Child Abuse & Neglect 33 (2009) 518–523

Contents lists available at ScienceDirect

Child Abuse & Neglect

Brief communication

A validation study of the Dutch Childhood Trauma Questionnaire-Short Form: Factor structure, reliability, and known-groups validity☆

Brett D. Thombs[a,*], David P. Bernstein[b], Jill Lobbestael[b], Arnoud Arntz[b]

[a] Department of Psychiatry, McGill University – Jewish General Hospital, Montreal, Quebec, Canada
[b] Department of Clinical, Medical, and Experimental Psychology, University of Maastricht, Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 4 September 2008
Received in revised form 10 March 2009
Accepted 10 March 2009
Available online 15 September 2009

Keywords:
Child abuse
Child maltreatment
Childhood Trauma Questionnaire
Physical Abuse
Sexual Abuse
Validation studies

Introduction

The Childhood Trauma Questionnaire-Short Form (CTQ-SF) is a 28-item retrospective self-report questionnaire designed to assess five dimensions of childhood maltreatment: (1) Physical Abuse, (2) Emotional Abuse, (3) Sexual Abuse, (4) Physical Neglect, and (5) Emotional Neglect. A March 2009 MEDLINE search found 141 references to the CTQ-SF, whereas no other self-report measure of childhood maltreatment listed in a recent review of childhood trauma assessment tools (Roy & Perry, 2004) was referenced more than 15 times. The CTQ-SF (Bernstein & Fink, 1998; Bernstein et al., 2003) has been translated into German (Bader, Schafer, Schenkel, Nissen, & Schwander, 2007; Driessen et al., 2000; Driessen, Schroeder, Widmann, von Schonfeld, & Schneider, 2006; Gast, Rodewald, Nickel, & Emrich, 2001; Kersting et al., 2003; Krischer & Sevecke, 2008; Moehler, Biringen, & Poustka, 2007; Schafer et al., 2006, 2007; Spitzer, Barnow, Gau, Freyberger, & Grabe, 2008; Woller, Hartkamp, & Tress, 2007), Norwegian (Fosse & Holen, 2002, 2006, 2007), Turkish (Aslan & Alparslan, 1999; Mirsal, Kalyoncu, Pektas, Tan, & Beyazyurek, 2004; Sar, Akyuz, Kundakci, Kiziltan, & Dogan, 2004; Sar, Akyuz, Kugu, Ozturk, & Ertem-Vehid, 2006; Sar, Unal, & Ozturk, 2007; Ucok & Bikmaz, 2007), Spanish (Basurte, Diaz-Marsa, Martin, & Carrasco, 2004), French (Collin-Vezina, Cyr, Pauze, & McDuff, 2005; Perroud et al., 2008), Haitian Creole (Martsolf, 2004), Portuguese (Grassi-Oliveira, Stein, & Pezzi, 2006), Italian (Sarchiapone, Carli, Cuomo, & Roy, 2007), and Dutch (Arntz, 1999; Arntz, Dietzel, & Dreessen, 1999; Arntz, Meeren, & Wessel, 2002; Giesbrecht, Merckelbach, Kater, & Sluis, 2007; Lobbestael, Arntz, & Bernstein, in

☆ Dr. Thombs is supported by a New Investigator Award from the Canadian Institutes of Health Research and an Établissement de Jeunes Chercheurs award from the Fonds de la Recherche en Santé Québec.

* Corresponding author address: Institute of Community and Family Psychiatry, Jewish General Hospital, 4333 Cote Ste Catherine Road, Montreal, Quebec, Canada H3T 1E4.

0145-2134/$ – see front matter
doi:10.1016/j.chiabu.2009.03.001

press; Lobbestael, Arntz, Harkema-Schouten, & Bernstein, 2009; Wessel, Meeren, Peeters, Arntz, & Merckelbach, 2001), but translated versions have generally not been validated.

The Dutch CTQ-SF has been found to have good internal consistency in a small sample of 16 patients with borderline personality disorder, 12 patients with a cluster-C personality disorder, and 15 control subjects (Cronbach's alpha $\geq$ .87 for all five scales), and patients with borderline personality disorder scored higher than other subjects on all scales (Arntz et al., 1999; Arntz, 1999). In addition, scores on the five CTQ-SF scales among Dutch respondents showed good convergent validity with responses from a semi-structured interview for childhood trauma (Lobbestael, Arntz, Harkema-Schouten, & Bernstein, 2009). No studies, however, have tested the structural validity of the translated Dutch CTQ-SF. The objective of this study was to test the factor structure, internal consistency reliability, and known-groups validity of a Dutch translation of the CTQ-SF in a large sample of clinical and non-clinical respondents.

Methods

Sample selection and procedure

The CTQ-SF was administered to participants in studies related to personality disorders, the association of child maltreatment history with other psychiatric disorders, and the validation of a childhood trauma interview at psychiatric hospitals, outpatient mental health clinics, forensic clinics and prisons in the Netherlands from 2004 to 2007. For all studies, individuals had to be between 18 and 70 years of age. Exclusion criteria included alcohol or drug intoxication, an IQ of <80, and the inability to independently complete the CTQ-SF. In addition, non-patient controls were recruited via advertisements, flyers, or personal contacts through the researchers. Written informed consent was obtained from all participants after a full explanation of study procedures. Participants were compensated 10 euro for completing the procedure. The protocols for each study were approved by the Medical Ethical Committee of the Academic Hospital in Maastricht.

The CTQ-SF

There are 5 items on each scale of the CTQ-SF (Bernstein & Fink, 1998; Bernstein et al., 2003) plus an additional 3-item minimization/denial scale. Item response categories are scored from 1 to 5 and are structured to reflect the frequency of maltreatment experiences (*never true*, *rarely true*, *sometimes true*, *often true*, *very often true*). Bernstein et al. (2003) used confirmatory analysis to demonstrate the validity of the five-factor structure of the original CTQ-SF and reported good internal consistency reliability for each of the CTQ-SF scales across 4 heterogeneous clinical samples from the USA: Physical Abuse = .83 to .86, Emotional Abuse = .84 to .89, Sexual Abuse = .92 to .95; Physical Neglect = .61 to .78, and Emotional Neglect = .85 to .91. Scher, Stein, Asmundson, McCreary, and Forde (2001) also reported adequate to good internal consistency reliability in a large community sample from the USA: Physical Abuse = .69, Emotional Abuse = .83, Sexual Abuse = .94, Physical Neglect = .58, and Emotional Neglect = .85. Reports on traumatic events on the CTQ-SF scales have good stability over time and good convergent and divergent validity when trauma histories have been ascertained via other methods, such as clinical interviews or independent case records (Bernstein, Ahluvalia, Pogge, & Handelsman, 1997; Bernstein, Fink, Handelsman, & Foote, 1994; Fink, Bernstein, Handelsman, & Foote, 1995; Lipschitz, Bernstein, Winegar, & Southwick, 1999). The Dutch CTQ-SF was generated using standard translation and back-translation procedures to ensure linguistic equivalence (Brislin, 1970).

Data analyses

Patient sociodemographic and clinical characteristics were compared between clinical and non-clinical respondents. Differences between the groups were assessed using t tests for continuous data and the χ^2 statistic or Fisher's Exact test for categorical data. Confirmatory factor analysis (CFA) models were conducted with Mplus (version 3.11) (Muthén & Muthén, 1998–2004) to test the validity of the five-factor structure of the CTQ-SF and to test the fit of individual items on each factor. Item responses were ordinal Likert data, so the weighted least squares estimator with a diagonal weight matrix and robust standard errors and a mean- and variance-adjusted chi-square statistic (WLSMV) was used with delta parameterization (Muthén & Muthén, 1998–2004). Mplus procedures for estimating models that included patients with missing data were used. Modification indices were used to identify pairs of items within scales for which model fit would improve if error estimates were freed to covary, and for which there appeared to be theoretically justifiable shared method effects (McDonald & Ringo Ho, 2002). A chi-square goodness-of-fit test and 3-fit indices were used to assess model fit, including the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973), the comparative fit index (CFI; Bentler, 1990), and the root mean square error of approximation (RMSEA; Steiger, 1990). Since the chi-square test is highly sensitive to sample size and can lead to the rejection of well-fitting models, practical fit indices were emphasized (Reise, Widaman, & Pugh, 1993). Guidelines proposed by Hu and Bentler (1999) suggest that models with TLI and CFI close to .95 or higher, RMSEA close to .06 or lower and SRMR close to .08 or lower are representative of good fitting models. A CFI of .90 or above (Kline, 2005) and a RMSEA of .08 or less (Browne & Cudeck, 1993), however, are also considered to represent reasonably acceptable model fit.