al., 2007). Scores on the *Negative Urgency* scale of the UPPS-P have been demonstrated to be associated with measures of external behavior responses such as problematic substance use, self-harm behaviors, and problematic eating behaviors in response to negative affect in undergraduate populations (Kaiser, Milich, Lynam, & Charnigo, 2013; Dir, Karyadi, & Cyders, 2013). *Negative Urgency* scores have also been supported to be positively associated with measures of affective lability and negatively associated with measures of self-control (Dir, Karyadi, & Cyders, 2013), making it a conceptually relevant measure to use for the construct of impatient urgency. Scores on the *Premeditation* scale have been supported to be reflective of tendencies to act without forethought without the presence of affective contexts in undergraduate samples (Billieux, Gay, Rochat, & van der Linden, 2010). Specifically, low scores on this scale have been associated with smoking behaviors (Miller et al., 2003) as well as behavioral measures of delayed discounting (Lynam & Miller, 2004), both of which are thought to be indicative of impulsive behavior without planning for future consequences. This validity evidence supports the use of this measure for the construct of planful control as described by Krueger and colleagues (2007).

Persistence Facet Scale of the Emotionality, Activity, Sociability, and Impulsivity III Scale (EASI-III; Buss & Plomin, 1975). The EASI-III consists of 50 self-report items assessing the temperament facets of emotionality, activity, sociability, and impulsivity.

Participants responded to items on a five-point Likert scale ranging from *1 – Uncharacteristic (Not at ALL like you)* to *5 – Characteristic (Very much like you).* Total and facet scale scores are calculated by summing normally and reverse-scored items to create a score suggesting temperamental tendencies. The 20-item impulsivity scale consists of four facet scales: *Inhibitory Control, Decision Time, Sensation Seeking,* and *Persistence.* The *Persistence* subscale was used in the current study to measure the construct of Dependability as defined by Krueger and colleagues (2007). Scale scores on *Persistence* have been previously demonstrated in a sample of young adults to be moderately high in internal consistency reliability, $\alpha = .66$ (Whiteside & Lynam, 2001). Previous research has supported the association between scores on this scale and scores on other well-validated measures of lack of perseverance and conscientiousness among young adults (Whiteside & Lynam, 2001), suggesting it is a sound measure of dependability for this study.

Boredom Susceptibility Subscale of the Sensation-Seeking Scale (SSS; Zuckerman, Eyesenck, & Eysenck, 1978). The SSS is a forced-choice 40-item self-report measure assessing individual differences in optimal level of stimulation. There are four 10-item subscales: *Thrill and Adventure Seeking, Experience Seeking, Disinhibition,* and *Boredom Susceptibility.* Participants responded to items by picking the option – A or B – that best describes the way they feel or act. Scores are compiled by taking a sum of the answers that are answered in the direction that suggests more sensation-seeking tendencies. For the current study, the *Boredom Susceptibility* scale was used as a measure of the construct boredom proneness as described by Krueger and colleagues (2007). Scores on this scale have been previously supported in a sample of undergraduates to demonstrate high internal consistency reliability, $\alpha = .76$ (Roberti, Storch, & Bravata, 2003). Previous literature has supported the use of scores on *Boredom Susceptibility*

in undergraduate samples to assess proneness to boredom in the context of high need for sensation seeking or environmental stimulation (Vodanovich, 2003). Scores on this scale have been associated with a variety of measures of impulsive behavioral responses to boredom, such as substance use, risky sexual behavior, and conduct problems (Vodanovich, 2003), suggesting that it is a conceptually relevant measure to gauge boredom susceptibility in the current study.

Venturesomeness Scale of Eysenck's I7 Impulsivity Questionnaire (I7; Eysenck, Pearson, Easting, & Allsopp, 1985). The I7 Impulsivity Questionnaire is a 54-item self-report measure consisting of three subscales: 19 items measuring *Impulsivity,* 16 items measuring *Venturesomeness,* and 19 items measuring *Empathy.* Participants responded to items with "yes" or "no" in correspondence to whether or not the item describes their behaviors or feelings in everyday life. Scores are summed for the total and subscale scores with number of items that are coded in the direction of each construct of interest. For the current study, *Venturesomeness* was used to measure the construct of excitement seeking as described by Krueger and colleagues (2007). Scores on this subscale have been demonstrated to display high internal consistency reliability ($\alpha > .80$) in samples of undergraduates (Aluja & Blanch, 2007). Scores on this subscale have been supported to be associated with measures of sensation seeking in undergraduate samples (Aluja et al., 2013), indicating its reflectiveness of a need for novel or exciting stimuli and making it a conceptually appropriate measure of excitement seeking for the current study.

Procedure

Participants were recruited from Ball State University's undergraduate research pool. They signed up for two sessions, exactly one week apart, which were facilitated by trained undergraduate and graduate research assistants. After agreeing to participate, participants were

asked to complete computerized administrations of the MMPI-2-RF as well as self-report and behavioral measures of externalizing difficulties. Self-report measures aside from the MMPI-2-RF were administered anonymously via an online survey website. The MMPI-2-RF was also given anonymously via its publishers' computerized administration software. All behavioral measures were administered on-screen with E-Prime. See Appendix A for a list of all measures included in the larger study. After study completion, all participants were provided with written debriefing information about the study as well as research credit for participation. All data was compiled on a weekly basis by a trained graduate assistant.

Data Analysis. All necessary assumptions of confirmatory factor analysis (CFA) were assessed for the models prior to their estimation. These assumptions included multicollinearity, assessed by evaluating scale correlations, as well as multivariate normality, assessed through the examination of univariate indicators such as skewness, kurtosis, and outliers. Because of the current study's large sample size, its data met one of the key assumptions of the maximum likelihood (ML) estimation method that is most frequently used in CFA analyses. However, the other primary assumption of this method is multivariate normality, an assumption that is often violated when measuring psychopathological traits due to their skewed distribution in the general population. For this reason, normality was required to be examined for the data prior to choosing an estimation method.

After all assumptions were assessed, the four models were estimated using the statistical software Mplus. For each model, the metric of latent variables was first defined by fixing the variance of all factors to 1.0. Error variances were all set to be freely estimated, and all error covariances, indicator cross-loadings, and factor covariances were fixed to 0.0. The appropriate

number of factors for a single or for a bifactor structure were inputted to fit the data in Mplus for each of the four models.

The two previously discussed models of externalizing psychopathology – the single factor model and the bifactor model – were compared in the current study. The single factor model, depicted as a part of *Figure 1,* suggests one underlying liability toward *Externalizing* that best accounts for risk for disorders associated with disinhibition. The bifactor model, depicted in *Figure 2,* suggests the same general *Externalizing* factor as well as two distinct subfactors, *Aggression* and *Substance Use,* that provide additional explanation of risk for subgroups of disorders within those accounted for by the *Externalizing* factor.

Each model was specified first with the ESI-bf and, then, with a set of alternative scales that measure externalizing psychopathology. Comparison of these models measured with the ESI-bf helps to determine whether the previously suggested single factor model or Krueger and colleagues' bifactor model provides a better explanation of externalizing psychopathology when using the measure developed by Krueger and colleagues (2007). In order to further determine which model is most effective at explaining externalizing psychopathology, the two models were again compared using alternative measures of this construct. This was done to potentially provide further support by investigating whether the same results would be found when using different measures of externalizing psychopathology than the one developed alongside the bifactor model.

The models that were estimated for the ESI-bf are depicted in *Figures 4* and *5*. The first includes one single latent construct, *Externalizing,* that gives rise to all 23 indicators of externalizing psychopathology as suggested by Krueger and colleagues (2007). In the first measurement model, displayed in *Figure 4, Externalizing* is depicted as a circle. The

Externalizing factor has 23 indicator variables, depicted by squares in the figure, which are the 23 facets of Krueger and colleagues' (2007) model and are measured with the subscales of the ESI-bf, such as *Relational Aggression, Alcohol Problems,* and *Planful Control.*

The second measurement model estimated for the ESI-bf is depicted in *Figure 5*. The *Externalizing* general factor has the same 23 indicator variables as represented by all of the 23 subscales of the ESI-bf. This model also includes two subfactors, *Aggression* and *Substance Use*, which are depicted as circles in the figure, are uncorrelated with and distinct from the general factor, and account for unique variance within each subset of indicators. The *Aggression* factor gives rise to observable behaviors associated with aggression and antisocial tendencies, and the *Substance* factor gives rise to substance use difficulties. Aside from the observable behaviors accounted for by these two specific factors, there are others which are best explained by the *Externalizing* liability alone and are largely characterized by impulsivity. Each of the two subfactors has six indicator variables which are represented by squares in the figure. The indicator variables for *Aggression* include the *Relational Aggression, Physical Aggression, Destructive Aggression, Empathy, Blame Externalization, and Alienation* scales of the ESI-bf. The indicator variables for *Substance Use* include the *Alcohol Problems, Alcohol Use, Marijuana Problems, Marijuana Use, Drug Problems, and Drug Use* scales of the ESI-bf.

The models that were estimated for the alternative measures of externalizing psychopathology are depicted in *Figures 6* and *7*. The first of these is shown in *Figure 6*. In this model, one latent variable of *Externalizing* is depicted with a circle and 13 indicator variables are represented by squares. These include the MMPI-2-RF scales *RC3, AGG,* and *SUB,* the MMPI-2-RF Triarchic Psychopathy scales *Meannness* and *Boldness,* the UPPS-P *Premeditation* and *Negative Urgency* scales, the EASI-III *Persistence* scale, the I7 *Venturesomeness* scale, the SSS

Boredom Susceptibility scale, and the *AUDIT, CUDIT*, and *DUDIT* scales of substance use difficulties.

The second measurement model estimated for the alternative measures is shown in *Figure 7.* This model is specified with a bifactor structure in which there is a general *Externalizing* factor as well as *Aggression* and *Substance Use* subfactors. The *Externalizing* factor includes the 13 indicator variables previously described in the first model using alternative measures of externalizing psychopathology. The *Aggression* subfactor's four indicator variables for this model include the MMPI-2-RF Triarchic Psychopathy scales *Meannness* and *Boldness* and the MMPI-2-RF scales *AGG* and *RC3*. The *Substance Use* subfactor's four indicator variables include the *AUDIT, CUDIT,* and *DUDIT* scales and the MMPI-2-RF *SUB* scale.

Following estimation, each model was examined for model fit. This was evaluated based on three features: overall goodness of fit, presence or absence of localized areas of strain in the solution, and the model's parameter estimates' interpretability, size, and statistical significance (Brown, 2015). Overall goodness of fit was determined by examining the initial test of fit as well as absolute fit, parsimony corrections, and comparative fit statistics for the model. The initial test of fit was the Chi Square Goodness of Fit value, which was required to be non-significant in order to determine that the model has overall good fit. Next, an absolute fit statistic, SRMR, was examined in order to assess model fit at an absolute level. The lower an SRMR value a model displays, the better its evidence of good fit is; values .08 or lower were considered in the range of good fit. Next, a parsimony correction test, root mean square error of approximation (RMSEA), was used to determine whether there were a large number of freely estimated parameters within the model. Lower values of RMSEA are preferred for the model because they indicate parsimony of fit within the model; values .06 or less were deemed

supportive of good fit. Finally, a comparative fit statistic, CFI, was examined in order to evaluate the model being specified based on theoretical preference with a more restricted, nested baseline model. CFI values between .90 and .95 were used as the range for good comparative fit.

After overall goodness of fit was assessed, localized areas of strain were examined in order to determine whether specific relationships among indicators in the data were reproduced adequately. This is specifically important given the complex nature of these models. The two statistics which were used for assessing localized areas of strain were residuals and modification indices. Residuals were examined to determine whether there were an appropriate number of parameters to estimate the covariance among the indicators. Specifically, the residual variances were observed because they indicate the amount of variance for each indicator that is not explained by the latent factors. Computing modification indices for both fixed and constrained parameters in the model determined how much the overall model fit would decrease if one of the parameters was instead freely estimated. Both residuals and indices small in magnitude were observed as supporting evidence of good fit for the model being examined.

Finally, models were contrasted for best fit using the Satorra-Bentler scaled difference Chi Square (SDCS) test. This test allows for approximation of Chi Square values in non-normal data by using scaling correction factors and degrees of freedom from both nested and comparison models (Brown, 2015). First, the single-factor externalizing psychopathology model was compared with the bifactor model when both were estimated using the ESI-bf. If the SDCS value was significant, this indicated that the bifactor model fit the data better when measured with the ESI-bf because it was the model with more freely estimated parameters. If the SDCS value was non-significant, this suggested that the single-factor model fit the data better when using the ESI-bf. Next, the models were compared when estimated with alternative measures.

The single-factor model of externalizing psychopathology was compared with the bifactor model, and an SDCS value was examined for significance for this comparison. If the value was significant, it suggested that the bifactor model fit the data best when estimated with alternative measures of externalizing psychopathology. Alternatively, if the SDCS value was not significant, this indicated that the single-factor model fit the data better than the bifactor model using these measures.

Results

Assumption Testing

Before measurement models were estimated, assumptions of CFA were first tested. Univariate normality was examined first by looking at skew and kurtosis statistics for each scale; these can be observed in *Tables 1* and *2*. Using the criteria of |1| for skewness for each scale, it was found that more than half of the ESI subscales were highly skewed, while only the AUDIT, CUDIT, and DUDIT measures were highly skewed in the alternative measures. A visual examination of q-q plots and histograms for each scale suggested that the sample's data for a majority of the scales for both measurement models displayed deviations from normality. Although these measures of normality were univariate, looking at these indicators of single scale distribution lends evidence to suggest concern for a violation of multivariate normality as well. Because the behavioral traits associated with externalizing psychopathology are not distributed normally within the general population, this violation of multivariate normality is theoretically understandable.

To examine multicollinearity for the data, an examination of correlations among scales was first conducted. In order to reduce Type I error due to the high number of correlational analyses performed on the data, a Bonferroni calculation was performed to determine

significance of Pearson *r* correlation values for all scale associations. The corrected significance value for the ESI correlation matrix was $p \leq .002$, and $p \leq .004$ for the correlation matrix of alternative measures. In examining these corrected correlations, it was observed that all correlations met the standard criteria of being below $r = .8$ (Kline, 2011). Given that the correlations between scales did not exceed the criterion range and that they are theoretically associated, an assumption of lack of multicollinearity was made.

Next, reliability for the scales was measured with Cronbach's alpha, a test of internal consistency reliability. Although most scales displayed alpha values of at least 0.70, a commonly used cutoff for psychological constructs (Kline, 1999), there were ten scales which displayed lower alpha values than 0.70 (see *Table 1*). Data for these scales were examined to determine whether there were any outliers, and it was determined that there were no significant outliers. These low reliability values could theoretically be due to low base rates of these externalizing traits in college populations. Due to the archival nature of this project, this issue could not be addressed methodologically but was noted for later consideration.

In summary, it was gathered from these tests of CFA assumptions that multivariate normality was violated, while multicollinearity absence was not. Due to the non-normal distribution of the data, the frequently used ML estimation method could not be used to analyze these CFA models. To account for this distribution of data, the MLM estimation method was used along with a robust estimator – the Sartorra-Bentler Chi Square Difference Test – to compare models because it accounts for non-normally distributed data by using a scaling correction factor for each model (Brown, 2015). The sample size of 513 meets the rule-of-thumb sample size recommendation of having at least 250 participants to complete MLM analyses with sufficient power (Kyriazos, 2018).

The four proposed models estimated in analyses met minimum requirements for measurement model identification (Kline, 2011). First, the factor variance of all latent variables was fixed to 1.00 to scale the metric for these factors. Next, it was assured that the model was both theoretically and statistically identified. The models were theoretically identified based on the previous research that has supported their use in explaining externalizing psychopathology. The models were additionally deemed statistically identified because there were an adequate number of indicators for each latent variable across measurement models. The models meet identification criteria because it is possible to obtain a unique set of parameter estimates for each parameter in the model of unknown value (Brown, 2015). Specifically, the number of unknown parameters, such as factor loadings and error variances, never exceeds the number of known parameters, such as set factor variances and set covariances.

Model Testing

The first model estimated was the single-factor model with 23 indicators from the ESI-bf (*Table 6)*. The model successfully converged but displayed poor global fit when looking at fit statistics, which can be examined in *Table 5*. Specifically, this was first determined by an initial test of goodness of fit, the Chi Square, which was significant and therefore not an initial sign of good overall fit for the model. Next, the absolute fit statistic SRMR was examined and it was found that the value did not meet the criteria of < .08. The parsimoniousness of the model was examined with RMSEA, which did not yield a value less than .06, which was the previously deemed criteria for good parsimonious fit. Finally, comparative fit was examined for the model in comparison with a more nested baseline model and it was found that the value was very low at 0.48, when it should have been between 0.90 and 0.95 to meet criteria for good fit. In examining

local fit for this model, it was observed that indicators Lack of Empathy and Alienation displayed poor fit in the model, with standardized loadings of less than 0.30 (Brown, 2015).

Next, the bifactor model fit with the same 23 indicators from the ESI-bf. The loadings for this model are displayed alongside those for the single-factor model in *Table 6.* Although fit statistics neared acceptable ranges for this model, they were not in the range for optimal or good local fit in terms of absolute fit, parsimoniousness, comparative fit, or absolute fit. These values can be observed in *Table 5.* In examining local fit for the model, the indicator Empathy displayed a loading below 0.30 for the EXT factor. Additionally, Blame Externalization and Alienation displayed loadings below this cutoff for the AGG factor. These poor loadings suggest that these indicators were not adding significant variance to the model.

Beginning with relative comparison of the single and bifactor model, it was found that the bifactor model displayed slightly improved global fit over the single-factor model; however, the fit statistics only yielded acceptable fit and the model was not optimal either. These relative fit values can be compared in *Table 5.* A visual comparison of local fit for each model suggested that factor loadings were more parsimonious in the bifactor model than in the single factor model, suggesting that the variance was better accounted for with the bifactor structure. Standardized factor loading comparison values are listed in *Table 6.* The next step in analyses was to examine statistical comparison of the single factor and bifactor model estimated with the ESI subscales. This was completed by calculating the Sartorra-Bentler Chi square difference test, which can be examined in *Table 5*. The test yielded that this scaled difference test value was significant, and therefore the model with more freely estimated parameters – the bifactor model – did indeed fit better.

Next, the single-factor model was estimated with the alternative measures. An examination of the model's fit statistics displayed poor fit similarly to the single factor model previously estimated with the ESI-bf indicators. All values were again out of range for each category of global fit, suggesting the model did not fit well overall. These global fit values can be observed in *Table 5.* In examining local fit, it was observed that EASI Persistence, MMPI-2-RF RC3, and Tri-Boldness displayed standardized loadings below 0.03 (*Table 7*). This again suggested that these indicators may not have been adding substantial variance or improving the fit of this model for externalizing psychopathology.

When the bifactor model was estimated with the alternative measures, it was found that the model mostly displayed acceptable global fit. Fit values can be observed in *Table 5*. Although not all values fell into the optimal ranges, the values for absolute fit, parsimoniousness, and comparative fit were very close to the criterion ranges for good fit previously described. Local fit for the model was next examined, and there were several areas of poor local fit observed. These loadings can be observed in *Table* 7 alongside those of the single-factor model. For the EXT factor, UPPS-P Negative Urgency, EASI Persistence, MMPI-2-RF RC3, Tri-Boldness, DUDIT, and CUDIT displayed standardized loadings below 0.03. Additionally, Tri-Boldness displayed a loading below this value for the AGG factor as well. This suggests that although the model displayed better global fit, there were a variety of issues with local fit that might suggest modification to the model is necessary.

A relative comparison of fit for the alternative measurement models was investigated by examining differences in previously described fit statistics; these values can be observed in *Table 5.* Relative fit values for the bifactor model were much closer to the range of good fit than the values for the single factor model. Next, local fit for each model was compared; these values are

listed in *Table 7.* There was a higher number of areas of local strain in the bifactor model; six indicators displayed loadings below the criteria of 0.03 for the bifactor model, while this issue was only present in three indicators in the single factor model. In terms of statistical comparison, the better fit of the bifactor model was again supported by a significant Satorra-Bentler Chi square difference test value, but this was not determined in the same way as with the first set of models. The Satorra-Bentler value was negative for the alternative measures, a frequently observed issue with this test that is typically due to poor estimation of the nested model (Brown, 2015). As the single-factor model was the nested model in this comparison, the bifactor model was again supported as the better model due to the negative Satorra-Bentler comparison test value. However, it should be noted that neither model provided optimal fit and neither should be determined as "good" fitting models based on the previously described fit statistics.

Discussion

The goal of the current study was to investigate the structure of externalizing psychopathology. Specifically, the purpose of the study was to compare two previously supported models of risk for externalizing difficulties and tendencies – the single factor model and the bifactor model – to determine which model provided better explanation of variance among the externalizing spectrum. The single factor model suggests that similarities among the externalizing difficulties are best explained by one general underlying risk factor known as the externalizing liability which gives rise to all difficulties with disinhibition. Alternatively, the bifactor model suggests that, in addition to the externalizing liability, there are also two additional, independent pathways of risk that contribute alongside the externalizing liability to best explain the similarities and differences among externalizing difficulties. In addition to determining which model fits best, the current study also aimed to investigate this comparison

with two discrete models of measurement. First, the two models were to be compared using ESI scales that were created alongside the bifactor model suggested by Krueger and colleagues (2007). Next, the comparison was done again using scales that measure externalizing psychopathology other than the ESI. Using measures other than the ESI, which was originally developed alongside the bifactor model, was important in determining whether the model would remain most efficient at capturing the spectrum of externalizing psychopathology when using alternative measures to represent the same constructs.

Analyses supported that the bifactor model displayed better fit in the data when measured with the ESI-bf. This is consistent with Krueger and colleagues' (2007) study that used CFA analyses to compare several models of externalizing psychopathology, including the single factor model, to find that the bifactor model had the best comparative fit. Analyses for the current study additionally yielded that the bifactor model again displayed best comparative fit when estimated with alternative measures of externalizing psychopathology. Before the current study, Sellbom (2016) had previously been the only known study to have attempted to compare models of externalizing psychopathology with measures other than the ESI. Sellbom's study also supported the bifactor structure as the best fitting model for externalizing psychopathology when measured with relevant MMPI-2-RF items. The current study therefore provides further support, in conjunction with previous research, for the notion that there is a bifactor pathway of risk for externalizing difficulties that includes both an underlying liability toward disinhibition as well as specific underlying liabilities toward substance use and aggressive tendencies.

The current study's support for the bifactor model is important to the advancement of the understanding of difficulties that lie on the spectrum of externalizing psychopathology. Knowledge of specific underlying liabilities that give rise to these difficulties helps to guide

research investigating the genetic correlates and associated cognitive and behavioral deficits they are associated with, allowing for better development of prevention and treatment strategies to support individuals with problematic externalizing tendencies. If the bifactor model does indeed provide the best explanation for the pathway of risk for externalizing psychopathology as the current and previous studies suggest, this allows for a more nuanced understanding of what makes difficulties associated with disinhibition both similar and different. Specifically, the bifactor model suggests that there might be specific genetic predispositions that make some individuals liable for substance use or aggressive difficulties above and beyond the risk contributed by the externalizing liability alone. This helps explain why some individuals who are high on the externalizing liability develop difficulties with substance use only, while others only experience problems with impulsivity or aggression.

Although support for the bifactor model was demonstrated in the current study, there are notable considerations yielding caution for interpretation of results. Specifically, neither the single-factor or bifactor models provided acceptable fit to the data according to commonly used guidelines for fit when interpreting CFA model estimation indices. Although the bifactor models showed improved fit over the single factor models in global fit, the bifactor models still had a variety of fit statistics that were out of range for good fit. In addition to global fit, all of the models estimated also displayed several areas of strain locally. These issues with fit being present even in the best fitting model of comparison suggests that there may be potential issues with the ability of either of the models to optimally explain the variance among the externalizing spectrum. This might suggest that the actual model of risk for the externalizing psychopathology spectrum of difficulties is neither of the two models compared in the current study, and that this optimal model has not yet been correctly identified in current research.

There were a few specific indicators that highlighted both problems and strengths within the ESI models. First, it was observed that Alienation had poor loadings for both the single and bifactor models, suggesting it was a poor indicator for the model generally. It was also observed that any variance being offered by Alienation was subsumed by the *EXT* and not the *AGG* factor, suggesting that this facet of externalizing psychopathology is best explained by general disinhibition rather than both disinhibition and aggression. In examining the indicator Lack of Empathy, it was observed that although it did not load well in the single factor model, its fit was much more parsimonious in the bifactor model where it was able to contribute more substantially to the *AGG* factor. This highlights the importance of the bifactor model's subfactor *AGG* in explaining the risk for the Lack of Empathy facet of Krueger and colleagues' (2007) model.

Although the indicators for the alternative measurement models were specifically chosen to best align with Krueger and colleagues' (2007) facets of externalizing psychopathology, a variety of the indicators were problematic in terms of loadings. EASI-III Persistence, the indicator representative of Krueger and colleagues' Dependability facet, displayed poor loadings for the *EXT* factor in both the single and bifactor model. This might be suggestive of this measure's lack of ability to capture the construct of dependability as defined by Krueger and colleagues; it could also mean that dependability itself is not an essential aspect of the externalizing psychopathology model. RC3, the scale chosen to represent Krueger and colleagues' Alienation facet, displayed poor loadings for the *EXT* factor for both models, but this was explained by the high loading displayed for this scale on the *AGG* subfactor of the bifactor model. This finding provides support for the importance of the bifactor model in its ability to capture variance for some indicators above and beyond what is captured by the *EXT* liability alone. Tri-Boldness, the indicator used for Kruger and colleagues' Relational Aggression facet,

displayed problems with loadings across the board in both models. One potential explanation for this might be its significant, moderate correlation ($r = 0.34$) with Tri-Meanness, another indicator that was fit to the *EXT* and *AGG* factors. This association is expected given the origin of these two scales in the triarchic model of psychopathy (Patrick, Fowles, & Krueger, 2009). The relationship between these two variables is also theoretically understandable because these two scales are representative of two distinct facets of the same personality construct, psychopathy. This association may have diminished the importance of the Tri-Boldness indicator in the model in comparison with Tri-Meanness.

There are a few possible methodological explanations for the generally poor fit found for the models that were estimated and compared in this study. Primarily, there was likely restriction of range for externalizing psychopathology and traits in this sample, because externalizing characteristics such as impulsivity and risk-taking behaviors are negatively associated with academic success and college enrollment (Spinella & Miley, 2003). In examining loading errors for specific indicators in the estimated models, it can be theoretically assumed that difficulties with many of these scales could have potentially been caused by the low base rates of these externalizing difficulties in a college population. This issue was a limitation of performing analyses for the current study on a previously collected dataset from a study in which a college sample was used. The likely restriction of range in the data may have contributed to the lack of normal distribution observed for the measures in this study. Restriction of range is an issue when using CFA because it leads to reduced variability in the sample, therefore decreasing the ability of the models to fully capture the psychological constructs being measured as they exist in the general population (Brown, 2015). If a sample characterized by higher tendencies toward the externalizing liability, such as a sample of incarcerated prisoners,

had been used, these results could potentially have been clearer. Future research should include samples such as these in order to observe competing modes within a full range of the externalizing spectrum.

Another methodological consideration when analyzing the poor model fit for the current study's analyses was the length of scales used in the study. Many of these were short in length, which has been demonstrated to have a decreasing effect on the internal consistency of scales when used in research (Tavakol & Dennick, 2011). The previously described likely restriction of range could have also had an effect on the internal consistency reliability of the scales used, as this has been demonstrated to be associated with lower Cronbach's alpha values (Fife, Mendoza, & Terry, 2012). Scales with lower internal consistency are likely to increase error variance and lead to poor loadings in CFA models such as those observed in the current study's analyses. As such, future research should aim to use measures with more adequate length to measure these constructs when possible.

In addition to the limitations of the current study already mentioned, one additional limitation is worthy of note and is related to current controversies in research involving factor analysis. This inherent limitation of the current study is the use of a bifactor model in comparison with a nested single-factor model. The use of bifactor models in research has been a topic of recent debate in psychological research due to the pros and cons associated with their use. The primary concern with the use of these models is that, when comparing a previously deemed "optimal" model with a newly suggested bifactor model, a bifactor model typically provides better fit through the estimation of numerous additional parameters (Sellbom & Tellegen, 2019). This frequently leads to the declaration of bifactor models as optimal when their better comparative fit could simply be due to the statistical leverage they have to begin

with. Because of this difficulty, it is important to be sure of conceptual relevance for using bifactor models before deeming them an appropriate comparison model. Further research is necessary to determine whether the *AGG* and *SUB* factors are theoretically necessary in explaining externalizing psychopathology. This should be addressed by exploring models of externalizing psychopathology other than the two compared in the current study to determine whether there is a more optimal explanation of variance among the observed difficulties.

Given the number of methodological shortcomings and limitations of analyses that might have accounted for poor fit in the estimated models, it is easy to assume poor fit in the current study resulted from these influences alone. However, there are also potential theoretical explanations that warrant consideration. Primarily, these issues might suggest that there is something missing from the bifactor model that inhibits its ability to explain the variance among externalizing difficulties. Specifically, the facets of externalizing psychopathology as described by Krueger and colleagues (2007) might be better explained by a different model with either rearranged or alternative underlying liabilities. For example, recent HiTOP research has suggested the externalizing liability to be instead split into two distinct liabilities – disinhibited externalizing and antagonistic externalizing (Kotov et al., 2017). This was suggested based on their findings that suggest two personality counterparts to the externalizing spectrum: disinhibition and antagonism. Kotov and colleagues (2017) suggest that disinhibition is primarily associated with substance use difficulties, while antagonism and disinhibition are both prominent in problems with aggression. This presents a potential method of examining the role of disinhibition and antagonism more broadly in their risk for externalizing difficulties, and it might provide insight for why poor fit was observed for the bifactor model in the current study. Generally, future research in externalizing difficulties should aim to parse apart more clearly the

differences as well as the similarities between these traits and problems to determine whether there is a more theoretically sound model to explain the variance among these disorders. Given the generally poor fit of both the single factor and the bifactor models found in this study, future studies should investigate the construct validity of the suggested factors that contribute risk for externalizing psychopathology. Specifically, this work should aim to pursue construct validation by comparing a variety of combinations of these factors' ability to correctly predict theoretically appropriate external criteria. These efforts are vital in advancing our knowledge of the risks likely to give rise to problems associated with disinhibition, risky behaviors, and impulsivity so that more appropriate methods of assessment and intervention can be developed.

As a final note, there are several characteristics of CFA analyses that are important to note when interpreting the results obtained from the current study's analyses. Primarily, these models are purely statistical and do not provide tangible insight into the specific mechanisms of risk for various facets of externalizing psychopathology (Brown, 2015). The sole ability of these models is to allow for a visualization of the patterns of variance accounted for by pre-specified groupings of indicators and factors. This means that, if an optimal model of risk for externalizing psychopathology were to be identified and repeatedly supported in future research, it would not provide a direct link to prevention or treatment strategies, or even a theoretical understanding of specific mechanisms that give rise to these disorders. However, if the appropriate factor structure of this spectrum were to be identified, this would guide future research to investigate specific downstream effects of these underlying liabilities so that a practical understanding of the distinct mechanisms that give rise difficulties on the externalizing spectrum can be identified and utilized for prevention and treatment.

Another characteristic of the analyses used that is important to consider when interpreting results is that the models examined used indicators with theoretically similar difficulties that are highly skewed. Skewness present in data when CFA analyses are used allows for the potential to give rise to an artificial "difficulty factor" (Sellbom & Tellegen, 2019) that groups indicators based on their sharing of severe psychopathology rather than the externalizing difficulties specifically. The notion of this difficulty factor is based in educational research which had suggested a difficulty factor for items measuring intelligence for which "high" scores could only be obtained by individuals who were able to correctly solve a majority of problems in the scale. This is translated to psychological research to be a "severity factor" due to "high" scores on psychological measures only being obtained by individuals who endorse a majority of the psychological difficulties or psychopathological constructs being surveyed for in the measure. In line with this thinking, it should be questioned whether the current study's sample had an appropriate distribution of severity level of the externalizing traits surveyed. Specifically, the college sample used might not have displayed the full spectrum of severity necessary to assume that the factors found were representative of the externalizing psychopathology models suggested. Previously, the bifactor model was supported by both Krueger and colleagues (2007) and Sellbom (2017) in samples with assumingly broader spectrums of severity, such as samples of incarcerated prisoners. It should be noted that this model might not have been as optimal in the current study based on a lack of severe difficulties with externalizing psychopathology in the undergraduate sample used.

In summary, the current study aimed to investigate the structure of externalizing psychopathology by comparing two previously suggested models of the pathways of risk for this spectrum of difficulties. The two models compared were the single-factor model and the bifactor

model, and the bifactor model was suggested through CFA estimation methods to have provided best comparative fit for the data. This finding was supported both when estimated with the ESI, the measure developed alongside the bifactor model, and when estimated with alternative relevant measures of the same constructs. These results suggested that there are important differences in the various facets of externalizing psychopathology that are not well explained by the single-factor model and better explained by a bifactor pathway of risk, and that these differences hold up across measurement methods. However, although the bifactor model seemed to fit best of the two, it did not display acceptable fit, suggesting that further research needs to be done in this area. This future research should aim to address the methodological limitations of the current study, as well as to investigate the potential theoretical limitations of the bifactor model by comparing it with other suggested models of externalizing psychopathology not examined in the current study. Developing an understanding of the model that best explains the variance in this spectrum of difficulty allows for specification of the underlying risk factors that give rise to difficulties with disinhibition so that a better understanding of how to prevent and treat them can be achieved.

References

Adamson, S. J., & J. Sellman, D. (2003). A prototype screening instrument for cannabis use disorder: The Cannabis Use Disorders Identification Test (CUDIT) in an alcohol-dependent clinical sample. *Drug and Alcohol Review, 22*(3), 309-315. doi:10.1080/0959523031000154454

Aluja, A., & Blanch, A. (2007). Comparison of impulsiveness, venturesomeness and empathy (I7) structure in English and Spanish samples: Analysis of different structural equation models. *Personality and Individual Differences, 43,* 2294-2305. doi:10.1016/j.paid.2007.07.019

Aluja, A., Escorial, S., Garcia, L. F., Garcia, O., Blanch, A., & Zuckerman, M. (2013). Reanalysis of Eysenck's, Gray's and Zuckerman's psychobiological personality models based in a new Zuckerman's measure: The Zuckerman-Kuhlman-Aluja Personality Questionnaire (ZKA-PQ). *Personality and Individual Differences, 54*(2), 192-196. doi:10.1016/j.paid.2012.08.030

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Arlington, VA: American Psychiatric Publishing.

Annaheim, B., Rehm, J., & Gmel, G. (2008). How to screen for problematic cannabis use in population surveys. *European Addiction Research, 14,* 190-197. doi:10.1159/000141643

Ben-Porath, Y. S., & Tellegen, A. (2008). Minnesota Multiphasic Personality Inventory – 2 – restructured form: Manual for administration, scoring, and interpretation. Minneapolis, MN: University of Minnesota Press.

Ben-Porath, Y. S., & Tellegen, A. (2011). MMPI–2–RF (Minnesota Multiphasic Personality Inventory–2 Restructured Form) manual. Minneapolis: University of Minnesota Press.

Berman, A. H., Bergman, H., Palmstierna, T., & Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *European Addiction Research, 11,* 22-31. doi:10.1159/000081413

Billieux, J., Gay, P., Rochat, L., & Van der Linden, M. (2010). The role of urgency and its underlying psychological mechanisms in problematic behaviours. *Behaviour Research and Therapy, 48,* 1085-1096. doi:10.1016/j.brat.2010.07.008

Blashfield, R. K., Flanagan, E., & Raley, K. (2010). Themes in the evolution of the 20th-century DSMs. In T. Millon, R. F. Krueger, & E. Simonsen (Eds.), *Contemporary directions in psychopathology: Scientific foundations of the DSM-5 and ICD-11* (53-72). New York, NY: Guilford Press.

Bobova, L., Finn, P. R., Rickert, M. E., Lucas, J. (2009). Disinhibitory psychopathology and delay discounting in alcohol dependence: Personality and cognitive correlates. *Experimental and Clinical Psychopharmacology, 17,* 51-61. doi:10.1037/a0014503

Brown, T. A. (2015). Confirmatory factor analysis for applied research: Second edition. New York, NY: The Guilford Press.

Buss, A., & Plomin, R. (1975). A temperament theory of personality development. New York: Wiley.

Crighton, A. H., Marek, R. J., Dragon, W. R., & Ben-Porath, Y. S. (2017). Utility of the MMPI-2-RF validity scales in detection of simulated underreporting: Implications of incorporating a manipulation check. *Assessment, 24*(7), 853-864. doi:10.1177/1073191115627011

Cyders, M. A., Smith, G. T., Spillane, N. S., Fischer, S., Annus, A. M., & Peterson, C. (2007). Integration of impulsivity and positive mood to predict risky behavior: Development and validation of a measure of positive urgency. *Psychological Assessment, 19,* 107-118. doi:10.1037/1040-3590.19.1.107

De Meneses-Gaya, C., Zuardi, A. W., Loureiro, S. R., & Crippa, J. A. S. (2009). Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of psychometric properties. *Psychology and Neuroscience, 2*(1), 83-97. doi:10.3922/j.psns.2009.1.12

Decker, H. S. (2013). *The making of DSM-III: A diagnostic manual's conquest of American psychiatry.* New York, NY: Oxford University Press.

Dir, A. L., Karyadi, K., & Cyders, M. A. (2013). The uniqueness of negative urgency as a common risk factor for self-harm behaviors, alcohol consumption, and eating problems. *Addictive Behaviors, 38*(5), 2158-2162. doi:10.1016/j.addbeh.2013.01.025

Eaton, N. R., South, S. C., & Krueger, R. F. (2012). The meaning of comorbidity among common mental disorders. In T. Millon, R. F. Krueger, & E. Simonsen (Eds.), *Contemporary directions in psychopathology: Scientific foundations of the DSM-V and ICD-11* (223-241). New York, NY: Guilford Press.

Eysenck, S. B. G., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences, 6*(5), 613-619. doi:10.1016/0191-8869(85)90011-X

Fife, D. A., Mendoza, J. L., & Terry, R. (2012). The assessment of reliability under range restriction: A comparison of α, ω, and test–retest reliability for dichotomous data. *Educational and Psychological Measurement, 72*(5), 862-888. doi:10.1177/0013164411430225

Fleming, M. F., Barry, K. L., & Macdonald, R. (1991). The Alcohol Use Disorders Identification Test (AUDIT) in a college sample. *International Journal of the Addictions, 26*(11), 1173-1185. doi:10.3109/10826089109062153

Freudenberg, N., Daniels, J., Crum, M., Perkins, T., & Richie, B. E. (2008). Coming home from jail: The social and health consequences of community reentry for women, male adolescents, and their families and communities. *American Journal of Public Health, 98,* 191-202. doi:10.2105/AJPH.2004.056325

Handel, R. W., Ben-Porath, Y. S., Tellegen, A., & Archer, R. P. (2010). Psychometric functioning of the MMPI-2-RF VRIN-r and TRIN-r scales with varying degrees of randomness, acquiescence, and counter-acquiescence. *Psychological Assessment, 22*(1), 87-95. doi:10.1037/a0017061

Hann, D. M., & Borek, N. (2001). *Taking stock of risk factors for child/youth externalizing behavior problems.* Rockville, MD: Department of Health and Human Services.

Hicks, B. M., Bernat, E., Malone, S. M., Iacono, W. G., Patrick, C. J., Krueger, R. F., & McGue, M. (2007). Genes mediate the association between P3 amplitude and externalizing disorders. *Psychophysiology, 44,* 98-105. doi:10.1111/j.1469-8986.2006.00471.x

Ingram, P. B., Kelso, K. M., & McCord, D. M. (2011). Empirical correlates and expanded interpretation of the MMPI-2-RF Restructured Clinical Scale 3 (Cynicism). *Assessment, 18*(1), 95-101. doi:10.1177/1073191110388147

Kaiser, A. J., Milich, R., Lynam, D. R., & Charnigo, R. J. (2013). Negative urgency, distress tolerance, and substance abuse among college students. *Addictive Behaviors, 37*(10), 1075-1083. doi:10.1016/j.addbeh.2012.04.017

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry, 60,* 929-937. doi:10.1001/archpsyc.60.9.929

Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopath: History, neuroscience, treatment, and economics. *Jurimetrics, 51,* 355-397.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guildford.

Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A., … Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454-477. doi:10.1037/abn0000258

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology, 2,* 111-133. doi:10.1146/annurev.clinpsy.2.022305.095213

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology, 116,* 645-666. doi:10.1037/0021-843X.116.4.645

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology, 111,* 411-424. doi:10.1037//0021-843X.111.3.411

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 9,* 2207-2230. doi:10.4236/psych.2018.98126

Linden, D. E. J. (2005). The P300: Where in the brain is it produced and what does it tell us? *The Neuroscientist, 11*(6), 563-576. doi:10.1177/1073858405280524

Lynam, D. R., & Miller, J. D. (2004). Personality pathways to impulsive behavior and their relations to deviance: Results from three samples. *Journal of Quantitative Criminology, 20,* 319-341. doi:10.1007/s10940-004-5867-0

Lynam, D. R., Smith, G. T., Whiteside, S. P., Cyders, M. A. (2006). *The UPPS-P: Assessing five personality pathways to impulsive behavior (technical report).* West Lafayette, IN: Purdue University.

Miller, J., Flory, K., Lynam, D. R., & Leukefeld, C. (2003). A test of the four-factor model of impulsivity-related traits. *Personality and Individual Differences, 34,* 1403-1418. doi:10.1016/S0191-8869(02)00122-8

Miller, T., & Hendrie, D. *Substance Abuse Prevention Dollars and Cents: A Cost-Benefit Analysis,* DHHS Pub. No. (SMA) 07-4298. Rockville, MD: Center for Substance Abuse Prevention, Substance Abuse and Mental Health Services Administration, 2008.

Newcomb, M. D., & Bentler, P.M. (1988). *Consequences of adolescent drug use: Impact on the lives of young adults.* Thousand Oaks, CA: Sage Publications, Inc.

NIAAA. (2010). Alcohol facts and statistics. *NIH: National Institute on Alcohol Abuse and Alcoholism.* Accessed July 13, 2018.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220-246. doi:10.1037/0033-2909.126.2.220

Nolen-Hoeksema, S., & Watkins, E. R. (2011). A heuristic for developing transdiagnostic models of psychopathology: Explaining multifinality and divergent trajectories. *Perspectives on Psychological Science, 6*(6), 589-609. doi:10.1177/1745691611419672

Patrick, C. J., Curtin, J. J., & Krueger, R. F. (2010). The externalizing spectrum: Structure and mechanisms. In D. Barch (Ed.), *Cognitive and Affective Neuroscience of Psychopathology*. New York: Oxford University Press.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology, 21,* 913-938. doi:10.1017/S0954579409000492

Patrick, C. J., Kramer, M. D., Krueger, R. F., & Markon, K. E. (2013). Optimizing efficiency of psychopathology assessment through quantitative modeling: Development of a brief form of the Externalizing Spectrum Inventory. *Psychological Assessment, 25,* 1332-1348. doi:10.1037/a0034864

Roberti, J. W., Storch, E. A., & Bravata, E. (2003). Further psychometric support for the Sensation Seeking Scale-Form V. *Journal of Personality Assessment, 81*(3), 291-292. doi:10.1207/S15327752JPA8103_12

Saunders, J. M., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. E. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction, 88*(6), 791-804. doi:10.1111/j.1360-0443.1993.tb02093.x

Sellbom, M. (2016). Elucidating the validity of the externalizing spectrum of psychopathology in correctional, forensic, and community samples. *Journal of Abnormal Psychology, 125*(8), 1027-1038. doi:10.1037/abn0000171

Sellbom, M., & Arbisi, P. (2017). Introduction to the special section: Linking the MMPI–2–RF to contemporary models of psychopathology. *Journal of Personality Assessment, 99,* 337-340. doi:10.1080/00223891.2016.1267642

Sellbom, M., & Bagby, R. M. (2010). Detection of overreported psychopathology with the MMPI-2 RF form validity scales. *Psychological Assessment, 22*(4), 757-767. doi:10.1037/a0020825

Sellbom, M., & Phillips, T. R. (2013). An examination of the triarchic conceptualization of psychopathy in incarcerated and nonincarcerated samples. *Journal of Abnormal Psychology, 122*(1), 208-214. doi:10.1037/a0029306

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment,* 1-14. doi:10.1037/pas0000623

Sellbom, M., Drislane, L. E., Johnson, A. K., Goodwin, B. E., Phillips, T. R., & Patrick, C. J. (2016). Development and validation of MMPI-2-RF scales for indexing triarchic psychopathy constructs. *Assessment, 23*(5), 527-543. doi:10.1177/1073191115590853

Spinella, M., & Miley, W. M. (2003). Impulsivity and academic achievement in college students. *College Student Journal, 37*(4), 545-555.

Swendsen, J., Conway, K. P., Degenhardt, L., Glantz, M., Jin, R., Merikangas, K. R., … & Kessler, R. C. (2010). Mental disorders as risk factors for substance use, abuse and

dependence: Results from the 10-year follow-up of the national comorbidity survey. *Addiction, 105*(6), 1117-1128. doi:10.1111/j.1360-0443.2010.02902.x

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55. doi:10.5116/ijme.4dfb.8dfd

Vodanovich, S. J. (2003). Psychometric measures of boredom: A review of the literature. *The Journal of Psychology, 137*(6), 569-595. doi:10.1080/00223980309600636

Voluse, A. C., Gioia, C. J., Sobell, L. C., Dum, M., Sobell, M. B., & Simco, E. R. (2012). Psychometric properties of the Drug Use Disorders Identification Test (DUDIT) with substance abusers in outpatient and residential treatment. *Addictive Behaviors, 37*(1), 36-41. doi:10.1016/j.addbeh.2011.07.030

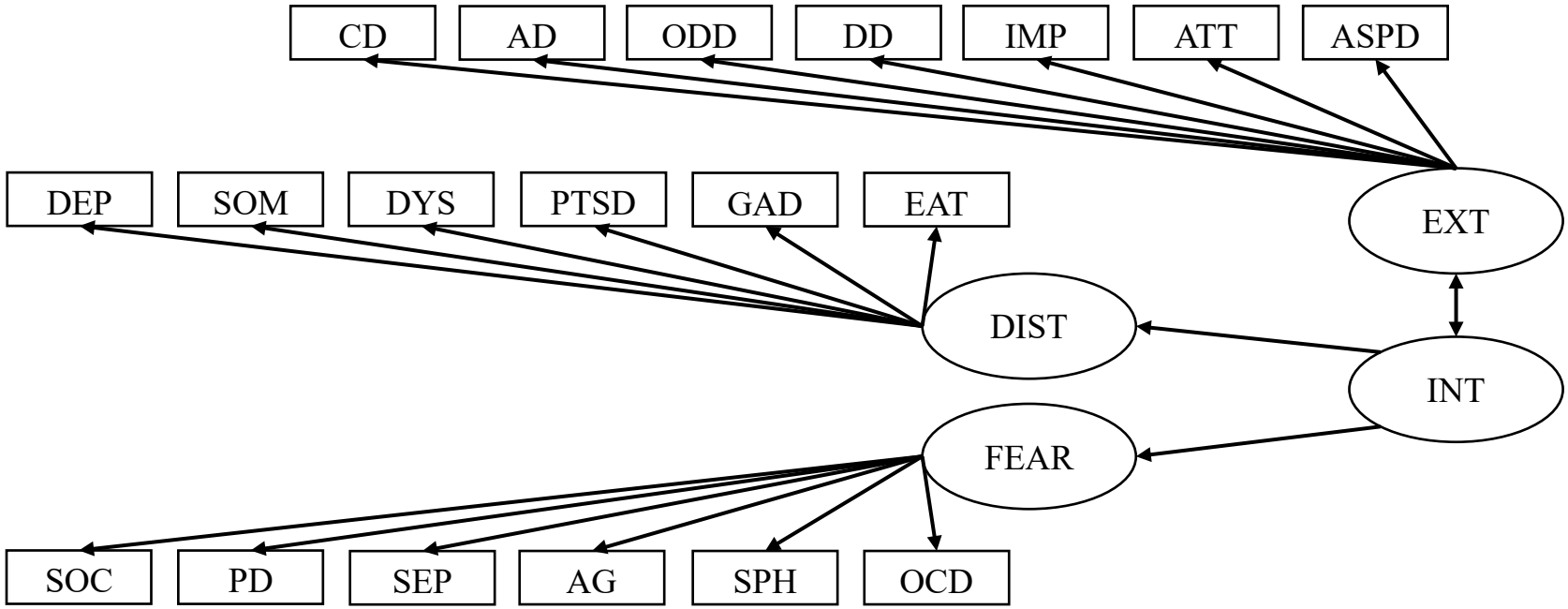
Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30*(4): 669-689. doi:10.1016/S0191-8869(00)00064-7

Whiteside, S. P., Lynam, D. R., Miller, J. D., Reynolds, S. K. (2003). Understanding the role of impulsivity and externalizing psychopathology in alcohol abuse: Application of the UPPS impulsive behavior scale. *Experimental and Clinical Psychopharmacology, 11*(3), 210-217. doi:10.1037/1064-1297.11.3.210

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition. *Journal of Abnormal Psychology, 114*(4), 494-504. doi:10.1037/0021-843X.114.4.494

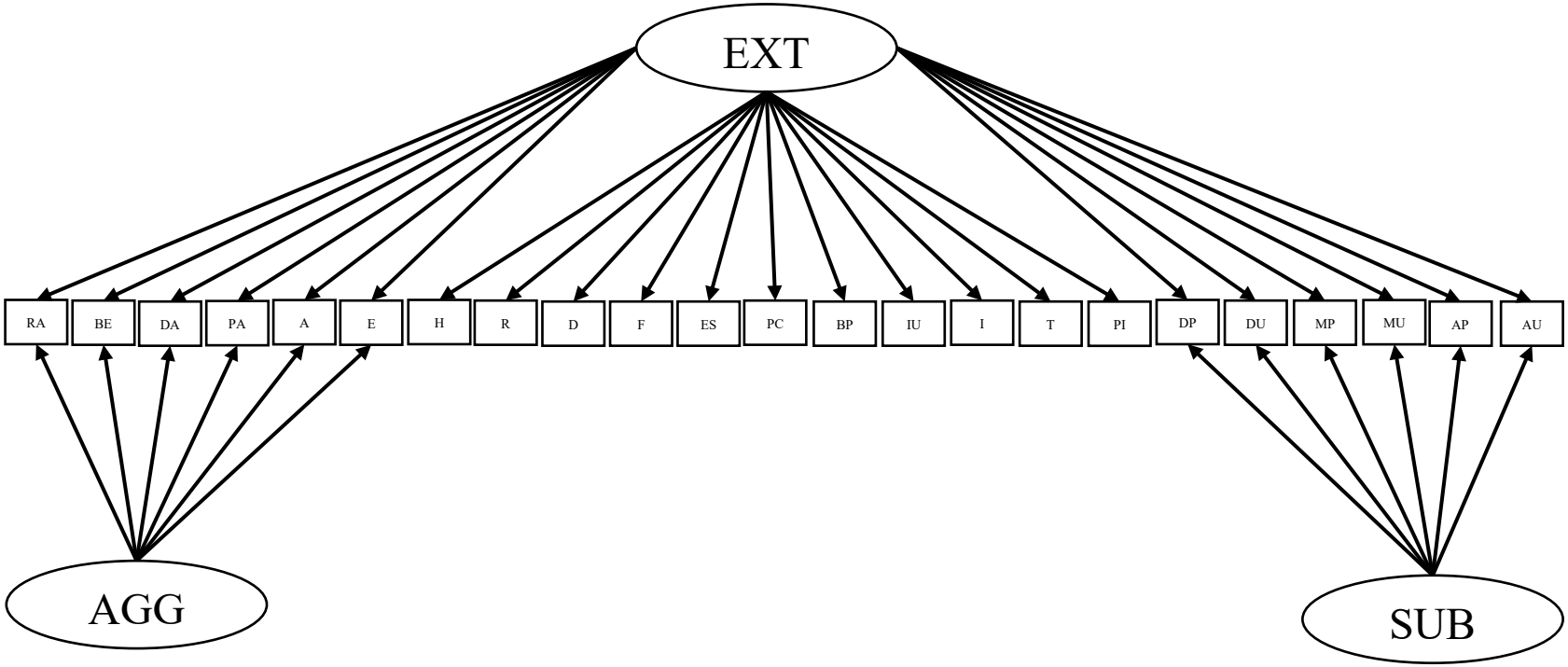
Zuckerman, M., Eysenck, S., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology, 46*(1), 139-149. doi:10.1037/0022-006X.46.1.139

Figure 1. The Multivariate Correlated Liabilities Model (MCLM; Eaton, South, & Krueger, 2012; Krueger & Markon, 2006).



INT = Internalizing Liability; EXT = Externalizing Liability; DIST = Distress Sub-dimension; FEAR = Fear Sub-dimension; DEP = Major Depressive Disorder; DYS = Dysthymia; GAD = Generalized Anxiety Disorder; SOM = Somatization Disorders; PTSD = Post-Traumatic Stress Disorder; EAT = Bulimia/Binge Eating Disorder; SOC = Social Phobia; SEP = Separation Anxiety Disorder; SPH = Specific (Simple) Phobias; PD = Panic Disorder; AG = Agoraphobia; OCD = Obsessive-Compulsive Disorder; ATT = Inattention; IMP = Impulsivity/Hyperactivity; ODD = Oppositional Defiant Disorder; CD = Conduct Disorder; ASPD = Antisocial Personality Disorder; DD = Drug Dependence; AD = Alcohol Dependence.

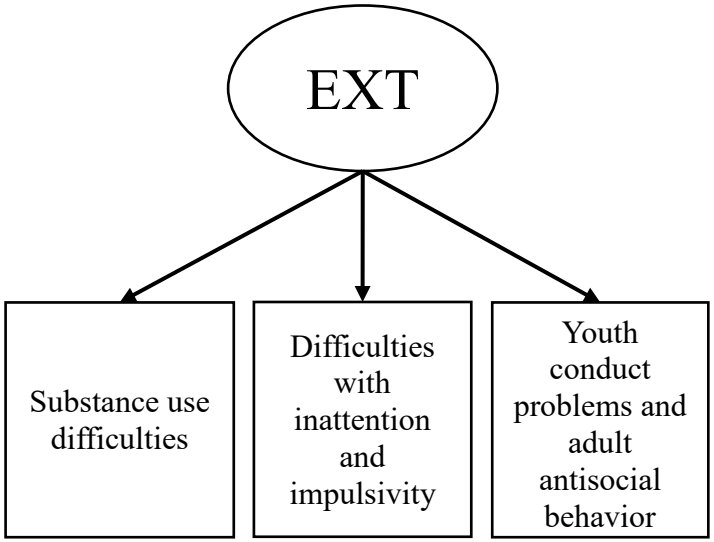
Figure 2. Expanded Model of the Adult Externalizing Spectrum (Krueger et al., 2007).



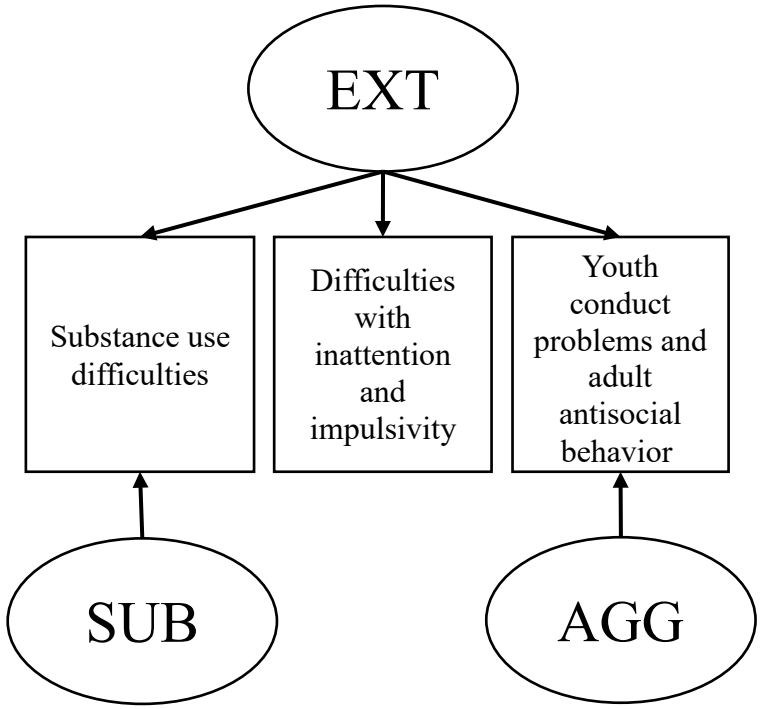
EXT = Externalizing Liability; AGG = Aggression Factor; SUB = Substance Use Factor; RA = Relational Aggression; BE = Blame Externalization; DA = Destructive Aggression; PA = Physical Aggression; A = Alienation; E = Empathy; H = Honesty; R = Rebelliousness; D = Dependability; F = Fraud; ES = Excitement Seeking; PC = Planful Control; BP = Boredom Proneness; IU = Impulsive Urgency; I = Irresponsibility; T = Theft; PI = Problematic Impulsivity; DP = Drug Problems; DU = Drug Use; MP = Marijuana Problems; MU = Marijuana Use; AP = Alcohol Problems; AU = Alcohol Use.

Figure 3. The Single Liability Model of Externalizing Psychopathology, adapted from the Multivariate Correlated Liabilities Model (Eaton, South, & Krueger, 2012; Krueger & Markon, 2006), and the Bifactor Model of Externalizing Psychopathology (Krueger et al. (2007).

(A) Single liability model



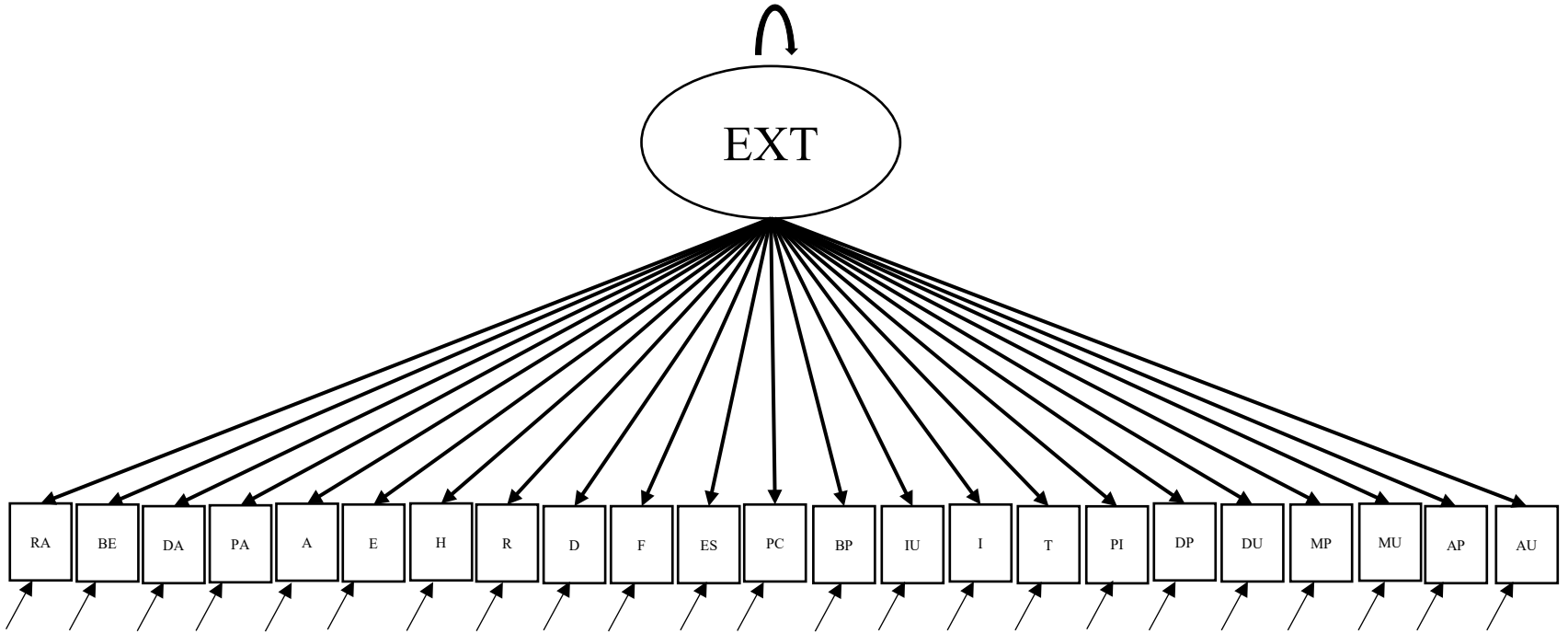
(B) Bifactor model



EXT = Externalizing Liability; AGG = Aggression Factor; SUB = Substance Use Factor;

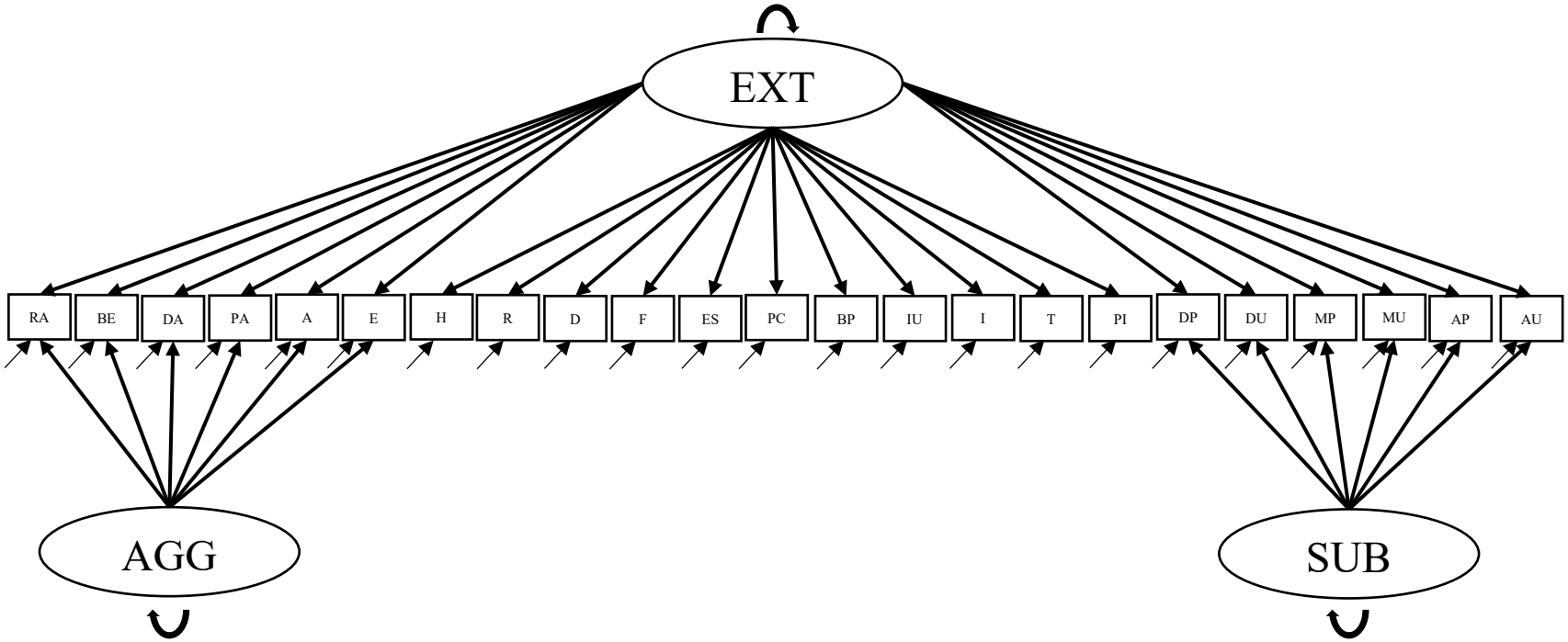
Indicators involve multiple observable behaviors and tendencies.

Figure 4. Proposed Single-Factor Measurement Model Specified with ESI-bf.



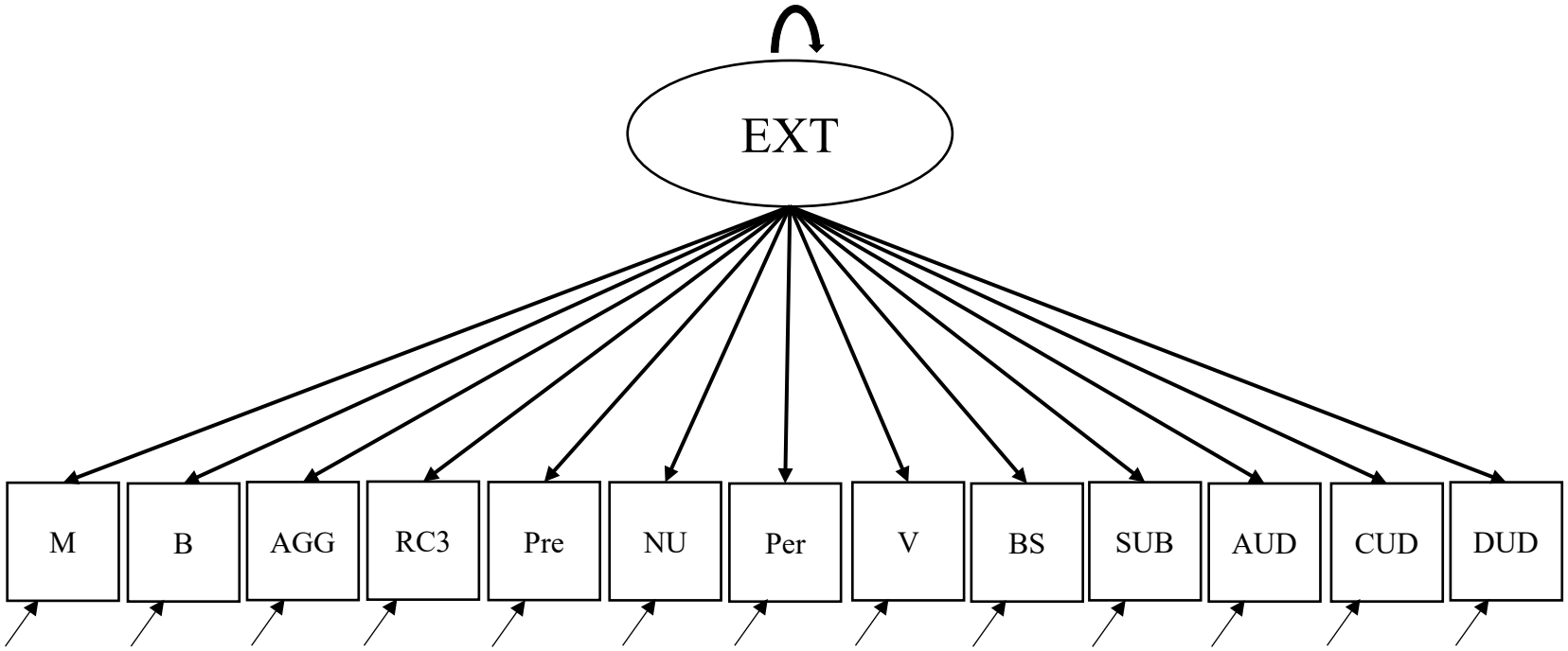
All indicators are subscales from the ESI-bf. EXT = Externalizing Liability; RA = Relational Aggression; BE = Blame Externalization; DA = Destructive Aggression; PA = Physical Aggression; A = Alienation; E = Empathy; H = Honesty; R = Rebelliousness; D = Dependability; F = Fraud; ES = Excitement Seeking; PC = Planful Control; BP = Boredom Proneness; IU = Impulsive Urgency; I = Irresponsibility; T = Theft; PI = Problematic Impulsivity; DP = Drug Problems; DU = Drug Use; MP = Marijuana Problems; MU = Marijuana Use; AP = Alcohol Problems; AU = Alcohol Use.

Figure 5. Proposed Bifactor Measurement Model Specified with ESI-bf.



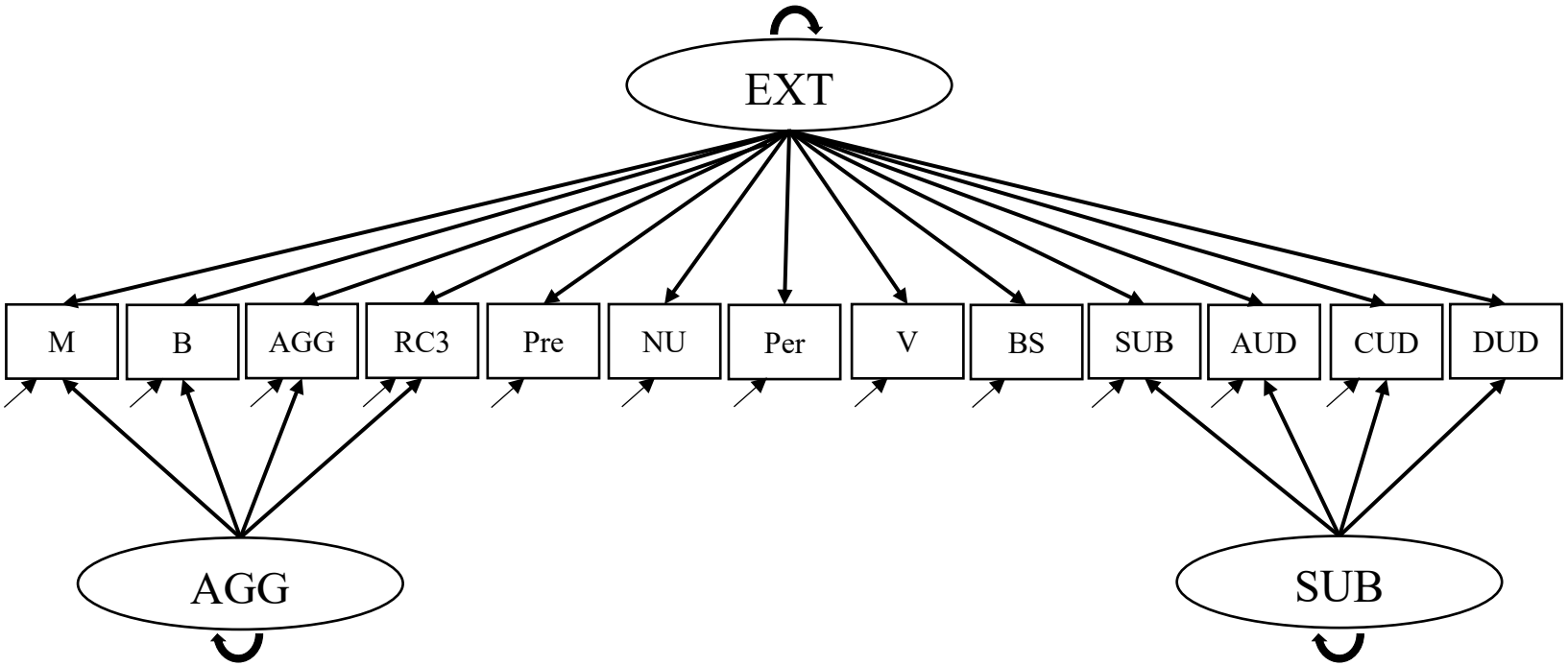
All indicators are subscales from the ESI-bf. EXT = Externalizing Liability; AGG = Aggression Factor; SUB = Substance Use Factor; RA = Relational Aggression; BE = Blame Externalization; DA = Destructive Aggression; PA = Physical Aggression; A = Alienation; E = Empathy; H = Honesty; R = Rebelliousness; D = Dependability; F = Fraud; ES = Excitement Seeking; PC = Planful Control; BP = Boredom Proneness; IU = Impulsive Urgency; I = Irresponsibility; T = Theft; PI = Problematic Impulsivity; DP = Drug Problems; DU = Drug Use; MP = Marijuana Problems; MU = Marijuana Use; AP = Alcohol Problems; AU = Alcohol Use.

Figure 6. Proposed Single Factor Measurement Model Specified with Alternative Measures.



All indicators are subscales from various measures of externalizing psychopathology. EXT = Externalizing Liability; M = Triarchic *Meanness* Scale; B = Triarchic *Boldness* Scale; AGG = MMPI-2-RF *AGG* Scale; RC3 = MMPI-2-RF *RC3* Scale; Pre = UPPS-P *Premeditation* Scale; NU = UPPS-P *Negative Urgency* Scale; Per = EASI-III *Persistence* Scale; V = I7 *Venturesomeness* Scale; BS = SSS *Boredom Susceptibility* Scale; SUB = MMPI-2-RF *SUB* Scale; AUD = *AUDIT;* CUD = *CUDIT;* DUD = *DUDIT.*

Figure 7. Proposed Bifactor Measurement Model Specified with Alternative Measures.



All indicators are subscales from various measures of externalizing psychopathology. EXT = Externalizing Liability; AGG = Aggression Factor; SUB = Substance Use Factor; M = Triarchic *Meanness* Scale; B = Triarchic *Boldness* Scale; AGG = MMPI-2-RF *AGG* Scale; RC3 = MMPI-2-RF *RC3* Scale; Pre = UPPS-P *Premeditation* Scale; NU = UPPS-P *Negative Urgency* Scale; Per = EASI-III *Persistence* Scale; V = I7 *Venturesomeness* Scale; BS = SSS *Boredom Susceptibility* Scale; SUB = MMPI-2-RF *SUB* Scale; AUD = *AUDIT;* CUD = *CUDIT;* DUD = *DUDIT.*

Table 1

ESI Subscale Reliability and Descriptive Statistics (N = 513)

Scale	*α*	*M*	*SD*	Min.	Max.	Skew	Kurtosis
ESI Problematic Impulsivity	0.73	3.94	3.06	1.00	21.00	1.61	3.41
ESI Irresponsibility	0.51	3.80	3.25	0.00	19.00	1.35	2.49
ESI Theft	0.60	1.51	2.62	0.00	17.00	2.23	5.90
ESI Fraud	0.52	1.67	2.07	0.00	11.00	1.52	2.15
ESI Impatient Urgency	0.83	7.16	3.63	0.00	15.00	-0.09	-0.63
ESI Planful Control	0.87	3.26	2.98	0.00	15.00	0.83	0.40
ESI Dependability	0.65	2.90	2.39	0.00	12.00	0.79	0.26
ESI Boredom Proneness	0.89	6.44	3.46	0.00	12.00	-0.21	-0.84
ESI Honesty	0.81	3.89	3.03	0.00	15.00	0.66	-0.03
ESI Rebelliousness	0.85	3.52	3.59	0.00	18.00	1.09	0.70
ESI Excitement Seeking	0.87	5.74	4.39	0.00	18.00	0.49	-0.60
ESI Alienation	0.80	4.14	2.64	0.00	9.00	-0.03	-1.08
ESI Blame Externalization	0.94	4.56	3.82	0.00	12.00	0.22	-1.26

Scale	α	*M*	*SD*	Min.	Max.	Skew	Kurtosis
ESI Physical Aggression	0.68	3.04	3.29	0.00	18.00	1.50	2.23
ESI Destructive Aggression	0.70	0.86	2.03	0.00	17.00	3.63	16.85
ESI Relational Aggression	0.77	4.44	3.95	0.00	20.00	1.13	1.09
ESI Empathy	0.89	5.03	5.26	0.00	31.00	1.75	3.64
ESI Marijuana Use	0.93	7.12	7.76	0.00	21.00	0.54	-1.23
ESI Marijuana Problems	0.84	1.56	3.39	0.00	20.00	2.80	8.00
ESI Drug Use	0.85	4.54	5.18	0.00	18.00	1.00	-0.09
ESI Drug Problems	0.79	1.02	2.91	0.00	25.00	4.29	22.76
ESI Alcohol Use	0.93	12.75	9.02	0.00	27.00	-0.21	-1.42
ESI Alcohol Problems	0.73	2.36	3.39	0.00	18.00	1.88	3.90

* α = Cronbach's Alpha, *M* = Mean, *SD* = Standard Deviation, *Min.* = Minimum Value, *Max.* = Maximum Value.

Table 2

Alternative Measure Subscale Reliability and Descriptive Statistics (N = 513)

Scale	*α*	*M*	*SD*	Min.	Max.	Skew	Kurtosis
UPPS-P Negative Urgency	0.80	26.97	6.29	15.00	44.00	0.17	-0.66
UPPS-P Premeditation	0.87	20.04	5.46	10.00	43.00	0.55	0.16
SSS Boredom Susceptibility	0.50	2.32	1.74	0.00	9.00	0.69	0.10
EASI Persistence	0.55	10.79	3.19	5.00	23.00	0.73	0.46
EIV Venturesomeness	0.72	10.25	3.56	2.00	17.00	-0.26	-0.70
MMPI-2-RF RC3	0.75	7.88	3.11	0.00	15.00	-0.27	-0.43
MMPI-2-RF AGG	0.61	1.85	1.65	0.00	7.00	0.86	0.24
Tri-Boldness	0.74	10.56	3.73	2.00	20.00	-0.14	-0.36
Tri-Meanness	0.60	4.93	2.41	0.00	12.00	0.24	-0.42
DUDIT	0.76	0.71	2.36	0.00	19.00	4.24	20.65
AUDIT	0.79	3.93	4.11	0.00	21.00	1.23	1.36
CUDIT	0.81	2.47	4.34	0.00	26.00	2.32	5.87
MMPI-2-RF SUB	0.69	1.43	1.58	0.00	7.00	0.94	0.02

* *α* = Cronbach's Alpha, *M* = Mean, *SD* = Standard Deviation, *Min.* = Minimum Value, *Max.* = Maximum Value.

Table 3

ESI Subscale Intercorrelations (N = 513)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. PI	-																						
2. IRR	.54*	-																					
3. THF	.35*	.47*	-																				
4. FRD	.38*	.36*	.27*	-																			
5. IU	.35*	.24*	.12	.29*	-																		
6. LPC	.68*	.37*	.20*	.30*	.29*	-																	
7. LD	.34*	.35*	.22*	.29*	.15*	.43*	-																
8. ALI	.24*	.18*	.04	.15*	.26*	.16*	.05	-															
9. BP	.32*	.23*	.14*	.26*	.45*	.27*	.18*	.29*	-														
10. BE	.33*	.26*	.13	.26*	.26*	.20*	.13	.52*	.34*	-													
11. LH	.29*	.15*	.13	.41*	.17*	.43*	.35*	.12	.24*	.15*	-												
12. REB	.58*	.45*	.39*	.29*	.33*	.47*	.28*	.21*	.27*	.26*	.29*	-											
13. PA	.23*	.25*	.24*	.23*	.25*	.10	.12	.16*	.20*	.27*	.13	.29*	-										
14. DA	.24*	.21*	.31*	.28*	.13	.10	.17*	.09	.15	.20*	.15*	.28*	.38*	-									
15. RA	.38*	.26*	.20*	.52*	.39*	.29*	.24*	.17*	.32*	.27*	.41*	.44*	.48*	.46*	-								
16. LE	.16*	.08	.11	.27*	.06	.16*	.25*	.01	.19*	.19	.32*	.21*	.33*	.29*	.45*	-							

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
17. ES	.52*	.38*	.33*	.28*	.36*	.44*	.17*	.20*	.32*	.29*	.20*	.64*	.33*	.25*	.39*	.18*	-						
18. MU	.21*	.34*	.38*	.04	.18*	.14	.05	.10	.13	.09	.11	.47*	.14*	.14*	.10	.04	.33*	-					
19. MP	.31*	.42*	.41*	.08	.11	.16*	.16*	.07	.10	.08	.10	.40*	.12	.18*	.09	.02	.26*	.66*	-				
20. DU	.28*	.37*	.40*	.07	.19*	.21*	.09	.08	.17*	.08	.12	.50*	.14*	.16*	.13*	.06	.38*	.88*	.70*	-			
21. DP	.30*	.34*	.38*	.04	.09	.18*	.18*	.05	.13*	.04	.10	.37*	.07	.18*	.07	.07	..22*	.50*	.80*	.58*	-		
22. AU	.19*	.23*	.18*	.07	.18*	.21*	.02	.01	.12	.08	.14*	.45*	.07	.11	.15*	.03	.36*	.61*	.35*	.57*	.26*	-	
23. AP	.31*	.32*	.25*	.12	.20*	.27*	.12	.16*	.14*	.18*	.15*	.47*	.08	.16*	.21*	.05	.35*	.51*	.43*	.51*	.43*	.68*	-

* $p \leq .002$. Note: PI = Problematic Impulsivity, IRR = Irresponsibility, THF = Theft, FRD = Fraud, IU = Impatient Urgency, LPC = Lack of Planful Control, LD = Lack of Dependability, ALI = Alienation, BP = Boredom Proneness, BE = Blame Externalization, LH = Lack of Honesty, REB = Rebelliousness, PA = Physical Aggression, DA = Destructive Aggression, RA = Relational Aggression, LE = Lack of Empathy, ES = Excitement Seeking, MU = Marijuana Use, MP = Marijuana Problems, DU = Drug Use, DP = Drug Problems, AU = Alcohol Use, AP = Alcohol Problems.

Table 4

Alternative Measure Scale Intercorrelations (N = 513)

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13
1. UPPS-N	-												
2. UPPS-P	.51*	-											
3. SSS-BS	.24*	.29*	-										
4. DUDIT	.15*	.15*	.09	-									
5. AUDIT	.26*	.28*	.18*	.23*	-								
6. CUDIT	.18*	.12	.12	.51*	.42*	-							
7. EASI-P	.32*	.33*	.14*	.14*	.05	.09	-						
8. EIV-V	.16*	.35*	.26*	.12	.27*	.19*	.10	-					
9. RC3	.23*	.05	.14*	.06	.11	.15*	.06	.07	-				
10. AGG	.36*	.13*	.24*	.12	.17*	.18*	.08	.18*	.40*	-			
11. SUB	.26*	.22*	.17*	.37*	.65*	.65*	.10	.29*	.17*	.25*	-		
12. Tri-B	-.18*	.07	.15*	-.04	.16*	.01	-.20*	.32*	-.02	.07	.10	-	
13.Tri-M	.26*	.18*	.27*	.10	.12	.09	.01	.28*	.42*	.52*	.21*	.34*	-

* $p \leq .004$. Note: UPPS-N = UPPS-P Negative Urgency, UPPS-P = UPPS-P Premeditation, SSS-BS = SSS Boredom Susceptibility, EASI-P = EASI Persistence, EIV-V = EIV Venturesomeness, RC3 = MMPI-2-RF Cynicism, AGG = MMPI-2-RF Aggression, SUB = MMPI-2-RF SUB, Tri-B = MMPI-2-RF Tri-Boldness, Tri-M = MMPI-2-RF Tri-Meanness.

Table 5

Confirmatory Factor Model Fit Statistics (N = 513)

ESI Models	**SB χ^2**	***df***	**c**	**RMSEA**	**CFI**	**SRMR**	**ESI Comparison**			
Single-Factor	2418.08	230	1.28	0.14	0.48	0.12	**Δ *df***	**cd**	**TRd**	**CV**
Bifactor	1325.97	218	1.26	0.10	0.74	0.08	12	1.47	958.11	21.03
Alternative Models	**SB χ^2**	***df***	**c**	**RMSEA**	**CFI**	**SRMR**	**Alternative Comparison**			
Single-Factor	850.85	65	1.11	0.15	0.50	0.12	**Δ *df***	**cd**	**TRd**	**CV**
Bifactor	170.95	57	2.60	0.06	0.82	0.08	8	-9.54	-52.01	15.51

* SB χ^2 = Satorra-Bentler Adjusted Chi Square Goodness of Fit, *df* = Degrees of Freedom, c = Scaling Correction Factor, RMSEA = Root Mean Square Error of Approximation, CFI = Comparative Fit Index, SRMR = Standardized Root Mean-Squared Residual, Δ *df* = Change in Degrees of Freedom, cd = Difference Test Scaling Correction, TRd = Sartorra-Bentler Chi Square Difference, CV = Critical Value of Comparison at $p < .05$. Lower fit criterion values indicate better fit. All χ^2 and TRd values significant at $p < .001$.

Table 6

ESI Model Standardized Parameter Estimates

	Single Factor Model	**Bifactor Model**		
ESI Subscale	*EXT*	*EXT*	*AGG*	*SUB*
ESI Problematic Impulsivity	0.67	0.79	-	-
ESI Irresponsibility	0.62	0.63	-	-
ESI Theft	0.53	0.47	-	-
ESI Fraud	0.40	0.50	-	-
ESI Impatient Urgency	0.42	0.48	-	-
ESI Planful Control	0.54	0.67	-	-
ESI Dependability	0.35	0.43	-	-
ESI Boredom Proneness	0.39	0.45	-	-
ESI Honesty	0.36	0.42	-	-
ESI Rebelliousness	0.78	0.76	-	-
ESI Excitement Seeking	0.66	0.69	-	-
ESI Alienation	0.27	0.31	0.01	-

	Single Factor Model	**Bifactor Model**		
ESI Subscale	*EXT*	*EXT*	*AGG*	*SUB*
ESI Blame Externalization	0.34	0.41	0.11	-
ESI Physical Aggression	0.36	0.36	0.50	-
ESI Destructive Aggression	0.37	0.34	0.46	-
ESI Relational Aggression	0.49	0.55	0.59	-
ESI Empathy	0.25	0.26	0.49	-
ESI Marijuana Use	0.64	0.38	-	0.85
ESI Marijuana Problems	0.61	0.40	-	0.63
ESI Drug Use	0.68	0.44	-	0.84
ESI Drug Problems	0.55	0.37	-	0.50
ESI Alcohol Use	0.54	0.37	-	0.51
ESI Alcohol Problems	0.60	0.46	-	0.41

* EXT, AGG, and SUB indicate standardized estimates for latent factors. See Appendix C for unstandardized estimates and standard errors. Loadings listed as - were fixed at 0.

Table 7

Alternative Measure Model Standardized Parameter Estimates

	Single Factor Model	**Bifactor Model**		
Alt. Measure Scale	*EXT*	*EXT*	*AGG*	*SUB*
UPPS-P Negative Urgency	0.40	0.21	-	-
UPPS-P Premeditation	0.37	0.57	-	-
SSS Boredom Susceptibility	0.30	0.49	-	-
EASI Persistence	0.18	0.19	-	-
EIV Venturesomeness	0.39	0.60	-	-
MMPI-2-RF RC3	0.26	0.16	0.54	-
MMPI-2-RF AGG	0.37	0.34	0.57	-
Tri-Boldness	0.13	0.27	0.13	-
Tri-Meanness	0.33	0.44	0.66	-
DUDIT	0.45	0.18	-	0.41
AUDIT	0.69	0.41	-	0.56
CUDIT	0.68	0.24	-	0.69

	Single Factor Model	**Bifactor Model**		
Alt. Measure Scale	*EXT*	*EXT*	*AGG*	*SUB*
MMPI-2-RF SUB	0.84	0.41	-	0.81

* EXT, AGG, and SUB indicate standardized estimates for latent factors. See Appendix D for unstandardized estimates and standard errors. Loadings listed as - were fixed at 0.

Appendix A

Complete List of Measures Included in Larger Study

1. Participant Description Questionnaires
    a. Study Criteria Questionnaire
    b. Demographics Questionnaire
2. Minnesota Multiphasic Personality Inventory – 2 – Restructured Form (Ben-Porath & Tellegen, 2008/2011)
3. Self-Report Measures of Impulsivity
    a. Control Vs. Impulsivity Scale (Tellegen & Waller, 2008)
    b. I7 Impulsivity Subscale (Eysenck et al., 1985)
    c. UPPS-P Premeditation Subscale (Lynam et al., 2006)
    d. SSS Boredom Susceptibility Scale (Zuckerman, 1979)
    e. EASI-III Persistence Facet Scale (Buss & Plomin, 1975)
    f. UPPS-P Perseverance Subscale (Lynam et al., 2006)
    g. UPPS-P Sensation Seeking Subscale (Lynam et al., 2006)
    h. I7 Venturesomeness Subscale (Eysenck et al., 1985)
    i. EASI-III Sensation Seeking Facet Scale (Buss & Plomin, 1975)
    j. UPPS-P Negative Urgency Subscale (Lynam et al., 2006)

k. IPIP Immoderation Scale (Goldberg, 1999)

l. EASI-III Inhibitory Control Facet Scale (Buss & Plomin, 1975)

m. UPPS-P Positive Urgency Subscale (Lynam et al., 2006)

4. Behavioral Measures of Impulsivity

    a. Eriksen Flanker Task (Eriksen & Eriksen, 1974)

    b. Cued Go/No-Go Task (Marczinski & Fillmore, 2003)

    c. Brown-Peterson Task (Kane & Engle, 2000)

    d. Delay Discounting Task (Petry & Casarella, 1999)

    e. TIME Paradigm (Dougherty et al., 2005)

5. Measures of Externalizing Behaviors

    a. Externalizing Spectrum Inventory – Brief Form (Patrick, Kramer, Krueger, & Markon, 2013)

    b. Alcohol Use Disorder Identification Test (World Health Organization, 2001)

    c. Alcohol Use Survey (Sobell & Sobell, 1995)

    d. Cannabis Use Disorder Identification Test – Revised (Adamson et al., 2010)

    e. Drug Use Disorder Identification Test (Berman, Bergman, Palstierna, & Schlyter, 2003)

Appendix B

List of ESI Scales with Descriptions

ESI Subscale	*Traits or Behaviors Measured*	*ESI Subscale*	*Traits or Behaviors Measured*
Blame Externalization	Belief that one has been wrongly and unfairly blamed or accused of something wrong	**Boredom Proneness**	Tendency to become bored easily and to have difficulty remaining entertained
Relational Aggression	Spreading rumors or lies about others, insulting them, or interfering in their relationships as a way to aggress	**Irresponsibility**	Irresponsibility as manifested in various ways, including failure to meet formal agreements or social or vocational expectation
Destructive Aggression	Aggression in the form of vandalism, property destruction, fire setting	**Impatient Urgency**	Impatience and a sense that desires must be satisfied immediately
Physical Aggression	Hitting and striking others, participating in fights, using physical restraint aggressively	**Theft**	Various forms of theft, including burglary and robbery involving property or money
Alienation	Feelings that one has been manipulated, used, or otherwise mistreated; feelings of betrayal by or mistrust toward others	**Problematic Impulsivity**	Impulsivity with negative consequences due to a subjective sense of lacking control over one's own actions
Empathy	Lack of sympathy for others and lack of sensitivity to the effects of one's actions on other people	**Drug Problems**	Problematic drug use that has legal and social consequences and is associated with dependency, withdrawal, and tolerance
Honesty	General tendency toward honesty and truthfulness	**Drug Use**	Experience with drugs, without reference to problematic consequences or specific drug type
Dependability	General conscientiousness and sense of commitment; ability to complete goals set or projects started	**Marijuana Problems**	Problematic marijuana use that has legal and social consequences and is associated with dependency, withdrawal, and tolerance
Planful Control	Thoughtful and inclined to weigh consequences before acting; ability to use foresight and restraint	**Marijuana Use**	Experience with marijuana use, without reference to problematic consequences

Fraud	General tendency to lie or otherwise misrepresent oneself to receive benefits of various sorts; swindling, conning	**Alcohol Problems**	Problematic alcohol use that has legal and social consequences and is associated with dependency, withdrawal, and tolerance
Excitement Seeking	Thrill- or excitement-seeking behavior, including enjoyment of arousal, risk, and excitement	**Alcohol Use**	Experience with alcohol use, without reference to problematic consequences
Rebelliousness	Tendency to violate rules and to disobey others		

Appendix C

Table of Unstandardized Estimates: ESI Model

	Single Factor Model		**Bifactor Model**					
ESI Subscale	*EXT*	*S.E.*	*EXT*	*S.E.*	*AGG*	*S.E.*	*SUB*	*S.E.*
ESI Problematic Impulsivity	2.05*	0.17	2.42*	0.16	-	-	-	-
ESI Irresponsibility	2.01*	0.17	2.04*	0.16	-	-	-	-
ESI Theft	1.39*	0.15	1.22*	0.143	-	-	-	-
ESI Fraud	0.83*	0.11	1.04*	0.12	-	-	-	-
ESI Impatient Urgency	1.51*	0.15	1.72*	0.16	-	-	-	-
ESI Planful Control	1.60*	0.14	1.99*	0.13	-	-	-	-
ESI Dependability	0.842*	0.110	1.04*	0.11	-	-	-	-
ESI Boredom Proneness	1.34*	0.15	1.54*	0.14	-	-	-	-
ESI Honesty	1.09*	0.14	1.28*	0.14	-	-	-	-
ESI Rebelliousness	2.81*	0.16	2.72*	0.16	-	-	-	-
ESI Excitement Seeking	2.91*	0.16	3.01*	0.16	-	-	-	-
ESI Alienation	0.71*	0.12	0.83*	0.12	0.02	0.12	-	-

	Single Factor Model		**Bifactor Model**					
ESI Subscale	*EXT*	*S.E.*	*EXT*	*S.E.*	*AGG*	*S.E.*	*SUB*	*S.E.*
ESI Blame Externalization	1.32*	0.16	1.57*	0.16	0.40*	0.19	-	-
ESI Physical Aggression	1.19*	0.17	1.17*	0.16	1.65*	0.17	-	-
ESI Destructive Aggression	0.75*	0.12	0.70*	0.10	0.92*	0.18	-	-
ESI Relational Aggression	1.93*	0.20	2.18*	0.19	2.33*	0.20	-	-
ESI Empathy	1.30*	0.28	1.37*	0.27	2.56*	0.33	-	-
ESI Marijuana Use	4.97*	0.25	2.93*	0.33	-	-	6.57*	0.18
ESI Marijuana Problems	2.08*	0.25	1.36*	0.24	-	-	2.13*	0.16
ESI Drug Use	3.54*	0.20	2.27*	0.24	-	-	4.33*	0.15
ESI Drug Problems	1.60*	0.25	1.09*	0.22	-	-	1.45*	0.15
ESI Alcohol Use	4.86*	0.30	3.30*	0.37	-	-	4.61*	0.29
ESI Alcohol Problems	2.04*	0.18	1.56*	0.18	-	-	1.38*	0.14

* indicates significance at $p < .05$. Values of "-" were fixed at 0.

Appendix D

Table of Unstandardized Estimates: Alternative Measures Model

	Single Factor Model		**Bifactor Model**					
Alt. Measure Scale	*EXT*	*S.E.*	*EXT*	*S.E.*	*AGG*	*S.E.*	*SUB*	*S.E.*
UPPS-P Negative Urgency	2.54*	0.27	13.72	12.79	-	-	-	-
UPPS-P Premeditation	2.01*	0.25	3.09*	0.29	-	-	-	-
SSS Boredom Susceptibility	0.52*	0.08	0.95**	0.28	-	-	-	-
EASI Persistence	0.56*	0.15	0.62**	0.20	-	-	-	-
EIV Venturesomeness	1.40*	0.15	2.15*	0.19	-	-	-	-
MMPI-2-RF RC3	0.82*	0.14	0.49**	0.17	1.69*	0.16	-	-
MMPI-2-RF AGG	0.61*	0.07	0.57*	0.09	0.95*	0.09	-	-
Tri-Boldness	0.50***	0.20	1.02*	0.25	0.48*	0.19	-	-
Tri-Meanness	0.79*	0.11	1.06*	0.12	1.60*	0.12	-	-
DUDIT	1.06*	0.21	0.41*	0.12	-	-	0.98*	0.19
AUDIT	2.83*	0.18	1.69*	0.23	-	-	2.31*	0.22
CUDIT	2.95*	0.28	1.02*	0.23	-	-	3.00*	0.29

	Single Factor Model		**Bifactor Model**					
Alt. Measure Scale	*EXT*	*S.E.*	*EXT*	*S.E.*	*AGG*	*S.E.*	*SUB*	*S.E.*
MMPI-2-RF SUB	1.32*	0.06	0.65*	0.08	-	-	1.28*	0.08

* indicates significance at $p < .05$. Values of "-" were fixed at 0.