



Validation of the full and brief Externalizing Spectrum Inventory in Dutch forensic inpatients

Sabrina E. Soe-Agnie, Christopher J. Patrick, Henk L.I. Nijman & Cor A.J. De Jong

To cite this article: Sabrina E. Soe-Agnie, Christopher J. Patrick, Henk L.I. Nijman & Cor A.J. De Jong (2016) Validation of the full and brief Externalizing Spectrum Inventory in Dutch forensic inpatients, *The Journal of Forensic Psychiatry & Psychology*, 27:1, 77-91, DOI: 10.1080/14789949.2015.1090621

To link to this article: <http://dx.doi.org/10.1080/14789949.2015.1090621>



Published online: 28 Sep 2015.



Submit your article to this journal [↗](#)



Article views: 53




View related articles [↗](#)



View Crossmark data [↗](#)

Validation of the full and brief Externalizing Spectrum Inventory in Dutch forensic inpatients

Sabrina E. Soe-Agnie^{a,c,*}, Christopher J. Patrick^d, Henk L.I. Nijman^{a,b} and Cor A.J. De Jong^{a,c} 

^a*Behavioral Science Institute, Radboud University, Nijmegen, The Netherlands;* ^b*Aventurijn, Altrecht GGZ, Den Dolder, The Netherlands;* ^c*Nijmegen Institute for Scientist-Practitioners in Addiction (NISPA), Behavioral Science Institute, Radboud University, Nijmegen, The Netherlands;* ^d*Department of Psychology, Florida State University, Tallahassee, USA*

(Received 6 August 2014; accepted 28 August 2015)

The Externalizing Spectrum Inventory (ESI) was designed for assessing a broad spectrum of externalizing problems, including impulsive-aggressive behavior and substance abuse. After translation of the ESI into Dutch, a computerized version of the full Dutch ESI (ESI-NL) was administered to a mixed sample consisting of inpatients in forensic and addiction care ($n = 99$) and non-psychiatric community participants ($n = 104$). Internal consistencies, test–retest reliabilities, ‘predictive’ validity of the full and 160-item ESI-NL total scores and subscales were examined, along with the correlations between these two versions. The results indicated high reliability and predictive validity for both versions and a strong similarity between the two in direct comparisons. The 160-item Dutch ESI is recommended for clinical studies on violence proneness and externalizing problem behavior.

Keywords: externalizing; violence proneness; substance-related disorders; co-occurrence; test validation

Introduction

Both substance abuse and antisocial behavioral syndromes (e.g. antisocial personality disorders [APD], conduct disorder, aggression) are not only highly prevalent in forensic populations, they are also strongly associated with criminal recidivism (Abraham et al., 2015; Jaffe, Du, Huang, & Hser, 2011; Mir et al., 2015; Tikkanen, Holi, Lindberg, Tiihonen, & Virkkunen, 2009). This issue, particularly the high prevalence of these comorbid disorders (between 50 and 85% in incarcerated populations) and their strong relationship with violent recidivism after treatment (20% higher recidivism rates than incarcerated subjects without this type of comorbidity), is considered a major problem in The Netherlands (Oliemeulen, Vuijk, Rovers, & VandenEijnden, 2007; Van Horn,

*Corresponding author. Email: s.soeagnie@tactus.nl

Eisenberg, Van Kuik, & Van Kinderen, 2012). The comorbidity of antisocial behavioral disorders and substance use disorders (SUDs), has been shown to reflect systematic covariance (Cerda, Sagdeo, & Galea, 2008; Kessler, 2004; Krueger, 1999; Merikangas & Kalaydjian, 2007) and research has shown these disorders –including APD, violence proneness, and different types of SUDs – to be more accurately viewed as phenotypic manifestations of an underlying so-called externalizing spectrum (Cerda et al., 2008; Krueger & South, 2009; Krueger et al., 2002). This view offers an evidence-based starting point for integrated care of co-occurring externalizing disorders, since the externalizing spectrum conception points to the disinhibitory tendencies that link patients' vulnerability to substance abuse and antisocial behavioral syndromes as major targets for treatment. This premise leads us to the conclusion that these mutual disinhibitory tendencies should be targeted in all clinical populations who exhibit substance abuse and antisocial behaviors, regardless of the type of care they receive (e.g. addiction care or forensic psychiatry). This is particularly relevant in the Netherlands where it is not uncommon for patients with both antisocial tendencies and substance abuse problems to receive treatment for either one of the disorders and to be transferred back and forth between addiction care and forensic care facilities, instead of being offered integrated treatment.

Krueger, Markon, Patrick, Benning, and Kramer (2007) modeled the specific subordinate dimensions of this externalizing spectrum through a multiscale questionnaire inventory, the Externalizing Spectrum Inventory (ESI). This inventory was developed to provide for self-report-based dimensional measurement of this broad spectrum of externalizing psychopathology consisting of differing forms of aggression (physical, relational, destructive), impulsive-reckless tendencies, irresponsibility and deceitfulness, blame externalization, and various types of substance use and abuse (Krueger et al., 2007). The benefits of this type of clinical assessment relative to classification systems in which manifest disorders are scored as being either 'present/not present' is that the severity of each specific externalizing expression and the relations among them can be quantified, yielding information about profiles of externalizing pathology. For instance, two forensic patients exhibiting aggressive behavior may require different intervention approaches if one patient has high sensation seeking and high blame externalization tendencies, while the other has low sensation seeking and blame externalization, but high substance abuse tendencies.

A disadvantage of the ESI, however, is that the inventory is lengthy (415 items), and therefore differing brief versions have been developed for differing assessment purposes. Most validation studies to date have utilized a 100-item screening version that provides for more efficient measurement of the general externalizing factor than the lengthy version, while another 159-item version has been used to estimate scores on the ESI's Callous-Aggression and Substance Abuse subfactors, along with scores on the ESI's general factor (labeled Disinhibition; Patrick, Kramer, Krueger, & Markon, 2013). These

shorted versions of the ESI have proven effective in predicting relevant criterion measures of various types including indicated incidence of rule-breaking behaviors in childhood and adulthood, heightened levels of alcohol dependence and drug abuse, and reduced amplitude of the error-related negativity (a brain response that normally occurs following behavioral errors) along effectively predicting alcohol and drug problems, aggressive symptoms of APD, and affective-interpersonal features of psychopathy as assessed by Hare's (1991) or Lilienfeld and Widows (2005) PPI (Blonigen et al., 2010; Hall, Bernat, & Patrick, 2007; Nelson, Patrick, & Bernat, 2011; Venables & Patrick, 2012)

Extending this work, a newer brief form of the ESI (ESI-bf) containing 160 items (Patrick et al., 2013) was developed that provides for efficient but fine-grained measurement, with quantification of the general and subfactor factor levels of assessment complemented by assessment at the subscale level. Thus, this latest brief version provides a more equivalent representation of the full version. Validity data for this new version in undergraduate and incarcerated samples show highly similar correlations with normal range personality traits, such as lack of control, aggression and lack of harm avoidance, relative to the full-form ESI (Patrick et al., 2013). As such, this 160-item brief form may be the most usable version for forensic clinical practice, since clinical diagnostics require efficient but simultaneously detailed information gathering in order to achieve an optimal understanding of the psychopathology underhand. By administering a relatively brief inventory, which renders comprehensive information, the assessment of individual differences in externalizing pathologies will also be enhanced which will contribute to customized forensic care and improved risk management.

Until now, however, this inventory has not been investigated in any clinical patient sample, even though externalizing psychopathology is of particular relevance to treatment planning and evaluation of risk for criminal recidivism in clinical populations (Grella, Joshi, & Hser, 2003; Lund, Forsman, Anckarsäter, & Nilsson, 2011; Martínéz-Raga, Marshall, Keany, Ball, & Strang, 2002). Furthermore, none of the abbreviated ESI versions have been validated outside the United States, meaning no data yet exist regarding the any cross-national validity of brief versions of this inventory.

The principal objective of the current study was the translation, cultural adaptation, and initial validation of the ESI for use with Dutch research participants. For the initial validation, we investigated the reliability and validity of the ESI cross-sectionally in a Dutch clinical sample and a non-clinical control sample, and examined the extent to which the brief 160-item Dutch ESI could be used for similar assessment purposes as the full-form Dutch ESI. In other words, we investigated to what extent the brief form could be used for similar fine-grained assessment as the full version, which we consider necessary for (forensic) clinical assessment.

Method

Translation procedure

For the translation, we sought to extend the traditional back-translation method by also evaluating the translated version for semantic equivalence. We chose this more extensive methodology because the literature cumulatively indicates that the traditional back-translation procedure alone may not suffice for cross-cultural validation purposes (Maneesriwongul & Dixon, 2004; Sperber, 2004). The translation method was based on the recommendations of the ISPOR task force for translation and cultural adaptation (Wild et al., 2005) as well as the bilingual testing procedure of Jones and colleagues (Jones, Lee, Phillips, Zhang, & Jaceldo, 2001; Maneesriwongul & Dixon, 2004). The procedure consisted of a combination of a standard back-translation procedure, followed by a bilingual testing procedure in which the Dutch- and English-language paper-and-pencil versions were administered sequentially to a sample of 25 bilingual Dutch subjects. These subjects were recruited via a 'snowball' method (Biernacki & Waldorf, 1981) in which researchers at the psychology department of the Nijmegen University were asked to recruit bilingual subjects in their social surroundings. The original and translated versions of the ESI were administered four days apart in alternating order (Dutch–English, English–Dutch) to control for any effect of the order of administration. To control for any priming effects the English items were randomized. Next, 'weak' items (items for which the mean scores of the Dutch and English version differed more than two points) were identified through an item level Wilcoxon Signed Rank tests (Wilcoxon, 1945) of differences between the responses on the Dutch and the English items. Twenty-one items were revised after these analyses. Finally, the revised items were back-translated and evaluated by Christopher Patrick, one of the co-authors of the original paper on the American version of the ESI (Krueger et al., 2007).

Validation study

Participants

For the initial validation work, two participant groups were recruited from the East and South–West of the Netherlands (total $n = 203$): A non-psychiatric control sample ($n = 104$) consisting of 69 psychology undergraduates and 35 subjects who were not students, and a psychiatric inpatient sample with a prior history of manifest externalizing psychopathology ranging from substance abuse to antisocial behavior or both ($n = 99$), consisting of 62 drug rehab inpatients, 21 forensic drug rehab patients, and 16 forensic psychiatric inpatients. The psychiatric subjects were recruited in these three different types of institutions to ensure a broad range of scores on the items and subscales of the ESI, as was the case in the study by Krueger et al. (2007). The following exclusion criteria were applied: Psychotic disorders, severe brain damage, problems with

reading the Dutch language (e.g. due to illiteracy). Next, age and gender differences between groups were calculated (see Hicks et al., 2007 for age and gender effects on externalizing behavior). The inpatients ($M_{\text{inpatients}} = 36.5$, $SD = 11.1$) were significantly ($z = -6.8$, $p < .001$) older than the controls ($M_{\text{controls}} = 25.85$, $SD = 9.95$). The inpatient group also included significantly more ($\chi^2 = 26.29$, $p < .001$) male subjects (78.6%) than the community-based sample (43.3%). Of the control sample, only 1.0% had a criminal record, whereas 65.3% of the inpatient sample had one.

The Externalizing Spectrum Inventory

The 23 subscales of the ESI assess differing expressions (facets) of externalizing proneness in domains of aggression, irresponsibility and deceitfulness, impulsivity/sensation-seeking, blame-externalization, and substance use/abuse. Structurally, all subscales operate as indicators of a general Externalizing (or Disinhibition) factor, with some scales also loading on two subsidiary factors reflecting Callous Aggression (e.g. 'I have beaten someone up for bothering me') and Substance Abuse (e.g. 'Alcohol makes a good time even better'). In the full-form ESI (Krueger et al., 2007) all items (415) are completed using a 0–3 scale ranging from 'not true' to 'true', with total scores thus varying between a minimum of 0 and a maximum of 1245 points. Item Response Theory analyses in the (English language) ESI development sample of students and prisoners (Krueger et al., 2007) showed adequate to high reliability and good information coverage. Internal consistencies for the differing subscales were high, ranging from $\alpha = .85$ –.98 (Patrick et al., 2013). The test–retest reliability has not been investigated until this current study.

The ESI-bf (Patrick et al., 2013) contains 160 items completed in the same manner, with the range of total scores thus varying from 0 to 480. The English-language versions of the full ESI and the ESI-bf show a high degree of correspondence with the original 415-item version. Internal consistencies for individual subscales ranged from $\alpha = .74$ –.96, with each subscale exhibiting unidimensionality (Patrick et al., 2013). In the current study only the 415-item version was administered and item scores corresponding to the 160-item ESI-NL were derived from the 415-item data-set. Therefore, the terms 'brief ESI-NL' or 'brief version' in the following text, refer to the extracted data corresponding to the 160-item version. For the analyses, all items were coded in the direction of high scores indicating higher levels of externalizing.

Procedure

Data collection for validation of the final Dutch ESI (ESI-NL) was conducted between March 2009 and April 2010. Recruitment of the non-psychiatric control sample was conducted using a snowball sampling method, through distribution of flyers. Patients were enrolled after they had provided written

informed consent. All participants received a maximum of €15.00 reimbursement for participation (€7.50 for each administration).

Demographic variables and the full 415-item version of the ESI-NL were administered via a web-based computerized version to the control subjects ($n = 104$), while psychiatric subjects ($n = 99$) completed the inventory using an equivalent offline computerized version. Of the 203 subjects, a subsample of 109 (consisting of 83 controls and 26 psychiatric subjects, of which 18 subjects were forensic) completed the full inventory twice across an average time interval of two weeks in order to evaluate test–retest reliability.

Data analysis

Regarding the data of the 202 subjects who completed the inventory only once, analyses showed a randomly distributed missing data percentage of 0.01%. Given the low proportion of missing data and the relatively random distribution of scores, missing item values were imputed using individual means per case per subscale.

Internal consistency (Cronbach's α) and a relatively conservative test–retest evaluation of reliability (intra-correlation coefficients, single measure) were conducted for the full and brief versions separately.

The 'predictive' validity of each version of the ESI was examined using a Receiver Operating Characteristic analysis. With this analysis, the chance of any inpatient scoring significantly higher on externalizing than any control subject and the cut-off score for an optimal discrimination between these two groups, were examined. The patient group was applied as a state variable in this analysis.

In addition, a multiple regression analysis was carried out on the full and brief ESI-NL in order to establish a standardized measure of the association between the assumed factors and each subscale. For this latter analysis the model as presented in Krueger and South (2009) was applied, meaning that the nine substance use and substance problems scales were used as predictor variables of the Substance Abuse criterion variable and the remaining subscales were used as predictors of the Callous Aggression criterion variable.

The equivalence between the Dutch ESI and the original English ESI as well as similarities between the 160-item and the 415-item ESI-NL were first evaluated by computing intercorrelations of (1) the subscales of the full- and brief-form ESI versions separately; (2) the total and subscale scores of the brief ESI-NL with the total and subscale scores of the full ESI-NL; (3) the total and subscale scores of the brief ESI-NL with corresponding total and subscale scores for the full-form ESI based on items not included in the brief ESI-NL (in order to control for any overestimation of intercorrelations between the full and brief versions). The latter aggregated subscale scores will be designated as $\text{Sum}_{\text{NotIncl}}$ in the text that follows. All statistical tests were conducted using SPSS 18.

Results

Translation procedure

By testing for 'weak items' 37 items with significantly differing score were identified ($z = 3.9, p < .000$ – $z = 2.0, p < .05$). Of these weak items 21 Dutch–English item pairs with extreme response reversals (respondents scoring 0 on one version and 3 on the other) were identified and revised.

Descriptive statistics

Mean scores for the inpatients and controls are depicted separately for each sample for descriptive purposes in Table 1. Due to the skewness of the data, both mean and median scores are presented.

Score reliability

In Table 2 the results indicate high reliability for both full- and brief-form subscales and total scores with alphas ranging from $\alpha = .87$ – $.99$ and intraclass correlation coefficients ranging from $ICC = .69$ – $.92$. Differences in test–retest reliability were modest to negligible for the full and brief total scores, and also for corresponding subscale scores. Internal consistencies of total scores were also highly similar between the brief and full ESI-NL versions. For all individual subscales, the internal consistency (α) of the brief versions was highly similar to that of the full version (average difference = $.06$). Since internal consistency depends on the number of items, average inter-item correlations are also shown, in Table 2. For all but two of the brief subscales, average inter-item correlations exceeded those for corresponding full-length subscales.

'Predictive' validity

Next, the validity of the brief- and full-form for predicting whether participants belonged to the externalizing inpatients subgroup vs. the non-clinical control group was evaluated. The two versions showed comparably high Area Under the Curves (AUCs) with highly similar confidence intervals (see Table 3). Parallel to results for the total scores, AUCs were very similar for individual full- and brief-form subscales (i.e. the largest difference in AUCs was $.77$ vs. $.75$ for the full vs. brief versions of the Excitement Seeking scale). Seventeen (full version) to 18 (brief version) of the 23 ESI subscales showed moderate to excellent AUC values ($.75$ or higher), with six (full version) to five (brief version) showing AUCs below $.75$.

Relations between the subscales and the subfactors

As one indicator of the resemblance between the Dutch ESI and the original English ESI, the intercorrelations among subscales for the two versions were

Table 1. Mean scores of inpatients and controls.

Subscales	Patients				Controls			
	<i>M</i>	<i>Mdn</i>	SD	Range	<i>M</i>	<i>Mdn</i>	SD	Range
Relational aggression	15.39	19.00	11.70	54	19.84	9.00	9.36	44
Physical aggression	17.54	24.00	16.67	61	25.15	7.00	10.23	59
Destructive aggression	5.65	5.00	9.69	41	8.62	0.00	5.55	27
Empathy (–)	21.11	25.00	15.58	69	26.05	12.00	12.05	62
Blame externalization	11.33	16.00	10.32	40	17.56	2.5	6.61	35
Alienation	9.63	14.00	6.34	27	13.86	4.00	5.29	21
Alcohol problems	25.40	41.00	29.26	88	41.04	5.00	13.79	59
Alcohol use	40.27	43.00	15.71	63	41.63	44.00	16.92	66
Marijuana problems	12.25	23.00	17.76	52	22.36	0.00	7.44	48
Marijuana use	21.52	38.00	15.01	50	32.12	9.00	10.40	45
Drug problems	23.60	48.00	22.11	75	43.16	0.00	12.74	70
Drug use	15.38	25.50	8.82	38	23.19	5.50	8.13	36
Theft	10.87	19.00	12.76	45	19.06	0.00	6.61	36
Fraud	9.01	12.00	10.30	39	13.67	2.5	6.22	38
Honesty (–)	11.12	12.00	7.91	45	12.32	9.00	5.68	27
Irresponsibility	21.96	35.50	17.36	67	34.84	6.00	11.88	67
Dependability (–)	16.85	20.00	10.37	56	20.64	12.00	9.04	46
Problematic impulsivity	21.72	36.00	13.79	57	34.58	7.00	10.17	57
Planful control (–)	10.97	13.00	7.24	33	13.89	8.00	5.37	29
Impatient urgency	16.25	23.00	7.86	35	21.54	11.00	7.50	33
Rebelliousness	15.47	21.00	10.32	39	21.95	7.50	7.78	41
Boredom proneness	14.02	19.50	9.22	36	19.39	8.00	7.24	36
Excitement seeking	19.92	27.00	13.91	51	26.62	12.00	9.76	54
<i>Total</i>	<i>969.79</i>	<i>981.44</i>	<i>178.63</i>	<i>828</i>	<i>648.20</i>	<i>624.50</i>	<i>144.70</i>	<i>912</i>

compared. For the Substance Abuse scales, all showed sufficient to high correlations with similar patterns in the full and brief versions ($r = .36-.86$ for the full version, and $r = .31-.86$ for the brief version) except Alcohol Use, which showed low correlations ($r < .25$) with all subscales aside from Alcohol Problems ($r = .49$ and $.60$, respectively). Regarding the Callous Aggression subscales all subscales, except Honesty and Blame Externalization correlated sufficient ($r = .30$ for both brief and full subscales) to good ($r = .77-.76$ for the full and brief scales respectively). In the multiple regression analysis (presented in Table 4) on the Callous Aggression and Substance Abuse subscales, both the full and brief subscales predicted the Substance Abuse and Callous Aggression factors significantly ($p < .001$). Of the Callous Aggression subscales Physical Aggression has the largest incremental value to the

Table 2. Internal consistency, average inter-item correlation ($N = 202$) and test-retest reliability ($N = 109$) full and brief ESI-NL.

Subscales	415-item version				160-item version			
	Items	α	M_r	ICC	Items	α	M_r	ICC
Relational aggression	19	.91	.34	.86	8	.80	.34	.83
Physical aggression	21	.94	.44	.90	8	.89	.51	.90
Destructive aggression	15	.92	.43	.90	7	.89	.55	.86
Empathy (-)	31	.93	.30	.85	11	.83	.32	.80
Blame externalization	14	.93	.50	.81	4	.91	.71	.70
Alienation	9	.88	.45	.86	3	.79	.55	.80
Alcohol problems	30	.98	.58	.81	9	.94	.64	.71
Alcohol use	23	.92	.35	.91	9	.82	.33	.86
Marijuana problems	18	.97	.65	.92	7	.92	.62	.88
Marijuana use	17	.95	.52	.92	7	.94	.70	.92
Drug problems	25	.98	.65	.88	11	.96	.69	.88
Drug use	13	.91	.43	.85	6	.89	.57	.85
Theft	15	.94	.51	.91	8	.91	.56	.90
Fraud	14	.91	.41	.84	6	.82	.43	.86
Honesty (-)	15	.87	.32	.79	5	.71	.35	.77
Irresponsibility	25	.95	.41	.85	10	.90	.47	.84
Dependability (-)	23	.89	.28	.80	7	.76	.34	.73
Problematic impulsivity	20	.96	.52	.87	7	.91	.59	.86
Planful control (-)	11	.89	.43	.72	6	.86	.51	.69
Impatient urgency	12	.91	.45	.79	5	.85	.53	.73
Rebelliousness	15	.93	.45	.86	6	.90	.60	.83
Boredom proneness	12	.93	.52	.83	4	.90	.70	.82
Excitement seeking	18	.93	.44	.86	6	.88	.55	.82
<i>Total</i>	<i>415</i>	<i>.99</i>	<i>-</i>	<i>.87</i>	<i>160</i>	<i>.99</i>	<i>-</i>	<i>.89</i>

Note: α = Cronbach's alpha; M_r = average inter-item correlation; ICC = intraclass correlation coefficient; $p < .001$.

subfactor score followed by Empathy in both the full and brief version. The Substance Abuse subscales, however, show differences between the full and brief inventory. In the full version Alcohol Problems emerged as the subscale with the highest incremental value to the Substance Abuse subfactor, while in the brief-form inventory, Drug Problems show the highest value.

Correlations full and brief ESI items

Regarding the resemblance between the full and brief ESI-NL versions (Table 5), correlations between the 160-item and the 415-item total scores, and between the 160-item ESI and the corresponding Sum_{NotIncl} were both extremely high ($r = .97-.96$, respectively), indicating that the first correlation was not the result of overestimation due to correlations of the same brief ESI-NL items. Correlations across versions for specific subscales were lower, as shown in Table 5, but still respectable (correlations brief-full subscales: $r = .87-.97$ and correlations brief and Sum_{NotIncl} subscales: $r = .72-.93$).

Table 3. 'Predictive' validity of the full and brief ESI-NL ($N = 202$).

Subscales	415-item			160-item		
	AUC	CI 95%	Max	AUC	CI 95%	Max
Relational aggression	.73***	.66–.80	57	.69***	.61–.76	24
Physical aggression	.77***	.71–.84	63	.76***	.69–.83	24
Destructive aggression	.73***	.66–.80	45	.70***	.63–.77	21
Empathy (–)	.69***	.62–.76	93	.69***	.62–.76	33
Blame externalization	.84***	.79–.90	42	.80***	.74–.86	12
Alienation	.84***	.78–.89	27	.82***	.77–.88	9
Alcohol problems	.82***	.76–.88	90	.78***	.72–.85	27
Alcohol use	.53	.45–.61	69	.59*	.51–.67	27
Marijuana problems	.84***	.78–.90	54	.82***	.76–.88	21
Marijuana use	.83***	.77–.90	51	.84***	.78–.90	21
Drug problems	.91***	.86–.95	75	.91***	.86–.95	33
Drug use	.88***	.82–.93	39	.86***	.80–.92	18
Theft	.87***	.82–.92	45	.85***	.80–.91	24
Fraud	.78***	.72–.85	42	.77***	.71–.84	18
Honesty (–)	.58	.50–.66	45	.58	.50–.66	15
Irresponsibility	.90***	.85–.94	75	.91***	.87–.95	30
Dependability (–)	.71***	.64–.78	69	.73***	.66–.80	21
Problematic impulsivity	.92***	.88–.96	60	.92***	.89–.96	21
Planful control (–)	.74***	.67–.81	33	.76***	.69–.82	18
Impatient urgency	.83***	.77–.88	36	.78***	.72–.84	15
Rebelliousness	.84***	.79–.90	45	.83***	.77–.89	18
Boredom proneness	.81***	.75–.87	36	.78***	.72–.84	12
Excitement seeking	.77***	.71–.84	54	.75***	.68–.82	18
<i>Total</i>	.93***	.89–.96	1245	.92***	.88–.96	480

Note: AUC = Area Under the Curve; CI = Confidence interval; Max = Maximum (sub)scale score.

* $p < .05$; *** $p < .001$.

Discussion

The results indicated that both the 160-item and the 415-item Dutch ESI have high reliability and excellent 'predictive' validity on a total score level. On a subscale level, the 'predictive' validity was sufficient to high for 17–18 out of 23 subscales for both versions. We found no significant differences in the psychometric properties of the full and brief Dutch ESI at a total score level. At the level of subscales, no differences in 'predictive' validity were evident. Some differences did emerge between the brief- and full-form subscales in intercorrelations and reliability coefficients, but in most cases these differences were slight and will most likely have minimal impact on ESI-NL-based decision-making. We address these subscale differences in more detail below.

The modest reduction in observed internal consistency for the brief vs. full subscales was congruent with findings reported by Patrick et al. (2013). However, for two subscales the average inter-item correlation declined minimally, and for the remaining 21 subscales inter-item correlations actually increased, indicating increased coherence among the items of the brief subscales. This

Table 4. Standardized regression weights of subscales predicting subfactor scores.

Subscale	<i>b</i> *	
	Full	Brief
<i>Callous aggression</i>		
Relational aggression	.103***	.112***
Physical aggression	.141***	.146***
Destructive aggression	.075***	.106***
Empathy (-)	.132***	.133***
Blame externalization	.095***	.079***
Alienation	.064***	.057***
Fraud	.086***	.095***
Honesty (-)	.063***	.057***
Dependability (-)	.094***	.078***
Planful control (-)	.063***	.088***
Impatient urgency	.083***	.095***
Rebelliousness	.100***	.119***
Boredom proneness	.088***	.085***
Excitement seeking	.123***	.111***
<i>Substance abuse</i>		
Alcohol problems	.210***	.159***
Alcohol use	.126***	.117***
Marijuana problems	.128***	.113***
Marijuana use	.127***	.143***
Drug problems	.201***	.218***
Drug use	.088***	.111***
Theft	.099***	.132***
Irresponsibility	.149***	.152***
Problematic impulsivity	.134***	.118***

Note: Callous aggression = Callous aggression subfactor score; Substance abuse = Substance abuse subfactor score; *b** = standardized regression weights.

****p* < .001.

result suggests that the brief subscales assess the core characteristics of each specific externalizing expression, at the cost of the broader content of the full ESI. Regarding test-retest reliability, the reliability of the brief subscales was impressively high despite the aforementioned modest declines in internal consistency.

The lower correlations between the brief subscales and the corresponding Sum_{NotIncl} indicate some loss of information. This loss of information was also found in IRT results of Patrick et al. (2013). However, the correlations were still high, indicating that information captured by the brief subscales overlapped to an acceptable degree with the corresponding Sum_{NotIncl} scales.

Certain limitations should be acknowledged that warrant attention in follow-up research. To begin with, the inpatient group was heterogeneous, contained significantly more males and had a higher mean age than the control group. However, these differences did not appear to have biased the results since the inpatients group scored as expected on the ESI-NL. In future validity

Table 5. Correlations between the full and brief ESI-NL ($N = 202$).

	$r_{\text{brief-full}}$	$r_{\text{brief-SumNotIncl}}$
Relational aggression	.94**	.81**
Physical aggression	.94**	.86**
Destructive aggression	.92**	.73**
Empathy (-)	.95**	.87**
Blame externalization	.92**	.85**
Alienation	.90**	.79**
Alcohol problems	.96**	.89**
Alcohol use	.94**	.85**
Marijuana problems	.97**	.93**
Marijuana use	.95**	.88**
Drug problems	.97**	.93**
Drug use	.96**	.86**
Theft	.97**	.88**
Fraud	.88**	.76**
Honesty (-)	.88**	.73**
Irresponsibility	.95**	.89**
Dependability (-)	.87**	.72**
Problematic impulsivity	.96**	.91**
Planful control (-)	.94**	.76**
Impatient urgency	.93**	.80**
Rebelliousness	.95**	.85**
Boredom proneness	.95**	.86**
Excitement seeking	.94**	.87**
<i>Total</i>	.97**	.96**

Note: $r_{\text{brief-full}}$ = Pearson correlation between the 160-item ESI-NL and the 415-item ESI-NL; $r_{\text{brief-SumNotIncl}}$ = Pearson correlation between the 160-item ESI-NL and the corresponding items that were not included in the 160-ESI.

** $p < .01$.

studies with purposes beyond discriminating severe externalizing inpatients from controls, we recommend sampling from control populations other than students.

Second, the sample size in this study did not allow for factor analysis, while validation studies on the source ESI focused mainly on the factor structure. The intercorrelations between corresponding subscales for the two versions did, however, suggest high resemblance between the factor structures of the Dutch and source ESI. The multiple regression analysis also indicated significant contributions of each subscale to the factor scores. The main reason for the small sample size is that the full 415 item ESI posed too much of a burden to the larger part of the clinical subjects, in spite of the exclusion of patients with cognitive disorders or mental retardation. Since the 160-item version of the ESI appears so highly similar to the full-form ESI, we recommend use of this briefer version in clinical studies, including studies on the factor structure of the ESI and its relationship to other measures in patient samples.

Finally, although the ESI-NL showed promising results in this initial study, extended research on the brief version and its relationship with other measures of psychopathology (cf. Venables & Patrick, 2012) will be needed for it to contribute effectively to scientific research and clinical practice. In particular, the subscales that showed poor discrimination between controls and inpatients in the current study and the practicality of administering the 160 items in forensic populations should be examined further. Lastly, the predictive validity of the ESI measure should be investigated in a prospective design.

In sum, based on findings from the current work, it seems fair to conclude that the brief version of the ESI is a promising, time-efficient instrument for assessing the spectrum of externalizing problem behaviors in forensic patients, and may turn out to be an important predictor of key clinical outcomes such as amenability to treatment, readjustment upon release to the community, and proneness to recidivism.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the DJI from the Ministry of Justice of The Netherlands [grant number 5620076/09/] and by funding of Tactus Addiction Care in the Netherlands.

ORCID

Cor A.J. De Jong  <http://orcid.org/0000-0003-1824-7303>

References

- Abraham, K. M., Zwecker, N. A., Welty, L. J., Hershfield, J. A., Dulcan, M. K., & Teplin, L. A. (2015). Comorbidity and continuity of psychiatric disorders in youth after detention: A prospective longitudinal study. *Journal of the American Medical Association Psychiatry, 72*, 84–93.
- Biernacki, P., & Waldorf, D. (1981). Snowball sampling - problems and techniques of chain referral sampling. *Social Methods Research, 10*, 141–163.
- Blonigen, D. M., Patrick, C. J., Douglas, K. S., Poythress, N. G., Skeem, J. L., Lilienfeld, S. O., Krueger, R. F. (2010). Multimethod assessment of psychopathy in relation to factors of internalizing and externalizing from the personality assessment inventory: The impact of method variance and suppressor effects. *Psychological Assessment, 22*, 96–107.
- Cerda, M., Sagdeo, A., & Galea, S. (2008). Comorbid forms of psychopathology: Key patterns and future research directions. *Epidemiologic Reviews, 30*, 155–177.
- Grella, C. E., Joshi, V., & Hser, Y.-I. (2003). Follow-up of cocaine-dependent men and women with antisocial personality disorder. *Journal of Substance Abuse Treatment, 25*, 155–164.

- Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing psychopathology and the error-related negativity. *Psychological Science, 18*, 326–333.
- Hare, R. D. (1991). *The hare psychopathy checklist-revised*. Toronto: Multi-Health Systems.
- Hicks, B. M., Blonigen, D. M., Kramer, M., Krueger, R. F., Patrick, C. J., Iacono, W. G., & McGue, M. (2007). Gender differences and developmental change in externalizing disorders from late adolescence to early adulthood: A longitudinal twin study. *Journal of Abnormal Psychology, 116*, 433–447.
- Jaffe, A., Du, J., Huang, D., & Hser, Y.-I. (2011). Drug-abusing offenders with comorbid mental disorders: Problem severity, treatment participation, and recidivism. *Journal of Substance Abuse and Treatment, 43*, 244–250.
- 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.
- Kessler, R. C. (2004). The epidemiology of dual diagnosis. *Biological Psychiatry, 56*, 730–737.
- Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry, 56*, 921–926.
- 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: Modelling the externalizing spectrum. *Journal of Abnormal Psychology, 111*, 411–424.
- 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.
- 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*, 2061–2070.
- Lilienfeld, S. O., & Widows, M. (2005). *Manual for the psychopathic personality inventory revised (PPI-R)*. Lutz: Psychological Assessment Resources.
- Lund, C., Forsman, A., Anckarsäter, H., & Nilsson, T. (2012). Early criminal recidivism among mentally disorders offenders. *International Journal of Offender Therapy and Comparative Criminology, 56*, 749–468.
- Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation: A methods review. *Journal of Advanced Nursing, 48*, 175–186.
- Martínez-Raga, J., Marshall, E. J., Keany, F., Ball, D., & Strang, J. (2002). Unplanned versus planned discharges from in-patient alcohol detoxification: Retrospective analysis of 470 first-episode admissions. *Alcohol & Alcoholism, 37*, 277–281.
- Merikangas, K. R., & Kalaydjian, A. (2007). Magnitude and impact of comorbidity of mental disorders from epidemiologic surveys. *Current Opinion in Psychiatry, 20*, 353–358.
- Mir, J., Kastner, S., Priebe, S., Konrad, N., Ströhle, A., & Mundt, A. (2015). Treating substance abuse is not enough: Comorbidities in consecutively admitted female prisoners. *Addictive Behaviors, 46*, 25–30.
- Nelson, L. D., Patrick, C. J., & Bernat, E. M. (2011). Operationalizing proneness to externalizing psychopathology as a multivariate psychophysiological phenotype. *Psychophysiology, 48*, 64–72.
- Oliemeulen, O., Vuijk, P., Rovers, B., & VandenEijnden, R. (2007). *Problematische alcoholgebruikers, druggebruikers en gokkers in het gevangeniswezen* [Problematic alcohol users, drug users and gamblers in detention]. Rotterdam: IVO.

- 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.
- Sperber, A. D. (2004). Translation and validation of study instruments for cross-cultural research. *Gastroenterology, 126*, S124–S128.
- Tikkanen, R., Holi, M., Lindberg, N., Tiihonen, J., & Virkkunen, M. (2009). Recidivistic offending and mortality in alcohol violent offenders: A prospective follow-up study. *Psychiatry Research, 168*, 18–25.
- Van Horn, J. E., Eisenberg, M. J., Van Kuik, S., & Van Kinderen, G. M. (2012). Psychiatrische stoornissen en recidive bij geweldplegers met dubbele diagnose: Vergelijking met andere subgroepen geweldplegers [Psychiatric disorders and recidivism in violent offenders with dual diagnosis: A comparison with other subgroups of violent offenders]. *Tijdschrift voor Psychiatrie, 54*, 497–507.
- 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*, 88–100.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*, 80–83.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health, 8*, 94–104.