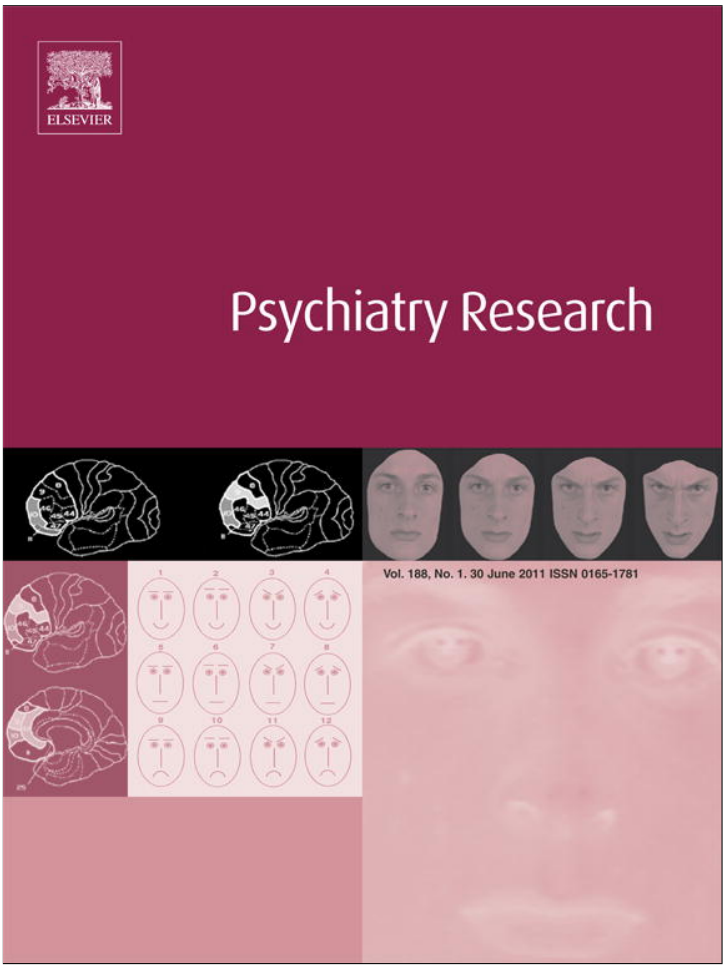

Psychiatry Research 188 (2011) 156–160

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Factorial validity and measurement equivalence of the Client Assessment of Treatment Scale for psychiatric inpatient care — A study in three European countries

Michelle Richardson [a], Christina Katsakou [a], Francisco Torres-González [b], George Onchev [c], Thomas Kallert [d], Stefan Priebe [a,*]

[a] Unit for Social and Community Psychiatry, Barts and the London School of Medicine, Queen Mary University of London, London, UK
[b] Centro de Investigación Biomedica en Red de Salud Mental (CIBERSAM), Departmental Section of Psychiatry and Psychological Medicine, University of Granada, Granada, Spain
[c] Department of Psychiatry, Medical University, Sofia, Bulgaria
[d] Park and Soteria Hospital Leipzig, Leipzig & Faculty of Medicine, Dresden University of Technology, Dresden, Germany

ARTICLE INFO

Article history:
Received 26 August 2010
Received in revised form 12 January 2011
Accepted 24 January 2011

Keywords:
CAT
Validity
Psychiatric patients

ABSTRACT

Patients' views of inpatient care need to be assessed for research and routine evaluation. For this a valid instrument is required. The Client Assessment of Treatment Scale (CAT) has been used in large scale international studies, but its psychometric properties have not been well established. The structural validity of the CAT was tested among involuntary inpatients with psychosis. Data from locations in three separate European countries (England, Spain and Bulgaria) were collected. The factorial validity was initially tested using single sample confirmatory factor analyses in each country. Subsequent multi-sample analyses were used to test for invariance of the factor loadings, and factor variances across the countries. Results provide good initial support for the factorial validity and invariance of the CAT scores. Future research is needed to cross-validate these findings and to generalise them to other countries, treatment settings, and patient populations.

1. Introduction

Treatment satisfaction is critical to treatment adherence (Chue, 2006) and among the most widely explored patient-reported outcomes (McCabe et al., 2007). Involuntary legal status has been identified as a precursor of lower treatment satisfaction (Greenwood et al., 1999) and the targeting of involuntary patients' satisfaction is an important target in mental health care but also an important ethical issue, as these patients cannot discontinue their treatment, even when they are displeased with it (Katsakou et al., 2010).

The Client Assessment of Treatment Scale (CAT; Priebe and Gruyters, 1995) provides a brief assessment of treatment satisfaction suitable for patients with severe mental health illnesses. Although originally developed in Germany, the CAT instrument has been translated into many languages and has been used in large scale European studies (e.g., Kallert et al., 2007; Priebe et al., in press). Items of the CAT were developed to assess patient's satisfaction with and perceptions of appropriateness of their treatment including different aspects of inpatient care deemed to be of clinical importance. Conceptually, items were designed to asses a single factor using 7 items (e.g., "is the treatment you are currently receiving right for you?"). Patient responses were self rated on a 100-mm-long visual analogue scale with endpoints ranging from 0 = not right at all to 1 = completely right. Each 10-mm interval was also marked thus combining the qualities of the visual analogue scale with those of an 11 point rating scale (Luria, 1975; Guyatt et al., 1987).

The CAT has been widely used and is shown to have good predictive utility. For example, research using the CAT is shown to predict global psychopathology among psychiatric patients in hospital and day hospital treatment settings (Priebe and Gruyters, 1994, 1995; Broker et al., 1995). The predictive effects of the CAT are reported to be independent of initial symptom improvement, psychiatrists' expectations of treatment success (Priebe and Gruyters, 1994) and more recently a range of potential clinical and demographic confounds including treatment setting, length of stay in hospital, gender, age, education, living status, employment status and diagnosis (Priebe et al., in press). The CAT has also been related to other outcomes including perceived coercion with higher treatment satisfaction linked to lower perceived coercion among patients with severe mental health illnesses (Katsakou et al., 2010). Additionally, Priebe et al. (2009) found higher satisfaction with treatment to be linked to lower involuntary admission rates and stronger patient reports of perceived justification of admission.

In these studies, treatment satisfaction was assessed within the first week of admission when clinicians might think that immediate patient satisfaction is not that relevant among patients compulsorily admitted. However, the findings emphasise that what patients think about their care within the first week has predictive utility even when

* Corresponding author at: Unit for Social and Community Psychiatry, Newham Centre for Mental Health, Newham, London, E13 8SP, UK. Tel.: +44 20 75406755; fax: +44 20 7540 2976.
E-mail address: s.priebe@qmul.ac.uk (S. Priebe).

0165-1781/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2011.01.020

symptom levels are often still high and the situation tense. Nonetheless, while the CAT items have good face validity and high internal consistency reliability (e.g., Priebe et al., 2009) the factorial validity of the CAT is yet to be established. Factorial or structural validity refers to the degree of fit or consistency between the a-priori conceptualisation of a construct and its statistical properties using confirmatory factor analysis. Specifically, a priori measurement models can be specified including items and their loading onto their respective conceptual dimension(s) or latent factor(s) and tested in terms of their model fit (Hoyle and Smith, 1994). Factorial invariance refers to similarity in the measurement models between groups or over time (Hoyle and smith, 1994) and is important for establishing generalisability and construct validity across settings.

The present study aims to examine the factorial validity and invariance of the CAT across samples from three European countries among involuntary patients with psychoses. Two hypotheses were tested.

i) the factorial validity of the CAT will be supported in each sample from each European country
ii) the factorial invariance of the CAT will be invariant across samples from each European country.

2. Methods

2.1. Participants

All potential participants had been admitted involuntarily to a psychiatric ward between July 2003 and October 2005 and were recruited for a larger multi-national study including 11 European countries for which detailed inclusion criteria and recruitment process have been described elsewhere (Priebe et al., 2010). Data collection for the initial study was approved by the multicentre research ethics committee and all participants gave written informed consent to take part. Compared to all patients from the initial study, ($n=2326$) only those diagnosed with schizophrenia or other psychosis, according to the ICD-10 categories (i.e., F20–29) were eligible for inclusion (1460). In terms of the n, only data from England ($n=90$), Spain ($n=128$) and Bulgaria ($n=186$) were sufficient for confirmatory factor analyses and were thus selected for inclusion. The data in England, Spain and Bulgaria were collected in samples from London, Malaga and Granada, and Sofia, respectively.

2.2. Measures

In baseline interviews, participants were asked to provide socio-demographic information including gender, age, age when leaving education, social contact (no social contacts versus some social contacts), past hospitalisation (yes versus no), employment (unemployed/pensioned versus employed/students), marital status (not married/divorced/widowed versus married), and living situation (living alone, versus living with others).

Treatment satisfaction was measured within the first week post admission using the Client Assessment of Treatment Scale (CAT; Priebe et al., 1995) which has been widely used in studies with psychiatric inpatients (Priebe et al., 1995; Kallert et al., 2007). The scale assesses patients' subjective satisfaction and perceptions of appropriateness of their treatment using 7 items (e.g., "do you believe you are receiving the right treatment/care for you here?", "does your therapist/case manager/key-worker understand you and is he/she engaged in your treatment/care", "are relations with other staff members pleasant for you?", "do you believe you are receiving the right medication for you?", "do you believe the other elements of the treatment/care here are right for you?", "do you feel respective and regarded well here?", and "has treatment/care here been helpful for you?"). Each item is rated on a visual analogue scale with 11 points with scale endpoints that ranged from 0 'not at all' to 10 'yes entirely'. Cronbach's alphas were acceptable in each country ($\alpha=0.91$, 0.79 and 0.95 respectively, in England, Spain and Bulgaria).

2.3. Analytic strategy

Hypotheses were tested in two analytic steps. First, following the procedure recommended by Byrne (1994) confirmatory factor analyses using structural equation modelling (SEM; Anderson and Gerbing, 1988) was used to assess the factorial validity of the CAT separately in each country. Second, to test for invariance between the countries multiple group comparisons were performed following the procedure recommended by Bentler (2006). The EQS 6 programme (Bentler, 2006) was used for all of the SEMs.

2.4. Model fit

Preliminary analyses revealed that Mardia's coefficient was large in each of the groups (18.72, 18.10, and 17.71 in England, Spain and Bulgaria respectively) indicating non-normality of the data. Consequently robust fit statistics and parameter estimates were used. As the chi-square goodness of fit statistic is sensitive to sample and correlation sizes (Marsh et al., 1988) additional recommended indexes for goodness of fit and cut offs (e.g., Hu and Bentler, 1999; Loehlin, 2004) were used to evaluate the adequacy of the models. Specifically, in addition to the Satorra–Bentler (S–B) corrected chi-square test statistic (Satorra and Bentler, 1994) the comparative fit index (CFI), non-normed fit index (NNFI) and the root mean square error of approximation (RMSEA) are reported. A non significant chi square value ($P>0.05$), CFI and NNFI values of 0.90 (or above) and a RMSEA of 0.10 (or lower) reflect adequate model fit. For the multi-sample analysis, Cheung and Rensfold (2002) recommend that a change of −0.01 or less in the fit indexes between baseline and subsequent nested models of invariance provide a robust test of multi-sample invariance (see too Hagger et al., 2007).

3. Results

3.1. Sample representativeness

Listwise deletion procedures were used to account for missing data. Thus, in order to assess the representativeness of our samples t test and χ^2 analyses were used to compare eligible participants in the focal countries (i.e., those meeting all of the current study's inclusion criteria) and those for whom complete data was available (i.e., responses on every item of the CAT). Table 1 reports the frequencies, means, and standard deviations for all the study measures where relevant comparing those with scores for every item of the CAT scale (Ns = 90, 128 and 186 for England, Spain and Bulgaria respectively) with those who were only present at baseline (Ns = 63, 97, and 63 in England, Spain and Bulgaria respectively). With the exception of more male patients in England χ^2 (1) = 10.03, $P<0.002$, eta = 0.256 and Bulgaria, χ^2 (1) = 14.41, $P<0.001$, eta = 0.241 and older patients in Spain, $t=2.20$ (223), $P<0.03$ there were no significant differences, indicating that the samples are fairly representative of the populations from which they were drawn.

3.2. Single confirmatory factor analyses

Parameter loadings for each separate item onto the treatment satisfaction latent factor were estimated. The Lagrange multiplier test indicated that model fit could be improved by allowing one pair of error terms to co-vary in each country. After allowing these error terms to co-vary Table 2 shows that the model fitted the data well in each country: England, S–Bχ^2 (13) = 20.107, CFI = 0.983, NNFI = 0.973, RMSEA = 0.078; Spain, S–Bχ^2 (13) = 22.544, CFI = 0.947, NNFI = 0.914, RMSEA = 0.076; Bulgaria, S–Bχ^2 (13) = 37.493, CFI = 0.980, NNFI = 0.967, RMSEA = 0.101. Table 3 presents the means, standard deviations and factor loadings in the CFAs. The factor loading of each indicator to its hypothesised latent factor was high and significant. Moreover, with exception of items 3 and 6 for Spain (factor loadings = 0.430 and 0.497, respectively, both $P<0.05$), each parameter estimated exceeded the recommended minimum of 0.50 (Ford et al., 1986) providing evidence of a stable structure in each group. Thus, in line with the first hypotheses, the factorial validity of the CAT in England, Spain and Bulgaria is supported.

3.3. Multi-sample analysis

To test for model invariance between samples, multiple group comparisons were performed. Using Vandenberg and Lance's (2000) terminology, a configural equivalence baseline model ($\nu^g = \nu^{g\prime}$) was specified that tested for factor structure equivalence across the groups. Following satisfactory fit, a more restricted tau-equivalent model, nested within the configural equivalence baseline model was tested for invariance of factor loadings ($\Lambda^g = \Lambda^{g\prime}$). Following satisfactory fit, a more restricted parallel model that included invariance of factor

Table 1
Descriptive statistics comparing patients at baseline and those with complete data.

	England baseline ($n=63$)	England ($n=90$)	Spain baseline ($n=97$)	Spain ($n=128$)	Bulgaria baseline ($n=63$)	Bulgaria ($n=186$)
Male *N%*	*40 (64%)*	*77 (86%)*	65 (67%)	92 (72%)	*13 (21%)*	*89 (47.8%)*
Age on admission	35.92	33.56	*39.03*	*35.94*	39.70	38.83
Mean (S.D.)	10.32	9.69	*10.07*	*(10.07)*	10.10	(10.20)
Ethnicity						
Age when leaving education	18.78	18.59	17.80	17.21	19.03	19.58
	4.77	4.55	(4.46)	4.75	2.77	2.80
Social contacts						
No social contacts	8 (13%)	7 (8%)	13 (13%)	15 (12%)	11 (20%)	28 (15%)
Some social contacts	53 (875)	81 (98%)	84 (87%)	113 (88%)	45 (80%)	151 (81%)
Not known		1 (2%)				7 (4%)
Past hospitalisation (inc day hosp)						
Yes	49 (77%)	68 (76%)	60 (63%)	90 (70%)	45 (78%)	141 (76%)
No	15 (15%)	22 (24%)	35 (37%)	35 (27%)	13 (22%)	38 (20%)
Not known				3 (2%)		7 (4%)
Employment						
Unemployed/pensioned	52 (84%)	77 (86%)	62 (65%)	88 (69%)	49 (86%)	163 (88%)
Employed/student	10 (16%)	11 (12%)	33 (35%)	39 (30%)	8 (14%)	15 (8%)
Not known		2 (2%)		1 (1%)		8 (4%)
Marital status						
Not married/divorced/widowed	49 (80%)	74 (82%)	79 (81%)	108 (84%)	44 (76%)	144 (77%)
Married	12 (20%)	16 (18%)	18 (19%)	20 (16%)	14 (24%)	33 (18%)
Not known						9 (5%)
Living situation						
Living alone	46 (74%)	72 (80%)	77 (79%)	107 (84%)	34 (60%)	108 (58%)
Living with others	16 (26%)	17 (19%)	20 (21%)	21 (16%)	23 (40%)	72 (39%)
Not known		1 (1%)				6 (3%)
Treatment satisfaction	4.79	5.64	6.38	6.46	4.05	4.47
Mean (S.D.)	3.21	2.98	2.12	2.57	3.24	2.56

Statistically significant findings at $P<0.05$ are shown in italics.

variances ($\Phi^{g}_{j}=\Phi^{g'}_{j}$) and factor loadings ($\Lambda^{g}=\Lambda^{g'}$) was tested. The invariance of error variances was not tested as this was considered to be a too stringent criterion (Byrne, 1994) Tables 4 and 5 present the model parameters and the goodness of fit indices for the multi-sample analyses, respectively.

The configural equivalence baseline model with no equality constraints that served as a comparison for subsequent models obtained acceptable fit statistics: S–B χ^2 (42) = 132.750, CFI = 0.945, NNFI = 0.923, RMSEA = 0.073. While the chi-square difference between the configural equivalence baseline model and the tau-equivalent model was statistically significant the fit indices indicate a change of −0.01 or less. These results indicate that the factor structure is invariant across countries enabling testing of equal factor variances. While the chi-square difference between the configural equivalence baseline model and the nested parallel model was statistically significant the NNFI and RMSEA indicate a change of −0.01 or less suggesting invariant factor structures and variances across the groups. Thus, in line with the second hypotheses, support for the invariance of the CAT across England, Spain and Bulgaria is presented.

4. Discussion

Little research has explored the psychometric properties of the CAT scale. This study examined the factorial validity and invariance of the CAT in three samples from three European countries among involuntary

Table 2
Fit indices for the single measurement models in each country.

Model	S–Bχ^2	d.f.	*n*	CFI	NNFI	RMSEA	ΔS–Bχ^2
England							
CFA	27.904	14	90	0.967	0.950	0.106	
CFA, correlated uniquenesses. [E6, E5]	20.107	13	90	0.983	0.973	0.078	7.797* (1)
Spain							
CFA	38.597	14	128	0.795	0.863	0.118	
CFA, correlated uniquenesses [E1, E4]	22.544	13	128	0.947	0.914	0.076	16.503* (1)
Bulgaria							
CFA	60.634	14	186	0.963	0.944	0.134	
CFA, correlated uniquenesses [E6, E3]	37.493	13	186	0.980	0.967	0.101	23.141* (1)

S–B = Satorra–Bentler; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root-mean-square error of approximation; E6, E5 = correlated error terms between items 6 and 5; E1, E4 = correlated error terms between items 1 and 4; E6, E3 = correlated error terms between items 6 and 3.
* $P<0.05$.

Table 3
Factor loadings of each CAT item for each country in the single measurement models.

Item	England		Spain		Bulgaria	
	Mean (S.D.)	Factor loading	Mean (S.D.)	Factor loading	Mean (S.D.)	Factor loading
1	5.27 (3.77)	0.874*	5.80 (3.49)	0.602*	4.31 (2.96)	0.952*
2	5.36 (3.89)	0.744*	7.13 (3.09)	0.639*	5.07 (2.83)	0.872*
3	6.67 (3.14)	0.763*	6.88 (2.91)	0.430*	4.45 (2.83)	0.805*
4	5.04 (4.10)	0.778*	5.55 (3.54)	0.593*	3.95 (3.12)	0.891*
5	5.63 (3.65)	0.624*	6.39 (3.34)	0.604*	4.18 (2.91)	0.836*
6	5.97 (3.52)	0.765*	7.07 (2.72)	0.497*	4.74 (2.69)	0.699*
7	5.57 (3.73)	0.819*	6.41 (3.45)	0.673*	4.48 (3.00)	0.919*

* $P<0.05$.

patients with psychotic disorders. The findings provide good initial support of the factorial validity of the CAT and its invariance across samples in England, Spain and Bulgaria, providing support for the study hypotheses.

4.1. Study strengths and limitations

The use of latent variable modelling in SEM has several advantages including the capacity to model error terms, and to test for invariance between different groups. Moreover, model fit can be assessed using a range fit indices. Nonetheless, data was only available for those patients willing to take part in academic research which may have introduced a selection bias. Moreover, although the data are from a homogenous patient population, it was collected from between 1 and 5 hospitals in specific areas of each country, thus the extent to which these findings are representative of the focal countries remains to be clarified. Also, although the sample size is impressive for this particular group of patients the large number required for confirmatory factor analyses meant that some of the fit indices may have underestimated model fit, especially in the single sample models. It is noteworthy that the factor loadings for items 3 and 6, (see Tables 3 and 4) were relatively low among the Spanish sample both in the single and multi-sample models. Finally, and perhaps most importantly, it is acknowledged that factorial structure is just one aspect of psychometric evaluation. Future research is needed to explore the item characteristic curves which imply item means, item standard deviations, higher order moments and item-total correlations by means of item-response modelling.

4.2. Error terms

Error terms are expected to be unique and uncorrelated unless there is a substantive reason for their association (Byrne, 1994). However, Bentler and Chou (1987) state that uncorrelated error terms are, in practise, unrealistic. A single pair of error terms were allowed to co-vary in the single sample models and the focal pair varied by country. In England and Bulgaria they involved items that referred to indirect or subsidiary aspects of care, such as relationships with staff while in Spain they concerned primary or more general aspects of care such as right treatment and right medication. Consequently, variance attributed to the meaning of particular items from specific cultures may have contributed to these correlations and the translation of items into different languages may have influenced their interpretation. Alternatively, it is possible that the error terms reflect a high degree of overlap between the focal items, the presence of small, unmeasured common variables or a causal relationship between the focal items (e.g., the answering of one item influences how another is answered). Nonetheless, it is noteworthy that the error terms in the more robust multi-sample model remained uncorrelated. Consequently, the presence of correlated errors in single sample models may be due to smaller sample sizes and the bias that this can have on the fit indices.

Table 4
Factor loadings and error variances of each item for each country in the multi-sample configural baseline model.

Item	England		Spain		Bulgaria	
	Factor loading	Error	Factor loading	Error	Factor loading	Error
1	0.859*	0.512	0.820*	0.573	0.942*	0.334
2	0.744*	0.669	0.599*	0.801	0.874*	0.486
3	0.774*	0.634	0.372*	0.928	0.824*	0.567
4	0.766*	0.642	0.808*	0.590	0.889*	0.457
5	0.664*	0.748	0.438*	0.899	0.845*	0.535
6	0.793*	0.609	0.370*	0.929	0.728*	0.685
7	0.813*	0.582	0.587*	0.810	0.916*	0.401

* $P<0.05$.

4.3. Research and clinical implications

The items of the CAT were shown to be highly internally reliable which is consistent with previous research. However, research to support the factorial validity of the CAT and its invariance across different samples in three European countries is new. The factor loading of each indicator to its hypothesised latent factor was high and significant, thus, based on classical test theory a meaningful unidimensional scale could be created. Thus, the research reported here seems to provide support for the CAT as a sufficiently valid brief method to assess patients' treatment in care. The CAT could be used in specific studies on inpatient care, but also in studies on other treatment methods in inpatient settings to capture and adjust for the variance of the non-specific factor and increase the specific variance.

The CAT is brief, simple, transparent and addresses issues of clinical relevance. As treatment satisfaction using the CAT is relatively easy to elicit and could be added easily to routine clinical practise these findings may have considerable practical application.

Hence, the CAT might usefully be used for identifying or flagging individuals unhappy with their treatment in cross-cultural settings. Once identified, these individuals might benefit from revised treatment plans. It might also provide a useful outcome measure for interventions designed to improve treatment.

However, future research is needed to cross-validate these findings with larger samples and in other European countries where it is used (e.g., Germany and Sweden). Similarly, in order to test the generalisability of these findings, future research is needed to replicate the current findings in samples with different diagnoses and for patients in different treatment settings. Moreover, it would be useful to examine the CATs convergent validity with other measures of treatment satisfaction and its discriminate validity with other patient reported outcomes such as therapeutic alliance (e.g., Catty et al., 2004) and quality of life (e.g., Oliver et al., 1997). Additionally, it would be useful to examine the stability of the factorial structure over time and its invariance between different genders and ethnicities. Also, as noted previously, evaluation based on item response theory is needed to explore the scaling properties of the CAT.

In summary, the CAT provides a brief one dimensional assessment of patients' treatment satisfaction that is suitable for clinicians typically faced with time pressures and patients with severe mental health

Table 5
Multi-sample analysis comparing invariance of factor loadings and variances across countries.

Invariance tests	S–Bχ^2	d.f.	*n*	CFI	NNFI	RMSEA	ΔS–Bχ^2
Configural equivalent baseline model	132.750	42	404	0.949	0.923	0.073	
Tau-equivalent model	161.972	54	404	0.939	0.929	0.071	29.222* (12)
Parallel model	173.340	56	404	0.934	0.926	0.070	11.368* (14)

S–B = Satorra–Bentler; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root-mean-square error of approximation; configural baseline model tests number of factors; tau-equivalent model tests equal factor loadings; parallel model examines equal factor loadings and factor variances.
* $P<0.05$.

illnesses who may find completing lengthy questionnaires challenging. The current study provides good initial support for the factor validity of the CAT among involuntary patients with psychotic disorders and its invariance across three samples from three European countries. Together with previous research demonstrating the CATs validity and reliability this study provides better support for the use of the CAT in multi-national research settings.

Acknowledgements

The multi-site research project European Evaluation of Coercion in Psychiatry and Harmonization of Best Clinical Practice (acronym: EUNOMIA) was funded by the European Commission (Quality of Life and Management of Living Resources Programme, contract no. QLG4-CT-2002-01036). We thank the EUNOMIA team for allowing us to use their data for secondary data analysis.

The views expressed in the publication are those of the authors and not necessarily those of the European Commission who had no further role in study design, collection, analysis and interpretation of data, in the writing of the report; and in the decision to submit the paper for publication.

References

Anderson, J.C., Gerbing, D.W., 1988. Structural equation modeling in practice: a review and recommended two-step approach. Psychological Bulletin 103, 411–423 doi:10.1037/0033-2909.103.3.411.

Bentler, P.M., 2006. EQS 6 Structural Equations Program Manual. Multivariate, Software, Inc., Encino, CA.

Bentler, P.M., Chou, C.P., 1987. Practical issues in structural modelling. Sociological Methods & Research 16, 78–117 doi:10.1177/0049124187016001004.

Broker, M., Rohricht, F., Priebe, S., 1995. Initial assessment of hospital treatment by patients with paranoid schizophrenia: a predictor of outcome. Psychiatric Research 58, 77–81 doi:10.1016/0165-1781(95)02718-C.

Byrne, B.M., 1994. Testing for the factorial validity, replication, and invariance of a measuring instrument: a paradigmatic application based on the Maslach Burnout Inventory. Multivariate Behavioral Research 29, 289–311 doi:10.1207/s15327906mbr2903_5.

Catty, J., 2004. The vehicle of success: theoretical and empirical perspectives on the therapeutic alliance in psychotherapy and psychiatry. Psychology and Psychotherapy: Theory Research and Practice 77, 255–272.

Cheung, G.W., Rensfold, R.B., 2002. Evaluating goodness of fit indexes for testing measurement invariances. Structural Equation Modeling 9, 233–255 doi:10.1207/S15328007SEM0902_5.

Chue, P., 2006. The relationship between patient satisfaction and treatment outcomes in schizophrenia. Journal of Psychopharmacology 20 (suppl. 6), 38–56 doi:10.1177/1359786806071246.

Ford, J.K., MacCallum, R.C., Tait, M., 1986. The application of exploratory factor analysis in applied psychology: a critical review and analysis. Personnel Psychology 2, 291–313 doi:10.1111/j.1744-6570.1986.tb00583.x.

Greenwood, N., Key, A., Burns, T., Bristow, M., Sedgwick, P., 1999. With in-patient psychiatric services—relationship to patient and treatment factors. The British Journal of Psychiatry 174, 159–163.

Guyatt, G.H., Townsend, M., Berman, L.B., Keller, J.L., 1987. A comparison of Likert and visual analogue scales for measuring change in function. Journal of Chronic Disease 40, 1129–1133 doi:10.1016/0021-9681(87)90080-4.

Hagger, M.S., Chatzisarantis, N.L.D., Barkoukis, V., Wang, J.C.K., Hein, V., Pihu, M., Soos, I., Karsai, I., 2007. Cross-cultural generalisazability of the theory of planned behaviour among young people in a physical activity context. Journal of Sport & Exercise Psychology 29, 1–19.

Hoyle, R.H., Smith, G.T., 1994. Formulating clinical research hypotheses as structural equation models: a conceptual overview. Journal of Consulting and Clinical Psychology 62, 429–440 doi:10.1037/0022-006X.62.3.427.

Hu, L., Bentler, P.M., 1999. Cutoff criterion for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling 6, 1–55.

Kallert, T.W., Priebe, S., McCabe, R., Kiejna, A., Rymaszewska, J., Nawka, P., Ocvar, L., Raboch, J., Starkova-Kalisova, L., Koch, R., Schützwohl, M., 2007. Are day hospitals effective for acutely ill psychiatric patients? A European multi-centre randomized controlled trial. The Journal of Clinical Psychiatry 68, 278–287.

Katsakou, C., Bowers, L., Amos, T., Morriss, R., Rose, D., Wykes, T., Priebe, S., 2010. Coercion and treatment satisfaction among involuntary patients. Psychiatric Services 61, 286–292 doi:10.1176/appi.ps.61.3.286.

Loehlin, J.C., 2004. Latent Variable Models, 4th ed. Lawrence Erlbaum Associates, Mahwah, N.J.

Luria, R.E., 1975. The validity and reliability of the visual analogue mood scale. Journal of Psychiatric Research 12, 51–57.

Marsh, H.W., Balla, J.R., McDonald, R.P., 1988. Goodness-of-fit indexes in confirmatory factor analysis: the effect of sample size. Psychological Bulletin 103, 391–410 doi:10.1037/0033-2909.103.3.391.

McCabe, R., Saidi, M., Priebe, S., 2007. Patient-reported outcomes in schizophrenia. The British Journal of Psychiatry 191 (suppl. 50), s21–s28 doi:10.1192/bip. 191.50.s21.

Oliver, J., Huxley, P., Priebe, S., Kaiser, W., 1997. Measuring the quality of life of severely mentally ill people using the Lancashire Quality of Life Profile. Social Psychiatry and Psychiatric Epidemiology 32, 76–83.

Priebe, S., Gruyters, T., 1994. Patients' and caregivers' initial assessments of day hospital treatment and course of symptoms. Comprehensive Psychiatry 35, 234–238.

Priebe, S., Gruyters, T., 1995. The importance of the first three days: predictors of treatment outcome in depressed inpatients. The British Journal of Clinical Psychology 34, 229–236.

Priebe, S., Gruyters, T., Heinze, M., Hoffmann, C., Jakel, A., 1995. Subjective criteria for evaluation of psychiatric care: methods for assessment in research and routine care. Psychiatric Praxis 22, 140–144.

Priebe, S., Katsakou, C., Amos, T., Leese, M., Morriss, R., Rose, D., Wykes, T., Yeeles, K., 2009. Patients' views and readmissions 1 year after involuntary hospitalisation. The British Journal of Psychiatry 194, 49–54 doi:10.1192/bjp.bp. 108.052266.

Priebe, S., Barnicot, K., McCabe, R., Kiejna, A., Nawka, P., Raboch, J., Schutzwohl, M., Kallert, T., in press. Patients' subjective initial response and the outcome of inpatient and day hospital treatment. European Psychiatry 2010 doi:10.1016/j.eurpsy.2010.03.009.

Priebe, S., Katsakou, C., Glockner, M., Dembinskas, A., Fiorillo, A., Karastergiou, A., Kiejna, A., Kjellin, L., Nawka, P., Onchev, G., Raboch, J., Schuetzwohl, M., Solomon, Z., Torres-González, F., 2010. Patients' views of involuntary hospital admission after 1 and 3 months: prospective study in 11 European countries. The British Journal of Psychiatry 196, 179–185 doi:10.1192/bjp.bp. 109.068916.

Satorra, A., Bentler, P.M., 1994. Corrections to test statistics and standard errors in covariance structure analysis. In: von Eye, A., Clogg, C.C. (Eds.), Latent Variables Analysis: Applications for Developmental Research. Sage, Thousand Oaks, CA, pp. 399–419.

Vandenberg, R.J., Lance, C.E., 2000. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Organizational Research Methods 3, 4–70 doi:10.1177/109442810031002.