



Research article

Revisiting the psychometric properties of the McArthur admission experience survey: Validating the Portuguese version using a bifactor approach

Deborah Oyine Aluh^{a,b,c,*}, Sofia Azeredo-Lopes^{b,d}, Barbara Pedrosa^{a,b},
Manuela Silva^{a,b}, Ugnė Grigaitė^{a,b}, Ana Rita Martins^e,
Maria Ferreira de Almeida Mousinho^f, Graça Cardoso^{a,b},
José Miguel Caldas-de-Almeida^{a,b}

^a Lisbon Institute of Global Mental Health, Lisbon, Portugal

^b Comprehensive Health Research Centre (chrc), NOVA Medical School, NOVA University of Lisbon, Lisbon, Portugal

^c Department of Clinical Pharmacy and Pharmacy Management, University of Nigeria Nsukka, Nigeria

^d Department of Statistics and Operational Research, Faculdade de Ciências, Universidade de Lisboa, Portugal

^e Centro Hospitalar De Vila Nova De Gaia/Espinho, E.P.E. | V. N. Gaia/Espinho Hospital Centre, Portugal

^f Department of Mental Health Unidade Local de Saúde do Baixo Alentejo, Beja, Portugal

ARTICLE INFO

Keywords:

McArthur admission experience survey
Perceived coercion
Psychometric
Validation
Portuguese
Bifactor model

ABSTRACT

Cultural factors play a significant role in shaping the perception of coercion during psychiatric admissions. The present study aimed to assess the psychometric properties of the Portuguese Admission Experience Survey (P-AES).

The study employed a cross-sectional approach in five psychiatric departments in three regions of Portugal. A total of 208 patients participated in the survey. Reliability was assessed through internal consistency and test-retest procedures. Internal validity was analyzed using a two-parameter logistic item response model, exploring three models, including a bifactor model. Convergent validity was determined by correlating AES scores with the Coercion Ladder (CL), Client Assessment of Satisfaction (CAT), and Global Assessment of Functioning (GAF) scale. Discriminatory power was assessed by comparing scores between patients with voluntary and involuntary admission status.

The P-AES demonstrated satisfactory internal consistency and test-retest reliability. The bifactor model exhibited superior fit compared to the one-factor and three-factor models. Correlations between P-AES and CL, as well as CAT scores, indicated good convergent validity. Additionally, P-AES scores were notably higher in patients with compulsory psychiatric hospital admission compared to those admitted voluntarily, confirming its discriminatory power.

The bifactor model suggests that all three domains of the AES should be used to measure the subjective experience of coercion.

* Corresponding author. Lisbon Institute of Global Mental Health Comprehensive Health Research Centre (CHRC) NOVA Medical School, Universidade Nova de Lisboa Rua do Instituto Bacteriológico n°5, 1150-082 Lisboa, Portugal.

E-mail addresses: deborah.aluh@nms.unl.pt, aluhdeborah@yahoo.com (D.O. Aluh).

<https://doi.org/10.1016/j.heliyon.2024.e24114>

Received 7 September 2023; Received in revised form 23 November 2023; Accepted 3 January 2024

Available online 9 January 2024

2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The significant attention given to the use of coercion in mental health care has prompted a global movement to minimize its use. Coercion encompasses both actions performed upon someone and the subjective experience resulting from those actions. So, it represents both objectively observable actions and subjectively perceived outcomes [1]. Perceived coercion is the subjective experience of coercion. It has been suggested that the power dynamics between patients and care providers may play a more significant role in shaping the perception of coercion than the actual coercive acts themselves [2].

There is a growing interest in perceived coercion, as it has been found to have significant detrimental effects on patients' prognoses [3]. It is not only related to legal status or the experience of coercive practices but also the level of patient involvement in treatment decisions and knowledge of legal issues [4]. High levels of perceived coercion can damage the therapeutic relationship with care providers [5], while lower levels are linked to higher satisfaction with treatment [6]. Several tools have been developed to measure the subjective experience of coercion by patients of psychiatric services, such as the Cantril Ladder [7–9], Coercion Experience Scale [10–15], the Experienced Coercion Scale [16], and Nordic Admission Interview [8,17]. Nevertheless, the Admission Experience Survey (AES), designed for the MacArthur Coercion Study, stands out as one of most commonly used tool [7,17–28].

The AES is a questionnaire consisting of 16 items based on the MacArthur Admission Experience Interview. It aims to examine the experience of hospitalization, patients' interactions with clinical staff and family members during admission, as well as their input in the decision to be admitted [23]. The experiences are categorized into three groups: (i) feelings of coercion or perceived coercion, (ii) negative pressures, and (iii) voice or procedural justice. The Perceived Coercion domain centers around concepts such as freedom, choice, initiative, control, and influence regarding admission to the hospital [29,30]. The Negative Pressure domain pertains to experiences of being forced, threatened, or physically compelled to enter the hospital [31]. Lastly, the Voice or Procedural justice domain relates to having an opportunity to express one's opinion regarding hospital admission [23,32]. The AES has demonstrated favorable psychometric properties and has been valuable in assessing the subjective experience of coercion in various psychiatric clinical environments [33,34]. Despite facing criticism for its focus solely on the hospital admission process and its limited ability to assess the impact of other coercive measures in the same situation [10], the AES still holds significant value because the admission experience plays a crucial role in shaping the overall impact of other coercive measures throughout admission and therapeutic relationships formed [35–37]. The perception of coercion during psychiatric admission can be influenced by context, culture, and beliefs [31,38]. Hence, it is crucial to examine the impact of diverse cultures on the perception of coercion during psychiatric admissions. Research in this area among Portuguese speaking populations seems to be hindered by the absence of a culturally adapted and standardized tool. Validating the Portuguese version of the Admission Experience Survey will facilitate meaningful comparisons of findings across cultures and pave the way for further investigation into factors that may influence the admission experience.

Psychometric analysis is of utmost importance for ensuring the instrument's reliability, validity, and appropriateness for its intended purpose. It allows for the verification of how well the included items align with the underlying theoretical framework of the instrument, facilitating the evaluation of the phenomenon of interest. These attributes contribute to the overall quality and credibility of research and provide a solid foundation for evidence-based interventions. Not many studies that have used the McArthur Admission Experience survey have validated the instrument in their context before use. The validation processes of the Swedish [26] Chinese [31], French [33] and Italian [34] versions of the AES have been reported. However, doubts have arisen regarding the actual factor structure of the AES due to concerns about the methodology used in the previous psychometric analyses. To ensure accurate assessment of internal validity, the two-parameter logistic (2 PL) item response models are preferred considering the dichotomous nature of the AES. This method was solely employed in the validation study of the French version of the AES. Earlier investigations have indicated a three-factor structure, and the French study also supports the same three-factor structure. However, the French study goes on to report that the one-factor structure also exhibits good fit properties. A bifactor model is recommended to explore complex constructs that consist of moderately related dimensions [39]. It allows for investigating whether there is a general factor present, and to what extent each specific dimension is distinct from this general factor [39]. Thus, it seems reasonable to hypothesize that the AES may be comprised of a general factor (i.e., subjective experience of coercion during admission) that accounts for the commonality shared by its dimensions.

Building upon previous studies, this article proposes a bifactor structure for examining the subjective experience of coercion during admission. Notably, as no previous research has explored the applicability of a bifactor model for the subjective experience of coercion during admission, the findings from this study will have significant implications for further investigations of the concept. The present study aimed to evaluate the validity and reliability of the Portuguese version of the AES, and to assess whether the bifactor model gives a better fit than the three-factor model or one-factor model.

2. Materials and methods

2.1. Study design and setting

The research employed a cross-sectional approach and took place in five selected public psychiatric departments located in Lisbon, Porto, and Alentejo representing districts from north and south of Portugal between February 2022 and May 2023. The study sample included patients on admission in the selected psychiatric departments who met the eligibility criteria of being at least 18 years old and being able to give informed consent to participate in the study. Patients with dementia were excluded. Ethical approval was granted by the Ethics committee of Nova medical school (100/2021/CEFCM) and the ethics committee of each psychiatric department.

2.2. Measures

The MacArthur Admission Experience survey was used to enable rapid assessment of patients' perceptions of psychiatric hospital admission [33]. The Perceived Coercion domain has a scoring range of 0–5 while the Negative Pressure domain has a scoring range from 0 to 6. The Voice domain has a scoring range from 0 to 3. Item 9 was discarded as in the original scale, and the 16th item, which asks, "How did being admitted to the hospital make you feel?" and gives four possible options, was not included in the scoring. The GAF (Global Assessment of Functioning) is a scale utilized to assess an individual's social, occupational, and psychological functioning [40]. It ranges from 1 corresponding to severely impaired to 100 corresponding to extremely high functioning and is based on the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition). The Coercion Ladder (CL) is a visual analogue tool that asks the patient to mark the level of their perceived coercion on a scale of 1 (indicating the minimum use of coercion – "I came totally on my own will and initiative") to 10 (indicating the maximum use of coercion) [41]. The Client's Assessment of Treatment Scale (CAT) is a patient-reported outcome measure consisting of seven items designed to assess the patient's appraisal of in-patient care [42]. These items explore the patient's satisfaction with the treatment received, the treating clinicians, other mental health professionals, medications, other received treatments, and the overall care received. Each item is rated on a 10-point Likert scale ranging from 0 ("not at all") to 10 ("yes entirely").

2.3. Procedure

The McArthur Admission Experience Survey was translated into Portuguese by a bilingual native speaker with clinical experience and subsequently reviewed by a team of experts who were also native speakers. Then the translated questionnaire underwent back-translation from Portuguese to English by an independent translator and was checked again by the team of experts for accuracy. The instrument was then piloted among five patients admitted to one of the psychiatric departments selected for the study to assess for any difficulties or misunderstandings related to the terminology used. No changes were required at this stage. The study instrument, consisting of the Portuguese-AES (P-AES), CL, and CAT, was administered to the participants as soon as they were stable within the first seven days after admission (T0). The same instrument was administered to the respondents a second time before discharge (T1). Sociodemographic information (such as gender, age, employment status, marital status, nationality, the legal status of admission, and previous involuntary admissions) was obtained from the patients' records. Additionally, GAF scores and diagnoses according to ICD-10 criteria were recorded at both T0 and T1. The reliability of the P-AES scores was assessed using a test–retest approach as the instrument was administered at two different times. The average length of time for the second assessment was 12.54 (SD: 10.37) days. Validity was estimated based on the first assessment, by testing the three-factor model (perceived coercion, negative pressure, and voice sub-domains) as reported in previous studies [18,33,34], a one-factor model [33], and a new bifactor model. To assess convergent validity, the following hypotheses were proposed: Firstly, the P-AES, including its perceived coercion and negative pressure domains, would exhibit positive correlations with the Coercion Ladder. Conversely, the Voice domain would demonstrate a negative correlation with the Coercion Ladder. Secondly, it was anticipated that the P-AES, perceived coercion domain, and negative pressure domain would display negative correlations with the GAF and CAT, while the Voice domain would show positive correlations with the GAF and CAT. This was based on the idea that individuals with higher functioning might experience less coercion during the admission process [33], and those who are more satisfied with their treatment could perceive lower levels of coercion. To evaluate discriminatory power, the AES's ability to distinguish between involuntarily admitted and voluntarily admitted patients was examined.

2.4. Statistical analysis

The data's normality was assessed using the Kolmogorov-Smirnov test, and a non-normal distribution was indicated. Consequently, non-parametric statistical methods were employed, namely the Wilcoxon Signed Rank test, for comparing paired data. To estimate the reliability of the MacArthur Admission Experience survey and its domains, McDonald's model-based Omega (ω) [43] and Cronbach's alpha coefficient were employed. Both these coefficients were obtained through the *psych* R package [44]. The test–retest reliability was assessed using Spearman's rank correlation and intraclass correlation coefficient (ICC). The cutoff criteria, based on Ko and Li's approach, were applied. If the ICC is below 0.50, it points to unreliable results. A value between 0.50 and 0.75 suggests a moderately reliable outcome, while a range of 0.75–0.90 indicates good reliability. Anything above 0.90 signifies excellent reliability [45]. The presence of significant changes in P-AES scores between the first and second assessments was evaluated using Wilcoxon tests.

To evaluate the reliability of the items, a two-parameter logistic (2 PL) item response model was utilized, given that the items were dichotomous. Item response theory (IRT) evaluates the validity of measurement scales by elucidating the connection between a latent trait (in this case, perceived coercion), the characteristics of the items within the scale, and how respondents answered each item individually [46]. This method estimates both item discrimination and difficulty. Additionally, a Confirmatory Factor Analysis (CFA) was conducted using maximum likelihood estimation to assess the adequacy of the factor structure of the P-AES. Three models were estimated: a three-factor model incorporating perceived coercion, negative pressure, and voice domains; a single-factor model with a general perceived coercion domain; and a Bifactor model, assuming all items load into a general domain, with subsets of items also loading into specific domains [47]. The models were obtained by fitting maximum likelihood factor analysis to dichotomous data under the item response theory paradigm using Cai's [48,49] Metropolis-Hastings Robbins-Monro (MHRM) algorithm and quasi-Monte Carlo integration for the parameters estimates. To evaluate the model fit, we considered several indicators, including the root mean square error of approximation (RMSEA), the comparison fit index (CFI), and the Tucker-Lewis fit index (TLI). A RMSEA value of ≤ 0.06 , and CFI and TLI values of ≥ 0.95 indicated that the model fit was good. Alternatively, RMSEA values of ≤ 0.08 , and CFI

and TLI values of ≥ 0.90 were deemed acceptable [50]. Bifactor models are susceptible to overfitting [51,52], leading to a biased preference for these models when using model selection indices. To prevent excessive model fitting, four commonly employed model selection indices were used to compare the model fitting. These indices include Akaike's information criterion (AIC), Bayesian information criterion (BIC), Hannan–Quinn criterion index (HQ), and sample size-adjusted BIC (SABIC) [53–56]. The computation of these indices aimed to ensure a reasonable and unbiased evaluation of the models.

To evaluate the convergent validity of the P-AES with other scales, Spearman correlation coefficients were calculated. Since the convergent validity coefficients do not have established criteria, values exceeding .30 were deemed satisfactory, corresponding to a medium effect size according to Cohen [57]. To determine the discriminatory power of the P-AES, an independent sample student t-test was conducted to compare the mean scores of patients admitted voluntarily and involuntarily. All statistical analyses were conducted using R software [58] with its *mirt* package being used to fit the Item Response Theory (IRT) models [59]. Two-tailed statistical tests were utilized, and the significance level was established at $\alpha = 0.05$.

3. Results

3.1. Characteristics of study participants

A total of 208 patients from the five psychiatric departments were recruited to complete the survey. The required socio-demographic and clinical information was completed for 152 patients. The mean age of the sample was 46.53 (SD = 16.85), and more than half (54.61 %, n = 83) of them were female. A significant proportion of the participants were not employed (48.68 %, n = 74), and approximately one-fifth of them were retired (22.37 %, n = 34). Most of the participants were single (53.95 %, n = 82), and a small portion of them were immigrants in Portugal (12.5 %, n = 19). Additionally, about one-third of the sample had an involuntary admission legal status (32.24 %, n = 49). The most prevalent diagnosis among the participants were psychotic disorders (44.08 %, n = 67).

3.2. Psychometric properties of the P-AES

The measures of internal consistency and test-retest reliability were found to be satisfactory, except for the Voice domain which has a Cronbach alpha of 0.654. The intraclass coefficients showed moderate levels of reliability overall, except the Voice domain, which exhibited poor reliability. Significant changes were observed between the scores obtained in the first and second assessments ($p < 0.05$) (Table 1).

Factor loadings ranged from 0.483 to 0.994 for the one-factor model, from 0.582 to 0.989 for the three-factor model, 0.392 to -0.951 in the general factor of the bifactor model, and 0.0506 to 0.851 for the specific factors of the bifactor model. The bifactor model had the best-fit properties compared to the one-factor model and the three-factor models (Table 2). The four model selection indices for the bifactor model were consistently lower than those of the other two models, with the three-factor model showing the highest values (Table 3).

As a supplementary addition to the factor loadings, the discrimination and difficulty of each item for the three models is presented using the Item Response Theory metric (Table 4).

Correlation coefficients were in the expected direction and significant for all the variables tested except the associations with the GAF which were neither significant nor in the expected direction. Differences in the scores P-AES and its subscales were significantly different between voluntary and involuntarily admitted patients $p < 0.05$ (Table 5).

4. Discussion

The main objectives of this study were twofold: first, to evaluate the psychometric properties of the Portuguese version of the Admission Experience Survey (P-AES), and second, to investigate the factor structure of the P-AES in a sample of patients admitted to five psychiatric wards across three regions of Portugal. Reliability refers to the extent of consistency in the observed scores on a specific test for a particular group of individuals [45]. Test-retest reliability allows for measurement of stability over time, internal consistency measures homogeneity across items instantaneously, while inter-rater reliability allows for measuring equivalence by comparison across different observers. Validity, which refers to how accurately the items in the questionnaire measure a concept, is often linked to reliability as well [60]. To ensure the reliability of test scores, it is advised to gather, combine, and incorporate various sources of

Table 1
Reliability of the P-AES.

Domains	Internal consistency		Test-retest reliability		Wilcoxon test
	McDonald's ω	Cronbach's α	Spearman rho	ICC (2,1) 95 % CI	V (p-value)
Perceived Coercion	0.850	0.809	0.701**	0.703 (0.591–0.747)	3167 (0.04)
Negative Pressures	0.853	0.756	0.526**	0.534 (0.41–0.636)	8136 (<0.001)
Voice	0.727	0.654	0.331**	0.286 (0.155–0.408)	8195 (<0.001)
Total Scale	0.910	0.880	0.6113**	0.539 (0.183–0.436)	7718 (<0.001)

* $p < 0.05$. ** $p < 0.01$.

Table 2
Standardized loadings of the three models.

Item*	One-factor model		Three-factor model				Bifactor model				
	F1	h2	F1	F2	F3	h2	F1	f1	f2	f3	h2
1	0.820	0.672	0.732			0.536	0.768	0.208			0.633
2	0.806	0.649		0.755		0.570	0.813		0.186		0.696
3	-0.929	0.864			0.987	0.975	-0.951			0.051	0.907
4	0.994	0.989	0.989			0.978	0.933	0.356			0.998
5	-0.750	0.563			0.766	0.586	-0.801			0.449	0.843
6	-0.647	0.419		0.926		0.858	0.640		0.583		0.750
7	0.928	0.862	0.983			0.966	0.792	0.574			0.957
8	0.680	0.463		0.789		0.623	0.684		0.370		0.605
10	0.483	0.234		0.771		0.595	0.392		0.851		0.878
11	0.512	0.262		0.685		0.470	0.470		0.543		0.516
12	0.610	0.372		0.582		0.338	0.591		0.178		0.381
13	-0.657	0.432			0.593	0.352	-0.713			-0.505	0.763
14	0.513	0.263	0.651			0.424	0.456	0.471			0.429
15	0.633	0.401	0.657			0.431	0.571	0.394			0.481

h2 = communality of the item.

Table 3
Comparisons of model fit.

Model	χ^2	df	p-value	RMSEA	90 % C.I. for RMSEA	CFI	TLI
One-factor model	205	77	<0.001	0.118	0.094-0.140	0.913	0.897
Three-factor model	282	77	0	0.149	0.126-0.171	0.860	0.835
2 PL Bifactor model	91.2	63	0.012	0.061	0.020-0.092	0.972	0.981
Direct comparison between the bifactor models with other models							
	AIC	SABIC	HQ	BIC	logLik	χ^2	p
One factor model	1804	1794	1836	1883	-874	68.786	0
Three factor model	1908	1898	1940	1987	-926	172.934	0
Bifactor model	1764	1748	1811	1881	-840		

* χ^2 = Chi-Square statistic; df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation, C.I. Confidence Interval.

Table 4
IRT parameterization for the P-AES.

Items	One-factor Model		Three-factor Model		Bifactor model (with 3 sub-factors)		
	Item difficulty	Item discrimination	Item difficulty	Item discrimination	Item difficulty	Item discrimination _(g)	Item discrimination _(s)
Item 1	-0.445	2.448	-0.375	2.070	-0.456	2.157	0.584
Item 2	0.110	2.340	0.104	1.915	0.107	2.512	0.575
Item 3	0.972	-4.410	4.498	13.424	1.233	-5.312	0.283
Item 4	-0.075	15.288	-0.037	9.646	-0.799	33.732	12.850
Item 5	0.611	-1.938	0.833	2.057	0.995	-3.436	1.926
Item 6	-1.533	1.468	-2.944	3.968	-2.393	2.177	1.985
Item 7	0.962	4.232	1.660	7.139	1.680	6.508	4.711
Item 8	-1.128	1.599	-1.311	2.135	-1.342	1.850	1.002
Item 10	-1.018	0.946	-1.415	2.074	-2.657	1.910	4.149
Item 11	-0.470	1.029	-0.545	1.619	-0.603	1.151	1.330
Item 12	0.333	1.318	0.341	1.224	0.334	1.279	0.384
Item 13	0.349	-1.518	0.439	1.289	0.554	-2.496	-1.767
Item 14	-0.468	1.015	-0.527	1.601	-0.555	1.028	1.060
Item 15	0.435	1.404	0.499	1.626	0.459	1.350	0.930

evidence [61]. In the case of the P-AES, satisfactory levels of internal consistency were observed based on both Cronbach’s alpha and McDonald’s omega, except for the voice domain, which showed marginally acceptable values. This might be because there are only three items on the voice subscale. When the number of test items is limited, it can violate the tau-equivalent assumption of the Cronbach alpha and lead to an underestimation of reliability [62].

Intraclass Correlation Coefficient (ICC) is a reliability measure that assesses both the correlation and agreement between measurements [45]. In this regard, the values were acceptable for the full scale and domains, except for the voice domain which had lower values than acceptable. The lower reliability observed in the voice domain highlights the greater risk of type 2 error in measurements of perceived procedural justice [63,64]. The low ICC observed for the voice domain has implications for the adjustment of sample size

Table 5
Convergent validity and discriminatory power of the P-AES scores.

Convergent validity						
	Coercion Ladder	p-value	GAF	p-value	CAT	p-value
Perceived coercion subscale	0.401	<0.001	0.089	0.200	-0.260	<0.001
Negative pressure subscale	0.475	<0.001	0.117	0.100	-0.342	<0.001
Voice subscale	-0.450	<0.001	-0.126	0.090	0.339	<0.001
Total scale	0.441	<0.001	0.115	0.100	-0.299	<0.001
Discriminatory power						
	Voluntary admission Mean score (SD)	Involuntary admission Mean score (SD)	2-Sample t-test	df	p-value	
Perceived coercion subscale	2.136 (1.807)	3.250 (1.643)	-4.2	125	<0.001	
Negative pressure subscale	1.750 (1.574)	3.183 (1.927)	-5	96	<0.001	
Voice subscale	1.803 (1.1151)	1.067 (0.9181)	4.8	137	<0.001	
Total scale	5.500 (2.099)	7.067 (2.246)	-4.6	107	<0.001	

and power calculations, more subjects are required for a given effect size to be statistically significant [65–67]. Significant changes were observed in the total P-AES and its domains over time. Two main factors influence the evaluation of test-retest reliability: the changing nature of the construct being measured over time and the length of the time interval between the two tests [68]. The changes observed in this study may be attributed to the variable length of time for the second assessment which was as long as 2 months in some cases. In test-retest reliability, it is assumed that the true score being measured remains consistent over a brief period [69]. It may also reflect changes in functioning and experiences during stay in the ward. Certain psychological phenomena, such as mood or perceived coercion in this context, can undergo rapid changes in a short period. Furthermore, the interpretation of test-retest reliability assumes that the two test administrations are identical and independent, and that individuals' performance remains consistent across different time points [70]. In real-life test settings, meeting these assumptions can be difficult.

In keeping with previous studies, associations were observed among the domains and the entire scale [21,33]. The perceived coercion domain was positively associated with the use of negative pressures such as threats and physical force and inversely related to the voice subscale which is the sense of procedural justice [4]. To confirm the internal validity of the P-AES, three models were evaluated, two of them having been previously discussed in earlier studies [18,21,33]. The French validation reported that while a single total coercion score sufficed, the comparison between one- and three-factor models implied that Perceived Coercion, Negative Pressures, and Voice should not be treated as identical and recommended adopting the three-factor solution [33]. In our study, the bifactor model consistently displayed the most favorable fit characteristics when compared to the one-factor and three-factor models using multiple model selection indices. The bifactor modeling approach assumes that the interrelationships among items are accounted for by both a general factor (a shared underlying variable) and specific group factors. The general factor, denoting the shared aspect of the subjective experience of coercion during admission, can be understood as the general perceived coercion, a broadly defined concept. On the other hand, the three group factors represent the distinct elements of the three domains and can be interpreted as domain-specific facets of the subjective experience of coercion, each with a more specific definition. The bifactor structure captures both the uniqueness and commonality of the subjective experience of coercion. Therefore, the observed bifactor model in this study indicates that perceived coercion, negative pressures, and voice share certain commonalities under the general subjective experience of coercion factor, while also having separate and unique components independent from the general factor. This aligns with Hodge et al.'s original concept of a core construct, where the experience of coercion is comprehensively described through other group latent factors, such as perceived fairness, relationships with individuals exerting pressure, and the aspect of voice or validation [1]. Based on this finding, it is recommended that when assessing the subjective experience of coercion, all three domains should be used, and the total score of the AES along with the scores of its domains should be reported.

There is some concern regarding Item #13 on the AES (My opinion about coming into the hospital didn't matter), as it exhibited negative item discrimination parameters in both the one-factor and bifactor models. This is problematic since it implies that respondents with higher levels of the underlying trait (perceived coercion) are less inclined to endorse a more severe response option [46]. Previous research suggests that reverse-worded items in mixed-worded scales may be problematic [71]. This can be especially so in negative-concord languages like Portuguese. The reversal of wording has been found ineffective in addressing bias and can even be counterproductive in some cases [72]. Researchers translating the instrument should reconsider reversing the wording of this item, especially in languages that use negative concord.

Correlations between the P-AES scores and its domains, with the Coercion Ladder and the CAT were largely in line with our hypotheses, suggesting that the P-AES is a valid measure of experience of coercion. The AES scale and its domains did not exhibit the anticipated associations with the GAF, and these associations were not statistically significant. This suggests that functioning, as assessed by the GAF, did not show any significant associations with perceived coercion. These results align with earlier studies reporting that perceived coercion is not strongly linked to the severity of psychiatric symptoms [37,73], but differ from the findings of the French validation study [33]. Additionally, the P-AES demonstrated good discriminatory power, as evidenced by significant mean score differences between voluntarily and involuntarily admitted patients, which supports our initial hypothesis.

4.1. Strengths and limitations

The validation study utilized the two-parameter logistic (2 PL) item response theory approach, which is deemed more suitable for the AES due to its dichotomous item format. An advantage of Item Response Theory over Classical Test Theory is that item parameter estimates remain unbiased and independent of the sample, even with unrepresentative samples [74]. Additionally, we extended the current understanding of the AES factor structure by exploring a bifactor model, a novel approach not previously explored in validation studies. While this study makes a significant contribution to the existing literature, further research is necessary to validate whether the bifactor model accurately represents the subjective experience of coercion during admission samples from other countries. Nevertheless, several methodological limitations need to be acknowledged. Firstly, convenience sampling was used, which could impact the generalizability of the findings. Secondly, to enhance accuracy, the second assessment for the test-retest should have been conducted within 48 h. Thirdly, reliance on self-report ratings for concurrent measures introduces the potential risk of artificially high correlations due to mono-method bias [75]. Moreover, it is essential to acknowledge that prior research has shown bifactor models to excel in terms of fit statistics compared to other models. Consequently, there is a chance that the bifactor model might have captured some irrelevant information and demonstrated a better fit due to an inclination to overfit the data. However, the study took steps to address this issue by using multiple model selection criteria to minimize potential biases.

5. Conclusion

The study introduces a bifactor structure to model the subjective experience of coercion during admission. While the domains of perceived coercion, negative pressure, and voice represent distinct concepts, they also share common elements as part of the broader concept of the subjective experience of coercion. This finding would be valuable in future research aimed at gaining a deeper understanding of the subjective experience of coercion. Furthermore, our findings affirm that the P-AES serves as a valid instrument for assessing the subjective experience of coercion during admission among individuals with mental health conditions in Portugal. Validating this questionnaire is a crucial initial step, enabling the testing of additional hypotheses and formulation of interventions to improve the admission experience. When combined with qualitative approaches like on-site observations and interviews, the P-AES can function as a tool for self-evaluation, encouraging more extensive conversations about optimal procedures during the admission process. It is expected to foster the advancement of future research endeavors in Portuguese-speaking countries and stimulate cross-cultural comparisons on this subject.

Ethics declarations

- This study was reviewed and approved by ethics committee of Nova medical school, with the approval number: 100/2021/CEFCM.
- All participants provided informed consent to participate in the study.

Data availability statement

The data analyzed during the current study are available from the corresponding author on reasonable request.

Funding details

The study did not receive any specific funding. Deborah Oyine Aluh is a PhD student receiving the support of a PhD fellowship from “la Caixa” Foundation (LCF/BQ/DI20/11780013). Barbara Pedrosa and Ugnė Grigaitė are PhD students receiving financial support from the FCT – Fundação para a Ciência e a Tecnologia (UI/BD/151073/2021 and UI/BD/151072/2021)

CRediT authorship contribution statement

Deborah Oyine Aluh: Conceptualization, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Sofia Azeredo-Lopes:** Formal analysis, Software, Writing – review & editing. **Barbara Pedrosa:** Methodology, Writing – review & editing. **Manuela Silva:** Data curation, Writing – review & editing. **Ugnė Grigaitė:** Writing – review & editing. **Ana Rita Martins:** Data curation, Writing – review & editing. **Maria Ferreira de Almeida Mousinho:** Data curation, Writing – review & editing. **Graça Cardoso:** Supervision, Writing – review & editing. **José Miguel Caldas-de-Almeida:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the study participants who generously shared their time for this study. We also acknowledge

Miguel Cardoso Caldas-de-Almeida and Ana Rita Dias for their help with the translation process.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e24114>.

References

- [1] S.K. Hoge, C. Lidz, E. Mulvey, L. Roth, N. Bennett, L. Siminoff, et al., Patient, family, and staff perceptions of coercion in mental hospital admission: an exploratory study, *Behav. Sci. Law* 11 (1993) 281–293, <https://doi.org/10.1002/BSL.2370110306>.
- [2] E. Verbeke, S. Vanheule, J. Cauwe, F. Truijens, B. Froyen, Coercion and power in psychiatry: a qualitative study with ex-patients, *Soc. Sci. Med.* 223 (2019) 89–96, <https://doi.org/10.1016/J.SOCSCIMED.2019.01.031>.
- [3] K. Prebble, K. Thom, E. Hudson, Service users' experiences of voluntary admission to mental hospital: a review of research literature, *Psychiatr. Psychol. Law* (2015), <https://doi.org/10.1080/13218719.2014.959156>.
- [4] M. Jaeger, W. Rossler, Enhancement of outpatient treatment adherence: patients' perceptions of coercion, fairness and effectiveness, *Psychiatr. Res.* 180 (2010) 48–53, <https://doi.org/10.1016/j.psychres.2009.09.011>.
- [5] A. Theodoridou, F. Schlatter, V. Ajdacic, W. Rössler, M. Jäger, Therapeutic relationship in the context of perceived coercion in a psychiatric population, *Psychiatr. Res.* (2012), <https://doi.org/10.1016/j.psychres.2012.04.012>.
- [6] C. Katsakou, L. Bowers, T. Amos, R. Morriss, D. Rose, T. Wykes, et al., Coercion and treatment satisfaction among involuntary patients, *Psychiatr. Serv.* (2010), <https://doi.org/10.1176/ps.2010.61.3.286>.
- [7] A. Fiorillo, D. Giacco, C. De Rosa, T. Kallert, C. Katsakou, G. Onchev, et al., Patient characteristics and symptoms associated with perceived coercion during hospital treatment, *Acta Psychiatr. Scand.* 125 (2012) 460–467, <https://doi.org/10.1111/J.1600-0447.2011.01809.X>.
- [8] L. Kjellin, K. Andersson, E. Bartholdson, I.L. Candefjord, H. Holmström, L. Jacobsson, et al., Coercion in psychiatric care - patients' and relatives' experiences from four Swedish psychiatric services, *Nord. J. Psychiatr.* 58 (2004) 153–159, <https://doi.org/10.1080/08039480410005549>.
- [9] K.W. Sørgaard, Patients' perception of coercion in acute psychiatric wards. An intervention study, *Nord. J. Psychiatr.* 58 (2004) 299–304, <https://doi.org/10.1080/08039480410005819>.
- [10] J. Bergk, E. Flammer, T. Steinert, "Coercion Experience Scale" (CES) - validation of a questionnaire on coercive measures, *BMC Psychiatr.* 10 (2010) 1–10, <https://doi.org/10.1186/1471-244X-10-5/TABLES/5>.
- [11] J. Bergk, B. Einsiedler, E. Flammer, T. Steinert, A randomized controlled comparison of seclusion and mechanical restraint in inpatient settings, *Psychiatr. Serv.* 62 (2011) 1310–1317, https://doi.org/10.1176/PS.62.11.PSS6211_1310/ASSET/IMAGES/LARGE/PSS6211_1310_FIG005.JPG.
- [12] F. Borgeat, D. Zullino, Attitudes concerning involuntary treatment of mania: results of a survey within self-help organizations, *Eur. Psychiatr.* 19 (2004) 155–158, <https://doi.org/10.1016/J.EURPSY.2003.08.002>.
- [13] J. Guzmán-Parra, C. Aguilera-Serrano, J.A. García-Sánchez, E. García-Spínola, D. Torres-Campos, J.M. Villagrán, et al., Experience coercion, post-traumatic stress, and satisfaction with treatment associated with different coercive measures during psychiatric hospitalization, *Int. J. Ment. Health Nurs.* 28 (2019) 448–456, <https://doi.org/10.1111/INM.12546>.
- [14] J. Mielau, J. Altunbay, J. Gallinat, A. Heinz, F. Bermpohl, A. Lehmann, et al., Subjective experience of coercion in psychiatric care: a study comparing the attitudes of patients and healthy volunteers towards coercive methods and their justification, *Eur. Arch. Psychiatr. Clin. Neurosci.* 266 (2016) 337–347, <https://doi.org/10.1007/S00406-015-0598-9>.
- [15] T. Steinert, M. Birk, E. Flammer, J. Bergk, Subjective distress after seclusion or mechanical restraint: one-year follow-up of a randomized controlled study, *Psychiatr. Serv.* 64 (2013) 1012–1017, <https://doi.org/10.1176/APPL.PS.201200315>.
- [16] O. Nyttिंगnes, J. Rugkåsa, A. Holmén, T. Ruud, The development, validation, and feasibility of the experienced coercion scale, *Psychol. Assess.* 29 (2017) 1210–1220, <https://doi.org/10.1037/PAS0000404>.
- [17] K.I. Iversen, G. Hoyer, H.C. Sexton, Coercion and patient satisfaction on psychiatric acute wards, *Int. J. Law Psychiatr.* 30 (2007) 504–511, <https://doi.org/10.1016/J.LJLP.2007.09.001>.
- [18] W. Gardner, S.K. Hoge, N. Bennett, L.H. Roth, C.W. Lidz, J. Monahan, et al., Two Scales for Measuring Patients' Perceptions for Coercion during Mental Hospital Admission, *Behavioral Sciences & the Law*, 1993, <https://doi.org/10.1002/bsl.2370110308>.
- [19] A. Anestis, M. Daffern, S.D.M. Thomas, T. Podubinski, Y. Hollander, S. Lee, et al., Predictors of Perceived Coercion in Patients Admitted for Psychiatric Hospitalization and the Stability of these Perceptions over Time 20 (2013) 492–503, <https://doi.org/10.1080/13218719.2012.712833>.
- [20] J. Bindman, Y. Reid, G. Szmukler, J. Tiller, G. Thornicroft, M. Leese, Perceived coercion at admission to psychiatric hospital and engagement with follow-up—a cohort study, *Soc. Psychiatr. Psychiatr. Epidemiol.* 40 (2005) 160–166, <https://doi.org/10.1007/S00127-005-0861-X>.
- [21] J.C.K. Fu, P.P.L. Chow, L.C.W. Lam, The experience of admission to psychiatric hospital among Chinese adult patients in Hong Kong, *BMC Psychiatr.* 8 (2008) 1–11, <https://doi.org/10.1186/1471-244X-8-86/TABLES/4>.
- [22] G.S. Gowda, E.O. Noorthoorn, C.N. Kumar, R.B. Nanjegowda, S.B. Math, Clinical correlates and predictors of perceived coercion among psychiatric inpatients: a prospective pilot study, *Asian J Psychiatr* 22 (2016) 34–40, <https://doi.org/10.1016/J.AJP.2016.04.004>.
- [23] W. Gardner, S.K. Hoge, N. Bennett, L.H. Roth, C.W. Lidz, J. Monahan, et al., Two scales for measuring patients' perceptions for coercion during mental hospital admission, *Behav. Sci. Law* 11 (1993) 307–321, <https://doi.org/10.1002/BSL.2370110308>.
- [24] J. Mielau, J. Altunbay, J. Gallinat, A. Heinz, F. Bermpohl, A. Lehmann, et al., Subjective experience of coercion in psychiatric care: a study comparing the attitudes of patients and healthy volunteers towards coercive methods and their justification, *Eur. Arch. Psychiatr. Clin. Neurosci.* 266 (2016) 337–347, <https://doi.org/10.1007/S00406-015-0598-9>.
- [25] B. O'Donoghue, J. Lyne, M. Hill, L. O'Rourke, S. Daly, C. Larkin, et al., Perceptions of involuntary admission and risk of subsequent readmission at one-year follow-up: the influence of insight and recovery style, *J. Ment. Health* 20 (2011) 249–259, <https://doi.org/10.3109/09638237.2011.562263>.
- [26] K. Seigel, T. Wallsten, G. Torsteinsdottir, E. Lindström, Perception of coercion: A pilot study using the Swedish version of the Admission Experience Scale 51 (2009) 49–54, <https://doi.org/10.3109/08039489709109084>.
- [27] K.M. Strack, S.E. Schulenberg, Understanding empowerment, meaning, and perceived coercion in individuals with serious mental illness, *J. Clin. Psychol.* 65 (2009) 1137–1148, <https://doi.org/10.1002/JCLP.20607>.
- [28] M. Svindseth, A. Dahl, T. Hatling, Patients' experience of humiliation in the admission process to acute psychiatric wards, *Nord. J. Psychiatr.* 61 (2007) 47–53, <https://doi.org/10.1080/08039480601129382>.
- [29] S. Klag, P. Creed, F. O'Callaghan, Development and initial validation of an instrument to measure perceived coercion to enter treatment for substance abuse, *Psychol. Addict. Behav.* 20 (2006) 463–470, <https://doi.org/10.1037/0893-164X.20.4.463>.
- [30] G. Sampogna, M. Luciano, V. Del Vecchio, B. Pocaí, C. Palumbo, G. Fico, et al., Perceived coercion among patients admitted in psychiatric wards: Italian results of the EUNOMIA study, *Front. Psychiatr.* 10 (2019) 316, <https://doi.org/10.3389/FPSYT.2019.00316>.
- [31] J.C.K. Fu, P.P.L. Chow, L.C.W. Lam, The experience of admission to psychiatric hospital among Chinese adult patients in Hong Kong, *BMC Psychiatr.* 8 (2008) 1–11, <https://doi.org/10.1186/1471-244X-8-86/TABLES/4>.

- [32] C.W. Lidz, S.K. Hoge, W. Gardner, N.S. Bennett, J. Monahan, E.P. Mulvey, et al., Perceived coercion in mental hospital admission: pressures and process, *Arch. Gen. Psychiatr.* 52 (1995) 1034–1039, <https://doi.org/10.1001/ARCHPSYC.1995.03950240052010>.
- [33] P. Golay, I. Semlali, H. Beuchat, V. Pomini, B. Silva, L. Loutrel, et al., Perceived coercion in psychiatric hospital admission: validation of the French-language version of the MacArthur Admission Experience Survey, *BMC Psychiatr.* 17 (2017) 1, <https://doi.org/10.1186/s12888-017-1519-4>.
- [34] G. Mandarelli, G. Parmigiani, F. Trobia, G. Tessari, P. Roma, M. Biondi, et al., The Admission Experience Survey Italian Version (I-AES): a factor analytic study on a sample of 156 acute psychiatric in-patients, *Int. J. Law Psychiatr.* 62 (2019) 111–116, <https://doi.org/10.1016/J.JLPL.2018.12.006>.
- [35] B. Silva, M. Bachelard, J.R. Amoussou, D. Martinez, C. Bonalumi, C. Bonsack, et al., Feeling coerced during voluntary and involuntary psychiatric hospitalisation: a review and meta-aggregation of qualitative studies, *Heliyon* 9 (2023) e13420, <https://doi.org/10.1016/J.HELIYON.2023.E13420>.
- [36] A. Theodoridou, F. Schlatter, V. Ajdacic, W. Rössler, M. Jäger, Therapeutic relationship in the context of perceived coercion in a psychiatric population, *Psychiatr. Res.* (2012), <https://doi.org/10.1016/j.psychres.2012.04.012>.
- [37] A.K. O'Callaghan, R. Plunkett, B.D. Kelly, The association between perceived coercion on admission and formal coercive practices in an inpatient psychiatric setting, *Int. J. Law Psychiatr.* 75 (2021) 101680, <https://doi.org/10.1016/J.JLPL.2021.101680>.
- [38] A. Wertheimer, A philosophical examination of coercion for mental health issues, *Behav. Sci. Law* 11 (1993) 239–258, <https://doi.org/10.1002/BSL.2370110303>.
- [39] S.P. Reise, W.E. Bonifay, M.G. Haviland, Scoring and modeling psychological measures in the presence of multidimensionality, *J. Pers. Assess.* 95 (2013) 129–140, <https://doi.org/10.1080/00223891.2012.725437>.
- [40] A.P. Association, *Diagnostic and Statistical Manual of Mental Disorders, (DSM IV), fourth ed.*, 1994.
- [41] G. Hoyer, L. Kjellin, M. Engberg, R. Kaltiala-Heino, T. Nilstun, M. Sigurjónsdóttir, et al., Paternalism and autonomy: a presentation of a Nordic study on the use of coercion in the mental health care system, *Int. J. Law Psychiatr.* 25 (2002) 93–108, [https://doi.org/10.1016/S0160-2527\(01\)00108-X](https://doi.org/10.1016/S0160-2527(01)00108-X).
- [42] S. Priebe, T. Gruyters, Patients' assessment of treatment predicting outcome, *Schizophr. Bull.* 21 (1995) 87–94, <https://doi.org/10.1093/SCHBUL/21.1.87>.
- [43] A.M. Gadermann, M. Guhn, B.D. Zumbo, Estimating ordinal reliability for likert-type and ordinal item response data: a conceptual, empirical, and practical guide, *Practical Assess. Res. Eval.* 17 (2012) 1–13.
- [44] W.R. Revelle, *Psych: Procedures for Personality and Psychological Research*, 2017.
- [45] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, *J Chiropr Med* 15 (2016) 155–163, <https://doi.org/10.1016/J.JCM.2016.02.012>.
- [46] F.M. Yang, S.T. Kao, Item response theory for measurement validity, *Shanghai Arch Psychiatry* 26 (2014) 171, <https://doi.org/10.3969/J.ISSN.1002-0829.2014.03.010>.
- [47] C.E. DeMars, A Tutorial on Interpreting Bifactor Model Scores 13 (2013), <https://doi.org/10.1080/15305058.2013.799067>, 354–78.
- [48] L. Cai, Metropolis-Hastings Robbins-Monro Algorithm for Confirmatory Item Factor Analysis 35 (2010), <https://doi.org/10.3102/1076998609353115>, 307–35.
- [49] L. Cai, High-dimensional exploratory item factor analysis by a Metropolis-Hastings Robbins-Monro algorithm, *Psychometrika* 75 (2010) 33–57, <https://doi.org/10.1007/S11336-009-9136-X/METRICS>.
- [50] L.T. Hu, P.M. Bentler, Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification, *Psychol. Methods* 3 (1998) 424–453, <https://doi.org/10.1037/1082-989X.3.4.424>.
- [51] A. Rodriguez, S.P. Reise, M.G. Haviland, Evaluating bifactor models: calculating and interpreting statistical indices, *Psychol. Methods* 21 (2016) 137–150, <https://doi.org/10.1037/MET0000045>.
- [52] M. Sellbom, A. Tellegen, Factor analysis in psychological assessment research: common pitfalls and recommendations, *Psychol. Assess.* 31 (2019) 1428–1441, <https://doi.org/10.1037/PAS0000623>.
- [53] H. Akaike, Factor analysis and AIC, *Psychometrika* 52 (1987) 317–332, <https://doi.org/10.1007/BF02294359/METRICS>.
- [54] S. G, Estimating the dimension of a model, *Ann. Stat.* 6 (1978) 461–464, <https://doi.org/10.2307/2958889>.
- [55] E.J. Hannan, B.G. Quinn, The determination of the order of an autoregression, *J. Roy. Stat. Soc. B* 41 (1979) 190–195, <https://doi.org/10.1111/J.2517-6161.1979.TB01072.X>.
- [56] S.L. Sclove, Application of model-selection criteria to some problems in multivariate analysis, *Psychometrika* 52 (1987) 333–343, <https://doi.org/10.1007/BF02294360/METRICS>.
- [57] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. *Statistical Power Analysis for the Behavioral Sciences*, 2013, <https://doi.org/10.4324/9780203771587>.
- [58] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [59] R.P. Chalmers, Mirt: a multidimensional item response theory package for the R environment, *J. Stat. Software* 48 (2012) 1–29, <https://doi.org/10.18637/JSS.V048.I06>.
- [60] R.M. Furr, V.R. Bacharach, *Psychometrics: an Introduction*, second ed., Sage, Thousand Oaks, CA, 2013.
- [61] S. Messick, Validity, in: R.L. Linn (Ed.), *Educational Measurement*, Macmillan Publishing Co, Inc; American Council on Education, 1989, pp. 13–103.
- [62] J.M. Graham, Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability 66 (2006), <https://doi.org/10.1177/0013164406288165>, 930–44.
- [63] R.A. Charter, Methodological commentary: Effect of measurement error on tests of statistical significance 19 (2008), <https://doi.org/10.1080/01688639708403872>, 458–62.
- [64] D.O. Perkins, R.J. Wyatt, J.J. Bartko, Penny-wise and pound-foolish: the impact of measurement error on sample size requirements in clinical trials, *Biol. Psychiatr.* 47 (2000) 762–766, [https://doi.org/10.1016/S0006-3223\(00\)00837-4](https://doi.org/10.1016/S0006-3223(00)00837-4).
- [65] J. Fleiss, *Design and Analysis of Clinical Experiments*, 2011.
- [66] P.E. Shrout, Measurement reliability and agreement in psychiatry 7 (1998), <https://doi.org/10.1177/096228029800700306>, 301–17.
- [67] Reliability. *Encyclopedia of Quality of Life and Well-Being Research*, Dordrecht: Springer, Dordrecht, 2014, https://doi.org/10.1007/978-94-007-0753-5_103496, 5454–5454.
- [68] Y. Shou, M. Sellbom, H.F. Chen, *Fundamentals of Measurement in Clinical Psychology*, 4, *Comprehensive Clinical Psychology*, 2022, pp. 13–35, <https://doi.org/10.1016/B978-0-12-818697-8.00110-2>. Second Edition.
- [69] R. William, D.M. Condon, Reliability. *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*, Wiley-Blackwell, 2017, pp. 709–749.
- [70] K.F. Geisinger, Reliability. *APA handbook of testing and assessment in psychology*, *Test Theory and Testing and Assessment in Industrial and Organizational Psychology* 1 (2013) 21–42, <https://doi.org/10.1037/14047-002>.
- [71] P.E. Spector, P.T. Van Katwyk, M.T. Brannick, P.Y. Chen, When two factors don't reflect two constructs: how item characteristics can produce artifactual factors, *J. Manag.* 23 (1997) 659–677, <https://doi.org/10.1177/014920639702300503>.
- [72] E. van Sonderen, R. Sanderman, J.C. Coyne, Ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain, *PLoS One* 8 (2013) 68967, <https://doi.org/10.1371/JOURNAL.PONE.0068967>.
- [73] J. Horvath, T. Steinert, S. Jaeger, Antipsychotic treatment of psychotic disorders in forensic psychiatry: patients' perception of coercion and its predictors, *Int. J. Law Psychiatr.* 57 (2018) 113–121, <https://doi.org/10.1016/J.JLPL.2018.02.004>.
- [74] S.E. Embretson, S.P. Reise, Item response theory for psychologists, *Item Response Theory for Psychologists* (2013) 1–371, <https://doi.org/10.4324/9781410605269/ITEM-RESPONSE-THEORY-SUSAN-EMBRETSON-STEVEN-REISE>.
- [75] D.T. Campbell, D.W. Fiske, Convergent and discriminant validation by the multitrait-multimethod matrix, *Psychol. Bull.* 56 (1959) 81–105, <https://doi.org/10.1037/H0046016>.