



Research article

Is the Brief Child Abuse Potential Inventory (BCAPI) a valid measure of child abuse potential among mothers and fathers of young children in Germany?



Christoph Liel^{a,1,*}, Franziska Meinck^{b,c,1}, Janina I. Steinert^b, Heinz Kindler^a,
Katrin Lang^{a,d}, Andreas Eickhorst^{a,e}

^a German Youth Institute, Department of Families and Family Policies, Munich, Germany

^b Centre for Evidence-Based Interventions, Department of Social Policy & Intervention, University of Oxford, Oxford, United Kingdom

^c OPTENTIA, School of Behavioural Sciences, North-West University, Vanderbijlpark, South Africa

^d Child Guidance Center Ingolstadt, Ingolstadt, Germany

^e Hochschule Hannover, Fakultät V – Diakonie, Gesundheit und Soziales, Hannover, Germany

ARTICLE INFO

Keywords:

Child abuse
Parent-child relationships
Child abuse potential
Psychometrics
Measurement

ABSTRACT

Background: In order to prevent child abuse, instruments measuring child abuse potential (CAP) need to be appropriate, reliable and valid.

Objective: This study aimed to confirm the 6-factor structure of the Brief Child Abuse Potential Inventory (BCAPI) in a German sample of mothers and fathers, and to examine longitudinal predictors of CAP.

Participants and setting: Two waves of data were collected from 197 mothers and 191 fathers of children aged 10–21 months for the “Kinder in Deutschland – KiD 0–3” in-depth study. Families were stratified based on prior self-report data for screening purposes.

Methods: 138 fathers and 147 mothers were included in the analysis (invalid: 25% mothers, 30% fathers). First, validity of reporting was examined. Second, confirmatory factor analysis (CFA) was employed to assess factor structure. Third, internal reliability and criterion validity were examined. Finally, multivariate poisson regressions investigated longitudinal predictors of CAP in mothers.

Results: A previously established six-factor structure was confirmed for mothers but not fathers. CFA failed for fathers due to large numbers of variables with zero variance. For mothers, internal consistency and criterion validity were good. BCAPE score at follow-up was associated with baseline BCAPE score ($\beta = 0.08$), stress ($\beta = 0.06$), education ($\beta = -0.19$) and alcohol use ($\beta = .58$).

Conclusions: Findings confirm the six-factor structure of the BCAPE among German mothers. The clinical use of the BCAPE in fathers is not recommended as it might produce data that are hard to interpret. Further research with fathers is needed to establish if this is due to limitations with this dataset or with the questionnaire.

* Corresponding author at: German Youth Institute, Department of Families and Family Policies, Research Department on Early Childhood Intervention, Nockherstr. 2, D-81541 Munich, Germany.

E-mail address: liel@dj.de (C. Liel).

¹ Both authors contributed equally to this manuscript.

<https://doi.org/10.1016/j.chiabu.2018.11.008>

Received 3 May 2018; Received in revised form 22 October 2018; Accepted 13 November 2018

Available online 05 January 2019

0145-2134/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Background

Child abuse is a major public health concern and associated with long-term negative effects on physical and emotional health and behaviour (Norman et al., 2012). Globally, an estimated 1 billion children experience violence annually (Hillis, Mercy, Amobi, & Kress, 2016) but there is a growing interest in child abuse surveillance (Meinck et al., 2016). Existing studies investigate either or both of the following: abusive incidents or risk factors shown to increase child abuse potential. One of the measures widely used in research and clinical practice is the Child Abuse Potential Inventory.

1.1. Measuring child abuse potential – the Child Abuse Potential Inventory (CAPI)

The (CAPI) is a risk screening tool for the detection of potential child abuse cases (Milner & Ayoub, 1980). It has undergone psychometric testing and is widely used in epidemiologic and intervention studies in multiple contexts (Chan, Lam, Chun, & So, 2006; Miragoli, Camisasca, & Blasio, 2015) but most of the evidence stems from studies in the U.S. Rather than measuring prevalence or incidence of actual child abuse cases, the inventory aims to assess abuse risk. The CAPI has been often used with at-risk caregivers such as substance users (Kelley, Lawrence, Millettich, Hollis, & Henson, 2015) or those with a history of childhood abuse (Henschel, de Bruin, & Möhler, 2013; Romero-Martínez, Figueiredo, & Moya-Albiol, 2014). The measure consists of 160 items in an “agree/disagree” format, 77 of which are summarized in a child physical abuse scale. The rest of the items are part of three validity scales, focusing on lies, random responses and inconsistency. The CAPI also has a response distortion index.

A recent review of the psychometric properties of the CAPI found high percentages (> 70%) of accurate classification and discrimination between abusive parents and non-abusive controls across multiple contexts (Walker & Davies, 2010). It also found high levels of internal consistency for the abuse scale across all but one U.S. study (Walker & Davies, 2010). Research on the factor structure of the CAPI has been limited thus far. As the CAPI aims to measure the unobservable construct of ‘child abuse potential’ it is important that consistently demonstrates a stable factor structure across the populations it is applied in. An unstable factor structure would suggest that it cannot be consistently reduced into a total CAP score as it may mean that it measures child abuse potential differently across different populations (DeVellis, 2012).

Milner et al suggested a six-factor structure for the instrument consisting of the dimensions distress, rigidity, unhappiness, problems with the child and oneself, problems with the family, and problems with others (1994, Milner & Ayoub, 1980; Milner, 1986). However, evidence across the world varies: some studies are able to replicate the six-factor structure (in China, Chan et al., 2006) while others are not (in Japan, Kawamura, Takahashi, Akiyama, Sasaki, & Kako, 2009 and Greece, Diareme, Tsiantis, & Tsitoura, 1997; Haz & Ramirez, 1998). The Chinese study tested the factor structure in a sample of 897 mothers recruited through child centres. The Greek study, which found a 5-factor structure, included 320 parents (77% mothers) of outpatient pediatric patients from a hospital. The Japanese study, which found a 3-factor structure, used a sample of 1809 fathers and mothers. These findings give indicative evidence that the instrument is less stable when used in populations that include fathers.

1.2. Use of the CAPI with mothers and fathers

The CAPI is used for the detection of child abuse potential in mothers and fathers. Research so far is limited on differences in psychometric properties of the CAPI comparing mothers and fathers. Some studies have investigated differences in mean CAPI scores in mothers and fathers with conflicting results: in Portugal, fathers had lower child abuse potential than mothers (Romero-Martínez et al., 2014), in Malaysia, fathers had higher CAPI mean scores (Chua & Abdul, 2001), and in the U.S., no differences were observed (Brewster, Nelson, McCanne, Lucas, & Milner, 1998). While predictors of child abuse potential were similar among fathers and mothers in Portugal (Romero-Martínez et al., 2014), a study from the U.S. found higher risk of distress, unhappiness and problems from people outside the family among mothers while fathers reported more parenting rigidity (Pittman & Buckley, 2006). It is not clear whether differences in scores between mothers and fathers reflect actual differences in child abuse potential or whether they are the result of reporting biases (e.g. questions may be phrased in a way that reflects mothers’ child abuse potential better than fathers’). Differences relating to reporting bias can affect the factor structure of the instrument. Much of the published research also focuses on samples consistent predominantly of mothers. No studies thus far have investigated differences in factor structure and in prospective validity of the CAPI. Given these conflicting results, there is insufficient evidence about differences between mothers and fathers in measurement functioning of the CAPI.

1.3. The Brief Child Abuse Potential Inventory (BCAPI)

The BCAPE is a short form of the CAPI developed by Ondersma, Chaffin, Mullins, and LeBreton, (2005). It consists of 33 items, 24 of which are summarized in a child physical abuse scale. The rest of the items form validity scales: a three-item Random Response Scale and a six-item Lie Scale. Ondersma et al. developed the BCAPE using two independent development samples recruited from families in the U.S. involved with child protective services. Each item of the CAPI Abuse Risk Scale was ranked by its predictive importance of the full CAPI and future child protective reports. Item total correlations were then checked for each of the sub-groups (White, Native American, African American and Hispanic women and men). 39 items in total were identified. Items were then examined for content. Those with complex phrasing were removed, resulting in 33 items which were subjected to principal axis factoring (PAF) using oblimin rotation. This resulted in a seven-factor solution with 24 items. Cross-validation was then performed on two additional independent samples (Ondersma et al., 2005).

The BCAP has been psychometrically evaluated in three studies to date. Ondersma et al. suggested a stable 7-factor solution in their U.S. samples which explained 66.2% of the variance in BCAP abuse scores. Their two cross-validation samples ($n = 713$) consisted of parents (74% & 64% mothers) of diverse ethnic backgrounds (Caucasian, Native American, African American and Hispanic) who were referred for prevention and treatment of child abuse. They established good internal consistency ($KR20 = .89$) and high correlation of the BCAP abuse score with the CAPI abuse score ($r = .96$). Both CAPI & BCAP Scores correlated highly with the Beck Depression Inventory Scores ($r = .67$), suggesting overlap between the two in measuring depression. The BCAP scores predicted reports of neglect more strongly than reports of physical abuse (Ondersma et al., 2005). Walker and Davies (2012) examined the validity of the BCAP in a convenience sample recruited through schools in the U.K. ($n = 358$). This sample was less ethnically diverse than the U.S. sample consisting of 88% mothers, 92% White. Principal axis factoring with oblique rotation resulted in a six-factor solution. Internal consistency was good ($\alpha = .82$) but criterion validity was not examined. The authors found that one item on the random response scale increased potential invalid protocols from 30.3% to 94%, possibly suggesting a need for adaptation for the U.K. population. Dawe, Taplin, and Mattick, (2017) examined the psychometric properties of the BCAP in a sample of Australian mothers on opioid substitution therapy. PAF with oblimin rotation resulted in a six-factor solution which accounted for 53% of the variance in BCAP Scores. Strong correlations were found between the BCAP sub-scales and depression and anxiety. Further, strong correlations between all sub-scales except for rigidity were shown, suggesting that rigidity items may be tapping into an independent construct outside of other BCAP sub-scales. Internal consistency of the BCAP Abuse Risk Scale was high ($KR20 = .90$) and tests for criterion validity showed weak correlations between other theoretically-related constructs, such as substance use or childhood history of abuse, but showed strong correlations with psychological distress (Dawe et al., 2017).

The above studies are subject to a number of limitations. First, the vast majority of participants in the four samples were mothers and only those studies with higher percentages of mothers were able to replicate the six-factor structure. Only one of the studies investigated some psychometric properties specifically for fathers. Second, data analyses were conducted in SPSS for three studies, but no mention is made about the distribution of the data and the treatment of binary indicator variables which may be problematic when using that specific software without add-ons. Third, sampling biases for all four samples cannot be excluded. In particular, Walker & Davies achieved a participation rate of only 24.7%, with an additional 10% excluded due to missing information (Walker & Davies, 2012). Fourth, three studies recorded high numbers ($> 30\%$) of invalid protocols, thus excluding a large part of the sample from further analysis, and one study did not report validity indices. To interpret the findings, these limitations must be taken into consideration.

1.4. Gender differences in BCAP child abuse measures

No published study to date has examined differences in factor structure of the BCAP among mothers as compared to fathers, nor has there been a systematic assessment of the factor structure solely in fathers. This type of research is important because differences in psychometric properties between men and women may indicate either actual differences in child abuse potential or measurement problems. Most research has focused on differences in mean BCAP abuse scores and internal consistency. Mean scores have been found to be either higher among fathers (Kelley et al., 2015; Tucker, 2014) or not different between parents (Rodriguez, Baker, Pu, & Tucker, 2017).

A recent study in the U.S. found acceptable internal consistency of the BCAP among fathers ($\alpha = .78$) and mothers ($\alpha = .80$) (Rodriguez et al., 2017) in a sample of 88 dyads. Studies investigating differences in predictors of BCAP Abuse scores among fathers and mothers are rare. For mothers, stress was found to be a significant predictor of BCAP Abuse Score while demographic variables and couple functioning were not significant. For fathers, ethnicity, stress, couple functioning, and the interaction between stress and couple functioning were significant predictors of BCAP Abuse score. A cross-over effect of the other parent's risk factors on self-reported BCAP Abuse Score could not be found (Tucker, 2014). Other studies have examined correlations between BCAP Abuse Scores and depressive symptoms for both fathers ($r = .67$) and mothers ($r = .80$) (Kelley et al., 2015), maternal depression ($\beta = .31$), and experience of social support among mothers ($\beta = -0.21$) (Whitson, Martinez, Ayala, & Kaufman, 2011) and fathers ($\beta = -1.69$) (Vasquez Guerrero, 2009). One study examined correlations between BCAP Abuse Scores and parenting attitudes (mothers $r = .26$; fathers $r = .38$), parental over-reactivity ($r = .27$; $r = .27$), use of physical punishment (ns; $r = .30$), use of aggression for child non-compliance (ns; $r = .32$), perceived stress ($r = .62$; $r = .49$), daily hassles ($r = .72$; $r = .43$), mental health symptoms ($r = .65$; $r = .64$), negative mood regulation ($r = -.25$; $r = -.60$) and frustration tolerance (ns; ns) (Rodriguez et al., 2017).

Thus, the evidence base of the BCAP is limited and of varying quality. Very little research has examined differences in BCAP Abuse Scores and psychometric properties of the instrument in mothers and fathers. No studies have tested for measurement invariance between mothers and fathers or have investigated whether the BCAP is a suitable measure for both parents. Furthermore, no evidence is available on the use of the BCAP outside predominantly English-speaking countries. To respond to these gaps, the current study therefore had three aims: i) To investigate whether the 6-factor structure found in previous studies could be confirmed for mothers and fathers in this German sample respectively; ii) To examine psychometric properties of the BCAP for mothers and fathers respectively; and iii) To investigate longitudinal predictors of child abuse potential for mothers and fathers in a non-English-speaking sample.

Table 1

Sample size at both data collection points and according to risk groups.

Sample size	Psychosocial risk factors for child maltreatment			
	Low (0-1 RF) T1/T2	Medium (2-3 RF) T1/T2	High (4 + RF) T1/T2	Total T1/T2
Children aged 10-14 months/ 17-21 months	36/35	39/37	23/21	98/93
Children aged 17-21 months/ 24-28 months	38/37	32/31	29/21	99/89
Total	74/72	71/68	52/42	197/182

RF = risk factor T1 = baseline T2 = follow-up.

2. Methods

2.1. Sample

Data were collected from September 2014 to February 2015 (baseline) and seven months later (follow-up) as part of the Kinder in Deutschland aged zero to three years (KiD 0–3) in-depth study. The KiD 0–3 is a nationally representative study examining the epidemiology of psychosocial burden and risk for child abuse and neglect in Germany (Eickhorst et al., 2016). A pilot study with a sample of N = 6000 caregivers of toddlers aged 0–3 was carried out in two large German cities to test the risk inventory and two different types of sampling (Eickhorst et al., 2015). Families were recruited via phone and mail (city 1) or at child development reviews (child-well visits) in pediatric surgeries (city 2). Ethical approval for the pilot study was granted by the General Medical Council in the North-Rhine region (No 2,013,247). The in-depth study utilizes a sub-sample of the pilot study population of primary (n = 197) and secondary caregivers (n = 197) of children aged 10–14 months or 17–21 months (Zimmermann et al., 2016). In-depth study participants were recruited according to the amount of distal and proximal risk factors for child maltreatment they disclosed in the pilot study. Distal risk factors included family receipt of welfare benefit, overcrowding, > 2 children, maternal unemployment, low maternal education, maternal history of child abuse or neglect, mental disorder or substance abuse of any caregiver. Proximal risk factors included parental conflict, domestic violence, maternal depression, negative attitudes during pregnancy. High-risk families had a mean of 5.4 risk factors, medium-risk families a mean of 2.3 risk factors, and low-risk families a mean of 0.5 risk factors. Equal numbers of parent dyads from all three risk groups were recruited for the in-depth study but the high-risk group was slightly under-represented (Table 1).

2.2. Procedure

Primary and secondary caregivers filled in a posted standardized self-report questionnaire. Primary caregivers also completed an interviewer-assisted questionnaire during a home visit. All completed questionnaires were retrieved during the home visit. In 97% of cases, the primary caregiver was the mother, and in 94% of cases, the secondary caregiver was the father.

All caregivers gave written informed consent prior to participation in the study. Participants received small financial incentives (€40 per family) and were told that they could withdraw at any time without declaration of reasons. Confidentiality was maintained throughout the study. Completed questionnaires were placed in sealed envelopes by participants and only linked via participant ID number. Researchers in contact with participants were not aware of families' risk group categorization. Where researchers had concerns about a child's risk of harm, procedures were put in place for self-referral to existing services with follow-up support.

2.3. Measures

Data from all measures were collected at both baseline and follow-up except where explicitly stated.

Child abuse potential was measured using the Brief Child Abuse Potential Inventory (BCAPI) (Ondersma et al., 2005). The BCAPI includes two validity scales, the lie scale and the random response scale. The actual child abuse risk scale includes 24 items measuring child abuse risk factors in an agree/disagree format, which are summed to a total score.

Socio-demographic characteristics on child age and gender, parental age and gender, parental relationship, training, and employment were collected with single-item questions from both fathers and mothers. Household information, education and immigration history of both parents were collected with single-item questions only from the primary caregivers. Immigration history was based on §6 of the German MighEV act: being a foreign national, born outside Germany or having any parent that immigrated to Germany. Overcrowding followed a definition used by Eurostat: less than one room per family, one room for both parents, one room per two children under 12 years, one room per children aged 12 years and older, and one room per other adult.

Psychiatric symptoms were measured using the Patient Health Questionnaire (PHQ-D) (Löwe, Spitzer, Zipfel, & Herzog, 2002). These included the PHQ-9 for depression, the GAD-7 for anxiety, and a six-item questionnaire on alcohol consumption. The PHQ-D has previously been used in German samples (Löwe, Kroenke, Herzog, & Graefe, 2004, 2008). Items of the depression and anxiety sub-scales that were measured in 4-point Likert-scales were summed up into a total scale score and then dichotomized based on validated clinical cut-offs (both > 9). Items on the alcohol consumption sub-scale were summed up into a total score and dichotomized (cut-off > 1).

Parental childhood adverse experiences were measured using the ten-item Adverse Childhood Experiences (ACE) measure, which has been psychometrically tested in Germany and showed satisfactory internal consistency and good construct validity (Wingenfeld et al., 2011). This measured physical, emotional and sexual child abuse victimization; neglect; parental separation; mental health problems; drug abuse; domestic violence and imprisonment. A total sum score of exposures to childhood adverse experiences (dichotomous scaled in yes/no) was created (max. 10), and participants were dichotomized into those with multiple exposures (> 3 ACES) and those with no or fewer exposures based on existing studies on health outcomes (Hughes et al., 2017). This measure was only employed at follow-up.

Family violence was measured using six items from the national prevalence study by the National Society for the Prevention of Cruelty to Children in the United Kingdom (NSPCC; Radford, Corral, Bradley, & Fisher, 2013), which are based on the Juvenile Victimization Questionnaire (Finkelhor, Hamby, Ormrod, & Turner, 2005). Items were translated and adapted to parent report on abusive behaviors experienced by child. All items had a yes/no answering format. Three items measured child maltreatment (hitting, shaking or neglecting the child), and three items assessed domestic violence (destruction caused by argument, threat or violence) since birth of the child. A dichotomous variable was created for no abusive incidents versus any abusive incidents.

Perceived stress was measured using the four-item Perceived Stress Scale (PSS-4) (Cohen, Kamarck, & Mermelstein, 1983). This assesses the general state of perceived stress in a person's life (i.e., "In the past month, how often have you felt that you were unable to control the important things in your life?"). For the purposes of this study, the five-option response code was extended by an additional option that signified "always." The PSS was created for use in the general public and has been used successfully in Germany (Chu, Jahn, Khan, & Kraemer, 2016). A psychometric evaluation of the German ten-item PSS showed good internal consistency and construct validity (Klein et al., 2016). Two items were reverse coded. Items were summed with higher scores demonstrating higher levels of perceived stress. Internal consistency in the KiD 0–3 in-depth study was $\alpha = .80$ for mothers and $\alpha = .77$ for fathers (Liel, 2018).

Satisfaction in the parental relationship was measured using the four-item short form of the Dyadic Adjustment Scale (DAS) (Spanier, 1976). The DAS-4 showed good reliability and measured the couple satisfaction and relationship quality (Sabourin, Valois, & Lussier, 2005). Items were summed to a total score with higher scores showing lower levels of happiness in the relationship.

Division of parental tasks was measured using an adapted version of the Who Does What (WDW) (Cowan & Cowan, 1988) sub-scale on Child Related Tasks (6/18 months). 15 items assessed the distribution of parental tasks between both parents. One item "bringing the child to bed" and two items on childcare on weekdays and weekends were added to the original 12 parental tasks. A mean score of items on nine-point Likert scales was created where higher scores signified higher involvement by the father (1 = mother does it all, 5 = both parents do it equally, 9 = father does it all). One additional item measured *satisfaction with division of parenting tasks* on a five-point Likert scale. A higher score signified lower satisfaction. The WDW was only employed at baseline with an internal consistency of $\alpha = .84$ for mothers and $\alpha = .81$ for fathers (Liel, 2018).

Co-parenting was measured using five items on a six-point Likert-scale adapted from the Family Panel studies "Growing up in Germany" (AiD: A II; German Youth Institute, 2013) and "Panel Analysis of Intimate Relationships and Family Dynamics" (pairfam; Wilhelm et al., 2016). The items measured three dimensions of co-parenting: caregivers' collaboration in relation to parenting issues, problem solving and triangulation with the child. Items were summed with higher scores demonstrating lower levels of co-parenting. The internal consistency of total score in this study was $\alpha = .80$ for both parents (Liel, 2018).

Parental self-efficacy was measured using a German version of the 16-item post-natal Self Efficacy and Nurturing Role Questionnaire (SENR) (Pedersen, Suwalsky, Cain, & Zaslow, 1987), which is adapted from the Parenting Sense of Competence Scale (Gibaud-Wallaston & Wandersman, 1978). It measures expectations in relation to parental competences, such as caring tasks for the child or understanding the child's needs, and has been utilized with mothers and fathers (Solmeyer & Feinberg, 2011). A total sum score was created with higher scores reflecting higher self-efficacy.

Missing data ($< 10\%$) were handled by mean score replacement in CAPI, ACE, WDW and SENR. Given small numbers of items in the other measures, missing data have not been imputed there. Internal consistency of risk scales in this study using mother and father self-report ranged from $\alpha = .75$ to $\alpha = .89$ with one outlier: DAS (father-report) $\alpha = .64$ (Liel, 2018).

2.4. Data analysis

Data analysis followed nine steps and used baseline data unless otherwise stated. Analyses were carried out separately for mothers and fathers. First, following Little's approach for testing whether data are missing completely at random, a missing values analysis was carried out in SPSS. Missing values for all BAPI items were below 5%, which suggests that missing data are likely not a concern in terms of results bias. Second, characteristics of the sample were examined. Prevalence rates for binary variables were compared using a χ^2 test. Mean values for continuous variables were tested for significant differences between mothers and fathers using an independent samples *t*-test or Mann-Whitney U Test (non-normal distribution). Third, following Walker and Davies (2012), a validity filter was used for all BAPI items, composed of the Lie Scale (six items) and Random Response Scale (three items). Fourth, differences between frequencies/means of valid and invalid protocols were examined using χ^2 test/*t*-test or Mann-Whitney U Test. Fifth, correlations between items were examined, tests for multicollinearity were conducted and distribution of the data was examined. Sixth, Confirmatory Factor Analysis (CFA) in MPLUS 8 using the Weighted Least Square Means and Variance Adjusted (WLSMV) estimator for categorical data was conducted to examine whether the six-factor structure identified in previous research (Walker & Davies, 2012) could be confirmed for fathers and mothers using valid protocols only. Where factor structure could not be confirmed, inadequate data or unsuitability of the BAPI was assumed and no further analyses carried out. As missing data were less than 5%, a decision was made to use maximum likelihood estimation with pairwise deletion for the factor analysis component.

Seventh, internal consistency of the BCAPE Abuse Risk Scale was tested using the Kuder-Richardson Formula 20 (KR20) for dichotomous data, and correlations using Spearman's ρ were conducted with measures thought to be theoretically and empirically associated with child abuse potential to establish concurrent validity. Finally, predictors of child abuse potential were tested using two separate longitudinal regression analyses (full BCAPE Abuse Risk Scale and CFA reduced BCAPE Abuse Risk Scale²) with BCAPE Abuse Score at T2 as outcome and BCAPE Abuse Score at T1 as control variable using the valid protocols only. All risk factors were measured at T1 or reported retrospectively (e.g., childhood experience of abuse). Following Hosmer and Lemeshow (2000), all independent variables were first examined in univariate regressions and retained for multivariate regressions if $p < 0.25$ (Step 1). In multivariate regressions, all independent variables with $p > 0.1$ were subsequently removed (Step 2), and then all independent variables with $p > 0.05$ were removed until all independent variables were significant at $p > 0.05$ (Step 3). These models were estimated using poisson regression to accommodate for the use of count variables. Zero-inflation with Vuong's test was used to adjust for the skewed distribution of some of the BCAPE scores where necessary (Fig. 1) (Li et al., 1999).

3. Results

3.1. Characteristics of the sample

Characteristics of the sample are presented in Table 1 and 2. Differences between mothers and fathers were observed with regards to migration history, employment, caregiving arrangements for children, mental health, adverse experiences in childhood, and BCAPE Abuse Scores (Table 2).

3.2. Validity indices

Validity indices were calculated for both mothers and fathers. Those with a score > 3 on the Lie Scale and a score > 2 on the Random Response Scale were excluded (Walker & Davies, 2012). Participants who scored > 3 on the Lie Scale and > 12 on the BCAPE Abuse Risk Scale were retained (Milner, 1986). 25% ($n = 50$) of mothers and 30% ($n = 59$) of fathers had invalid protocols. This resulted in a sample of 147 mothers and 138 fathers with valid protocols that was included in the analyses. The difference in invalid protocols between mothers and fathers was not significant. Using the Random Response Cut-Off > 0 from Ondersma et al. (2005) would have produced much higher rates of invalid protocols in mothers (41%) and fathers (43%).

Mothers with invalid protocols were more likely to have a history of migration ($\chi^2(1,189) = 3.77, p < .05$). No statistically significant differences between valid and invalid protocols for mothers were found on the BCAPE Abuse Risk Score ($t(190) = .56, p = .58$), parental role distribution ($t(174) = .42, p = .77$), and dissatisfaction with parental role distribution ($t(174) = -.30, p = .77$), parental relationship unhappiness ($t(179) = .51, p = .61$), conflicts in co-parenting ($t(161) = -.78, p = .44$), perceived stress ($t(189) = -.23, p = .82$), parental self-efficacy ($t(190) = -1.91, p = .06$), maternal anxiety ($\chi^2(1,191) = 1.39, p = .18$), depression ($\chi^2(1,187) = .07, p = .53$), alcohol abuse ($\chi^2(1,191) = .02, p = .59$), adverse childhood experiences ($\chi^2(1,176) = .10, p = .46$) or reported family violence ($\chi^2(1,178) = .113, p = .21$).

Fathers with invalid protocols were more likely to have higher child abuse potential (BCAPE Abuse Risk Scale; $U(130/59) = 2279.50, p < .001$), dissatisfaction with parental role distribution ($t(183) = 3.00, p < .01$), parental relationship unhappiness ($U(122/54) = 2097.50, p < .001$), conflicts in co-parenting ($U(115/53) = 2162.00, p < .01$), perceived stress ($t(183) = 3.40, p < .01$), less parental self-efficacy ($U(127/59) = 2381.00, p < .001$) and paternal depression ($\chi^2(1,182) = 9.36, p < .01$). No statistical differences between valid and invalid protocols in fathers were found on parental role distribution ($t(96) = -.70, p = .48$), migration history ($\chi^2(1,163) = 3.24, p < .10$), anxiety ($\chi^2(1,183) = .98, p = .28$), alcohol abuse ($\chi^2(1,187) = 1.05, p = .24$), adverse childhood experiences ($\chi^2(1,171) = .97, p = .24$) and reported family violence ($\chi^2(1,174) = 2.23, p = .11$).

These results suggest that fathers at higher risk for abusing their child may be more likely to have invalid protocols. According to testing by psychosocial risk groups, fathers in the group at high risk were more likely to have invalid protocols than those in the medium- or low-risk groups (44.6% versus 33.8% versus 21.5%, $\chi^2(2,189) = 7.65, p < .05$). There was no statistically significant difference for mothers. All subsequent analyses presented data using valid protocols only.

3.3. Characteristics of the data

Prior to factor analysis testing, data were assessed for multicollinearity using the variance inflation factor (VIF). For fathers and mothers, no variables fulfilled the assumptions for multicollinearity ($VIF > 10$). However, 11 variables for fathers and two variables for mothers correlated very highly with each other ($r > .8$), and thus these were kept under observation in the analyses (Field, 2009). Distribution of the data across the BCAPE Abuse Risk Scale was assessed. The data were non-normally distributed with a high number of zero scores for both mothers and fathers, particularly at baseline.

3.4. Confirming the six-factor solution

Walker and Davies (2012) established a six-factor structure in a sample of British parents (92.3% female) that was confirmed in a

² The reduced BCAPE score is a sum score of all the items confirmed in the CFA.

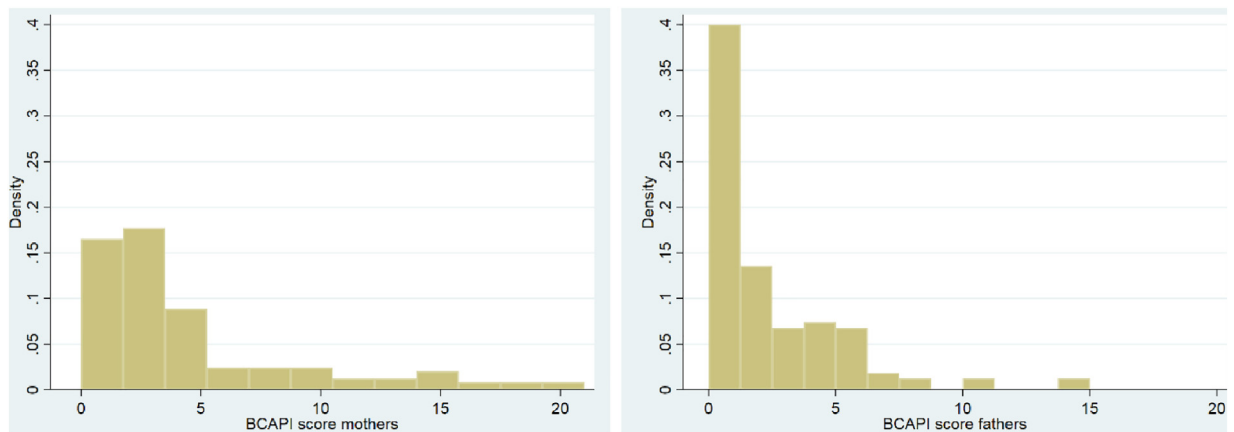


Fig. 1. Histograms for BCAPI (valid protocols).

sample of Australian drug-using mothers (Dawe et al., 2017). Thus, Confirmatory Factor Analysis in MPLUS using the WLSMV robust estimator was employed to confirm the factor structure in this current sample. Analyses were run separately for fathers and mothers.

For mothers, two items were removed due to empty cells related to selection of the response option (all participants selected “disagree”), and one item was removed due to low factor loadings ($< .4$) (Bowen & Guo, 2012). After removal of these items, the six-factor structure suggested by Walker and Davies could be confirmed. Model fit for the final model was $\chi^2/df = 1.17$, $p = .062$ for CMIN 203.586, $df = 174$; RMSEA .034 ($p = .914$), WRMR .727, CFI .992, TLI .991 (Table 3). High correlations between all factors except rigidity could be observed (Table 3).

For fathers, 13 items had to be removed due to empty cells and low variance (all participants selected “disagree”). This eliminated two factors and resulted in a non-positive definite covariance matrix. The six-factor structure identified by Walker and Davies (2012) could not be confirmed among fathers in this sample and thus no further analyses were conducted on the paternal data.

With regards to psychometric properties, these analyses confirmed the six-factor structure of the BCAPI for mothers but not for fathers. This demonstrates potential differences in suitability of the measure among parents. Taking into account that fathers reported lower abuse scores than mothers, this could indicate that BCAPI items are not gender-sensitive for fathers. It is possible that the wording results in inner rejection by fathers. While the measure has been successfully used in mothers in other populations and factor structure could be replicated, use in fathers is questionable based on these data and requires further research.

3.5. Internal consistency and concurrent validity

KR20 was computed to test internal consistency. For mothers, internal consistency of the complete BCAP abuse risk scale (KR20 = .79), the reduced BCAP abuse scale (KR20 = .78), loneliness (KR20 = .72), distress (KR20 = .73), happiness (KR20 = .80), and family conflict ($\alpha = .79$) was good. Internal consistency was poor for impact of others (KR20 = .66), rigidity ($\alpha = .42$) and the lie scale (KR20 = .50), and it was unacceptable for the random response scale (KR20 = .11).

To investigate concurrent criterion validity of the BCAPI Abuse Risk Scale, correlations between the BCAPI reduced scale, the BCAPI sub-scales, and empirically and theoretically associated measures were conducted for mothers (Table 4). For mothers, moderate to strong correlations were observed between BCAPI scores and all hypothesized factors except for family violence and adverse childhood experience, both of which showed small correlations.

3.6. Predictors of child abuse potential in mothers

As the BCAPI Abuse scale is often used in clinical practice, we present two sets of regression results here: those using the full BCAPI Abuse Scale and those using the reduced BCAPI Abuse Score resulting from the CFA. Results for Hosmer & Lemeshow's Step 2, Step 3 and Step 4 (where applicable) of the zero-inflated poisson regression are presented for mothers (Table 5). Step 2 of the analysis includes sub-sets of significant risk factors for mothers identified through correlations.

For mothers, baseline BCAPI Abuse Score indicated that perceived stress and alcohol abuse increased child abuse potential at follow-up. An additional risk factor for increased child abuse potential using the full BCAPI score was dissatisfaction with parental roles. Higher education status at baseline decreased child abuse potential at follow-up (Table 5).

4. Discussion

This study described the first psychometric evaluation of the BCAPI in a sample of mothers and fathers that was not predominately English speaking. It found significantly higher BCAPI mean scores among mothers than fathers, and a higher number of invalid protocols among fathers than mothers. The hypothesized 6-factor structure could be confirmed for mothers but not fathers. Internal

Table 2
Socio-demographic characteristics of the German KiD 0–3 in-depth study sample.

N = 197			Study sample n (%)
Child Characteristics			
Age [M (SD)]			[1.27 (.30)]
Gender female			98 (49.7)
Household Characteristics			
Social welfare receipt			28 (17.7)
Overcrowded home			9 (6.3)
Child cared in day nursery			41 (20.9)
Household members [M (SD)]			[3.59 (.86)]
Housing space in m ² [M (SD)]			[103.70 (35.28)]
N = 197/191	Mothers n (%)	Fathers n (%)	Difference p
Caregiver Characteristics			
Age [M (SD)]	[32.90 (5.01)]	[36.07 (9.31)]	< .001
Migration history	62 (32.1)	42 (25.0)	ns
Primary caregiver	191 (97.0)	6 (3.1)	< .001
Cohabiting parents	183 (93.9)	181 (94.8)	ns
Single parent	4 (2.0)		ns
Homosexual parent	2 (1.0)		ns
Non-biological parent		2 (1.0)	ns
Educational Level			
University degree	84 (43.1)	79 (42.0)	ns
High school degree or equivalent	49 (25.1)	43 (22.9)	
Lower degree	39 (19.8)	56 (29.8)	
None	23 (11.8)	10 (5.3)	
Employment			
Economically active	63 (33.2)	151 (81.2)	
Parental leave	90 (47.4)	19 (10.2)	< .001
Non-working/unemployed	37 (19.4)	16 (8.6)	
Psychiatric Symptoms			
Anxiety	15 (7.9)	6 (3.2)	< .10
Depression	17 (9.1)	13 (7.1)	ns
Alcohol Abuse	7 (3.7)	11 (5.9)	ns
Adverse Experience			
Parents' Childhood (ACE)	28 (15.7)	16 (9.2)	< .10
Family Violence	22 (12.4)	22 (12.6)	ns
Wellbeing			
Perceived Stress [M (SD)]	[10.44 (3.58)]	[10.05 (3.15)]	ns
Couple Dissatisfaction [M (SD)]	[4.03 (3.27)]	[3.43 (2.62)]	< .10
Parenting			
Parenting Role Distribution [M (SD)]	[3.37 (.89)]	[3.65 (.84)]	< .01 ^a
Dissatisfaction Parenting Roles [M (SD)]	[2.22 (1.02)]	[1.95 (.91)]	< .01
Co-parenting [M (SD)]	[3.98 (3.60)]	[4.11 (3.39)]	ns
Parenting Self-Efficacy [M (SD)]	[90.44 (10.95)]	[89.75 (11.03)]	ns
Child Abuse Potential			
BCAPI Abuse Score at time 1 [M (SD)]	[4.46 (4.41)]	[2.40 (2.94)]	< .001
BCAPI Abuse Score at time 2 [M (SD)]	[4.13 (4.39)]	[2.62 (3.09)]	< .001

^a dependent samples testing.

reliability and concurrent criterion validity were adequate for mothers. Longitudinal analyses examining risk factors for child abuse potential in mothers showed some hypothesized results.

Only one study had examined differences in BCAPE mean scores between fathers and mothers and reported no differences (Rodríguez et al., 2017). Previous research on the CAPI found that mothers who were married or in relationships had higher CAPI abuse mean scores than fathers, whereas single fathers had higher CAPI scores than single mothers (Miragoli et al., 2015). Further research is needed with samples containing both caregivers and different population groups to investigate under which circumstances fathers or mothers may have a higher risk for child abuse potential.

This study found lower percentages of invalid study protocols compared to Walker and Davies (2012), but the number of invalid protocols was still high (25% for mothers; 30% for fathers). Among mothers, those with invalid protocols were more likely to have a history of migration. Among fathers, those with invalid protocols were found to have higher risk for child maltreatment. These fathers had higher BCAPE scores, elevated stress and depression rates, and were more dissatisfied with their relationship with the mother, including co-parenting and role distribution. This is concerning as the BCAPE is often used by clinicians and thus does not identify fathers at the highest risk of maltreating their children unless the validity indices are taken into consideration during screening. As no other studies have used the BCAPE with a population of fathers, future research is needed to examine whether these findings are

Table 3

Six-factor solution of the BCAP as per Walker and Davies (2012) amongst German mothers and correlations among factors (n = 147).

Factor structure	Loneliness	Distress	Happiness	Conflict	Impact of Others	Rigidity
I often feel lonely inside	.875*					
Sometimes I feel all alone in the world	.910*					
I often feel worthless	.943*					
I often feel very alone	.988*					
I often feel alone	.986*					
I often feel very upset		.795*				
I am easily upset by my problems		.812*				
I am often upset and I don't know why		.611*				
I am often upset		.740*				
I am often depressed		.904*				
My life is good			-.914*			
I am a happy person			-.981*			
My life is happy			-.945*			
My family fights a lot				.956*		
My family has problems getting along				.842*		
My family has many problems				.961*		
Other people have made my life hard					.883*	
I often feel that my needs are not met					.722*	
Other people have made my life unhappy					.979*	
Everything should be in its place at my home						.892*
Children should never disobey						.778*
Correlations among factors	Loneliness	Distress	Happiness	Conflict	Impact of Others	Rigidity
Loneliness	1	.887*	-.910*	.818*	.830*	-.168
Distress		1	-.845*	.706*	.592*	-.291*
Happiness			1	-.759*	-.755*	-.129
Conflict				1	.747*	-.268
Impact of Others					1	-.225
Rigidity						1

*significance at $p < .05$ CFA reports standardized coefficients.Model fit: $\chi^2/df = 1.17$, $p = .062$ for CMIN 203.586, $df = 174$; RMSEA 0.034 (.00–.053), WRMR 0.727, CFI .992, TLI .991.**Table 4**

Correlations between BCAP scores, subscales and empirically and theoretically associated measures for mothers.

	r_s	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 BCAP Score	1																		
2 BCAP CFA	.97**	1																	
3 BCAP Distress	-.74**	.77**	1																
4 BCAP Happiness	-.55**	-.44**	-.48**	1															
5 BCAP Family Conflict	.59**	.60**	.42**	-.44**	1														
6 BCAP Impact of others	.61**	.65**	.37**	-.36**	.37**	1													
7 BCAP Rigidity	.18*	.14	-.15	-.06	-.13	-.16	1												
8 BCAP Loneliness	.73**	.74**	.53**	-.63**	.47**	.57**	-.08	1											
9 Depression	.40**	.43**	.33**	-.33**	.45**	.37**	-.14	.48**	1										
10 Anxiety	.41**	.42**	.37**	-.33**	.41**	.28**	-.02	.36**	.73*	1									
11 Alcohol Abuse	.02	.03	.04	-.04	.09	.03	-.07	.06	.20*	.07	1								
12 Perceived Stress	.63**	.63**	.57**	-.43**	.47**	.43**	.02	.50**	.36*	.41*	.00	1							
13 Couple Dissatisfaction	.44**	.43**	.33**	-.32**	.35**	.30**	.04	.36**	.30**	.25*	-.04	.46**	1						
14 Dissatisfaction with Parenting	.44**	.40**	.31**	-.33**	.26*	.24*	.08	.39**	.19*	.15	-.02	.39**	.54**	1					
15 Roles	.44**	.44**	.47**	-.36**	.37**	.30**	-.11	.35**	.15	.14	-.02	.52**	.61**	.42**	1				
16 Co-Parenting	-.49**	-.48**	-.49**	.30**	-.30**	-.31**	.03	-.40**	-.24*	-.25*	-.09	-.59**	-.36**	-.32**	-.47**	1			
17 Efficacy	.38**	.41**	.29**	-.31**	.50**	.29**	-.05	.36**	.31*	.45*	.14	.20*	.07	.14	-.01	-.09	1		
18 Adverse Experiences	.28**	.29**	.23*	-.11	.28**	.29**	-.17	.25**	.16	.14	.06	.26*	.18*	.17*	.16	-.20*	.21*	1	

* $p < .05$.** $p < .01$.

replicable.

This study was able to confirm the six-factor structure of the BCAP among German mothers previously identified in the literature (Dawe et al., 2017; Walker & Davies, 2012). The six-factor structure could not be confirmed for fathers, mostly because insufficient variance for several items required removal of these. In the worst cases, insufficient variance led to empty cells, which means that for some variable combinations, certain dimensions have zero observations. These issues can make the use of robust methods, such as factor analysis for extracting the dimensions of child abuse potential, impossible. While this problem might be specific to this dataset, it might also point to an insufficient sensitivity of the questionnaire items in detecting underlying risk factors that contribute to child abuse potential, particularly among fathers. Future research should investigate whether insufficient variance is a specific problem with fathers also in other cultural contexts and languages and requires adaptation of the measure to make it suitable for clinical use with fathers.

Internal consistency of the BCAP and its sub-scales among mothers was good except for the impact of others and rigidity sub-

Table 5
Risk and Protective Factors of BCAPi Abuse Score at follow-up in mothers using multivariate (zero-inflated) poisson regressions.

Predictors	Mothers BCAPi-Abuse Score at follow-up (n = 142) (zero-inflated)			Reduced mothers BCAPi Abuse Score at follow-up (not zero-inflated) (n = 142)		
	Step 2	Step 3	Step 4	Step 2	Step 3	Step 4
	β , SE, 95% CI	β , SE, 95% CI	β , SE, 95% CI	β , SE, 95% CI	β , SE, 95% CI	β , SE, 95% CI
BCAPI Abuse Score at baseline	.04 (SE .02) (–.00 to .08) +	.05 (SE .01) (.03 to .08)***	.05 (SE .01) (.02 to .08)***	.05 (SE .02) (.01 to .09)***	.08 (SE .16) (.09 to .70)*	
Adverse Childhood Experiences	.18 (SE .17) (–.16 to .52)	Removed in Step 2	Removed in Step 2	.17 (SE .15) (–.12 to .46)	Removed in Step 2	
Perceived Stress	.10 (SE .02) (.06 to .14)***	.10 (SE .02) (.06 to .14)***	.08 (SE .02) (.05 to .25)***	.07 (SE .02) (.03 to .11)***	.06 (SE .01) (.04 to .09)***	
Anxiety	Removed in Step 1			Removed in Step 1		
Depression	Removed in Step 1			Removed in Step 1		
Alcohol Abuse	.71 (SE .23) (.26 to 1.16)**	.78 (SE .20) (.39 to 1.17)	.69 (SE .19) (.32 to 1.05)***	.47 (SE .21) (.06 to .88)*	.58 (SE .15) (.28 to .89)***	
Parenting Self-Efficacy	.01 (SE .00) (–.00 to .02) +	–.01 (SE .01) (–.00 to .02)	Removed in Step 3	–.00 (SE .00) (–.01 to .01)	Removed in Step 2	
Co-Parenting: Conflict	Removed at Step 1			Removed at Step 1		
Parenting Role Distribution	.15 (SE .10) (–.05 to .35)	Removed in Step 2		–.10 (SE .07) (–.22 to .02)	Removed in Step 2	
Unhappiness with Partnership	.01 (SE .02) (–.03 to .06)	Removed in Step 2		.02 (SE .02) (–.02 to .06)	Removed in Step 2	
Dissatisfaction with Parenting Roles	.23 (SE .09) (.09 to .44)**	.15 (SE .05) (.05 to .25)***	.15 (SE .01) (.05 to .25)**	Removed at Step 1		
Education: Degree	–.51 (SE .13) (–.76 to –.26)**	–.35 (SE .11) (–.56 to –.14)**	–.34 (SE .11) (–.55 to –.13)**	–.23 (SE .11) (–.44 to –.02)*	–.19 (SE .09) (–.27 to –.02)*	
Employed	Removed at Step 1			Removed at Step 1		
Household Income	.00 (SE .00) (–.00 to 0.00)	Removed in Step 2		–.00 (SE .00) (–.00 to 0.00)	Removed in Step 2	
Child Gender	Removed at Step 1			Removed at Step 1		
Child Age	.02 (SE .13) (–.23 to 27)	Removed at Step 2		Removed at Step 1		
Mother Age	Removed at Step 1			Removed at Step 1		
LR χ^2	163.23***	221.12***	219.40***	143.53***	230.74***	
AIC	376.97	434.60	428.05	415.92	492.53	
BIC	418.49	458.22	452.44	444.35	506.39	
Vuong/Pseudo r^2	3.33**	3.16**	3.16**	Pseudo r^2 .27	Pseudo r^2 .32	

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

scales. Internal consistency for the validity indices in this study was unacceptable. Previous research on the BCAP and the CAPI has identified low levels of internal consistency on the validity scales. This is not surprising given the small number of items and the phrasing of the questions, which do not suggest associations between items (Haz & Ramírez, 2002; Walker & Davies, 2012). Poor internal validity on some of the other sub-scales may be related directly to low numbers of items within the sub-scales despite moderate inter-item correlations (Tavakol & Dennick, 2011).

Concurrent validity was also assessed by examining associations among commonly correlated constructs with child abuse potential. Concurrent validity was found to be adequate for mothers. Notably, the association of BCAP with family violence was small among mothers, suggesting either social desirability bias on the family violence section of the questionnaire or a weakness of the BCAP in relation to accurate identification of abusive parents. Future research may valuably use measures which measure actual incidents of abusive behavior rather than underlying risk factors.

4.1. Risk factors for child abuse potential using sum scores

This paper identified significant and substantive differences with regards to risk factors for child abuse potential using the traditionally calculated BCAP abuse scale sum score versus the reduced sum score based on CFA results. Both scores were examined in this study as clinicians in practice use the full sum score of the whole 24-item measure.

For mothers, we find clear associations between BCAP Abuse Scores at baseline and BCAP Abuse Scores at T2 (for both full BCAP and reduced BCAP), suggesting that these scores tend to remain more or less stable over time without intervention (here, 7 months). For mothers, dissatisfaction with parental role distribution (risk), alcohol abuse (risk) and higher education level (protective) were predictive of full BCAP score at follow-up over and above baseline BCAP score. Baseline BCAP score, alcohol abuse and higher education level were predictive of the reduced BCAP score at follow-up.

Interestingly, presence of more than three adverse childhood experiences did not predict any BCAP scores at follow-up. A large body of previous research has assumed linkages between childhood victimization and perpetration of abuse in parenthood (Deater-Deckard, Lansford, Dodge, Pettit, & Bates, 2003; McCloskey & Bailey, 2000). This view has become more nuanced in recent years, however, with research demonstrating surveillance or detection bias for families with prior involvement in child protective services as well as the moderating effects of safe, stable and nurturing relationships in breaking the intergenerational transmission of child abuse (Schofield & Lee, 2013; Widom et al., 2015). Adverse childhood experiences have been found to be associated in particular with poor mental health outcomes (Hughes et al., 2017), which again have been identified as a risk factor for child maltreatment (Meinck, Cluver, Boyes, & Mhlongo, 2015).

Alcohol abuse at T1 was identified as a specific risk factor for mothers in the present sample, replicating a large corpus of existing research (e.g., Walsh, MacMillan, & Jamieson, 2002; Widom & Hiller-Sturmhofel, 2001). A further predictor specific to maternal child abuse potential was higher education status. This was shown to have an inverse relationship and thus associated with lower BCAP Abuse Scores. Thus far, little research has been conducted on the impacts of education on child abuse potential, but research has consistently shown the protective effects of secondary education on harsh parenting in adulthood (World Health Organization, 2002). A recent study from Italy found that fathers with a university degree had markedly lower CAPI Abuse Scores than those who had completed high school and those with only some schooling (Miragoli et al., 2015). Further research is needed to examine the role of parental education for child abuse prevention.

4.2. Limitations

This study is subject to a number of limitations. First, this study is based on a single dataset. Due to small sample size it was impossible to split the sample into a dataset to conduct EFA and CFA processes. As CFA is considered superior and more robust when testing a hypothesized factor structure, a CFA approach was used (Byrne, 2005). Further, this study had a sample size which exceeded that of previous studies (Dawe et al., 2017; Tucker, 2014). Second, the psychometric testing on a reduced dataset could not take the dyadic nature of the data into account. The number of invalid protocols for this study was so high that analyses were conducted on valid protocols for fathers and mothers independent of each other. Third, zero variance was particularly high among paternal BCAP reports effectively making the data unsuitable for further analyses. Whether this is a general problem with paternal reporting on the BCAP or specific to this dataset needs to be examined in further research. Fourth, no other validated measures of child abuse potential were part of the study; thus, no criterion validity testing could be carried out against another child abuse potential measure. As this study was part of a larger study on children at risk, participant burden had to be taken into account during data collection limiting the amount of measures included. Fourth, data collection utilized parent self-report which may be subject to desirability bias. However, the BCAP was specifically developed to measure less stigmatizing indicators of child abuse potential rather than actual child abuse occurrence.

Despite these limitations, this is the first study investigating the properties of the BCAP in a German sample. While the investigation showed adequate psychometric properties for mothers, a number of measurement problems could be identified in relation to fathers which necessitates future research.

Taking into account the results of the present study, a clinical use of the BCAP in fathers is not recommended as it might produce data that are hard to interpret. As the BCAP is a relatively new instrument, research thus far has focused on predominantly female samples and has not investigated gender differences with regards to internal consistency or construct validity. Future research is needed to establish whether these poor results for fathers are due to differences in measurement functioning for fathers in general or just for this specific population.

Author contributions: CL, KL and AE were involved in the overall design and management of the study and data collection. CL and FM had responsibility for conceptualizing and writing the paper. FM led the analyses. FM and JS conducted the analyses. CL, HK and AE contributed to the analyses and interpretation of findings. FM and JS wrote the manuscript with input from CL, HK, KL and AE. All authors reviewed and approved the final version.

Acknowledgments

The KiD 0-3 prevalence studies on psychosocial burden were conducted by the National Centre for Early Prevention (NZFH) and funded by the Federal Ministry of Family Affairs, Senior Citizens, Women and Youth (BMFSFJ) as part of the Federal Initiative on Early Prevention from 2013 to 2015. The KiD 0-3 In-Depth study was conducted in cooperation with the University of Erlangen-Nuremberg (Gottfried Spangler), University of Wuppertal (Peter Zimmermann) and Bielefeld University (Marc Vierhaus). The authors wish to thank the participants, researchers from NZFH (Christian Brand, Andrea Schreier and Alexandra Sann) and University of Oxford (Janina Jochim), the involved partner organisations, and Jennifer Rabedeau for proof-reading and editing. FM received additional support from the Economic and Social Research Council (ESRC; Grant No ES/N017447/1) and the University of Oxford's ESRC Impact Acceleration Account (1602-KEA-189; 1311-KEA-004 & 1609-GCRF-227).

References

- Bowen, N., & Guo, S. (2012). *Structural equation modeling*. Oxford: Oxford University Press.
- Brewster, A. L., Nelson, J. P., McCanne, T. R., Lucas, D. R., & Milner, J. S. (1998). Gender differences in physiological reactivity to infant cries and smiles in military families. *Child Abuse & Neglect*, 22(8), 775–788. [https://doi.org/10.1016/S0145-2134\(98\)00055-6](https://doi.org/10.1016/S0145-2134(98)00055-6).
- Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment*, 85(1), 17–32. https://doi.org/10.1207/s15327752jpa8501_02.
- Chan, Y. C., Lam, G. L. T., Chun, P. K. R., & So, M. T. E. (2006). Confirmatory factor analysis of the Child Abuse Potential Inventory: Results based on a sample of Chinese mothers in Hong Kong. *Child Abuse & Neglect*, 30(9), 1005–1016. <https://doi.org/10.1016/j.chiabu.2006.05.005>.
- Chu, J. J., Jahn, H. J., Khan, M. H., & Kraemer, A. (2016). Alcohol consumption among university students: A Sino-German comparison demonstrates a much lower consumption of alcohol in Chinese students. *Journal of Health, Population, and Nutrition*, 35(1), 25. <https://doi.org/10.1186/s41043-016-0062-0>.
- Chua, B., & Abdul, H. (2001). The influence of ethnicity, gender, personality, life satisfaction and number of children in the family on child abuse potential. *Hong Kong Journal of Paediatrics*, 6(1), 40–43.
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385–396. <https://doi.org/10.2307/2136404>.
- Cowan, C. P., & Cowan, P. A. (1988). Who does what when partners become parents. *Marriage & Family Review*, 12(3–4), 105–131. https://doi.org/10.1300/J002v12n03_07.
- Dawe, S., Taplin, S., & Mattick, R. P. (2017). Psychometric investigation of the Brief Child Abuse Potential Inventory in mothers on opioid substitution therapy. *Journal of Family Violence*, 32(3), 341–348. <https://doi.org/10.1007/s10896-016-9821-3>.
- Deater-Deckard, K., Lansford, J. E., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2003). The development of attitudes about physical punishment: An 8-year longitudinal study. *Journal of Family Psychology*, 17(3), 351–360. <https://doi.org/10.1037/0893-3200.17.3.351>.
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. SAGE Publications.
- Diareme, S., Tsiantis, J., & Tsioura, S. (1997). Cross-cultural validation of the child abuse potential inventory in Greece: A preliminary study. *Child Abuse & Neglect*, 21(11), 1067–1079. [https://doi.org/10.1016/S0145-2134\(97\)00066-5](https://doi.org/10.1016/S0145-2134(97)00066-5).
- Eickhorst, A., Brand, C., Lang, K., Liel, C., Neumann, A., Schreier, A., ... Sann, A. (2015). [The KiD 0-3 study on the prevalence of psychosocial burden and early prevention in families with children aged 0-3: Study design and potential of analyses]. *Soziale Passagen*, 7(2), 381–387. <https://doi.org/10.1007/s12592-015-0212-z>.
- Eickhorst, A., Schreier, A., Brand, C., Lang, K., Liel, C., Renner, I., et al. (2016). [Knowledge and use of different support programs in the context of early prevention in relation to family-related psychosocial burden]. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 59(10), 1271–1280. <https://doi.org/10.1007/s00103-016-2422-8>.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE Publications.
- Finkelhor, D., Hamby, S. L., Ormrod, R., & Turner, H. (2005). The Juvenile Victimization Questionnaire: Reliability, validity, and national norms. *Child Abuse & Neglect*, 29(4), 383–412. <https://doi.org/10.1016/j.chiabu.2004.11.001>.
- German Youth Institute (2013). *Questionnaire templates of the AID: A II Study. Target population 0 to 8 years*. Munich.
- Gibaud-Wallaston, J., & Wandersman, L. (1978). *Development and utility of the parenting sense of competence scale*. *The Annual Meeting of the American Psychological Association*.
- Haz, A., & Ramirez, V. (1998). Preliminary validation of the child abuse potential inventory in Chile. *Child Abuse & Neglect*, 22(9), 869–879. [https://doi.org/10.1016/S0145-2134\(98\)00071-4](https://doi.org/10.1016/S0145-2134(98)00071-4).
- Haz, A., & Ramirez, V. (2002). Adaptation of Child Abuse Potential Inventory in Chile: Analysis of the difficulties and challenges in the application in Chilean studies. *Child Abuse & Neglect*, 26(5), 481–495. [https://doi.org/10.1016/S0145-2134\(02\)00324-1](https://doi.org/10.1016/S0145-2134(02)00324-1).
- Henschel, S., de Bruin, M., & Möhler, E. (2013). Self-control and child abuse potential in mothers with an abuse history and their preschool children. *Journal of Child and Family Studies*. <https://doi.org/10.1007/s10826-013-9735-0>.
- Hillis, S., Mercy, J., Amobi, A., & Kress, H. (2016). Global prevalence of past-year violence against children: A systematic review and minimum estimates. *Pediatrics*, 137(3), e20154079. <https://doi.org/10.1542/peds.2015-4079>.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/0471722146>.
- Hughes, K., Bellis, M. A., Hardcastle, K. A., Sethi, D., Butchart, A., Milton, C., ... Dunne, M. P. (2017). The effect of multiple adverse childhood experiences on health: A systematic review and meta-analysis. *The Lancet Public Health*, 2(8), e356–e366. [https://doi.org/10.1016/S2468-2667\(17\)30118-4](https://doi.org/10.1016/S2468-2667(17)30118-4).
- Kawamura, Y., Takahashi, Y., Akiyama, T., Sasaki, T., & Kako, M. (2009). Reliability and validity of the Japanese version of the child abuse potential inventory abuse scale. *Asia-Pacific Psychiatry*, 1(3), 138–142. <https://doi.org/10.1111/j.1758-5872.2009.00033.x>.
- Kelley, M. L., Lawrence, H. R., Millettich, R. J., Hollis, B. F., & Henson, J. M. (2015). Modeling risk for child abuse and harsh parenting in families with depressed and substance-abusing parents. *Child Abuse & Neglect*, 43, 42–52. <https://doi.org/10.1016/j.chiabu.2015.01.017>.
- Klein, E. M., Brähler, E., Dreier, M., Reinecke, L., Müller, K. W., Schmutzer, G., ... Tosti, N. (2016). The German version of the Perceived Stress Scale – psychometric characteristics in a representative German community sample. *BMC Psychiatry*, 16(1), 159. <https://doi.org/10.1186/s12888-016-0875-9>.
- Li, C.-S., Lu, J.-C., Park, J., Kim, K., Brinkley, P. A., & Peterson, J. P. (1999). Multivariate zero-inflated poisson models and their applications. *Technometrics*, 41(1), 29–38. <https://doi.org/10.2307/1270992>.
- Liel, C. (2018). *Väter und familiäre Gewalt [Fathers and family violence]*. Munich: Ludwig-Maximilians-Universität.
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., ... Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274. <https://doi.org/10.1097/MLR.0b013e318160d093>.

- Löwe, B., Kroenke, K., Herzog, W., & Graefe, K. (2004). Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders*, 81, 61–66. [https://doi.org/10.1016/S0165-0327\(03\)00198-8](https://doi.org/10.1016/S0165-0327(03)00198-8).
- Löwe, B., Spitzer, R., Zipfel, S., & Herzog, W. (2002). *PHQ-D Gesundheitsfragebogen für Patienten*. Heidelberg: Manual Komplettversion und Kurzform.
- McCloskey, L., & Bailey, J. (2000). The intergenerational transmission of risk for child sexual abuse. *Journal of Interpersonal Violence*, 15, 1019–1035. <https://doi.org/10.1177/088626000015010001>.
- Meinck, F., Cluver, L., Boyes, M., & Mhlongo, E. (2015). Risk and protective factors for physical and sexual abuse of children and adolescents in Africa: A review and implications for practice. *Trauma, Violence & Abuse*, 16(1), 81–107. <https://doi.org/10.1177/1524838014523336>.
- Meinck, F., Steinert, J. I., Sethi, D., Gilbert, R., Bellis, M., Mikton, C., ... Baban, A. (2016). *Measuring and monitoring national prevalence of child maltreatment: A practical handbook*. Copenhagen: World Health Organization Regional Office for Europe.
- Milner, J. (1986). *The Child Abuse Potential Inventory: Manual*. Webster, NC.
- Milner, J. (1994). Assessing physical child abuse risk: The child abuse potential inventory. *Clinical Psychology Review*, 14(6), 547–583. [https://doi.org/10.1016/0272-7358\(94\)90017-5](https://doi.org/10.1016/0272-7358(94)90017-5).
- Milner, J., & Ayoub, C. (1980). Evaluation of “At Risk” parents using the child abuse potential inventory. *Journal of Clinical Psychology*, 36(4), 945–948. [https://doi.org/10.1002/1097-4679\(198010\)36:4](https://doi.org/10.1002/1097-4679(198010)36:4).
- Miragoli, S., Camisasca, E., & Blasio, P. D. I. (2015). Validation of the Child Abuse Potential Inventory in Italy. *SAGE Open*, 5(3).
- Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: A systematic review and meta-analysis. *PLoS Medicine*, 9(11), e1001349. <https://doi.org/10.1371/journal.pmed.1001349>.
- Ondersma, S. J., Chaffin, M. J., Mullins, S. M., & LeBreton, J. M. (2005). A Brief Form of the Child Abuse Potential Inventory: Development and validation. *Journal of Clinical Child & Adolescent Psychology*, 34(2), 301–311. https://doi.org/10.1207/s15374424jccp3402_9.
- Pedersen, F., Suwalsky, J., Cain, R., & Zaslow, M. (1987). Paternal care of infants during maternal separations: Associations with father-infant interaction at one year. *Psychiatry*, 50(3), 193–205. <https://doi.org/10.1080/00332747.1987.11024352>.
- Pittman, J. F., & Buckley, R. R. (2006). Comparing maltreating fathers and mothers in terms of personal distress, interpersonal functioning, and perceptions of family climate. *Child Abuse & Neglect*, 30(5), 481–496. <https://doi.org/10.1016/j.chiabu.2004.10.017>.
- Radford, L., Corral, S., Bradley, C., & Fisher, H. L. (2013). The prevalence and impact of child maltreatment and other types of victimization in the UK: Findings from a population survey of caregivers, children and young people and young adults. *Child Abuse & Neglect*, 37, 801–813. <https://doi.org/10.1016/j.chiabu.2013.02.004>.
- Rodriguez, C., Baker, L., Pu, D., & Tucker, M. (2017). Predicting parent-child aggression risk in mothers and fathers: Role of emotion regulation and frustration tolerance. *Journal of Child and Family Studies*. <https://doi.org/10.1007/s10826-017-0764-y>.
- Romero-Martínez, A., Figueiredo, B., & Moya-Albiol, L. (2014). Childhood history of abuse and child abuse potential: The role of parent's gender and timing of childhood abuse. *Child Abuse & Neglect*, 38(3), 510–516. <https://doi.org/10.1016/j.chiabu.2013.09.010>.
- Sabourin, S., Valois, P., & Lussier, Y. (2005). Development and validation of a brief version of the dyadic adjustment scale with a nonparametric item analysis model. *Psychological Assessment*, 17(1), 15–27. <https://doi.org/10.1037/1040-3590.17.1.15>.
- Schofield, T. J., & Lee, R. D. (2013). Safe, stable, nurturing relationships as a moderator of intergenerational continuity of child maltreatment: A meta-analysis. *Journal of Adolescent Health*, 53(4), 32–38. <https://doi.org/10.1016/j.jadohealth.2013.05.004>.
- Solmeyer, A. R., & Feinberg, M. E. (2011). Mother and father adjustment during early parenthood: The roles of infant temperament and coparenting relationship quality. *Infant Behavior & Development*, 34(4), 504–514. <https://doi.org/10.1016/j.infbeh.2011.07.006>.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, 38(1), 15. <https://doi.org/10.2307/350547>.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 1, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>.
- Tucker, M. (2014). *Personal and couple level risk factors: Maternal vs. paternal child abuse risk*. University of North Carolina at Greensboro.
- Vasquez Guerrero, D. A. (2009). Hypermasculinity, intimate partner violence, sexual aggression, social support, and child maltreatment risk in urban, heterosexual fathers taking parenting classes. *Child Welfare*, 88(4), 135–155.
- Walker, C. A., & Davies, J. (2010). A critical review of the psychometric evidence base of the Child Abuse Potential Inventory. *Journal of Family Violence*, 25(2), 215–227. <https://doi.org/10.1007/s10896-009-9285-9>.
- Walker, C. A., & Davies, J. (2012). A cross-cultural validation of the Brief Child Abuse Potential Inventory (BCAP). *Journal of Family Violence*, 27(7), 697–705. <https://doi.org/10.1007/s10896-012-9458-9>.
- Walsh, C., MacMillan, H., & Jamieson, E. (2002). The relationship between parental psychiatric disorder and child physical and sexual abuse: Findings from the Ontario health supplement. *Child Abuse & Neglect*, 26, 11–22. [https://doi.org/10.1016/S0145-2134\(01\)00308-8](https://doi.org/10.1016/S0145-2134(01)00308-8).
- Whitson, M. L., Martinez, A., Ayala, C., & Kaufman, J. S. (2011). Predictors of parenting and infant outcomes for impoverished adolescent parents. *Journal of Family Social Work*, 14(4), 284–297. <https://doi.org/10.1080/10522158.2011.587173>.
- Widom, C., Czaja, S. J., DuMont, K. A., Berlin, L. J., Appleyard, K., Dodge, K. A., ... Heckman, J. J. (2015). Intergenerational transmission of child abuse and neglect: Real or detection bias? *Science (New York, NY)*, 347(6229), 1480–1485. <https://doi.org/10.1126/science.1259917>.
- Widom, C., & Hiller-Sturmhofel, S. (2001). Alcohol abuse as a risk factor for and consequence of child abuse. *Alcohol Research & Health*, 25(1), 52.
- Wilhelm, B., Thönnissen, C., Alt, P., Fiedrich, S., Walper, S., Buhr, P., ... Wetzel, M. (2016). *pairfam Scales Manual Waves 1-7*. Munich.
- Wingenfeld, K., Schäfer, I., Terfehr, K., Grabski, H., Driessen, M., Grabe, H., ... Spitzer, C. (2011). [The reliable, valid and economic assessment of early traumatization: First psychometric characteristics of the German version of the Adverse Childhood Experiences Questionnaire (ACE)]. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 61(1), e10–4. <https://doi.org/10.1055/s-0030-1263161>.
- World Health Organization (2002). *World Report on Violence and Health: Chapter 3 Child abuse and neglect by parents and other caregivers*. Geneva: WHO.
- Zimmermann, P., Vierhaus, M., Eickhorst, A., Sann, A., Egger, C., Förthner, J., ... Spangler, G. (2016). [Growing up under adversity in Germany: Design and methods of a developmental study on risk and protective mechanisms in families with diverse psychosocial risk]. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 59(10), 1262–1270. <https://doi.org/10.1007/s00103-016-2423-7>.