

Development and Aging

Psychometric properties of the Norwegian version of the Childhood Trauma Questionnaire in high-risk groups

ANDERS DOVRAN,^{1,2} DAGFINN WINJE,¹ SIMON N. ØVERLAND,^{3,4} KYRRE BREIVIK,⁵ KJERSTI AREFJORD,¹ ANITA S. DALSBØ,⁶ METTE B. JENTOFT,² ANITA L. HANSEN^{7,8} and LEIF WAAGE⁸

¹Faculty of Psychology, Department of Clinical Psychology, University of Bergen, Christiesgt 12, 50150, Bergen, Norway

²Division of Psychiatry, Haukeland University Hospital, Pb 23, 5812, Bergen, Norway

³Faculty of Psychology, Department of Health Promotion and Development Research Centre for Health Promotion, University of Bergen, Pb 7800, Bergen, Norway

⁴Division of Mental Health, Norwegian Institute of Public Health, Kalfarveien 31, 5018, Bergen, Norway

⁵Uni Research, Uni Health, Pb 7810, Bergen, Norway

⁶Department of Child and Adolescent Psychiatry, St Olav's University Hospital, Pb, 3250, 7006, Trondheim, Norway

⁷Faculty of Psychology, Department of Psychosocial Science, University of Bergen, Pb 7800, Bergen, Norway

⁸Centre for Research and Education in Forensic Psychiatry, Haukeland University Hospital, Pb 23, 5812, Bergen, Norway

Dovran, A., Winje, D., Øverland, S. N., Breivik, K., Arefjord, K., Dalsbø, A. S., Jentoft, M. B., Hansen, A. L. & Waage L. (2013). Psychometric properties of the Norwegian version of the Childhood Trauma Questionnaire in high-risk groups. *Scandinavian Journal of Psychology* 54, 286–291.

The Childhood Trauma Questionnaire - Short Form (CTQ-SF) is widely used to measure childhood abuse of all types. In the present study, we examined the psychometric properties of the Norwegian version of the instrument. The participants constituted four subsamples ($n = 517$): substance abusers ($n = 126$), psychiatric patients ($n = 210$), prisoners ($n = 109$) and adolescents in out-of-home placements ($n = 72$). Confirmatory factor analysis revealed a reasonable fit of the data to the original five-factor structure of the CTQ-SF. Measurement invariance was found across gender and the four subsamples. It was concluded that the Norwegian version of the CTQ-SF has acceptable psychometric properties, with good reliability and satisfactory accuracy, to assess different dimensions of childhood trauma.

Key words: Childhood abuse, CTQ-SF, psychometric properties, high-risk exposure sample.

Anders Dovran, Faculty of Psychology, Department of Clinical Psychology, University of Bergen, Christiesgt 12, 50150, Bergen, Norway.
E-mail: anders.dovran@psykp.uib.no

INTRODUCTION

After traumatic events, many children and young people may experience significant and long-term psychological reactions, such as anxiety disorders, depression, eating disorders, personality disorders and posttraumatic stress symptoms (Briere, Hodges & Godbout, 2010; Pine & Cohen, 2002). Studies have shown strong associations between childhood abuse and developmental difficulties, negative psychosocial consequences and interpersonal problems (Briere & Jordan, 2009) following exposure to trauma.

A number of trauma assessment scales are currently used and available (Elhai, Gray, Kashdan & Franklin, 2005; Strand, Sarmiento & Pasquale, 2005). Brewin (2005) summed up the requirements for screening instruments of trauma and posttraumatic stress among adults as follow: they should be short and easy to administer, with the minimum number of items necessary for accurate case identification; they should not require respondents to ponder over a large number of alternative scale points; they should be written in a language that is easy to understand; their purpose should be plain and acceptable to respondents; there should be simple decision rules for determining the presence or absence of clinically meaningful trauma severity levels; and they should be applicable to populations who have a varying prevalence of posttraumatic stress symptoms and are experiencing different traumas. Ohan, Myers and Collett (2002) reviewed rating scales of trauma and effects of trauma, and concluded that

most scales suitable for younger populations are insufficiently examined in terms of psychometric properties.

In a recent systematic review of 27 studies, we found a wide range of assessment instruments, including 17 different diagnostic interviews and 44 self-report instruments. The study showed a general lack of standardization or modification of the measurements used, which makes it difficult to compare research findings across studies (Dovran, Winje, Arefjord & Haugland, 2012).

There has been a growing awareness of the high prevalence of abuse and neglect in clinical populations, yet studies still suggest that there is a major under-recognition of childhood abuse in clinical settings (Read, Hammersley & Rudegeair, 2007). As childhood abuse is a highly sensitive issue for many, some may be reluctant to reveal information about abuse and neglect through direct interviews. An assessment scale may be a useful tool for initiating and facilitating dialogue about a history of abuse without the intimacy of conversation. Because types of childhood abuse often co-occur, it is vital that the scales comprise various aspects of childhood abuse (Scher, Stein, Asmundsen, McCreary & Forde, 2001). Research indicates that the accuracy of obtaining abuse histories retrospectively is greater for specific behaviors and events than the recall of subjective experiences (Brewin, Andrews & Gotlib, 1993).

The CTQ-SF (Bernstein & Fink, 1998; Bernstein, Stein, Newcomb *et al.*, 2003) is a 28-item self-report questionnaire that

retrospectively provides screening for a history of childhood abuse and neglect. It inquires about five types of maltreatment (each assessed by five items): physical, emotional, and sexual abuse, and physical and emotional neglect, plus an additional 3-item minimization scale (Bernstein & Fink, 1998). Most of the items on the CTQ-SF inquire about specific behavioral events reflecting each type of maltreatment (e.g. "I was punished with a belt, a board, a cord, or some other hard object") and as such target specific exposures. However, some of the items ask about an evaluation of general exposure to childhood abuse (e.g. "When I was growing up, someone molested me"). Such, a mix between general and specific items is in line with recommendations on test generation (Myers & Winters, 2002). Because the items do not specify details about the perpetrator(s), the CTQ-SF provides a less intrusive screening method for childhood maltreatment. The CTQ-SF has been translated into several languages (Thombs, Bernstein, Lobbestad & Arnz, 2009), and a number of studies with different national clinical and non-clinical samples have documented good psychometric properties, including well-established factorial structure (Bernstein *et al.*, 2003; Thombs *et al.*, 2009), and measurement invariance (MI) across gender and ethnicity (Thombs, Lewis, Bernstein, Medrano & Hatch, 2007). MI pertains to the consistency of measurements across some specified group demarcation. This means that individuals from different groups who have similar experiences with childhood maltreatment will respond similarly to the individual questionnaire items about maltreatment (Thombs *et al.*, 2007).

Multiple versions of the CTQ measure are available: the original 70-item version, and two subsequent versions with 53 and 34 items, respectively. The most recent and author-recommended version, CTQ-SF, contains 28 items, and has become a leading measure in the field of childhood abuse assessment (for an overview see: Baker & Maiorino, 2010).

In Norway, early and modified versions of the CTQ-SF have been used in several studies (Fosse & Holen, 2002, 2006, 2007; Ravndal, Lauritzen, Jansson & Larsson, 2001). These studies lack information regarding the validity and psychometric properties of their translated version, and the modifications of the scale make it difficult to compare research findings across samples. The present study investigate the psychometric properties, including factor structure and measurement invariance across gender and group of the authorized Norwegian translation of the CTQ-SF (Winje, Dovran & Murison, 2004) in groups at high risk of trauma exposure, such as substance users, psychiatric patients, prison inmates and adolescents in out-of-home placements.

METHOD

Participants and procedure

After being granted authorization by the original scale developers and the copyright owners, the CTQ-SF scale was translated into Norwegian using a back-translation procedure (Sperber, 2004). Two of the authors (DW and AD) translated the English version to Norwegian. To ensure that the target version was linguistically equivalent to the source version, a bilingual psychologist translated this initial Norwegian version back into English. Errors of meaning and concept nonequivalence between the translated versions were corrected in consensus meetings.

Different subsamples were recruited from exposed groups between September 2006 and May 2011. The participants constituted four subsamples: the out-of-home placement group ($n = 72$) was comprised of adolescents currently in foster care or young adults who had previously been in foster care; the sample of persons with drug abuse and/or mental health problems ($n = 126$) comprised patients receiving out-patient or in-patient treatment for substance abuse; the sample of mental health patients comprised patients seeking or receiving out-patient or in-patient treatment for psychiatric symptoms ($n = 210$); and the prisoner sample ($n = 109$) comprised male inmates from both high- and low-security prisons. Respondents were excluded if they were intoxicated or had psychotic symptoms.

The subsamples were approached in different ways. Individuals currently or previously in foster care received an introduction letter about the study from their executive officer in the child welfare system. The patient subsamples were invited by their therapists to participate in a trauma screening that could become a relevant part of the assessment of the patient, although they were notified that their screening results would be stripped of any identifying information for research purposes. The prison subsample was invited to participate in the research project by their contact officer.

All participants received written and oral information about this "study of trauma and mental health" before they signed the consent form. The participants received no economic or other type of compensation. The study was approved by the Regional Committee for Medical and Health Research Ethics, Western Norway (REK-Vest), and by the relevant authorities representing child welfare, mental health and drug abuse institutions and prison.

The CTQ-SF

The CTQ-SF (Bernstein & Fink, 1998) is a 28-item retrospective self-report questionnaire developed to assess five types of maltreatment: (1) emotional neglect (e.g., "felt loved," "was looked out for"), (2) physical neglect (e.g., "was taken care of," "not enough to eat"), (3) emotional abuse (e.g., "called names by family," "felt hated by family"), (4) sexual abuse (e.g., "was touched sexually," "made to do sexual things") and physical abuse (e.g., "hit hard enough to leave bruises," "punished with hard objects"). Five items each assess all five types of maltreatment, and there is a 3-item minimization scale (e.g., "best family in the world"). To reflect the frequency of maltreatment experiences, the item response categories are scored from 1 to 5 (1 = never true, 2 = rarely true, 3 = sometimes true, 4 = often true, 5 = very often true). Summing the scores of the 25 items creates individual scores, which results in a total score that ranges from 25 to 125. Individual subscale scores are created by summing the scores of the 5 items in each of the subscales, resulting in subscale scores that range from 5 to 25. In addition, there are established thresholds (none, low, moderate, severe), which allows description and evaluation of the severity and frequency of the different maltreatments and combinations of them.

Data analysis

All primary data analyses were done in SPSS, version 18. To examine if the original five-factor structure could be reproduced in the translated version, we ran confirmatory factor analysis (CFA) in Mplus (version 6.0; Muthén & Muthén, 1998–2010). Except when otherwise noted, the mean and variance adjusted weighted least squares estimator (WLSMV) appropriate for categorical data (Flora & Curran, 2004) were used. Missing data was handled by the use of the default method when using this method (pairwise deletion). Within this overall framework, the five-factor model of the CTQ-SF was initially tested for the whole sample (model A). To evaluate different aspects of model fit, a number of goodness-of-fit indexes were used across our analyses, 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). The results were examined according to Hu and Bentler's (1999) suggestion that models with TLI and CFI values of

0.95 or higher represent an acceptable fit, and RMSEA values less than 0.06 are representative of a close fit of the model. Values close to 0.08 or less are also representative of moderate fit and may be considered to be reasonable (Browne & Cudeck, 1993). To test for measurement invariance across groups (gender and risk sample), we conducted multi group analyses where we compared the equality of the loadings and thresholds between groups. This was done by comparing the fit of (a) a model where the all the loadings and thresholds (except the anchors) were allowed to be freely estimated in each of the groups with (b) a model where the same the same parameters were constrained to be equal between the groups. For the assessment of differential fit the Mplus DIFFTEST (Muthén & Muthén, 1998–2010, p. 553) was used.

RESULTS

Subjects and descriptive statistics

A total of 517 participants had sufficient CTQ-SF data to be used in the confirmatory factor analysis. The total sample comprised 278 men (54%) and 239 women (46%). The total CTQ-score difference between men and women was not significant, $F(1, 509) = 1.55, p = 0.21$. Participant characteristics and maltreatment subscale group means are shown in Table 1.

Measurement invariance

Preliminary multigroup analyses (Muthén & Muthén, 1998–2010, pp. 432–435) revealed no significant sex or sample group difference in the adjusted chi-square for gender ($\chi^2 = 108.814, df = 90, p = 0.09$) or exposed group ($\chi^2 = 121.807, df = 192, p = 1.0$) when loadings and thresholds were constrained to be equal between the groups, compared to when they were free to vary. On the basis of the results from these analyses, it was therefore considered safe to conduct the remaining analyses on the total sample.

Confirmatory factor analysis

The alternative goodness-of-fit indices from the confirmatory analysis of the five-factor model of the CTQ-SF indicated an acceptable fit of the model to the data (CFI = 0.98, TLI = 0.98, RMSEA = 0.059, chi-sq = 746.82, $df = 265, p < 0.001$) suggesting that the original five-factor structure of the Norwegian CTQ-SF is plausible in this sample. Item means (SD), internal consistency, standardized factor loadings and frequencies of the observed variable are reported in Table 2. For all items, the

factor loadings were 0.60 or greater and loaded significantly on their latent factor ($p < 0.001$). Internal reliability for the five factors was assessed by the use of a latent variable approach as described by Raykov (2009). The maximum likelihood estimator (which assumes normal distribution) was used, as this is most appropriate when considering how reliable the simple unweighted sum scores score is (Gustafsson & Åberg-Bengtsson, 2010). The reliability coefficients were: emotional abuse = 0.86; physical abuse = 0.90; sexual abuse = 0.96; emotional neglect = 0.90; and physical neglect = 0.79.

DISCUSSION

The present study is the first to present the psychometric properties of a Norwegian version of the CTQ-SF (Winje *et al.*, 2004). As proposed by Baker and Maiorino (2010), we report both the total CTQ-SF score and its subscale scores to increase comparability across research studies and to assist researchers and clinicians in their use of the CTQ-SF.

Overall, the results of the confirmatory factor analysis indicate a reasonable fit of our data to the original five-factor structure of the English version of the CTQ-SF (Bernstein & Fink, 1998). Moreover, our findings suggest that the scale is measurement invariant across sex and various high-risk samples in Norway. This is consistent with a previous study, where the amount of bias from measurement noninvariance in the CTQ-SF across gender and ethnicity at the total scale level was minimal (Thombs *et al.*, 2007). However, it is important to emphasize, that both the present and Thombs *et al.* (2009) study explored whether the CTQ-SF is measurement invariant by the use of subgroups based on certain observed characteristics (gender, ethnicity). Future studies that use mixture modeling to explore whether there exist unobserved latent classes of individuals with distinct patterns of abuse might have more success and is encouraged. However, the consistent emergence of the original five constructs across different gender and maltreatments subsamples support the overall construct validity of the scale.

The subscales demonstrated satisfactory to excellent internal consistency, with reliability estimates ranging from 0.79 (physical neglect) to 0.96 (sexual abuse). This means that approximately 20% or less of the simple unweighted sum scores are due to residual (non-common) variance. For the sexual abuse subscale in particular, strong internal consistency has been reported in the literature (Bernstein *et al.*, 2003; Gerdner & Allgulander,

Table 1. Participant characteristics ($n = 517$) and subsample CTQ-SF scores (Mean, SD), by group

	Out-of home placement ($n = 72$)	SUD patients ($n = 126$)	Psychiatric patients ($n = 210$)	Prisoners ($n = 109$)
Age, mean (SD)	19.1 (3.5)	30.9 (11.6)	25.5 (10.7)	32.2 (9.9)
Male, n (%)	32 (44)	75 (60)	62 (30)	109 (100)
CTQ-SF scale scores, mean (SD)				
Emotional abuse	11.9 (6.1)	10.7 (4.5)	10.1 (5.5)	10.2 (4.9)
Physical abuse	9.4 (6.1)	7.2 (3.6)	6.8 (3.8)	8.2 (4.7)
Sexual Abuse	7.5 (5.6)	7.6 (4.9)	7.5 (5.1)	6.8 (3.9)
Physical neglect	14.6 (5.3)	12.6 (5.1)	11.2 (5.1)	12.6 (5.1)
Emotional neglect	12.0 (4.9)	8.1 (3.0)	7.1 (3.4)	8.2 (3.8)
Sum CTQ	55.5 (22.9)	46.2 (15.2)	42.27 (18.2)	45.7 (15.6)

Table 2. Item means (SD), standardized factor loading and percentage of item response categories

Item	Men (n = 278)		Women (n = 239)		Percentage of item response categories					
	Mean (SD)	Factor loading	Mean (SD)	Factor loading	1	2	3	4	5	Missing
Physical abuse										
PA9: Hit hard enough to see doctor	1.35 (0.95)	0.950	1.19 (0.66)	0.854	84.9	4.7	3.6	1.8	2.7	2.2
PA10: Hit hard enough to leave bruises	1.75 (1.20)	0.954	1.71 (1.23)	0.924	65.5	9.6	11.4	5.4	6.0	1.8
PA11: Punished with hard objects	1.58 (1.15)	0.892	1.36 (0.87)	0.812	76.8	5.1	9.1	3.3	4.0	1.8
PA14: Was physically abused	1.70 (1.25)	0.925	1.71 (1.28)	0.938	69.5	6.4	8.9	5.6	7.8	1.8
PA15: Hit badly enough to be noticed	1.36 (0.94)	0.903	1.29 (0.86)	0.824	82.8	5.8	4.2	2.5	2.7	2.0
Emotional abuse										
EA3: Called names by family	2.09 (1.28)	0.772	2.39 (1.43)	0.814	43.0	16.6	19.2	8.9	10.0	2.4
EA8: Parents wished subject was never born	1.75 (1.15)	0.755	1.95 (1.25)	0.753	57.0	15.6	13.8	5.4	6.2	2.0
EA13: Family said hurtful things	2.34 (1.23)	0.800	2.57 (1.32)	0.831	29.6	24.3	22.5	12.2	9.3	2.2
EA16: Felt hated by family	1.87 (1.22)	0.778	1.92 (1.36)	0.846	57.0	14.5	11.1	8.3	7.3	1.8
EA22: Was emotionally abused	1.95 (1.32)	0.895	2.20 (1.50)	0.939	54.3	12.7	11.6	8.5	10.9	2.0
Sexual abuse										
SA18: Was touched sexually	1.40 (0.94)	0.939	1.78 (1.32)	0.970	71.9	8.9	6.5	3.8	6.7	2.2
SA19: Hurt if subject did not do something sexual	1.16 (0.61)	0.950	1.43 (1.09)	0.937	85.3	2.9	3.8	1.8	4.0	2.2
SA20: Made to do sexual things	1.31 (0.78)	0.970	1.59 (1.21)	0.964	78.0	5.6	6.2	3.8	4.2	2.2
SA21: Was molested	1.36 (0.88)	0.954	1.71 (1.31)	0.978	75.1	6.5	6.9	2.7	6.5	2.2
SA24: Was sexually abused	1.32 (0.89)	0.972	1.78 (1.45)	0.979	76.6	6.7	3.5	1.8	9.3	2.2
Physical neglect										
PN1: Not enough to eat	1.43 (0.92)	0.778	1.39 (0.92)	0.735	77.7	9.1	6.5	2.7	2.2	1.8
PN2: Got taken care of (R)	2.01 (1.24)	0.781	1.93 (1.17)	0.904	49.7	18.9	14.3	9.8	5.1	2.2
PN4: Parents were high or drunk	1.76 (1.21)	0.625	1.74 (1.28)	0.595	66.3	8.2	10.0	8.2	5.6	1.8
PN6: Wore dirty clothes	1.45 (0.92)	0.834	1.40 (0.86)	0.785	74.8	10.9	7.1	3.3	1.8	2.2
PN23: Got taken to a doctor (R)	1.74 (1.13)	0.817	1.66 (1.06)	0.861	60.3	18.5	9.3	5.8	3.6	2.5
Emotional neglect										
EN5: Made to feel important (R)	2.45 (1.24)	0.657	2.47 (1.25)	0.741	28.1	26.7	20.1	16	6.9	2.2
EN7: Felt loved (R)	2.31 (1.23)	0.919	2.30 (1.24)	0.907	33.6	26.0	20.0	12.9	5.6	2.0
EN12: Was looked out for (R)	2.23 (1.15)	0.814	2.22 (1.21)	0.908	34.3	27.2	21.1	10.0	5.3	2.2
EN17: Family felt close (R)	2.63 (1.24)	0.814	2.71 (1.24)	0.841	19.8	26.5	25.8	15.8	9.8	2.4
EN25: Family was source of strength (R)	2.68 (1.31)	0.874	2.67 (1.33)	0.904	23.2	22.9	25.2	14.2	12.2	2.4

Note: R: reversed coded; Item response categories for overall sample: 1 = never true, 2 = rarely true, 3 = sometimes true, 4 = often true, 5 = very often true.

2009; Thombs *et al.*, 2009). The relative weakness of the physical neglect subscale compared to the other subscales has been reported in different exposed samples, such as street-based sex workers, prisoners and substance abusers, patients in different clinical settings and in community samples (Bernstein *et al.*, 2003; Gerdner & Allgulander, 2009; Scher, 2001; Thombs *et al.*, 2009; Vilano, 2004). Future revisions of this particular subscale might be considered, with the PN4 item –“parents were high or drunk” as a prime candidate for replacement: In our data it had a noticeable lower factor loading. It also refers to a specific risk for physical neglect, only relevant for a subgroup of children who is exposed to such maltreatment, more than to measure neglect itself.

In accordance with common practice, we have used a reflective measurement model on the CTQ-SF. Netland (2001) has strongly criticized the use of such modeling approach when analyzing traumatizing events and argues that one should treat items on potentially traumatizing events as causal indicators of composite variables. We are sympathetic to many of the objections raised by Netland (2001), but note that the use of causal (forma-

tive) indicator models is controversial as well (see Bollen, 2007; Edwards, 2011; Harding & Marcoulides, 2011; Howell, Breivik & Wilcox, 2007). One major limitation with causal indicator/formative models is the assumption of error-free measurement. For an instrument like the CTQ-SF, based on retrospective recall of childhood abuse, absence of measurement error is hard to defend (Hardt & Rutter, 2004). In addition to an actual history of severe abuse, the item responses are probably influenced by factors such as selective recollection of childhood events, shaped by more recent difficult circumstances, forgetfulness over time, willingness to report sensitive information, response styles among individuals, subjective interpretation of item content, and individual or cultural thresholds of what constitutes “abuse.” As long as such influences affect the specific items differently, use of models with reflective indicators can potentially adjust for at least some the indicators measurement error. This is due to the fact that reflective models focus on common variance and leave what is unique out from the latent construct. Such considerations made us opt for the reflective indicator model, even though we believe that the CTQ-SF indicators probably also contain some

formative aspects (for a similar argument see: Dyregrov, Gupta, Gjestad & Mukanoheli, 2000). However, it is important to emphasize that the choice between using a reflective versus formative modeling approach might turn out to be relatively inconsequential when evaluating the dimensionality of a measure. It has been recommended that principal component analysis (PCA), instead of common factor analysis, should be used when extracting separate dimensions since PCA retains both the unique and common variance for each item in the measure (Petter, Straub & Rai, 2007; Chin, 1995). Previous studies that have used principal component analysis (PCA) have generally found support for the same five dimensions as modeled in the present study, as have previous studies with a common factor approach (Gerdner & Allgulander, 2009). The similarity of findings across the approaches is reasonable considering that the specific items within each factor of the CTQ-SF had relatively little unique variance. Future research should aim at furthering our understanding of the strengths and weaknesses associated with various measurement models when analyzing retrospective self-report questionnaires of childhood abuse experiences.

Several limitations should be considered when interpreting the results of this study. First, there are possible challenges regarding the validity of self-report of childhood maltreatment without any supplementary information of the reported traumatic exposures.

Second, the patient subsample was invited to participate by their therapists. This may have created a conflict of interest influencing the validity of their responses. Third, the present study does not provide information about the convergent and divergent validity of the Norwegian CTQ-SF. However, our study still makes several important contributions.

With the publishing of this study, a Norwegian version of the CTQ-SF, with acceptable psychometric properties, good reliability and satisfactory accuracy to assess different dimensions of childhood trauma, will now be available. We would encourage the replication of our study with other clinical and non-clinical samples to increase the utility of the CTQ-SF. Extended use of standardized and psychometrically sound instruments increases comparability across studies. Intervention, of course, requires detection. In routine clinical work with adults retrospective detection might help find individuals who need help to address the long-term emotional and behavioral consequences of childhood adversities that contribute to their ongoing elevated risk on new onsets (Kessler, McLaughlin, Green *et al.*, 2010). As a screening instrument for clinical use in Norway, the CTQ-SF can contribute to the identification of abused subjects; broaden diagnostic evaluations and guide treatment planning. However, it is important to remind potential users that the CTQ-SF remains a screening tool which main purpose is to collect information that may be used as a starting point for a dialogue with patients about probable childhood trauma. The retrospective questions in assessment instruments cannot differentiate between an actual history of serious abuse and selective recollection of childhood events. Previous studies have also argued that retrospective report of adverse childhood experiences may be hampered by false negatives as well as inconsistencies over longer time periods (Hardt & Rutter, 2004; Hepp, Gamma, Milos *et al.* 2006; Krinsley, Gallagher, Weathers, Kutter & Kaloupek, 2003). Threats to validity concerns both occurrence and intensity, and

careful interpretation of results when using the CTQ-SF remains very important.

This study was supported by "The National Program for Integrated Clinical Specialist and PhD-training for Psychologists" in Norway. This program is a joint cooperation between the Universities of Bergen, Oslo, Tromsø, The Norwegian University of Science and Technology (Trondheim), the Regional Health Authorities, and the Norwegian Psychological Association. The program is funded jointly by The Ministry of Education and Research and The Ministry of Health and Care Services. We also acknowledge the work of senior executive Bente Haktorsen, who has contributed to the coordination of data collection among our respondents.

REFERENCES

- Baker, A. J. L. & Maiorino, E. (2010). Assessments of emotional abuse and neglect with the CTQ: Issues and estimates. *Children and Youth Services Review*, *32*, 740–748.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.
- Bernstein, D. P. & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report manual*. San Antonio, TX: The Psychological Corporation.
- Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., Stokes, J., Medrano, M., Desmond, D. & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, *27*, 169–190.
- Brewin, C. R. (2005). Systematic review of screening instruments for adults at risk of PTSD. *Journal of Traumatic Stress*, *18*, 53–62.
- Brewin, C., Andrews, B. & Gotlib, I. (1993). Psychopathology and early experience: A reappraisal of retrospective reports. *Psychology Bulletin*, *113*, 82–98.
- Briere, J., Hodges, M. & Godbout, N. (2010). Traumatic stress, affect dysregulation, and dysfunctional avoidance: A structural equation model. *Journal of Traumatic Stress*, *23*, 767–774.
- Briere, J. & Jordan, C. E. (2009). The relationship between childhood maltreatment, moderating variables, and adult psychological difficulties in women: An overview. *Trauma, Violence, and Abuse: A Review Journal*, *10*, 375–388.
- Bollen, K. A. (2007). Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods*, *12*, 219–228.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 111–135). Newbury Park, CA: Sage.
- Chin, W. W. (1995). Partial least squares is to LISREL as principal components analysis is to common factor analysis. *Technology Studies*, *2*, 315–319.
- Dovran, A., Winje, D., Arefjord, K. & Haugland, B. S. M. (2012). Traumatic events and posttraumatic reactions among children and adolescents in out-of-home placement: A 25-year systematic literature review. *Journal of Child and Adolescent Trauma*, *5*, 16–32.
- Dyregrov, A., Gupta, L., Gjestad, R. & Mukanoheli, E. (2000). Trauma exposure and psychological reactions to genocide among Rwandan children. *Journal of Traumatic Stress*, *13*, 3–21.
- Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research Methods*, *14*, 370–388.
- Elhai, J. D., Gray, M. J., Kashdan, T. B. & Franklin, C. L. (2005). Which instruments are most commonly used to assess traumatic event exposure and posttraumatic effects? A survey of traumatic stress professionals. *Journal of Traumatic Stress*, *18*, 541–545.
- Flora, D. B. & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*, 466–491.
- Fosse, G. K. & Holen, A. (2002). Childhood environment of adult psychiatric outpatients in Norway having been bullied in school. *Child Abuse & Neglect*, *26*, 129–137.

- Fosse, G. K. & Holen, A. (2006). Childhood maltreatment in adult female psychiatric outpatients with eating disorders. *Eating Behaviors*, 7, 404–409.
- Fosse, G. K. & Holen, A. (2007). Reported maltreatment in childhood in relation to the personality features of Norwegian adult psychiatric outpatients. *Journal of Nervous & Mental Disease*, 195, 79–82.
- Gerdner, A. & Allgulander, C. (2009). Psychometric properties of the Swedish version of the Childhood Trauma Questionnaire–Short Form (CTQ-SF). *Nordic Journal of Psychiatry*, 63, 160–170.
- Gustafsson, J.-E. & Åberg-Bengtsson, L. (2010). Unidimensionality and interpretability of psychological instruments. In S. E. Embretson (Ed.), *Measuring psychological constructs: Advances in model-based approaches* (pp. 123–144). Washington, DC: American Psychological Association.
- Hardin, A. & Marcoulides, G.A. (2011). Commentary on the Use of Formative Measurement. *Educational and Psychological Measurement*, 71, 753–764.
- Hardt, J. & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *Journal of Child Psychology and Psychiatry*, 45, 260–273.
- Hepp, U., Gamma, A., Milos, G., Eich, D., Ajdacic-Gross, V., Rössler, W., Angst, J. & Schnyder, U. (2006). Inconsistency in reporting potentially traumatic events. *The British Journal of Psychiatry*, 188, 278–283.
- Howell, R. D., Breivik, E. & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods*, 12, 205–218.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Kessler, R. C., McLaughlin, K. A., Green, J. G., et al. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, 197, 378–385.
- Krinsley, K. E., Gallagher, J. G., Weathers, F. W., Kutter, C. J. & Kaloupek, D. G. (2003). Consistency of retrospective reporting about exposure to traumatic events. *Journal of Traumatic Stress*, 16, 399–409.
- McDonald, R.P. (2003). Behavior domains in theory and in practice. *The Alberta Journal of Educational Research*, 49, 212–230.
- Muthén, B. & Muthén, L. (1998–2010). Mplus user's guide. Los Angeles, CA: Muthén & Muthén.
- Myers, K. & Winters, C. N. (2002). Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 114–122.
- Netland, M. (2001). Assessment of exposure to political violence and other potentially traumatizing events. A critical review. *Journal of Traumatic Stress*, 14, 311–326.
- Ohan, J. L., Myers, K. & Collett, B. R. (2002). Ten-year review of rating scales. IV: Scales assessing trauma and its effects. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 1401–1422.
- Petter, S., Straub, D. & Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Quarterly*, 31, 623–656.
- Pine, D.S. & Cohen, J. A. (2002). Trauma in children and adolescents: Risk and treatment of psychiatric sequelae. *Biological Psychiatry*, 51, 519–531.
- Ravndal, E., Lauritzen, G., Frank, O., Jansson, I. & Larsson, J. (2001). Childhood maltreatment among Norwegian drug abusers in treatment. *International Journal of Social Welfare*, 10, 142–147.
- Raykov, T. (2009). Interval estimation of revision effect on scale reliability via covariance structure analysis. *Structural Equation Modeling*, 16, 539–555.
- Read, J., Hammersley, P. & Rudegeair, T. (2007). Why, when and how to ask about childhood abuse. *Advances in Psychiatric Treatment*, 13, 101–110.
- Scher, C. D., Stein, M. B., Asmundson, G. J. G., McCreary, D. R. & Forde, D. R. (2001). The Childhood Trauma Questionnaire in a community sample: Psychometric properties and normative data. *Journal of Traumatic Stress*, 14, 843–857.
- Sperber, A. D. (2004). Translation and validation of study instruments for cross-cultural research. *Gastroenterology*, 126, 124–128.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180.
- Strand, V. C., Sarmiento, T. L. & Pasquale, L. E. (2005). Assessment and screening tools for trauma in children and adolescents: A review. *Trauma, Violence & Abuse*, 6, 55–78.
- Thombs, B. D., Bernstein, D. P., Lobbstad, J. & Arntz, A. (2009). A validation study of the Dutch Childhood Trauma Questionnaire-Short Form: Factor structure, reliability, and known-groups validity. *Child Abuse & Neglect*, 33, 518–523.
- Thombs, B. D., Lewis, C., Bernstein, D. P., Medrano, M. A. & Hatch, J. P. (2007). An evaluation of the measurement equivalence of the Childhood Trauma Questionnaire-Short Form across gender and race in a sample of drug-abusing adults. *Journal of Psychosomatic Research*, 63, 391–398.
- Tucker, L. R. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10.
- Villano, C. L., Clealand, C., Rosenblum, A., Fong, C., Nuttbrock, L., Marthol, M. & Wallace, J. (2004). Psychometric utility of the Childhood Trauma Questionnaire with female street-based sex workers. *Journal of Trauma Dissociation*, 5, 33–41.
- Winje, D., Dovran, A. & Murison, R. (2004). Childhood Trauma Questionnaire (CTQ). Translated and reproduced by permission. The Psychological Corporation, Harcourt Assessment Company. Retrieved from <http://www.pearsonassessment.no>.

Received 19 December 2011, accepted 4 February 2013