BASIC RESEARCH ARTICLE

𝗮 OPEN ACCESS Check for updates

Translation and validation of the Chinese ICD-11 International Trauma Questionnaire (ITQ) for the Assessment of Posttraumatic Stress Disorder (PTSD) and Complex PTSD (CPTSD)

Grace W. K. Ho [a], Thanos Karatzias [b,c], Marylene Cloitre[d,e], Athena C. Y. Chan [a], Daniel Bressington [a], Wai Tong Chien [f], Philip Hyland [g] and Mark Shevlin [h]

aSchool of Nursing, The Hong Kong Polytechnic University, Hong Kong; bSchool of Health & Social Care, Edinburgh Napier University, Edinburgh, UK; cRivers Centre for Traumatic Stress, NHS Lothian, Edinburgh, UK; dSchool of Medicine, New York University, New York, NY, USA; eNational Center for PTSD, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, USA; fThe Nethersole School of Nursing, Chinese University of Hong Kong, Hong Kong; gDepartment of Psychology, Maynooth University, Ireland; hSchool of Psychology, Ulster University, Derry, UK

ABSTRACT

Background: Two stress-related disorders have been proposed for inclusion in the revised ICD-11: Posttraumatic Stress Disorder (PTSD) and Complex PTSD (CPTSD). The International Trauma Questionnaire (ITQ) is a bespoke measure of PTSD and CPTSD and has been widely used in English-speaking countries.

Objective: The primary aim of this study was to develop a Chinese version of the ITQ and assess its content, construct, and concurrent validity.

Methods: Six mental health practitioners and experts rated the Chinese translated and back-translated items to assess content validity. A sample of 423 Chinese young adults completed the ITQ, the WHO Adverse Childhood Experiences International Questionnaire, and the Hospital Anxiety and Depression Scale. Among them, 31 participants also completed the English and Chinese versions of the ITQ administered in random order at retest. Four alternative confirmatory factor analysis models were tested using data from participants who reported at least one adverse childhood experience (ACE; N = 314).

Results: The Chinese ITQ received excellent ratings on relevance and appropriateness. Test–retest reliability and semantic equivalence across English and Chinese versions were acceptable. The correlated first-order six-factor model and a second-order two-factor (PTSD and DSO) both provided an acceptable model fit. The six ITQ symptoms clusters were all significantly correlated with anxiety, depression, and the number of ACEs.

Conclusions: The Chinese ITQ generates scores with acceptable psychometric properties and provides evidence for including PTSD and CPTSD as separate diagnoses in ICD-11.

Traducción y validación de la versión China del Cuestionario Internacional de Trauma ICD-11 (CIT) para la Evaluación del Trastorno de Estrés Postraumático (TEPT) y TEPT Complejo (TEPTC)ABSTRACT

Antecedentes: Dos trastornos relacionados con estrés han sido propuestos para su inclusión en la CIE-11 revisada: Trastorno de estrés postraumático (TEPT) y TEPT complejo (TEPTC). El Cuestionario Internacional de Trauma (CIT) esta diseñado para medir TEPT y TEPTC, y ha sido utilizado ampliamente en los países de habla inglesa. Objetivo. El objetivo principal de este estudio fue desarrollar una versión china del CIT y evaluar su validez concurrente, de contenido y constructo.

Método: Seis profesionales de salud mental y expertos calificaron los items traducidos al chino y retraducidos para evaluar la validez de contenido. Una muestra de 423 adultos jóvenes chinos completó el CIT, el Cuestionario Internacional de Experiencias Adversas en la Infancia de la OMS y la Escala de Ansiedad y Depresión Hospitalaria. Entre ellos, 31 participantes también completaron las versiones en inglés y chino del CIT administrado en orden aleatorio al realizar la prueba nuevamente. Se probaron cuatro modelos confirmatorios de análisis factorial utilizando datos de los participantes que informaron al menos una experiencia adversa en la infancia (EAI; N = 314).

Resultados: El CIT chino recibió calificaciones excelentes en relevancia y pertinencia. La fiabilidad test-retest y la equivalencia semántica entre las versiones en inglés y chino fue aceptable. El modelo de seis factores de primer orden correlacionado y uno de dos factores de segundo orden (TEPT y DSO) proporcionaron un ajuste aceptable. Los seis grupos de síntomas de CIT se correlacionaron significativamente con ansiedad, depresión y número de EAI.

Conclusiones: El CIT chino genera cifras con propiedades psicométricas aceptables y proporciona evidencia para incluir el TEPT y el TEPTC como diagnósticos separados en la CIE-11.

CONTACT Grace W. K. Ho ✉ grace.wk.ho@polyu.edu.hk 🏛 The Hong Kong Polytechnic University, PQ426 Hung Hom, Hong Kong

评估创伤后应激障碍（PTSD) 和复杂PTSD (CPTSD) 的中文版ICD-11国际创伤问卷 (ITQ) 的翻译和验证

背景：两种与压力有关的疾病包括在修订的ICD-11中：创伤后应激障碍 (PTSD) 和复杂PTSD (CPTSD)。国际创伤问卷 (ITQ) 是PTSD和CPTSD的衡量标准，并已广泛应用于英语国家。

目的：本研究的主要目的是开发中文版的ITQ并评估其内容、结构和共时效度。

方法：六名心理健康从业者和专家对中文翻译和反译题目的内容效度进行了评估。423名中国年轻成人样本完成了ITQ，WHO负性童年经历国际问卷，以及医院焦虑和抑郁量表。其中，31名被试在重测中以随机排序完成了ITQ的英文和中文版本。根据报告了至少一种负性童年经历的数据 (ACE; N = 314)。考察了四种备选的验证性因子分析模型。

结果：中文ITQ在相关性和适当性方面获得了极好的评价。中英文版本的重测信度和语义等效性是可以接受的。相关的一阶六因子模型和二阶二因子 (PTSD和DSO) 都提供了可接受的模型拟合。六个ITQ症状簇都与焦虑、抑郁和ACE数量显著相关。

结论：中文ITQ生成具有可接受的心理测量属性的分数，并提供将PTSD和CPTSD纳入ICD-11中的进行单独诊断的支持证据。

1. Introduction

The 11th revision to the World Health Organisation's *International Classification of Diseases* (ICD-11) was recently published in June 2018, and proposes two distinct but related disorders, Posttraumatic Stress Disorder (PTSD) and Complex PTSD (CPTSD), under new grouping of '*Disorders* specifically *associated with stress*' (Maercker et al., 2013). The ICD-11 proposes the inclusion of PTSD symptoms reflecting three symptom clusters that arise as a result of trauma exposure (First, Reed, Hyman, & Saxena, 2015), namely: (1) re-experiencing of the trauma in the here and now (Re), (2) avoidance of traumatic reminders (Av), and (3) a persistent sense of current threat that is manifested by arousal and hypervigilance (Th). This proposed three-factor structure of ICD-11 PTSD (Re, Av, Th) is well supported in prior research (Forbes et al., 2015; Gluck, Knefel, Tran, & Lueger-Schuster, 2016; Hansen, Hyland, Armour, Shevlin, & Elklit, 2015; Tay, Rees, Chen, Kareth, & Silove, 2015). In addition to these core PTSD symptoms, the ICD-11 proposes additional symptoms that reflect 'disturbances in self-organisation' (DSO) in its diagnostic formulation for CPTSD. The DSO symptoms are represented by three symptom clusters: (1) affective dysregulation (AD), (2) negative self-concept (NSC), and (3) disturbances in relationships (DR), which are frequently associated with sustained, repeated, and multiple forms of traumatic exposures (e.g. genocide campaigns, childhood sexual abuse, child soldiering, severe domestic violence, torture, or slavery). The second-order factorial structure of CPTSD is also well established in the literature (Hyland et al., 2017).

The qualitative distinction between PTSD and CPTSD, where PTSD is essentially conceptualised as a fear condition and CPTSD includes additional features of DSO as result of trauma (Cloitre, Garvert, Brewin, Bryant, & Maercker, 2013), has been supported among different trauma samples (Cloitre et al., 2013; Elklit, Christiansen, Palic, Karsberg, & Eriksen, 2014; Knefel, Garvert, Cloitre, & Lueger-Schuster, 2015; Perkonigg et al., 2016). An important limitation with these studies is that they have been based on archival data gathered using measures not specifically designed to capture the content of the ICD-11 diagnoses of PTSD and CPTSD. More recently, the *International Trauma Questionnaire* (ITQ) (Cloitre et al., 2009) was developed to generate a self-report measure of the ICD-11 PTSD and CPTSD diagnoses. The most recent version of the ITQ includes 18 items that reflect the final composition of symptoms specified for ICD-11 PTSD and CPTSD; 12 of which measure the core symptoms of PTSD and CPTSD. In its current form, six items are included to represent the three PTSD symptom clusters: Re (items Re1-Re2), Av (items Av1-Av2), and Th (items Th1-Th2). Separately, six items are included to represent the three DSO clusters that make up the symptoms of CPTSD; two items measure the AD cluster, which encompass symptoms of hyper- and hypo-activation (items AD1-AD2), two items measure NSC (items NSC1-NSC2), and two items measure DR (items DR1-DR2). The remaining 6 items measure impairments in functioning.

Although the ITQ has been developed and validated in English-speaking samples (e.g. Karatzias et al., 2017, Hyland et al., 2017), it has not been translated or validated for use in Asian countries before. Further, no known study has examined the test–retest reliability of the ITQ. This study aimed to: (1) translate the ITQ into Chinese; (2) test the Chinese-translated items for content validity; (3) assess the test–retest reliability and the construct validity of the Chinese ITQ; and (4) examine the concurrent validity of the ITQ by testing its correlations with related criterion constructs (i.e. depression, anxiety, and exposure to childhood adversities). The overall goal is to provide a Chinese translation and initial validation of the ITQ using a non-clinical young adult sample to inform future research to widen its scope of use in Asian countries.

2. Materials and methods

2.1. Phase 1: translation and content validation of ITQ

The ITQ was translated and back-translated using the process suggested by Beaton, Bombardier, Guillemin, and Ferraz (2000); all items were translated from English to traditional Chinese by a bilingual technical writer, then back-translated by a bilingual study team member. Three other experts in mental health independently reviewed the initial forward and backward translations, and provided comments and wording suggestions for revision; two additional iterations of the translations were reviewed before a consensus was reached for the initial draft.

The Chinese-translated items were tested for content validity based on their relevance to the construct and their appropriateness in the Chinese culture (Polit & Beck, 2006; Polit, Beck, & Owen, 2007). An expert panel of two clinical psychologists, two mental health nurse researchers, and two social workers were invited to rate the relevance and appropriateness of each translated question on a 4-point Likert scale – *'highly relevant/appropriate' (4)*, *'quite relevant/appropriate' (3)*, *'somewhat relevant/appropriate' (2)*, and *'not relevant/appropriate' (1)*. Content validity indices were computed by item and for the overall scale to assess relevance and appropriateness. Item-level analysis was conducted using item content validity index (I-CVI), which is the proportion of experts who rated the item with a score of 3 or 4 (out of 4) (Polit & Beck, 2006). A modified kappa statistic (k*) was computed to correct for the chance agreement among experts that might artificially inflate the I-CVI ratings (Polit et al., 2007). Content validity index for the overall scale (S-CVI) was computed using an average I-CVI of all scale items (S-CVI$_{Ave}$) (Polit & Beck, 2006). For a panel of six raters, I-CVI and S-CVI$_{Ave}$ are considered good when coefficient exceeds 0.78 and 0.90, respectively (Lynn, 1986; Polit & Beck, 2006); while k* >0.74 is considered excellent (Polit et al., 2007).

All translated items of the ITQ received excellent ratings on relevance and appropriateness, with I-CVIs ranging between 0.83 and 1.0, and k* between 0.82 and 1.0. Scale-level content validity was also high for both PTSD and CPTSD subscales, with S-CVI$_{Ave}$ for relevance and appropriateness ranging between 0.92 and 1.00 for PTSD subscales and 1.00 for all CPTSD subscales. After content validation by expert panel, the Chinese ITQ was pilot tested with eight young adults recruited from a university setting; all participants gave positive comments on the clarity, understandability, and ease of answering the questions (Hinkin, 1998). The final Chinese ITQ was administered to a larger sample to assess its psychometric properties.

2.2. Phase 2: psychometric evaluation

Test–retest reliability, semantic equivalence, factorial structure, and concurrent validity of the ITQ were evaluated. Young adults between ages 18 and 24, who could read English and traditional Chinese, and were enrolled in an undergraduate degree program in Hong Kong were eligible to participate. Participants were recruited via convenience sampling from two major universities and their affiliated community colleges using flyers circulated around college campuses between April to June of 2017. MySurvey v1.1. (The Hong Kong Polytechnic University, 2016) was used to collect data online. To prevent multiple responses from the same participant, the survey site precluded repeat entry from the same electronic device. Participants entered the study via a website and provided their responses anonymously. However, they may provide their contact information if they agree to be contacted again for a study follow-up or if they would like to be entered into a prize draw to win an electronic tablet. At approximately two weeks after initial survey completion, participants who agreed to be contacted again were selected at random to complete both the English and Chinese-translated versions of the ITQ via an individualised study weblink; the order of these measures was administered at random.

This study was approved by the ethics committee of the first author's institution. Details of the research study were posted on the survey landing page to inform participants of the study procedures, their rights as research participants, and potential risks. Implied consent was obtained by way of survey completion; this is common a practice to protect participant privacy and anonymity in online surveys (Jacobson, 1999).

2.3. Measurement

2.3.1. ICD-11 PTSD and CPTSD

The *International Trauma Questionnaire* (ITQ) (Cloitre et al., 2013) is a self-report measure of ICD-11 PTSD and CPTSD symptoms. This validation study evaluated the psychometric properties of 18 core items of the ITQ. A total of six PTSD core symptoms and three symptoms of functional impairment were used to assess PTSD symptomatology in the ITQ. Respondents are instructed to indicate how much they have been bothered by each of the core symptom **in the past month**, considering their most traumatic event, using a five-point Likert scale ranging from 'Not at all' (0) to 'Extremely' (4). Two symptoms reflect the 'Re-experiencing' (Re) cluster (i.e. Re1 *upsetting dreams* and Re2 *feeling the experience is happening again in the here and now*). Two

core symptoms reflect the 'Avoidance' (Av) cluster (i.e. Av1 *internal reminders* and Av2 *external reminders*). Two core symptoms reflect the 'Sense of Threat' (Th) cluster (i.e. Th1 *hypervigilance* and Th2 *exaggerated startle response*). The internal reliability (Cronbach's alpha) of the six PTSD items used for diagnostic purposes was satisfactory (α = .89), as were the reliabilities for the Re (α = .80), Av (α = .87), and Th (α = .86) clusters. Three additional items screened for functional impairment associated with these symptoms (Func1-Func3): (1) relationships and social life, (2) work or ability to work, and (3) other important aspects of life, such as parenting, school/college work, or other important activities.

To assess CPTSD symptomatology, participants are asked to respond to a set of six questions reflecting how they **typically** feel, think about themselves, and relate to others, also using a five-point Likert scale ranging from 'Not at all' (0) to 'Extremely' (4). These symptom domains collectively represent disturbances in self-organisation (DSO) that is central to CPTSD diagnosis. Two items capture the 'Affective Dysregulation' (AD) cluster; one measures hyper-activation (AD1) (i.e. *When I am upset, it takes me a long time to calm down*) and another measures hypo-activation (AD2) (i.e. *I feel numb or emotionally shut down*). Two items capture the 'Negative Self-concept' (NSC) cluster (i.e. NSC1 *I feel like a failure* and NSC2 *I feel worthless*), and two items capture the 'Disturbed Relationships' (DR) cluster (i.e. DR1 *I feel distant or cut off from people* and DR2 *I find it hard to stay emotionally close to people*). The internal reliability of the six DSO items was satisfactory (α = .90), as were the reliability estimates for the AD (α = .67), NSC (α = .94), and DR (α = .87) clusters. As with the PTSD symptoms, there are three items that screen for functional impairment associated with CPTSD symptoms (Func4-Func6).

Diagnostic criteria for PTSD requires a score of $\geq$2 ('Moderately') for at least one of two symptoms from each of the Re, Av, and Th clusters. The diagnostic criteria for CPTSD includes satisfying PTSD criteria in addition to scoring $\geq$2 ('Moderately') for at least one symptom from each of the AD, NSC, and DR clusters. Diagnosis of PTSD and CPTSD also requires the endorsement of functional impairment. Based on the ICD-11 taxonomic structure, a person may only receive a diagnosis of PTSD or CPTSD, but not both.

2.3.2. Depression and anxiety

The Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) is a 14-item, self-report measure. Seven items measure depression (α = .73) and anxiety (α = .84), respectively, and each item is scored on a four-point Likert scale (0–3). Total scale scores can be calculated where higher scores reflect greater distress, and scores of 11 and above are used to indicate clinical cases of anxiety and depression, respectively (Zigmond & Snaith, 1983). The Chinese version of the HADS was used in this study; prior evaluation demonstrated sound psychometric properties in a community sample of young people in Hong Kong (Chan, Leung, Fong, Leung, & Lee, 2010).

### 2.3.3. Exposure to adverse childhood experiences (ACE)

The WHO ACE-International Questionnaire (ACE-IQ) (World Health Organization, 2016) measures exposure to 13 categories of ACE: physical abuse; sexual abuse; emotional abuse; physical neglect; emotional neglect; domestic violence; household member with mental illness; household member who is a substance abuser; household member who was imprisoned; parenting separation or death; bullying; exposure to community violence; and exposure to collective violence. Overall exposure to ACE was dichotomized into 'Non-exposed' (i.e. no ACE) and 'Exposed' (i.e. one or more ACE); the level of exposure was calculated by summing the total number of ACEs exposed (possible range = 0–13). The internal consistency of the Chinese ACE-IQ was satisfactory in this study sample (α = 0.82). Translation and validation of the Chinese ACE-IQ were reported elsewhere (Ho, Chan, Chien, Bressington, & Karatzias, 2019).

2.4. Statistical analysis

The initial dataset included 423 respondents who completed the ITQ. Test-rest reliability and semantic equivalence between English and Chinese versions of the ITQ from 31 participants were assessed using STATA SE14 (StataCorp, 2015) under three criteria. First, percentage agreement (PA) of paired responses was calculated; PA $\geq$ 70% is generally considered satisfactory (Kazdin, 1977). Second, weighted Kappa coefficients (Cohen, 1960) using quadratic weights were computed to assess item agreement between test-retest and across languages. Kappa coefficients were interpreted according to the following criteria to determine the strength of agreement: $\geq$0.81 almost perfect; 0.61–0.80 substantial; 0.41–0.60 moderate; 0.21–0.40 fair; $\leq$0.20 slight/poor (Richard & Koch, 1977). Third, correlations between symptom scores for each of six symptom clusters were examined using Spearman's *Rho* due to small sample size.

The latent structure of the ITQ was tested using confirmatory factor analysis (CFA) based on responses to the 12 core symptom items using data from participants who had endorsed one or more ACEs (N = 314). Brewin et al. (2017) described the three-factor analytic model that can be most directly derived from the ICD-11 description of CPTSD. These, along with a baseline comparison model
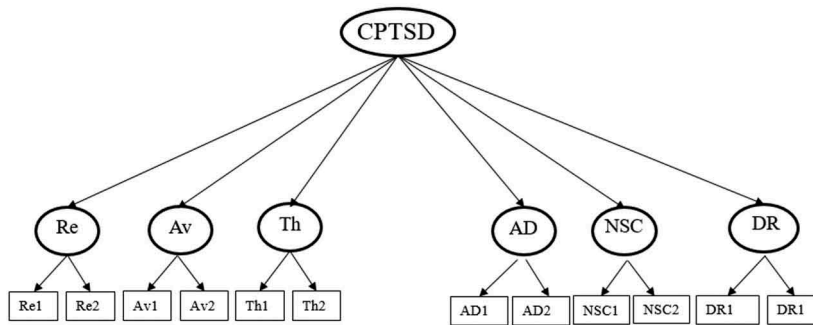
Model 1: Unidimensional CPTSD
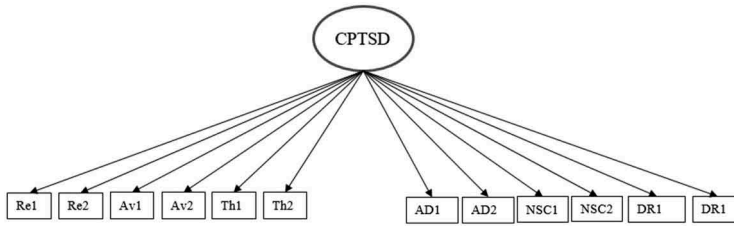
Model 2: Six Factor First-Order Model of CPTSD
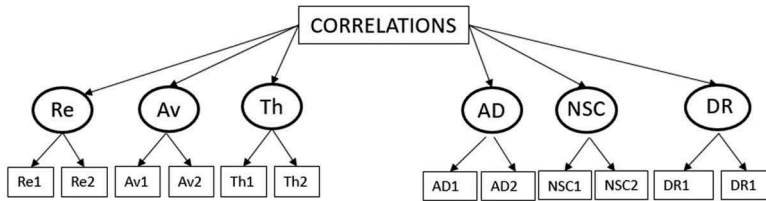
Model 3: Single-Factor Second-Order with Six First Order Factors

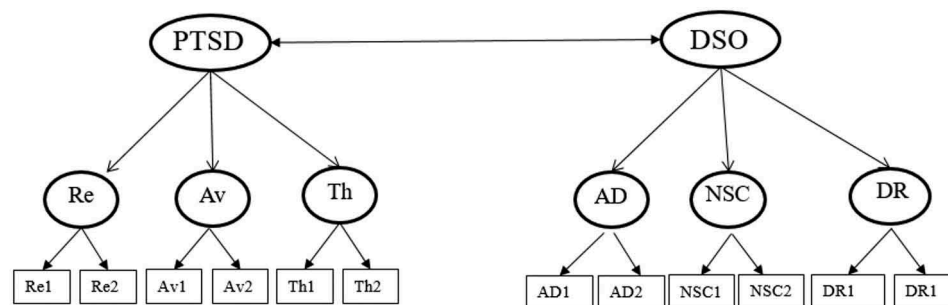Model 4: Two-Factor Second-Order Model, Each Measured by Three First-Order Factors

Figure 1. Loading patterns of models in confirmatory factor analysis.

(Model 1), were specified and tested as representations of CPTSD (see Figure 1). Model 1 is a one-factor model where all symptoms load on a single latent variable representing CPTSD. Model 2 is a correlated six-factor model. This model this based on the ICD-11 specification of three PTSD and three DSO symptom clusters each measured by their respective indicators. Model 3 replaced the factor correlations in Model 2 with a single second-order factor representing CPTSD. This model proposes that there is no distinction between PTSD and DSO at the second-order level. Model 4 specified two correlated second-order factors (PTSD and DSO) to

explain the covariation among the six first-order factors, with Re, Av and Th loading on the PTSD factor and AD, NSC and, DR loading on the DSO factor. For all models the error variances were uncorrelated.

Each model was estimated using robust maximum likelihood estimation (MLR) (Yuan & Bentler, 2000), which has been shown to produce correct parameter estimates, standard errors and test statistics (Rhemtulla, Brosseau-Liard, & Savalei, 2012) using Mplus 7.0 (Muthén 2012). Model fit was assessed using standard procedures: a non-significant chi-square ($\chi2$) test; Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values greater than .90; Root-Mean-Square Error of Approximation with 90% confidence intervals (RMSEA 90% CI); and Standardised Root-Mean-Square Residual (SRMR) values of .08 or less reflect acceptable model fit. Furthermore, the Bayesian Information Criterion (BIC) was used to evaluate alternative models, with the smaller value in each case indicating the best fitting model. Not all models were hierarchically nested so chi-square difference tests were not appropriate for all comparisons. Therefore, the Bayesian Information Criterion (BIC) was used as the main index for model comparison with the smaller value in each case indicating the best fitting model. A desirable property of the BIC is that it penalises less parsimonious models and is therefore suitable for comparing non-nested models that differ in complexity. A difference greater than 10 is considered to be indicative of a 'significant' difference (Raftery, 1995). Lastly, the concurrent validity of the best fitting model was further examined by calculating the correlations between latent factors with the total number of ACEs and two criterion variables – depression and anxiety.

3. Results

A total of 423 participants completed the Chinese version of the ITQ online. Among them, 31 were selected at random to complete a retest on Chinese and English versions of the ITQ. For the full sample (n = 423), the mean age of the participants was 20.17 (SD = 1.66). Over half of the participants were female (58.63%). The total number of ACEs reported ranged from 0 to 9, with almost three quarters (n = 314) of the participants reported exposure to at least one ACE (1 ACE = 28.6%, 2 ACEs = 14.7%, 3 ACEs = 12.5%, 4 ACEs = 11.6%, 5 or more ACEs = 6.9%). The group who reported 1 or more ACEs did not differ significantly to the non-ACE group in terms of age (t(421) = 1.130, p = .259) or gender (χ^2(1) = .185, p = .667). Rates for probable PTSD and CPTSD in the sample that reported at least one ACE was 4.14% (n = 13) and 7.64% (n = 24), respectively. For participants who

also completed the retest, their average age was 20.84 (SD = 1.72) and most were female (77.4%). Time between test and retest ranged between 14 and 31 days (M = 22.2, SD = 6.2).

3.1. Reliability and semantic equivalence

Findings on test–retest reliability and semantic equivalence of the ITQ are summarised in Table 1. Absolute percentage agreement (PA) for test-retest of the Chinese ITQ by item ranged from 35% to 74%. PA for PTSD items ranged from 42% to 68%; PA for DSO items ranged from 35% to 74%. Significant kappa coefficients for test-retest ranged from fair (0.29) to strong (0.81); one item did not have significant correlation at test and retest (Th1; K = 0.24, p = 0.087). Retest reliability by symptom clusters was generally good, with lowest r_s = 0.40 for the Re-experiencing subscale.

For semantic equivalence between Chinese and English versions of ITQ, an absolute agreement between responses by item across languages ranged between 59% and 84%. Weighted kappa coefficients for PTSD items ranged from 0.42 to 0.75, indicating moderate to strong item agreement across languages. All CPTSD items also had weighted Kappa coefficients above 0.51, indicating moderate to the strong agreement between languages. Moderate to strong correlations for symptom cluster scores across languages were also observed (r_s = 0.51–0.94).

Table 1. Test–retest reliability and semantic equivalence of ITQ by item and symptom cluster (n = 31).

Item	Test-Retest Reliability			Semantic Equivalence		
	PA (%)	K	r_s	PA (%)	K	r_s
PTSD						
Re1	52	0.36	0.40	59	0.52	0.51
Re2	65	0.46		59	0.58	
Av1	42	0.29	0.55	63	0.71	0.77
Av2	48	0.41		63	0.75	
Th1	61	0.24*	0.60	72	0.72	0.62
Th2	55	0.71		66	0.42	
Func1	55	0.30	–	63	0.43	–
Func2	65	0.60		72	0.68	
Func3	68	0.49		72	0.59	
CPTSD						
AD1	35	0.47	0.62	75	0.79	0.86
AD2	45	0.71		84	0.69	
NSC1	48	0.67	0.70	81	0.87	0.94
NSC2	52	0.62		66	0.87	
DR1	58	0.53	0.75	66	0.76	0.85
DR2	61	0.81		81	0.86	
Func4	58	0.69	–	59	0.75	–
Func5	74	0.67		84	0.86	
Func6	74	0.57		72	0.74	

PA = percentage agreement; K = weighted Kappa coefficient; r_s = Spearman's rho for symptom cluster score; All K and r_s were significant at p < 0.05 unless noted with *.

3.2. Construct validity

Results of the CFA based on participants who endorsed one or more adverse childhood experiences (N = 314) showed that the models with six correlated first-order factors (Model 2) and two correlated second-order factors (Model 4) were acceptable (see Table 2). Although the chi-square statistics were statistically significant, this should not lead to the rejection of the models as the large sample size increased the power of the test (Tanaka, 1987). Comparisons across model fit indices indicate Model 2, the first-order correlated 6-factor model, to be the best fitting solution given highest CFI and TLI, and lowest RMSEA, SRMR, and BIC. However, it should be noted that the differences in the BIC values for Models 2 and 4 did not exceed 10. For Model 2, all items loaded significantly positively onto factors representative of their respective symptom cluster (see Table 3). Significant correlations were found between all factors, ranging between $r = 0.386$ (Re and NSC) to $r = 0.868$ (AD and DR) (see Table 4).

3.3. Concurrent validity

For participants who endorsed one or more adverse childhood experiences, there was a positive correlation between the number of ACEs reported and total scores for PTSD ($r = .346$) and DSO ($r = .384$). Each of the six PTSD/DSO symptom clusters, and the total PTSD and DSO scores, correlated significantly positively with the two criterion variables, showing a weak to moderate correlation with anxiety and depression (see Table 5). Overall, correlations with PTSD/DSO symptom cluster scores were higher for depression ($r = .398-.556$) compared to anxiety ($r = .306-.519$), and depression correlated more highly than anxiety with the PTSD and DSO total scores.

4. Discussion

The present study provides the first Chinese translation and psychometric evaluation of the ITQ using a non-clinical student sample from Hong Kong. This study also produced novel evidence on test–retest reliability of the ITQ and its relation with exposure to ACEs. The overall findings show that the Chinese ITQ has sufficient scale reliability and validity, and good content validity and semantic equivalence with the original English version.

The test–retest reliability of the Chinese ITQ by symptom cluster scores was acceptable, especially for DSO symptoms. However, an absolute agreement between test-retest was not satisfactory (most below 70%), but this may be due to the long retest interval (i.e. average of 22 days). The retest reliability of the PTSD symptom subscales was also less stable than DSO symptoms, presumably because PTSD asks about 'past month,' whereas DSO refers to how one 'typically feels.' More research on the test–retest reliability of the ITQ using a clinical sample is needed to further establish the stability of the measure over time. Similarly, semantic equivalence by symptom cluster was generally acceptable, but further refinement of specific items, particularly those pertaining to Re cluster, is warranted.

The CFA findings indicated that a six-correlated first-factor model best represented the latent structure of the Chinese ITQ (Model 2). However, consistent with previous findings (Hyland et al., 2017; Karatzias et al., 2017), a two-factor second-order model that reflects ICD-11's distinction between PTSD and DSO symptomatology was also acceptable (Model 4). In fact, Model 2 was only marginally superior to Model 4 based on fit indices, and the difference in the BIC between the models did not exceed 10. There are two possible explanations for these findings. First, the low rates of PTSD and CPTSD found in the non-clinical sample of young adults precluded generating more unique differentiations between PTSD and DSO symptom clusters. This is consistent with previous research showing that the second-order model fitted better than the first-order model in clinical (Cloitre et al., 2018) and highly traumatised samples, such as refugees (Vallières et al., 2018), and the first-order model fitted better in population studies (Ben-Ezra et al., 2018; Shevlin et al., 2017). Second, it is possible that symptoms of PTSD and CPTSD are less clearly delineated in the Chinese population. For example, AD (i.e. hyper- or hypo-activation) was found to correlate highly with PTSD symptom clusters, which suggests that AD also closely reflects PTSD, a prerequisite for CPTSD diagnosis, within the

Table 2. Model fit statistics for alternative models of ICD-11 PTSD based on the ITQ (n = 314).

Model	χ^2	df	p	CFI	TLI	RMSEA (90% CI)	SRMR	BIC
Model 1	569.413	54	.000	.690	.621	0.174 (.162 − .187)	.090	10,309.943
Model 2	63.234	39	.000	.985	.975	0.044 (.023 −.064)	.029	9647.463
Model 3	149.362	48	.000	.939	.916	0.082 (.067 − .097)	.060	9713.464
Model 4	102.125	47	.000	.967	.953	0.061 (.045 − .077)	.042	9652.513

χ^2 = Chi-square Goodness of Fit statistic; df = degrees of freedom; p = probability value; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA (90% CI) = Root-Mean-Square Error of Approximation with 90% confidence intervals; SRMR = Standardised Square Root Mean Residual; BIC = Bayesian Information Criterion.

Table 3. Standardised Factor Loadings for Model 2 of PTSD and CPTSD Symptoms.

Item	Re	Av	Th	AD	NSC	DR
Re1	.902					
Re2	.744					
Av1		.909				
Av2		.853				
Th1			.871			
Th2			.875			
AD1				.690		
AD2				.718		
NSC1					.915	
NSC2					.939	
DR1						.849
DR2						.914

All loading statistically significant (p < .05). Re1 to Th2 are the PTSD items and AD1 to DR2 are the DSO items.

Table 4. Factor Correlations for Model 2 of ITQ Symptom Clusters.

	Re	Av	Th	AD	NSC
Av	.703				
Th	.716	.716			
AD	.641	.727	.859		
NSC	.386	.534	.531	.743	
DR	.454	.597	.530	.868	.734

All correlations significant (p < .001).

Table 5. Correlations between the ITQ symptom clusters with Anxiety and Depression.

ITQ Symptom Clusters	HADS-Anxiety	HADS-Depression
Re	.306	.398
Av	.348	.404
Th	.297	.552
AD	.360	.494
NSC	.519	.556
DR	.433	.468
PTSD	.371	.526
DSO	.508	.583

All correlations significant (p < .001).

Chinese culture. Indeed, prior studies have identified emotion dysregulation as a key dimension of developmental trauma in Chinese children exposed to repeated physical and/or sexual abuse (Ma & Li, 2014). Our findings form the basis to further test the utility of the ITQ as a self-report measure of ICD-11 PTSD and CPTSD diagnoses across different Chinese populations and settings. Future research using clinical samples is required to provide clarity on the distinctiveness of PTSD and DSO symptomatology in Chinese populations.

The Chinese ITQ demonstrated good concurrent validity with two criterion variables – depression and anxiety. All six symptom clusters were positively and significantly associated with scores on the measures of depression and anxiety. There are a plethora of studies that have demonstrated that PTSD is associated, and comorbid, with many other disorders, particularly anxiety and depression (Elklit & Shevlin, 2007; Pietrzak, Goldstein, Southwick, & Grant, 2011) and this study shows that, similarly, the DSO dimensions are also related to anxiety and depression.

However, the associations with the PTSD and DSO clusters were slightly stronger for depression than anxiety, which contradicts recent research based on the DSM-5 conceptualisation of PTSD showing stronger associations with anxiety than depression (Ito, Takebayashi, Suzuki, & Horikosh, 2019; Yang et al., 2017). This finding calls for more investigations into how different diagnostic formulations of post-traumatic stress disorders may vary in relation to other external psychopathological variables, particularly across different cultural groups.

Lastly, the proportion of participants with at least one ACE meeting probable diagnosis for either PTSD or CPTSD was low (12%), which was expected given this was a general student sample. However, a higher proportion of these students met diagnostic criteria for CPTSD than PTSD. In light of the high rate of participants with multiple ACE exposure in this sample (i.e. two or more ACEs) the present finding is largely consistent with prior studies using the ITQ (Karatzias et al., 2017), where higher cumulative exposure to traumatic childhood experiences confer higher likelihood of meeting diagnostic criteria for CPTSD. Results of this study, showing a positive correlations between total number of ACEs and PTSD and DSO scores, also corroborate with existing literature where dose–response relationships between cumulative ACE exposure and negative mental health outcomes in adulthood have been reported consistently across populations (Felitti et al., 1998; Herringa et al., 2013; National Scientific Council on the Developing Child, 2010). However, the diagnostic rates of probable PTSD and CPTSD should be treated with caution as a full trauma assessment was not undertaken, and it is unclear which population the rates of probable PTSD/CPTSD represents.

Our results require replication using larger clinical or community-based representative adult samples across cultural contexts. Validation of the Chinese ITQ using simplified Chinese is also warranted to enhance its generalisability for use across different Chinese populations. Additionally, we did not include a measure of adverse life events in adulthood in the current study, and we were only able to examine relationships between PTSD/CPTSD with cumulative exposure to 13 core categories of childhood adversities. It is possible that there were other potential traumatic childhood events that were not captured in this study (e.g. major accident or illness). Nonetheless, the continued empirical support for ICD-11 CPTSD should encourage clinicians to screen for this new condition in routine clinical practice. Concerns have been expressed about the availability of two diagnostic systems (e.g. DSM and ICD), particularly for patients and carers, as it is possible that one system may be used over another for the purposes of litigation, insurance coverage, and benefit

refusal (Bisson, 2013). However, these are unintended consequences of our continuing search for the most accurate understanding of trauma-related psychopathology. Finding the most appropriate classification of traumatic distress across cultural contexts will enable the development of effective treatments for survivors of psychological trauma.

5. Conclusions

This study provided the first Chinese translation of the ITQ and demonstrated acceptable psychometric properties from a large non-clinical student sample. However, refinement of items in specific symptom clusters is warranted. Further research on correlations between PTSD and DSO symptoms are also needed in order to fully understand how complex trauma is presented in Chinese populations. Our findings support the inclusion and evaluation of PTSD and CPTSD as separate but related diagnoses in the recently published ICD-11.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Disclosure statement

No potential conflict of interest was reported by the authors.

Ethics approval and consent to participate

This study received ethics approval and consent from the Human Subjects Ethics Application Review System from The Hong Kong Polytechnic University.

Funding

ORCID

Grace W. K. Ho ⓘ http://orcid.org/0000-0003-4703-5430
Thanos Karatzias ⓘ http://orcid.org/0000-0002-3002-0630
Athena C. Y. Chan ⓘ http://orcid.org/0000-0002-2870-4582
Daniel Bressington ⓘ http://orcid.org/0000-0003-0951-2208
Wai Tong Chien ⓘ http://orcid.org/0000-0001-5321-5791
Philip Hyland ⓘ http://orcid.org/0000-0002-9574-7128
Mark Shevlin ⓘ http://orcid.org/0000-0001-6262-5223

References

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191.

Ben-Ezra, M., Karatzias, T., Hyland, P., Brewin, C. R., Cloitre, M., Bisson, J. I., … Shevlin, M. (2018). Posttraumatic stress disorder (PTSD) and complex PTSD (CPTSD) as per ICD-11 proposals: A population study in Israel. *Depression and Anxiety*, 35(3), 264–274.

Bisson, J. I. (2013). What happened to harmonization of the PTSD diagnosis? The divergence of ICD11 and DSM5. *Epidemiology and Psychiatric Sciences*, 22(3), 205–207.

Brewin, C. R., Cloitre, M., Hyland, P., Shevlin, M., Maercker, A., Bryant, R. A., … Reed, G. M. (2017). A review of current evidence regarding the ICD-11 proposals for diagnosing PTSD and complex PTSD. *Clinical Psychology Review*, 58, 1–15.

Chan, Y. F., Leung, D. Y., Fong, D. Y., Leung, C. M., & Lee, A. M. (2010). Psychometric evaluation of the hospital anxiety and depression scale in a large community sample of adolescents in Hong Kong. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 19(6), 865–873.

Cloitre, M., Garvert, D. W., Brewin, C. R., Bryant, R. A., & Maercker, A. (2013). Evidence for proposed ICD-11 PTSD and complex PTSD: A latent profile analysis. *European journal of psychotraumatology*, 4(1), 20706. doi:10.3402/ejpt.v4i0.20706

Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., … Hyland, P. (2018). The international trauma questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, 138, 536–546.

Cloitre, M., Stolbach, B. C., Herman, J. L., van der Kolk, B., Pynoos, R., Wang, J., & Petkova, E. (2009). A developmental approach to complex PTSD: Childhood and adult cumulative trauma as predictors of symptom complexity. *Journal of Traumatic Stress*, 22(5), 399–408.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.

Elklit, A., Christiansen, D. M., Palic, S., Karsberg, S., & Eriksen, S. B. (2014). Impact of traumatic events on posttraumatic stress disorder among Danish survivors of sexual abuse in childhood. *Journal of Child Sexual Abuse*, 23(8), 918–934.

Elklit, A., & Shevlin, M. (2007). The structure of PTSD symptoms: A test of alternative models using confirmatory factor analysis. *British Journal of Clinical Psychology*, 46(3), 299–313.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., … Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. ThE ADVERSE CHILDHOOD EXPERIEnces (ACE) study. *American Journal of Preventive Medicine*, 14(4), 245–258.

First, M. B., Reed, G. M., Hyman, S. E., & Saxena, S. (2015). The development of the ICD-11 clinical descriptions and diagnostic guidelines for mental and behavioural disorders. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 14(1), 82–90.

Forbes, D., Lockwood, E., Creamer, M., Bryant, R. A., McFarlane, A. C., Silove, D., … O'Donnell, M. (2015). Latent structure of the proposed ICD-11 post-traumatic stress disorder symptoms: Implications for the diagnostic algorithm. *The British Journal of Psychiatry: the Journal of Mental Science*, 206(3), 245–251.

Gluck, T. M., Knefel, M., Tran, U. S., & Lueger-Schuster, B. (2016). PTSD in ICD-10 and proposed ICD-11 in elderly with childhood trauma: Prevalence, factor structure, and symptom profiles. *European Journal of Psychotraumatology*, 7, 29700.

Hansen, M., Hyland, P., Armour, C., Shevlin, M., & Elklit, A. (2015). Less is more? Assessing the validity of the ICD-11 model of PTSD across multiple trauma samples. *European Journal of Psychotraumatology*, 6, 28766.

Herringa, R. J., Birn, R. M., Ruttle, P. L., Burghy, C. A., Stodola, D. E., Davidson, R. J., & Essex, M. J. (2013). Childhood maltreatment is associated with altered fear circuitry and increased internalizing symptoms by late adolescence. *Proceedings of the National Academy of Sciences*, 110(47), 19119–19124.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1(1), 104–121.

Ho, G. W. K., Chan, A. C. Y., Chien, W.-T., Bressington, D. T., & Karatzias, T. (2019). Examining patterns of adversity in Chinese young adults using the adverse childhood experiences—international questionnaire (ACE-IQ). *Child Abuse & Neglect*, 88, 179–188.

Hyland, P., Shevlin, M., Elklit, A., Murphy, J., Vallieres, F., Garvert, D. W., & Cloitre, M. (2017). An assessment of the construct validity of the ICD-11 proposal for complex posttraumatic stress disorder. *Psychological Trauma: Theory, Research, Practice and Policy*, 9(1), 1–9.

Ito, M, Takebayashi, Y, Suzuki, Y, & Horikosh, M. (2019). Posttraumatic stress disorder checklist for dsm-5: psychometric properties in a Japanese population. *Journal of Affective Disorders*, 247, 11-19.

Jacobson, D. (1999). Doing research in cyberspace. *Field Methods*, 11(2), 127–145.

Karatzias, T., Shevlin, M., Fyvie, C., Hyland, P., Efthymiadou, E., Wilson, D., … Cloitre, M. (2017). Evidence of distinct profiles of posttraumatic stress disorder (PTSD) and complex posttraumatic stress disorder (CPTSD) based on the new ICD-11 trauma questionnaire (ICD-TQ). *Journal of Affective Disorders*, 207, 181–187.

Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis*, 10(1), 141–150.

Knefel, M., Garvert, D. W., Cloitre, M., & Lueger-Schuster, B. (2015). Update to an evaluation of ICD-11 PTSD and complex PTSD criteria in a sample of adult survivors of childhood institutional abuse by Knefel & Lueger-Schuster (2013): A latent profile analysis. *European Journal of Psychotraumatology*, 6, 25290.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–386.

Ma, E. Y. M., & Li, F. W. S. (2014). Developmental trauma and its correlates: A study of Chinese children with repeated familial physical and sexual abuse in Hong Kong. *Journal of Traumatic Stress*, 27(4), 454–460.

Maercker, A., Brewin, C. R., Bryant, R. A., Cloitre, M., van Ommeren, M., Jones, L. M., … Reed, G. M. (2013). Diagnosis and classification of disorders specifically associated with stress: Proposals for ICD-11. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, 12(3), 198–206.

Muthén, L.K, & Muthén, B.O. (2012). Mplus user's guide (Seventh edition). Los Angeles, CA: Muthén & Muthén.

National Scientific Council on the Developing Child. (2010). *Persistent fear and anxiety can affect young children's learning and development* (Working Paper No. 9). http://www.developingchild.net

Perkonigg, A., Hofler, M., Cloitre, M., Wittchen, H. U., Trautmann, S., & Maercker, A. (2016). Evidence for two different ICD-11 posttraumatic stress disorders in a community sample of adolescents and young adults. *European Archives of Psychiatry and Clinical Neuroscience*, 266(4), 317–328.

Pietrzak, R. H., Goldstein, R. B., Southwick, S. M., & Grant, B. F. (2011). Prevalence and axis I comorbidity of full and partial posttraumatic stress disorder in the United States: Results from wave 2 of the national epidemiologic survey on alcohol and related conditions. *Journal of Anxiety Disorders*, 25(3), 456–465.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354.

Richard, L. J., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.

Shevlin, M., Hyland, P., Karatzias, T., Fyvie, C., Roberts, N., Bisson, J. I., … Cloitre, M. (2017). Alternative models of disorders of traumatic stress based on the new ICD-11 proposals. *Acta Psychiatrica Scandinavica*, 135(5), 419–428.

StataCorp. (2015). *Stata statistical software: Release 14*. College Station, TX: Author.

Tanaka, J. S. (1987). How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, 58, 134–146.

Tay, A. K., Rees, S., Chen, J., Kareth, M., & Silove, D. (2015). The structure of post-traumatic stress disorder and complex post-traumatic stress disorder amongst West Papuan refugees. *BMC Psychiatry*, 15, 111.

The Hong Kong Polytechnic University. (2016). *mySurvey (Version 1.1) [Software]*. Retrieved from https://www.polyu.edu.hk/mysurvey/

Vallières, F., Ceannt, R., Daccache, F., Abou Daher, R., Sleiman, J., Gilmore, B., … Hyland, P. (2018). ICD-11 PTSD and complex PTSD amongst Syrian refugees in Lebanon: The factor structure and the clinical utility of the international trauma questionnaire. *Acta Psychiatrica Scandinavica*, 138(6), 547–557.

World Health Organization. (2016). *Adverse childhood experiences international questionnaire (ACE-IQ)*. Retrieved from www.who.int/violence_injury_prevention/violence/activities/adverse_childhood_experiences

Yang, H., Wang, L., Cao, C., Cao, X., Fang, R., Zhang, J., & Elhai, J. D. (2017). The underlying dimensions of DSM-5 PTSD symptoms and their relations with anxiety and depression in a sample of adolescents exposed to an explosion accident. *European Journal of Psychotraumatology*, 8(1), 1272789.

Yuan, K., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, 30(1), 165–200.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370.