university of groningen

University Medical Center Groningen

University of Groningen

Psychometric Properties of the Externalizing Spectrum Inventory

Soe-Agnie, Sabrina E; Paap, Muirne C S; Nijman, Henk L I; De Jong, Cornelis A J

Published in:
Journal of Personality Assessment

DOI:
10.1080/00223891.2020.1753752

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Soe-Agnie, S. E., Paap, M. C. S., Nijman, H. L. I., & De Jong, C. A. J. (2021). Psychometric Properties of the Externalizing Spectrum Inventory: Replication and Extension across Clinical and Non-Clinical Samples. *Journal of Personality Assessment*, *103*(3), 332-341. https://doi.org/10.1080/00223891.2020.1753752

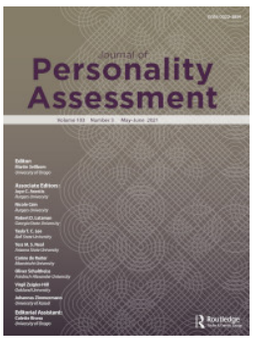
Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 05-06-2022



Routledge
Taylor & Francis Group

Journal of Personality Assessment

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hjpa20

Psychometric Properties of the Externalizing Spectrum Inventory: Replication and Extension across Clinical and Non-Clinical Samples

Sabrina E. Soe-Agnie, Muirne C. S. Paap, Henk L. I. Nijman & Cornelis A. J. De Jong

To cite this article: Sabrina E. Soe-Agnie, Muirne C. S. Paap, Henk L. I. Nijman & Cornelis A. J. De Jong (2021) Psychometric Properties of the Externalizing Spectrum Inventory: Replication and Extension across Clinical and Non-Clinical Samples, Journal of Personality Assessment, 103:3, 332-341, DOI: 10.1080/00223891.2020.1753752

To link to this article: https://doi.org/10.1080/00223891.2020.1753752

© 2020 The Author(s). Published with license by Taylor and Francis Group, LLC

Published online: 24 Apr 2020.

Submit your article to this journal

Article views: 982

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hjpa20

JOURNAL OF PERSONALITY ASSESSMENT
2021, VOL. 103, NO. 3, 332–341
https://doi.org/10.1080/00223891.2020.1753752

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Psychometric Properties of the Externalizing Spectrum Inventory: Replication and Extension across Clinical and Non-Clinical Samples

Sabrina E. Soe-Agnie[1], Muirne C. S. Paap[2,3], Henk L. I. Nijman[1,4], and Cornelis A. J. De Jong[1,5]

[1]Behavioural Science Institute (BSI), Faculty of Clinical Psychology, Radboud University, Nijmegen, The Netherlands; [2]Nieuwenhuis Institute for Educational Research, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands; [3]Clinic Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; [4]Forensic psychiatric institute Fivoor, Den Dolder, The Netherlands; [5]Nijmegen Institute for Scientist-Practitioners in Addiction (NISPA), Radboud University, Nijmegen, The Netherlands

ABSTRACT

The Externalizing Spectrum Inventory aims at assessing personality features that underlie externalizing disorders such as substance abuse and antisocial behaviors. The objective was to replicate the psychometric properties of the 160-item Externalizing Spectrum Inventory in Dutch clinical and non-clinical samples. First, Cronbach's alpha, test-retest reliability and the factor structure were analyzed on a mixed sample of inpatients (n = 149), undergraduates (n = 227), and community participants (n = 178). The factor structure was evaluated through confirmatory and exploratory factor analyses; for the latter Parallel Analysis was used, based on Minimum Rank Factor Analysis. Next, the criterion validity was analyzed using the Aggression Questionnaire and the NEO-Five Factor Inventory as external measures. The Dutch Externalizing Spectrum Inventory subscales showed sufficient reliability (α=.68-.94; ICC=.68-.91), except in the undergraduate sample (α=.49-.96; ICC=.43-.97). The factor structure of the Externalizing Spectrum Inventory was not confirmed and the exploratory analysis yielded different factor solutions across samples. The criterion validity was supported with regard to trait aggression and partly supported with regard to the Five Factor Model. The results suggest that the ESI-160 and its original factor model can be used for prediction purposes. However, further research of the factor structure is strongly recommended.

ARTICLE HISTORY
Received 23 June 2019
Accepted 22 February 2020

Substance use disorders (SUDs) and antisocial behavior disorders are associated with increased rates of crime, violence and recidivism (Durbeej et al., 2015; Dyjkstra, Schumacher, Mota, & Coffey, 2015; Simpson et al., 2015). The comorbidity between these disorders is high, and both substance abuse and disorders, such as antisocial personality disorder and conduct disorder, are currently conceptualized as expressions of a single externalizing spectrum of psychiatric disorders (e.g. Krueger & Markon, 2006; see APA, 2015 and Soe-Agnie et al., 2018). In the externalizing spectrum model, substance abuse and aggressive/antisocial behavior are considered to fall under an overall tendency to engage in antisocial impulsive behaviors. The externalizing spectrum approach to substance abuse and antisocial behavioral disorders has been adopted in the DSM-5 psychiatric classification system (APA, 2015), although little research has been done on this concept in clinical populations (Soe-Agnie et al., 2018). More recently, the Externalizing Spectrum model was adopted in the Hierarchical Taxonomy of Psychopathology (HiTOP) model (Kotov et al., 2017; Krueger et al., 2018). The HiTOP model aims to address shortcomings in traditional nosologies, such as the DSM-classification system. The HiTOP consortium have proposed a novel diagnostic system, which is based on empirical research of the covariation between symptoms and traits. Externalizing spectrum disorders have been associated with personality constructs as is described by, for instance, Krueger & Markon, 2006 and Sher & Trull, 1994. Krueger and South (2009) argued that the Five Factor model of personality is particularly relevant to externalizing disorders, specifically Neuroticism, Agreeableness and Conscientiousness. In a meta-analytic review of 53 research papers, Jones et al. (2011) showed that antisocial and aggressive behaviors had the most robust relations to Neuroticism (weighted mean effect size (*M* ES)=.09 and .17 respectively), Agreeableness (*M* ES= −.31 and -.33) and Conscientiousness (*M* ES= −.18 and -.23). Ruiz et al. (2008) reported adequate weighted mean effect sizes between SUD and Neuroticism (*M* ES=.26), Agreeableness (*M* ES= −.20) and Conscientiousness (*M* ES= −.32) in their meta-analysis based on 63 studies. They showed low to moderate weighted mean effect sizes between antisocial personality and Neuroticism (*M* ES=.10), Agreeableness (*M* ES= −.38) and Conscientiousness (*M* ES= −.30). Following prior research on the externalizing structure of psychiatric disorders, Krueger, Markon, Patrick, & Kramer, (2007) developed a self-report assessment instrument: The Externalizing Spectrum Inventory (ESI). This inventory contains items referring to different forms of personality traits and closely

CONTACT Sabrina Soe-Agnie sabrina.soeagnie@gmail.com Nijmegen Institute for Scientist-Practitioners in Addiction, Behavioral Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands.

© 2020 The Author(s). Published with license by Taylor and Francis Group, LLC
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

related symptoms of impulsive-recklessness, substance abuse and antisocial/aggressive tendencies. It is assumed that these features represent an underlying liability of externalizing spectrum disorders (Kotov et al., 2017; Krueger et al., 2007). This means that persons with a high level of an externalizing personality share a higher risk of developing externalizing disorders than persons with a low level of an externalizing personality. Different brief versions of the full ESI of 415 items have been developed and research has shown moderate to high associations (r=.22 to r=.57) between these versions and antisocial behaviors in adulthood and childhood (Hall et al., 2007; Venables & Patrick, 2012), heightened levels of alcohol dependence and drug abuse (Venables & Patrick, 2012) and neural networks associated with disinhibition and antisocial behavior (Abram et al., 2015; Hall et al., 2007; Nelson et al., 2011). Extending this work, a 160-item version was developed by Patrick and colleagues (Patrick et al., 2013). This version offers a more balanced representation of the full version, than the other brief versions did (Patrick et al., 2013). The ESI-160 was designed to provide efficient but fine-grained measurement of the impulsive-reckless, substance abuse and aggressive/antisocial tendencies, while the other versions only provide screening of the latter two problem areas. Concerning the factor structure of the full ESI, Krueger et al. (2007) reported that a bifactor model consisting of a general externalizing factor and two specific factors reflecting substance abuse and callous aggression was preferred over a one-factor and a two-subfactor higher order model. This model was replicated for the 160-item version by Patrick et al. (2013), who tested the same factor models as Krueger et al. (2007) in a sample derived from the Krueger et al. (2007) study. These analyses were conducted on a combined sample of 1,787 psychology students and prisoners (M_{age} total sample 26.8, $SD = 9.4$), of which 49% were males. The sample characteristics were not reported per subgroup. In addition, Patrick et al. (2013) reported high positive correlations (r=.60 to .63) between the general Externalizing and specific Callous Aggression factor and aggression as measured by the Multidimensional Personality Questionnaire (MPQ, Tellegen & Waller, 2008).

The present study

To date, the vast part of research on the ESI has been conducted using American prison and student samples. The only study of the ESI outside the American population was conducted on the full 415-item version in a Brazilian student sample by Carvalho, Patrick, Krueger, Markon, and Pinheiro (2010). In this study exploratory factor analysis revealed three specific factors, and the bifactor model of Krueger et al. (2007) was not investigated through confirmatory factor analysis. It is currently unclear to what degree the reliability and factor structure of the ESI-160 generalizes to other relevant populations, such as clinical populations and populations outside of the USA. Such studies are urgently needed, since the underlying liability to substance use disorders and antisocial/aggressive tendencies, which the ESI measures, is a highly relevant topic to other populations, as well. For instance, co-occurrence between antisocial behavior and substance abuse is found in 50% of Dutch forensic psychiatric patients residing in correctional institutions (Ministry of Justice and Security [Dienst Justitiële Inrichtingen], 2018). Chan et al. (2008) reported antisocial behaviors in 39.7% young adult patients and in 25.4% patients above 40 years of age, who were admitted to various substance abuse treatment settings in the US. Furthermore, although the externalizing spectrum model refers to psychiatric disorders and the ESI is developed to assess the liability to such disorders, the ESI-160 has not been investigated in psychiatric samples yet. Therefore, the main objective of the current study is to examine the reliability and factor structure of the Dutch ESI-160 in psychiatric inpatients from forensic and addiction care. The second objective is to seek further evidence of the criterion validity of the ESI-160 (Patrick et al., 2013) with regard to other measures of personality, specifically trait aggression and the Five Factor Model of personality. The third objective was to investigate the robustness of the psychometric properties of the Dutch ESI-160 across samples including healthy undergraduate students and community participants. Firstly (h1), we expect to confirm an adequate reliability and we expect to find support for the bifactor model with two specific Callous Aggression and Substance Abuse factors, as was reported in Patrick et al. (2013). Secondly (h2), we hypothesize that the current study will support the criterion validity of the Dutch ESI-160. Specifically, we expect to find that the general ESI-160 factor is strongly and positively correlated to trait aggression, in line with the results reported by Patrick et al. (2013). Based on the results reported by Jones et al. (2011) and Ruiz et al. (2008) we also expect to find that the Callous Aggression factor in particular is strongly and positively correlated to trait aggression and negatively correlated to Agreeableness and Conscientious. We expect that the Substance Abuse factor has an adequate positive correlation with Neuroticism and an adequate negative correlation with Agreeableness and Conscientiousness. Thirdly (h3), we expect that the psychometric properties of the Dutch ESI-160 are comparable across all subsamples, since as far as we know no differences between populations have been described in the literature yet.

Materials and method

Participants

Participants were inpatients, students and community participants.

The inpatient sample (n = 149) consisted for 84.4% of males. The mean age was M_{age}=39.3 (SD = 9.4) The majority (87.5%) was educated at a secondary or lower level), DSM-IV classifications were retrieved for 60.4% of the total patient group. Of these patients, 92.1% met the criteria of a substance use disorder. The patient sample was representative for Dutch forensic psychiatric population regarding age (M_{age} estimate =35.6 to 41.0) and gender (75.0% to 93% male) (see Ministry of Justice and Security [Dienst Justitiële Inrichtingen], 2018; Wisselink et al., 2014).

In the undergraduate psychology student group (n = 227) 12.8% were male, and the mean age was M_{age}=20.7 (SD = 2.9).

In the community sample (n = 178) 74.3% were male, and the mean age was M_{age}=36.6 (SD = 14.4). The majority (71.1%) was educated at a secondary or lower level.

Measures

The *Externalizing Spectrum Inventory* (ESI; Patrick et al., 2013) is a 160-item self-report inventory with four point item scales, scored 0-3. The ESI contains 23 subscales (also indicated as *facets*) assessing a range of expressions of externalizing proneness tapping into aggression, irresponsibility and deceitfulness, impulsivity/sensation-seeking, and substance use/abuse. In previous studies high Cronbach's alphas of $\alpha \geq .85$ (Patrick et al., 2013; Soe-Agnie et al., 2016) and sufficient to high test-retest reliability were found for the ESI-160 subscales, with intraclass correlations ranging from ICC=.69 to .90 (Soe-Agnie et al., 2016). The subscales are thought to be subsumed under one general factor of Disinhibition and two specific factors of Substance Abuse and Callous Aggression (Patrick et al., 2013). The results reported by Patrick et al. (2013) supported the criterion validity of the ESI-160, since the three factors (Patrick et al., 2013) showed negative correlations with Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008) subscales indicating well-being, community and harm avoidance (r=-.12 to -.59) and positive correlations with subscales indicating aggression and problem behaviors (r=.12 to .63). All items are coded in the direction of high scores indicating higher levels of externalizing. We used the authorized translation of the ESI-160 (Soe-Agnie et al., 2016). The translation was conducted following a standard back-translation procedure in combination with a bilingual testing procedure (see Jones, Lee, Philips, Zhang, & Jaceldo, 2001). The translation procedure is described in more detail in Soe-Agnie and colleagues (2016).

The *Aggression Questionnaire* (AQ; Buss & Perry, 1992; Meesters et al., 1996) measures trait aggression with 29 items rated on a 5-point scale ranging from "entirely disagree" to "entirely agree". The items are distributed over four subscales: Physical Aggression (9 items), Verbal aggression (5 items, Anger (7 items) and Hostility (8 items). In the current study only total AQ scores were used, with alpha's of .89, .92 and .90 in the undergraduate, community and patient sample, respectively.

The *NEO-Five Factor Inventory* (NEO-FFI; Costa & McCrae, 1992; Hoekstra et al., 1996) is a short-form version of the NEO-Personality Inventory. It measures the domains of the five-factor model of adult personality: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. It comprises 60 items which are rated on a 5-point scale ranging from "strongly disagree" to "strongly agree". Cronbach's alpha's in the current sample ranged between α= .74 to .88 in the undergraduate sample; α= .79 to .89 in the community sample and α=.64 to .82 in the patient sample.

Procedure

This study was approved by the medical ethical committee of the medical center in Twente, The Netherlands (ID number METC/10078.soe). The data collection was conducted between January 2011 and November 2012.

Recruitment

The exclusion criteria for all subgroups were: Psychotic disorders, severe brain damage, and problems with reading the Dutch language (e.g., due to illiteracy). Additional exclusion criteria in the undergraduate and community sample were: 1) current psychological or psychiatric treatment and 2) any current psychiatric disorder. The three subgroups were each enrolled in different manners. Inpatients were recruited in a low security forensic hospital, six drugs rehab clinics and a forensic drug rehab clinic. All patients who did not meet the exclusion criteria received a written informed consent letter. Undergraduate Psychology students were recruited from the Radboud University Nijmegen. The undergraduates were recruited through an electronic recruitment system of the Radboud University Nijmegen; all included students were required to participate in research projects to receive enough study credits needed to graduate. Community participants were recruited through staff members from the cooperating clinical facilities and neighbors of these staff members. The staff members were recruited by a research assistant who posted web-based invitations on the internal websites of the participating institutions. In these electronic flyers, employees were invited to participate and were asked to invite their neighbors to participate. Hospital employees who recruited neighbors were encouraged to approach persons with no more than a secondary vocational education, in order to recruit a community sample that is equivalent to the patient sample with regard to the level of education. Participation was voluntary for all subgroups. Patients and community subjects received an incentive of €7. The undergraduates received one study credit point.

Test administration

A self-report demographic variables questionnaire and the Dutch ESI-160, AQ and NEO-FFI were administered via a web-based form to the undergraduates and community subjects, while the inpatients completed either the same web-based version or an equivalent offline version on a computer when internet was not available at the test site. All questions were set to 'required', to prevent missing data. The ESI was administered twice, with a two weeks interval in a subset of each subgroup to assess the test-retest reliability.

Data analysis

Reliability analysis

We conducted the statistical analyses on two levels: The item level and the subscale level. On the item level, we examined the average inter-item correlations and Cronbach's alpha (α). On the subscale level we examined

the test-retest reliability by means of an intraclass correlation coefficient (ICC). ICC estimates along with their corresponding 95% confidence intervals were based on a single measurement, consistency, two-way mixed-effects model. All reliability analyses were conducted in SPSS 25.

Factor analysis

Confirmatory factor analysis (CFA) using the total scores of the 23 subscales as units of measurement was conducted separately for each sample (patients, community subjects, undergraduates) as well as on the combined data-set. The CFA's were conducted using maximum likelihood estimation. In this analysis, correlated residuals were not permitted. When the data did not meet the assumption of normality, the CFA was repeated using robust maximum likelihood estimation, as is recommended by Rosseel (2017). Three factor models were specified following Patrick et al. (2013): A unidimensional model; a three-factor model of Disinhibition, Callous Aggression and Substance Abuse, and a bifactor model containing two specific factors of Callous Aggression and Substance Abuse, alongside a general factor. The CFA analysis was conducted in R (R Core Team, 2018), using the Lavaan package (version 3.5.0; Rosseel, 2012). We chose similar fit indices as those reported in Patrick et al. (2013). We used the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR) as absolute fit indices. RMSEA values under .06 indicated good fit, and values under .08 indicated acceptable fit. SRMR values under .05 and .08 indicated good and adequate fit, respectively. We used the Tucker-Lewis index (TLI) and the comparative fit index (CFI) as relative fit indices. TLI and CFI equal or above .95 indicated good fit. The cutoff scores of the fit indices were based on the recommendations in Hooper et al. (2008). When poor fit was found for CFA models, the confirmatory analyses were followed by exploratory factor analysis (EFA) for each subsample and the full sample. Following Lenferink et al. (2016), EFA was conducted applying a Parallel Analysis (PA) based on Minimum Rank Factor Analysis (MRFA) using the software package FACTOR (Lorenzo-Seva & Ferrando, 2006). In this method, for each factor the empirical value of the proportion of explained common variance (ECV) is compared to corresponding factors ECV derived from random data assuming independence among items. To determine the optimal number of factors, for each successive factor the observed ECV was compared to the mean or the 95th percentile of the sampling distribution associated with the respective factor. We used the standard configuration for PA-MRFA: 500 random correlation matrices were generated based on "random permutation of sample values". We used a Pearson correlation matrix and oblique rotation. When a one-dimension solution was revealed, we computed the ECV and the unidimensional congruence (UniCo) as indices of item congruence. An ECV value larger than .85 and an UniCo value above .95 are thought to indicate support for a unidimensional solution. When more than one dimension was found in the exploratory factor analysis, we conducted an exploratory bifactor analysis, using the Schmid-Leiman orthogonalization (Schmid & Leiman, 1957). An explanation of the Schmid-Leiman procedure is provided in Mansolf and Reise (2016). For this exploratory analysis the Root Mean Square of Residuals (RMSR) and Kelley's criterion (Kelley, 1935) were used as fit indices. A RMSR above Kelley's criterion (Kelley, 1935) indicated poor fit.

Criterion validity

Pearson's correlations between the original factors of the ESI (Patrick et al. (2013) and the AQ and the NEO-FFI domains were computed per subsample in SPSS 25. We considered r<.20 inadequate, .20≤ r <.35 adequate, .35≤ r < .50 good and r> .50 excellent correlations, following Evers et al. (2008).

Results

Reliability Dutch ESI

Descriptive and reliability coefficient are presented in Table 1. The differences in mean subscale scores between gender are not presented in the Table, but male participants scored significantly higher than female participants on four out of 23 subscales in the patient group (Mdn_{male}=5.00 to 12.00, Mdn_{female}=2.00 to 8.00, U = 966.00 to 1037.5, z=-2.083 to -2.460); on 21 subscales in the community group (Mdn_{male}=0.00-18.0; Mdn_{female}=0.00-13.00; U = 973.00 to 2997.00, z= −.037 to -6.907, p<.001 to .017); and on 16 subscales in the undergraduate group (Mdn_{male}=0.00 to 20.00, Mdn_{female}=0.00 to 16.00, U = 1555.50 to 2855.00, z=-.049 to -4.373, p<.001 to .043). The Cronbach's alpha estimates of the subscales in the total sample (N = 554) were between α= .68 and α= .96. They ranged from α= .68 to α= .94 for patients, α= .73 to α= .94 for community participants and from α= .49 to α= .95 for undergraduates. Regarding the group of undergraduates, two items showed zero-variance and were deleted from the Cronbach's alpha analysis. The average inter-item correlations per subscale ranged from r= .25 to r=.69 in the patient group, r= .30 to r= .70 in the community group and r= .08 to r= .83 in the undergraduate group. The test-retest reliability was analyzed in 86 inpatients, 140 community participants and 31 undergraduates. The intraclass correlation coefficient estimates (single measure) ranged from r=.71 to r=.96 in the patient group, r=.68 to r=.91 in the community group and r=.43 to .97 in the undergraduates.

Confirmatory factor analysis

Confirmatory factor analyses and fit indices were computed for all subsamples separately as well as for the full sample (see Table 2). The absolute fit indices (RMSEA and SRMR) indicated poor to moderate fit for all the models that were tested. The fit indices of the 1-factor model were poor for all samples (RMSEA 90% confidence interval (CI)$_{patients}$=.181-.200, CI$_{community}$=.143-.161, CI$_{undergraduates}$=.121-.137, CI$_{full\ sample}$=.152-.162). Both absolute fit indices of the 3-factor model were poor in the

Table 1. Mean, standard error, range, Cronbach's α and test-retest reliability (ICC) of ESI-NL subscale scores for the three subsamples.

		Patients						Community						Undergraduates					
Subscales	n_{items}	*M*	SE	range	α	M_r	ICC	*M*	SE	range	α	M_r	ICC	*M*	SE	range	α	M_r	ICC
Alcohol problems	9	16.47	.79	0-27	.93	.61	.89	4.42	.32	0-19	.81	.32	.84	2.86	.19	0-14	.70	.20	.89
Alcohol use	9	18.34	.56	1-27	.81	.31	.85	15.35	.50	0-26	.85	.38	.92	14.70	.42	0-25	.86	.40	.94
Marijuana use	7	12.25	.68	0-21	.94	.69	.96	5.74	.55	0-21	.94	.68	.96	3.09	.37	0-21	.92	.63	.97
Marijuana problems	7	7.77	.63	0-21	.93	.65	.92	1.52	.27	0-21	.90	.57	.91	.36	.08	0-9	.71	.26	.84
Drug problems	11	18.10	.95	0-33	.93	.54	.91	2.17	.37	0-26	.90	.44	.91	.31	.07	0-7	.50	.08	.86
Drug use	6	11.56	.48	0-18	.84	.47	.89	5.33	.40	0-18	.83	.45	.86	3.12	.23	0-17	.72	.30	.91
Alienation	3	4.87	.21	0-9	.68	.42	.80	2.44	.17	0-9	.77	.52	.76	1.70	.12	0-8	.72	.46	.83
Blame externalization	4	5.62	.30	0-12	.87	.63	.77	1.74	.21	0-12	.90	.70	.77	1.11	.15	0-8	.95	.83	.57
Boredom proness	4	6.73	.33	0-12	.88	.64	.78	4.25	.27	0-12	.88	.69	.70	3.31	.18	0-12	.84	.56	.56
Dependability	7	6.91	.39	0-19	.81	.38	.80	3.66	.28	0-19	.81	.38	.68	2.79	.17	0-12	.73	.28	.67
Empathy (-)	11	9.94	.65	0-33	.90	.43	.76	7.25	.41	0-29	.82	.29	.83	3.85	.25	0-27	.75	.22	.74
Excitement seeking	6	8.05	.47	0-18	.90	.60	.85	4.55	.35	0-17	.87	.53	.87	3.01	.21	0-15	.83	.45	.92
Fraud	6	5.36	.41	0-18	.79	.39	.86	1.36	.18	0-15	.77	.36	.83	.62	.08	0-8	.49	.14	.40
Honesty	5	4.84	.31	0-15	.80	.45	.74	4.17	.25	0-15	.73	.36	.76	3.41	.18	0-12	.65	.27	.47
Impatient urgency	5	8.99	.35	0-15	.84	.51	.71	6.12	.30	0-15	.82	.47	.82	5.19	.22	0-14	.79	.43	.82
Irresponsibility	10	14.83	.61	0-30	.76	.25	.86	4.67	.42	0-27	.83	.33	.90	1.55	.16	0-17	.72	.20	.84
Physical aggression	8	10.60	.61	0-24	.87	.45	.87	3.47	.30	0-22	.84	.40	.86	2.00	.15	0-11	.66[a]	.22	.75
Destructive aggression	7	5.62	.50	0-21	.88	.52	.79	1.57	.22	0-19	.78	.34	.68	.40	.09	0-11	.67	.23	.74
Relation aggression	8	8.60	.51	0-24	.85	.41	.86	4.49	.34	0-19	.83	.37	.79	2.95	.19	0-15	.72	.24	.64
Planful control (-)	11	7.16	.38	0-18	.88	.55	.79	3.52	.26	0-18	.80	.40	.71	2.62	.15	0-11	.70	.40	.46
Problematic impulsivity	7	11.97	.50	0-21	.85	.44	.86	2.64	.27	0-16	.82	.39	.75	1.38	.13	0-12	.70	.25	.75
Rebelliousness	6	8.89	.44	0-18	.88	.56	.78	3.19	.26	0-17	.84	.46	.85	1.62	.15	0-12	.79	.39	.77
Theft	8	10.74	.68	0-24	.89	.50	.94	2.21	.27	0-18	.78	.30	.76	.68	.10	0-12	.50[a]	.12	.43

Note. Reversed-keyed subscales are indicated by a minus sign in parentheses following the subscale labels. $M_{r=}$ average inter-item correlation.
[a]One item deleted from subscale due to zero variance.

Table 2. Fit statistics for confirmatory factor models for subsamples and the full sample.

	AIC	SABIC	CFI	TLI	RMSEA	SRMR
1-factor (K = 44)						
Patients (n = 149)	20132.773	20264.947	.45	.40	.191	.156
Community (n = 178)	20227.659	20228.315	.68	.64	.152	.092
Undergraduates (n = 227)	21682.256	21693.505	.60	.56	.129	.108
Full sample (N = 554)	67394.022	67444.302	.72	.69	.157	.086
3-factor (K = 43)						
Patients (n = 149)	18209.846	18202.932	.62	.57	.169	.147
Community (n = 178)	18552.765	18553.406	.75	.71	.141	.085
Undergraduates (n = 227)	19809.071	19820.065	.76	.73	.105	.086
Full sample (N = 554)	61656.924	61706.061	.79	.76	.145	.080
Bifactor 2 specific factors (K = 69)						
Patient (n = 149)	20425.792	20414.698	.71	.64	.143	.117
Community (n = 178)	20807.732	21027.275	.78	.73	.129	.068
Undergraduates (n = 227)	22365.097	22382.739	.80	.75	.094	.070
Full sample (N = 554)	20155.074	20233.922	.83	.79	.126	.060

Note. AIC = Akaike Information Criterion; CFI = Comparable Fit Index; K = number of free parameters; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation; SABIC = Sample-size adjusted Bayesian; SRMR = Standardized Root Mean Square Residual.
Robust Maximum Likelihood estimates for the 1-, 3- and bifactor model: Full sample: RMSEA CI 90%= .94 to .102; SMSR-.060; CFI=.84; TLI=.80; Patients: RMSEA CI 90%=.123 to .143; SMSR= .117; CFI=.71; TLI=.65; Community: RMSEA CI 90%=.095-.111; SMSR=.068; CFI=.80; TLI=.75; Student: RMSEA CI 90%=.070 to .093; SMSR=.070; CFI=.83; TLI=.79.

separate subsamples (RMSEA $CI_{patients}$=.158-.180, $CI_{community}$= .131-.151, $CI_{undergraduates}$= .096-.114). In the full sample the SRMR showed acceptable fit, while the RSMEA indicated poor fit (RMSEA $CI_{full\ sample}$=.139-.150). Regarding the bifactor model (based on Patrick et al., 2013), the SRMR showed acceptable fit in the two non-clinical samples, but not in the patient sample. The RMSEA of the bifactor model was poor in all three samples ($CI_{patients}$=.133-.153, $CI_{community}$=.120-.138, $CI_{undergraduates}$=.086-.103 and $CI_{full\ sample}$ =.121-.131) [1]. Since the data did not meet the assumption of normality in all subscales in the student and community samples, the analyses were repeated using Robust Maximum Likelihood. These estimates offered no improvement of the fit indices (see Table 2).

Exploratory factor analysis

The Minimum Rank Factor Analysis revealed a different number of factors for each subsample (see Table 3). The analysis indicated one-dimensional patterns in the community sample and the full sample, a two-dimensional pattern in the undergraduate sample and a three-dimensional pattern in the patient sample. As a next step, an exploratory bifactor analysis (Schmid-Leiman, 1957) was conducted for the undergraduate and patient sample. As opposed to the confirmatory bifactor analysis, cross-loadings were allowed. For both subsamples the RMSR was equal or smaller than

[1]In order to replicate Patrick et al. (2013), modification indices were analysed and the analyses were repeated with correlated residuals of the Alcohol Use - Alcohol Problem and the Blame Externalization - Alienation subscales. Allowing correlated residuals did not improve the fit estimates, and the modification indices did not indicate any other correlated residuals between subscales.

Table 3. Results of the Minimum rank factor analysis of the Dutch Externalizing Spectrum inventory per subsample.

	ECV real data				Mean ECV random data				95th percentile ECV random data				Eigenvalue			
Factor	C	P	U	F	C	P	U	F	C	P	U	F	C	P	U	F
1	51.1	35.9	39.0	57.7	11.0	12.8	11.0	11.7	12.2	14.2	12.2	13.4	10.17	7.53	6.98	11.41
2	8.7	18.2	13.9	9.0	10.1	11.6	10.1	11.0	10.9	12.7	11.1	12.4	1.86	3.34	2.58	1.69
3	7.4	12.2	7.1	6.5	9.3	10.6	9.4	10.4	10.0	11.5	10.2	11.6	1.65	2.43	1.46	1.18
4	6.3	7.9	6.2	5.2	8.7	9.8	8.8	9.6	9.3	10.5	9.5	10.7	1.37	1.51	1.27	0.95
5	5.0	5.9	5.5	4.1	8.0	8.9	8.2	8.8	8.6	9.6	8.8	9.9	1.13	1.25	1.17	0.77
6	3.6	5.0	4.7	3.5	7.4	8.1	7.5	7.9	8.0	8.7	8.1	9.1	0.79	1.05	1.09	0.65
7	3.0	3.1	3.6	2.7	6.9	7.3	6.9	7.0	7.3	7.9	7.4	8.3	0.77	0.78	0.96	0.39
UniCo		–	–	.93												
UniCo CI 95%		–	–	.91-.96												
ECV		–	–	.87												
ECV CI 95%		–	–	.85-.89												
Schmid-Leiman																
RMSR	–	.08	.06	–												
RMSR CI 95%	–	.07-.08	.06-.06	–												
Kelly's criterion	–	.08	.07	–												

Note. C = Community; P = Patients; U = Undergraduates; F = Full sample; UniCo = unidimensional congruence; ECV = explained common variance.

Table 4. Factor loadings per subsample.

	Full	Community	Patients				Undergraduates		
Subscales	F1	F1	F1	F2	F3	G	F1	F2	G
1.Alcohol use	.34	.43	.05	.38	-.59	.18	-.05	.30	.46
2.Alcohol problems	.67	.77	-.20	. 39	-.37	.15	-.14	.33	.63
3.Marijuana use	.69	.67	.76	-.06	-.15	.41	-.14	.48	.69
4.Marijuana problems	.77	.79	.57	-.06	.17	.47	-.14	.39	.55
5.Drug use	.75	.68	.75	-.06	-.17	.38	-.11	.44	.65
6.Drug problems	.85	.81	.73	.01	-.10	.51	.04	.28	.48
7.Destructive aggression	.76	.79	.10	.09	.55	.55	.33	.03	.26
8.Physical aggression	.78	.57	.33	.14	.23	.60	.47	-.02	.27
9.Relational aggression	.78	.80	.10	.17	. 49	.62	.64	-.02	.36
10.Honesty(-)	.38	.52	-.06	-.14	.64	.18	.35	.03	.26
11.Empathy(-)	.49	.33	-.20	-.06	.77	.26	.32	.04	.26
12.Excitement Seeking	.72	.72	.19	.20	.26	.59	.42	.11	.45
13.Irresponsibillity	.89	.82	.53	.08	.08	.56	.58	.07	.48
14.Planful control(-)	.58	.43	-.31	.27	.28	.33	.54	-.09	.19
15.Problematic impulsivity	.89	.81	.27	.40	-.15	.67	.72	-.01	.43
16.Rebellious	.86	.78	.14	.35	.17	.70	.52	.09	.47
17.Alientation	.56	.43	-.14	.37	-.23	.26	.43	-.11	.09
18.Blame externalization	.65	.45	-.19	.35	.08	.40	.55	-.07	.23
19.Boredom proneness	.62	.68	.02	.39	-.09	.47	.46	.01	.31
20.Dependability(-)	.61	.57	-.11	-.04	.63	.27	.51	.06	.42
21.Fraud	.78	.76	.27	-.02	.47	.49	.53	-.02	.30
22.Impatient urgency	.64	.60	.08	.47	-.15	.62	.57	-.06	.26
23. Theft	.81	.61	.55	-.03	.24	.53	.29	.07	.29
Correlations factors									
F1xF2	–	–	.44				.45		
F2xF3	–	–	.33				–		
F1xF3	–	–	.49				–		

Note. Reversed-keyed subscales are indicated by a minus sign in parentheses following the subscale labels.

Kelley's criterion (Kelley, 1935), indicating adequate fit. Based on the factor loadings (see Table 4) the two-dimensional factor scales in the undergraduate sample were labeled as Substance Abuse and Impulsive Aggression. The first factor contained subscales of the Substance Abuse scales from Patrick et al. (2013) directly reflecting forms of alcohol or drug use, while subscales reflecting forms of irresponsibility and impulsivity were omitted. The second dimension was labeled as such, since both features of impulsivity and aggressive behavior showed high loadings. In the patient sample, the first dimension was labeled as Disinhibition/Alcohol Abuse, containing most subscales reflecting impulsivity, irresponsibility and the two alcohol subscales. The second dimension was labeled as Callous Aggression, containing most subscales of the Callous aggression scale from Patrick et al. (2013), except physical aggression. Lastly, the third dimension was labeled as Drug Abuse, containing all scales referring to drug abuse and different forms of antisocial/criminal behaviors, including physical aggression.

Criterion validity

The ESI bifactor structure as reported by Patrick et al. (2013) was used for the criterion validity analysis in all samples. In the patient sample, the general factor, showed adequate positive correlations to trait aggression and Neuroticism and adequate to good negative correlations to Agreeableness and Conscientiousness. The Callous Aggression factor showed a good positive correlation to trait aggression and good negative correlations to Agreeableness and Conscientiousness,

Table 5. Correlations of factor scores (Patrick et al., 2013) and self-report personality measures per subsample.

	Patients				Community				Undergraduates			
	n	AGG	SUB	g	n	AGG	SUB	g	n	AGG	SUB	g
NEO	145				178				35			
Neuroticism		.07	.08	.24		.33	.28	.49		−.09	−.00	−.05
Extraversion		−.13	.07	−.10		−.15	.01	−.17		.05	.11	.16
Openess		−.07	−.01	−.12		.16	.30	.24		.11	.05	.19
Agreeableness		−.38	−.01	−.30		−.59	−.28	−.57		−.37	.36	−.27
Conscientiousness		−.41	−.19	−.42		−.56	−.37	−.61		−.30	.16	−.63
AQ	142	.49	.17	.51	178	.65	.40	.66	227	.56	−.31	.57

Note. NEO = NEO-FFI; AQ = Aggression Questionnaire; AGG = Callous Aggression factor, according to Patrick et al. (2013); SUB = Substance Abuse factor, according to Patrick et al. (2013); g = general factor, according to Patrick et al. (2013).

while the correlation to Neuroticism was low. The Substance Abuse factor showed inadequate correlations to all external measures. In comparison to the patient sample, most correlations were larger in the community sample and the Substance Abuse factor showed a higher, good correlation to Conscientiousness. In the undergraduate sample the correlations between the general factor and Callous Aggression on one hand and Agreeableness, Conscientiousness and trait aggression on the other hand were similar to the correlations in the patient sample. The correlations between the Substance Abuse factor and Agreeableness and trait aggression were higher and adequate, and in the opposite direction of the correlations in the patient sample. The correlations are presented in Table 5.

Discussion

The objective of the current study was to extend the literature of the Externalizing Spectrum Inventory (ESI) by investigating its psychometric properties across Dutch clinical and non-clinical samples. To our knowledge, the current paper is the first to address the relations between the ESI and the Five Factor model. The reliability of the 23 subscales was satisfactory in the patient and community sample. In the undergraduate sample only the test-retest reliability was satisfactory, while the Cronbach's alpha estimates were not. The proposed bifactor model, reported by Patrick et al. (2013), was not replicated in our samples and exploratory analyses yielded models with a different number of specific factors across samples. The hypothesized relations between the factors proposed by Patrick et al. (2013) and trait aggression were confirmed in all samples. The expected relations between the factors (Patrick et al., 2013) and the Five Factor Model personality domains were fully confirmed in the community sample and partly in the patients and undergraduates. In these latter samples only the hypothesized correlations between the Callous Aggression factor and Agreeableness and Conscientiousness were confirmed.

The lack of support of the factor models as presented by Patrick et al. (2013), raises the question as to whether the original model is generalizable to other populations than the ones sampled by Patrick et al. (2013). Before drawing any conclusions on this, we would like to reflect on the possible role of the limitations of the current study. Firstly, even though the quality of the translated ESI was researched through bilingual testing (Soe-Agnie et al., 2016), a negative impact of weaknesses in the cultural adaptation cannot be ruled out. Cultural differences may cause respondents to interpret items differently across nations. For example, cannabis consumption was not illegal in The Netherlands during the test administration, while it was illegal in the US. Therefore, items such as "My marijuana use led to legal problems" are likely to have a different meaning for a Dutch respondent than an American respondent. Secondly, the current undergraduate sample consisted mostly of women, while the patient and community sample consisted mostly of men. Prior research has shown that women show higher prevalence rates on internalizing disorders, while men show higher prevalence rates on externalizing disorders (see Eaton et al., 2012). However, Carragher et al. (2014) and Kramer et al. (2008) found that the structure of the externalizing liability was invariant across gender. Furthermore, our patient and community samples were comparable when it came to gender ratios, but still differences in factor structure emerged. Therefore, we think it is unlikely that the differences in factor solutions across subgroups and in comparison to Patrick et al. (2013) are a result of the differences in gender ratios. Lastly, in the current non-patient samples, current diagnosed psychiatric disorders were used as exclusion criteria to ensure the subgroups were homogeneous for the subgroup comparisons, namely that the student and community sample did not include patients. This could be considered to be a limitation, since the ESI is designed to capture a model of psychopathology, which was now tested in two non-psychiatric samples. However, the absence of current psychiatric diagnoses does not exclude the presence of problem behaviors below the psychiatric threshold. The average ESI scores in the non-patient groups were above zero and in some subscales only slightly below the average scores of the inpatients. This implies that some externalizing tendencies were present in the non-patient groups. Furthermore, the ESI items were originally selected to target a maximum variety from severe to low levels of externalizing (Krueger et al., 2007) and the students and prisoners used in the study by Krueger et al. (2007) and Patrick et al. (2013) were selected to cover this variety. In the current sample, non-patients and (forensic) patients were selected with the same aim. In sum, cultural differences between the Dutch and the US sample may have had a negative impact on the current results, but we think it unlikely that the differences in gender ratio or the absence of current psychiatric disorders in the non-patient samples would have had a substantial impact.

To our knowledge, the original ESI-160 bifactor model is yet to be replicated in other (US) samples than the sample of Patrick et al. (2013), which originates from the study of Krueger et al. (2007). The only other research of a translated version of the ESI was conducted by Carvalho et al. (2010). However, these authors researched the full 415 item ESI and not the brief ESI-160. In addition, Carvalho et al. (2010) did not research the bifactor model with two specific factors through confirmatory analysis. Although, the original bifactor model (Patrick et al., 2013) has not been replicated in other samples than the ones from Krueger et al. (2007), the original general and Callous Aggression factors did show good predictive value in the current study and thus criterion validity of these factors was supported. Criterion validity support was weaker for the Substance Abuse factor, since the expected correlations between this factor and the Five Factor Model were only found in the community sample.

The factor models that were found in the exploratory analysis, indicate several weaknesses concerning the robustness of the original ESI-160 structure. For instance, the results showed specific weaknesses of the Physical Aggression and the alcohol subscales in the inpatient sample. In the exploratory model, Physical Aggression converged to the drug scales, while the alcohol subscales diverged from these drug scales. Both these results do not fit within the Externalizing Spectrum framework. Moreover, the results raise the questions how models so different can surface in different populations. The differences in factor models across subsamples (the 3- 2- and 1-factor structures for the inpatients, undergraduates and community participants respectively), may be partly due to a difference in severity of problem behaviors for the different samples. The higher number of specific factors in externalizing inpatients in comparison to the healthy non-clinical subjects, corresponds with the results reported by Paap et al. (2012), who demonstrated that the severity of psychopathology influenced the number of factors found for the SCL-90-R (Derogatis, 1994): A higher degree of multidimensionality was found in samples with a higher severity of psychopathology. This may explain the higher degree of multidimensionality found in the undergraduates in comparison to the community participants as well, since the literature supports relatively higher levels of externalizing behaviors in student populations as compared to community samples (see e.g. Armstrong et al., 2006; Burns et al., 2015). This could suggest that the factor solution found by Patrick et al. (2013) was not generalizable to our samples, because of differences in levels of psychopathology. Unfortunately, the prevalence of psychiatric disorders were not reported by Patrick et al. (2013), and it is unclear whether the samples used by Patrick et al. (2013) were representative of the US student and prisoner populations, which precludes a direct comparison of their sample characteristics with ours. Another possible explanation for finding different model solutions across samples may be that the factor structure was distorted by response bias. Self-report measures are known to be sensitive to response bias (see Achenbach, 2006). Navarro-González et al. (2016)) showed that response bias can distort the factor structure of personality questionnaires.

Implications and future directions

Researchers have tried to identify underlying personality traits and psychopathological patterns of substance use disorders and antisocial behavior for decades, with mixed success (e.g. Sher & Trull, 1994, Terracciano et al., 2008). A recent promising endeavor is the HiTOP model (Kotov et al. 2017), which integrates these underlying traits and patterns and has been put forward as a possible replacement of traditional nosologies, such as the DSM classification system. The ESI contributes to this particular research field by reflecting the complexity of comorbid externalizing disorders, while offering a framework to further research the underlying traits of externalizing disorders. However, although the DSM system may be flawed, it is a widespread used system across different populations, cultures, and nations. In order to replace such a system, the HiTOP model must prove to be a robust model across populations and research should be targeted to evidence this robustness. The current paper was the first to compare factor solutions of the ESI-160 across populations. Since little research has been done of the factor structure of the ESI, and the robustness of the ESI structure was not supported in our Dutch samples, future research of the robustness of the model in different populations is recommended. Particularly, cross-national studies are recommended, in order to establish cross-cultural stability of the ESI, identify and modify 'weak' items and to encourage cross-cultural research of the Externalizing Spectrum Model. This could improve the internal structure of the Dutch ESI, as well. Specifically, the problematic Alcohol subscales and the Physical Aggression subscale should be investigated further. Regarding patient populations, further research of the influence of response bias and symptom severity on the factor structure is recommended.

The apparent contradiction between the lack of support from the factor analyses and the present support from the criterion validity analyses, leads to the question which purpose the ESI can presently serve: Measurement or prediction? The current results suggest the latter, since accurate measurement of attributes requires high inter-item relations, while instruments aimed at prediction, require lower inter-item correlations and higher correlations with an external criterion (Smits et al., 2018).

In conclusion, the current Dutch ESI-160 is recommended in research of the relations between externalizing personality and aggression, in particular. However, the specific factors should be interpreted with care. Further research and adaptation of the Dutch ESI-160 is vital, to improve the precision of this inventory, and to enhance the understanding of Externalizing across nations and cultures.

Funding

This article was supported by Grant 5620076/09/DJI from the Dutch Ministry of Justice and Security and by Tactus addiction care.

ORCID

Sabrina E. Soe-Agnie http://orcid.org/0000-0001-5063-4691
Muirne C. S. Paap http://orcid.org/0000-0002-1173-7070
Cornelis A. J. De Jong http://orcid.org/0000-0003-1824-7303

References

Abram, S. V., Wisner, K. M., Grazioplene, R. G., Krueger, R. F., MacDonald, A. W., & DeYoung, C. G. (2015). Functional coherence of insula networks is associated with externalizing behaviour. *Journal of Abnormal Psychology*, *124*(4), 1079–1091. doi:10.1037/abn0000078

Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science*, *15*(2), 94–98.

American Psychiatric Association (2015)., Brief overview of the DSM-5 criteria. [Beknopt overzicht van de criteria (DSM-5).] Dutch translation of the Desk reference to the diagnostic criteria from DSM-5. Boom, Amsterdam. pp. 340, 455–456.

Armstrong, E. A., Hamilton, L., & Sweeney, B. (2006). Sexual assault on campus: A multilevel, integrative approach to party rape. *Social Problems*, *53*(4), 483–499.

Burns, S., Crawford, G., Hallett, J., Jancey, J., Portsmouth, L., Hunt, K., & Longo, J. (2015). Consequences of low risk and hazardous alcohol consumption among university students in Australia and implications for health promotion interventions. *Open Journal of Preventive Medicine*, *05*(01), 1–13.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*(3), 452–459.

Carragher, N., Krueger, R. F., Eaton, N. R., Markon, K. E., Keyes, K. M., Blanco, C., Saha, T. D., & Hasin, D. S. (2014). ADHD and the externalizing spectrum Direct comparison of categorical, continuous, and hybrid models of liability in a nationally representative sample. *Social Psychiatry and Psychiatric Epidemiology*, *49*(8), 1307–1317. doi:10.1007/s00127-013-0770-3

Carvalho de, H. W., Patrick, C. J., Krueger, R. F., Markon, K. E., & Pinheiro, A. M. V. (2010). Construct validity of the Brazilian version of the Externalizing Spectrum Inventory: Evidence from a university student sample. *Revista de Psiquiatria Clínica*, *37*(5), 206–201.

Chan, Y., Dennis, M. L., & Funk, R. R. (2008). Prevalence and comorbidity of major internalizing and externalizing problems among adolescents and adults presenting to substance abuse treatment. *Journal of Substance Abuse Treatment*, *34*(1), 14–24. doi:10.1016/j.jsat.2006.12.031

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO PERSONALITY INVENTORY (NEO-PI-R) and the Five Factor Inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources Inc.

Derogatis, L. R. (1994). *SCL-90-R: Administration, scoring and procedures manual*. National Computer Systems.

Durbeej, N., Palmstierna, T., Rosendahl, I., Berman, A. H., Kristiansson, M., & Gumpert, C. H. (2015). Mental health services and public safety: Substance abuse outpatient visits were associated with reduced crime rates in a Swedish cohort. *Plos One*, *10*(9), e0137780–17. doi:10.1371/journal.pone.0137780

Dykstra, R. E., Schumacher, J. A., Mota, N., & Coffey, S. F. (2015). Examining the role of antisocial personality disorder in inmate partner violence among substance use disorder treatment seekers with clinically significant trauma. *Violence against Women*, *21*(8), 958–974. doi:10.1177/1077801215589377

Eaton, N. R., Keyes, K. M., Krueger, R. F., Balsis, S., Skodol, A. E., Markon, K. E., Grant, B. F., & Hasin, D. S. (2012). An invariant dimensional liability model of gender differences in mental disorder prevalence: Evidence from a national sample. *Journal of Abnormal Psychology*, *121*(1), 282–288. doi:10.1037/a0024780

Evers, A., Hagemeister, C., Høstmælingen, A., Lindley, P., Muñiz, J., & Sjöberg, A. (2008). EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. Version 4.2.6. https://core.ac.uk/download/pdf/71863941.pdf

Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing psychopathology and the error-related negativity. *Psychological Science*, *18*(4), 326–333. doi:10.1111/j.1467-9280.2007.01899.x

Hoekstra, H. A., Ormel, J., & de Fruyt, F. (1996). Handleiding NEO persoonlijkheids-vragenlijsten NEO-PI-R en NEO-FFI. Lisse, Swets Test Services.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, *6*(1), 53– 60. Retrieved from www.ejbrm.com.

Jones, P.S., Lee, J. W., Phillips, L. R., Zhang, X. E., & Jaceldo, K. B. (2001). An adaptation of Brislin's translation model for cross-cultural research. *Nursing Research*, *50*, 300–304. doi:10.1097/00006199-200109000-00008

Jones, S. E., Miller, J. D., & Lynam, D. R. (2011). Personality, antisocial behaviour, and aggression: A meta-analytic review. *Journal of Criminal Justice*, *39*(4), 329–337. doi:10.1016/j.jcrimjus.2011.03.004

Kelley, T. L. (1935). *Essential Traits of Mental Life, Harvard Studies in Education* (Vol. 26, pp. 146). Harvard University Press..

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., … 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

Kramer, M. D., Krueger, R. F., & Hicks, B. M. (2008). The role of internalizing and externalizing liability factors in accounting for gender differences in the prevalence of common psychopathological syndromes. *Psychological Medicine*, *38*(1), 51–61. doi:10.1017/S0033291707001572

Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., Simms, L. S., Widiger, T. A., Achenbach, T. M., Bach, B., Bagby, R. M., Bornovalova, M. A., Carpenter, W. T., Chmielewski, M., Cicero, D. C., Clark, L. A., Conway, C., DeClercq, B., DeYoung, C. G., … Johannes Zimmermann, J. (2018). Progress in achieving quantitative classification of psychopathology. *World Psychiatry*, *17*(3), 282–293. doi:10.1002/wps.20566

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, *2*(1), 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*(4), 645–666. doi:10.1037/0021-843X.116.4.645

Krueger, R. F., & South, S. C. (2009). Externalizing disorders: Cluster 5 of the proposed meta-structure for DSM-V and ICD-11. *Psychological Medicine*, *39*(12), 2061–2070.

Lenferink, A., Effing, T., Harvey, P., Battersby, M., Frith, P., Van Beurden, W., Van der Palen, J., & Paap, M. C. S. (2016). Construct validity of the Dutch version of the 12-item partners in health scale: Measuring patient self-management behaviour and knowledge in patients with chronic obstructive pulmonary disease. *Plos One*, *11*(8), e0161595–16. doi:10.1371/journal.pone.0161595

Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: a computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, *38*(1), 88–91. doi:10.3758/BF03192753

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, *51*(5), 698–717. doi:10.1080/00273171.2016.1215898

Meesters, C., Muris, P., Bosma, H., Schouten, E., & Beuving, S. (1996). Psychometric evaluation of the Dutch version of the aggression questionnaire. *Behaviour Research and Therapy*, *34*(10), 839–843. doi:10.1016/0005-7967(96)00065-4

Ministry of Justice and Security [Dienst Justitiële Inrichtingen]. (2018, September). Gevangeniswezen [Infosheet]. Retrieved from https://www.dji.nl/binaries DEF %20%20infographic% 20gevangeniswezen% 20sep_2018_tcm41-352270.pdf

Navarro-González, D., Lorenzo-Seva, U., & Vigil-Colet, A. (2016). How response bias affects the factorial structure of personality self-reports. *Psicothema*, *28*(4), 465–470. doi:10.7334/psicothema2016.113

Nelson, L. D., Patrick, C. J., & Bernat, E. M. (2011). Operationalizing proneness to externalizing psychopathology as a multivariate psychophysiological phenotype. *Psychophysiology*, *48*(1), 64–72. doi:10.1111/j.1469-8986.2010.01047.x

Paap, M. C. S., Meijer, R. R., Cohen-Kettenis, P. T., Richter-Appelt, H., de Cuypere, G., Kreukels, B. P. C., Pedersen, G., Karterud, S., Malt, U. F., & Haraldsen, I. R. (2012). Why the factorial, structure of the SCL-90-R is unstable: Comparing patient groups with different levels of psychological distress using Mokken Scale Analysis. *Psychiatry Research*, *200*(2-3), 819–826. doi:10.1016/j.psychres.2012.03.012

Patrick, C. J., Kramer, M. D., Krueger, R. F., & Markon, K. E. (2013). Optimizing efficiency of psychopathology assessment through quantitative modelling: Development of a brief form of the Externalizing Spectrum Inventory. *Psychological Assessment*, *25*(4), 1332–1348.

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Rosseel, Y. (2017). Structural equation modelling with lavaan. Retrieved from https://users.ugent.be/~yrosseel/lavaan/zurich2017/INTRO/lavaan_intro_zurich2017.pdf

Ruiz, M. A., Pincus, A. L., & Schinka, J. A. (2008). Externalizing pathology and the five-factor model: A meta-analysis of personality traits associated with antisocial personality disorder, substance use disorder, and their co-occurrence. *Journal of Personality Disorders*, *22*(4), 365–388. doi:10.1521/pedi.2008.22.4.365

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*(1), 53–61. doi:10.1007/BF02289209

Sher, K. J., & Trull, T. J. (1994). Personality and disinhibitory psychopathology: Alcoholism and antisocial personality disorder. *Journal of Abnormal Psychology*, *103*(1), 92–102.

Simpson, A. I., Grimbos, T., Chan, C., & Penney, S. R. (2015). Developmental typologies of serious mental illness and violence: Evidence from a forensic psychiatric setting. *Australian & New Zealand Journal of Psychiatry*, *49*(11), 1048–1012. doi:10.1177/0004867415587745

Smits, N., Van der Ark, L. A., & Conijn, J. M. (2018). Measurement versus prediction in the construction of patient-reported outcome questionnaires: Can we have our cake and eat it? *Quality of Life Research*, *27*(7), 1673–1682. doi:10.1007/s11136-017-1720-4

Soe-Agnie, S. E., Paap, M. C. S., VanDerNagel, J. E. L., Nijman, H. J. M., & De Jong, C. A. J. (2018). The generalizability of the structure of substance abuse and antisocial behavioural syndromes: A systematic review. *Psychiatry Research*, *259*, 412–421. doi:10.1016/j.psychres.2017.11.007

Soe-Agnie, S. E., Patrick, C. J., Nijman, H. J. M., & De Jong, C. A. J. (2016). Validation of the full and brief externalizing spectrum inventory in Dutch forensic inpatients. *The Journal of Forensic Psychiatry & Psychology*, *27*(1), 77–91. 2015.1090621 doi:10.1080/14789949

Tellegen, A., & Waller, N. G. (2008). Exploring personality through test construction: Development of the multidimensional personality questionnaire. In Boyle, G.J., Matthews, G., Saklofske, D.H., editors. *Handbook of personality theory and testing: Personality measurement and assessment*. (Vol. II, pp 261–292). Sage.

Terracciano, A., Löckenhoff, C. E., Crum, R. M., Bienvenu, O. J., & Costa, P. T. (2008). Five-factor model personality profiles of drug users. *BioMed Central Psychiatry*, *8*, 22. doi:10.1186/1471-244X-8-22

Venables, N. C., & Patrick, C. J. (2012). Validity of the externalizing spectrum inventory in a criminal offender sample: Relations with disinhibitory psychopathology, personality, and psychopathic features. *Psychological Assessment*, *24*(1), 88–10. https://doi.org/doi:

Wisselink, D. J., Kuijpers, W. G. T., & Mol, A. (2014). *Kerncijfers verslavingszorg 2013. LADIS: landelijk alcohol en drugs informative system*. Stichting Informatie Voorziening Zorg.