




BMJ Open Psychometric properties of the German version of the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5) in clinical routine settings: study design and protocol of a multitrait-multimethod study

Jan-Peter Spies,¹ Marcella Lydia Woud ², Henrik Kessler,³ Heinrich Rau,⁴ Gerd Dieter Willmund,⁴ Kai Köhler,⁴ Stephan Herpertz,³ Simon E Blackwell ², Michelle Bovin,⁵ Brian P Marx,⁵ Jan Christopher Cwik ¹

To cite: Spies J-P, Woud ML, Kessler H, *et al.* Psychometric properties of the German version of the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5) in clinical routine settings: study design and protocol of a multitrait-multimethod study. *BMJ Open* 2020;**10**:e036078. doi:10.1136/bmjopen-2019-036078

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-036078>).

Received 29 November 2019
Revised 08 April 2020
Accepted 02 June 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Jan Christopher Cwik;
jcwik@uni-koeln.de

ABSTRACT

Introduction The aim of this study is to investigate the diagnostic accuracy, psychometric properties and clinical utility of the German version of the Clinician-Administered Post-Traumatic Stress Disorder (PTSD) Scale for Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) (CAPS-5) in routine clinical settings.

Methods and analysis This study is a non-interventional, multitrait-multimethod design, multicentre study that will be carried out at German civil and military inpatient and outpatient clinics. A total sample size of n=219 participants who have experienced at least one traumatic event according to criteria as defined in the DSM-5 will be recruited. For the investigation of the diagnostic accuracy and clinical utility of the CAPS-5, participants will be categorised into one of three groups, depending on their traumatic experiences and post-traumatic symptomatology: (1) monotraumatization with PTSD; (2) multiple traumatization with PTSD and (3) traumatization without PTSD. Interviews will be conducted face to face by interviewers in routine clinical settings. All participants will also be asked to complete a comprehensive set of questionnaires in order to investigate different facets of construct validity and clinical utility. First, differences between all three groups in CAPS-5 sum and subscale scores will be investigated. Test-retest reliability and inter-rater reliability will be determined. Internal consistency will be calculated using structural equation modeling (SEM) based internal consistency coefficients. Construct validity will be measured with Spearman's rank correlation analyses and multivariate analyses of variance with Holm-Bonferroni corrected post hoc analysis of variances. In order to test diagnostic accuracy, receiver operating characteristics and sensitivity and specificity analyses will be conducted. The model structure of the German CAPS-5 will be analysed using confirmatory factor analyses.

Ethics and dissemination The study received ethical approval by the Ethics Committees of the Faculty of Psychology at the Ruhr-Universität Bochum (reference numbers: 331 and 358). The results of the study will

Strengths and limitations of this study

- All diagnostic interviews will be conducted as face-to-face interviews with patients in their actual treatment settings.
- Interviewers will be blind to the results of former interviews, the results of the other assessments and to a previously assessed trauma index.
- The study covers a broad range of psychometric properties such as test-retest and inter-rater reliability, convergent and divergent validity.
- A sample encompassing both civilian and military participants will be included to ensure interpretability of the results in relation to both sample types and to allow generalisation of the study's results.
- Defining a trauma index as required by the Clinician-Administered Post-Traumatic Stress Disorder Scale-5 instruction can be very difficult for participants who have undergone multiple traumatic experiences.

be presented nationally and internationally at scientific conferences and will be published in scientific journals.

Trial registration number DRKS00015325

BACKGROUND

The diagnosis of post-traumatic stress disorder (PTSD) was first introduced as a codable diagnosis in the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III),¹ with extensive reporting during and after the Vietnam war a particularly strong driver for its inclusion. The inclusion of this diagnosis was intended to reduce the stigmatisation of war-related psychological reactions, facilitate compensation claims, improve treatment and to

stimulate disorder-specific research.²⁻⁴ From DSM-III to DSM-IV,⁵ the PTSD diagnosis was classified as an anxiety disorder and consisted of a stressor criterion (criterion A) in addition to re-experiencing symptoms (criterion B), numbing and avoidance symptoms (criterion C), and hyperarousal symptoms (criterion D). A number of changes made in moving from DSM-IV to the fifth edition of the DSM (DSM-5).⁶ First of all, PTSD is now part of the new category trauma-related and stressor-related disorders. With regard to the stressor criteria, the experienced events must now be qualified as 'traumatic', and the A2 criterion requiring a peritraumatic, subjective emotional response consisting of intense fear, helplessness and/or horror was eliminated as a result of studies showing that such emotional reactions should be seen as risk factors rather than a diagnostic criterion.⁷⁻¹¹ In DSM-5, sexual violence is now explicitly listed as potentially traumatic. Furthermore, besides directly experience of or witnessing an event, criterion A specifies that individuals can have been exposed to trauma via learning that it occurred to someone close to them or via repeated exposure to aversive details of a traumatic event. Additionally, in criterion A, a distinction between symptom onset and symptom exacerbation was added in order to highlight that the consequences of a trauma are not limited to symptom onset only.¹²

Due to an increasing amount of research based on cognitive theories¹³⁻¹⁵ and the fact that the DSM-IV three-factor model lacked robust empirical evidence, the symptom clusters were adapted and an additional cluster was added, reflecting the finding across the literature that four-factor models of PTSD symptoms consistently fit better and are unaffected by measure or sample type.^{16 17} This four-factor structure was based on the four-factor emotional numbing model,^{18 19} which was the first model to confirm the difference between avoidance and numbing. Consequently, the three symptom clusters of the DSM-IV entitled 're-experiencing', 'avoidance and numbing' and 'hyperarousal' were reformulated in the DSM-5 criteria into the clusters 're-experiencing' (criterion B), 'avoidance' (criterion C), 'negative alterations in cognitions and mood' (criterion D) and 'alterations in arousal and reactivity' (criterion E). The DSM-IV cluster 'avoidance and numbing' was split in DSM-5 into 'avoidance' and 'negative alterations in cognition and mood', and the former DSM-IV cluster 'hyperarousal' was renamed 'alterations in arousal and reactivity'.

In addition to the changes and reformulations of the PTSD criteria, the total number of symptoms increased from of 17 in DSM-IV¹² to 20 in DSM-5.²⁰ Two new specific symptoms were added to the new criterion D ('negative alterations in cognitions and mood'): (1) 'persistent distorted blame of self or others about the traumatic event(s)' and (2) 'persistent negative emotional state'. Following the redefinition of criterion E ('hyperarousal cluster') as 'alterations in arousal and reactivity', one new item 'reckless or self-destructive behaviour' was added, and the focus of the anger criterion was exclusively

shifted to behavioural aspects.⁶ Due to these changes in the diagnostic criteria, diagnostic instruments have had to be adapted accordingly.

Due to the changes made in PTSD diagnostic criteria from DSM-IV to DSM-5, several adaptations were made to the Clinician-Administered PTSD Scale (CAPS). First of all, the new DSM-5-based version of the CAPS (CAPS-5)²¹ reflects all adaptations of diagnostic criteria from DSM-5 (ie, omission of criterion A2, reformulations between and within criteria B-E, and inclusion of new symptoms).²² Furthermore, the CAPS-5 scoring was simplified with regard to the intensity and frequency ratings. That is, both ratings can be converted into a severity scale. This severity scale is used as the basic scoring rule in CAPS-5 for each symptom.²² The CAPS has traditionally been translated into several languages.²³ For each translated version, an investigation of psychometric properties is required, and the comparability of all versions should be tested to ensure robust CAPS-based clinical and scientific results.

Therefore, the aim of this study is to investigate the diagnostic accuracy, psychometric properties and clinical applicability of the German version of CAPS-5²⁴ under different routine clinical conditions. Furthermore, studies investigating the psychometric properties of the CAPS where often carried out in specific samples like veterans,^{22 25 26} or combined samples with different types of traumatic experience.²⁷⁻²⁹ However, investigations regarding the differentiation between specific traumatised samples are rare. Accordingly, this study also aims to investigate the diagnostic accuracy of the CAPS-5 between and within specific traumatised samples (persons with multiple-traumatisation and PTSD, persons with mono-traumatisation and PTSD, and traumatised persons without PTSD).

METHODS AND ANALYSIS

Participants, eligibility and procedure

As this study aims to test the applicability of the CAPS-5 in a broad German sample, traumatised civilian participants in addition to traumatised active German Armed Forces (GAF) soldiers and veterans will be recruited in both inpatient and outpatient units and in daily clinical routine settings.

To be included in this study: (1) participants have to be adults (≥ 18 years; all genders), (2) who have experienced at least one traumatic event according to the DSM-5-A-criterion. This criterion requires experiencing an event that comprises threatened death, serious injury or sexual violation. The traumatic situation can occur to the person herself/himself, can be witnessed by the person, can be occurred to a close family member or friend of the person, or can include a repeated or extreme exposure to aversive details of a traumatic event (for more details, see ref. 6). Furthermore, (3) the traumatic event(s) must have occurred more than a month before the application of the CAPS-5. There is only one exclusion criterion, that

is, an insufficient knowledge of the German language, in order to ensure both the feasibility of the interview and the validity of the patients' answers.

Participants will be interviewed at the very beginning of their treatment in order to reduce potential therapy-related effects on participants' answers. The results of the diagnostic assessment will be communicated to the participant's responsible therapist if the participant agrees to this.

Recruitment

Participants will be recruited from the outpatient therapy centre of the Mental Health and Research Centre (Forschungs-und Behandlungszentrum für psychische Gesundheit) at Ruhr-Universität Bochum, the inpatient and outpatient clinic of the Department of Psychosomatic Medicine and Psychotherapy of the Landschaftsverband Westfalen-Lippe (LWL) Universitätsklinikum Bochum at the Ruhr-Universität Bochum, the outpatient therapy centre of the university ambulance for psychotherapy at the University of Cologne (Hochschulambulanz für Psychotherapie Universität Köln), and the inpatient and outpatient clinic of the GAF Centre for Military Mental Health (Psychotraumazentrum Bundeswehrkrankenhaus Berlin). Advertisements for the study will be placed in all participating therapy facilities, and potential participants will also be identified by the clinicians at each facility. Patients who report traumatic events during admission to the respective inpatient or outpatient unit will be asked to participate in the study. Study participation is voluntary and participants will not receive any financial reimbursement.

Study design

This study is a non-interventional, multitrait-multimethod design multicentre study. Participants will be categorised into one of three groups, depending on their traumatic experience(s) and post-traumatic symptomatology: (1) monotraumatization (ie, a single traumatic event) with PTSD; (2) multiple traumatization (ie, multiple traumatic events) with PTSD or (3) traumatization without PTSD.

After enrolling in the study, participants will take part in two additional assessments. During each assessment, participants will be asked to fill out a questionnaire battery and take part in a CAPS-5 interview. During the first assessment, a German diagnostic short interview Diagnostisches Kurzinterview bei psychischen Störungen-Open Access (Mini-DIPS-OA)³⁰; will be administered in order to identify any potential comorbid diagnoses. For a detailed schedule of the study, see [figure 1](#).

All diagnostic interviews (CAPS-5 and Mini-DIPS-OA) will be audiotaped for post hoc randomly conducted quality checks of the interviews, and for the calculation of inter-rater reliability. The two interviews with each patient will be conducted by two different interviewers, and interviewers will be blind with respect to the results of the first interview and patient group. After completing all measurements, all patients will be reviewed in case

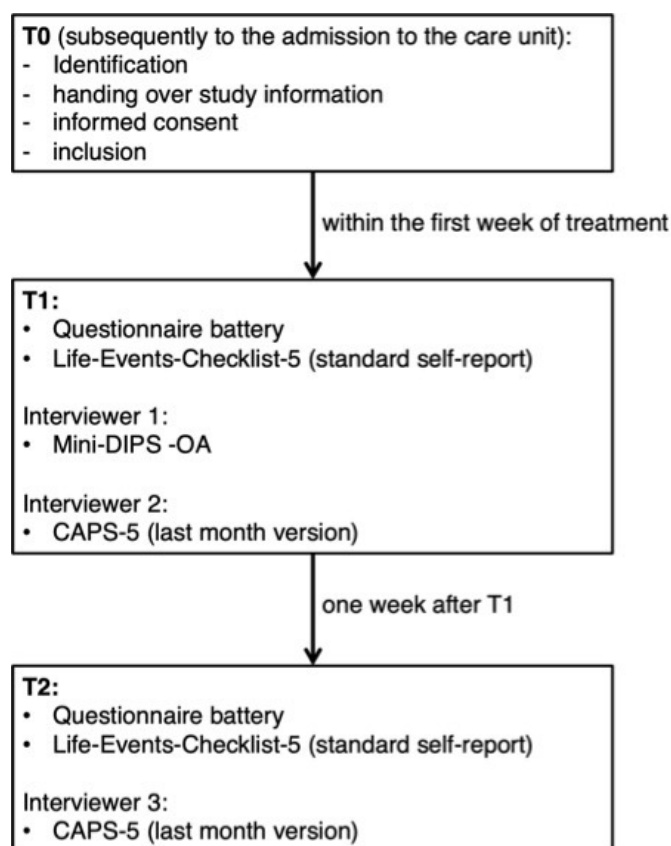


Figure 1 Study procedure of the inclusion and repeated measurement of participants. CAPS-5, Clinician-Administered Post-Traumatic Stress Disorder Scale for DSM-5; DSM-5, Diagnostic and Statistical Manual of Mental Disorders-5; Mini-DIPS-OA, Diagnostisches Kurzinterview bei psychischen Störungen-Open Access.

conferences in order to check the quality of the interviews and the integrity of the ratings. All interviewers involved in the study and trained interviewers not involved in the study will participate in these case conferences. Potential errors regarding the application of the CAPS-5 interview and the scoring rules of the CAPS-5 will be documented on case conference sheets and will be subjected to internal discussion and correction if necessary. The final consensus ratings of all symptoms as well as PTSD overall symptom severity will be used for final analyses.

Assessment of traumatic life events and diagnostic interviews

Life events checklist

For this study, the Life Events Checklist-5 (LEC-5)³¹ has been translated into German.³² The LEC-5 is available in three different versions: (1) the standard self-report version; (2) the extended self-report version and (3) the interview version. The standard self-report version lists 17 difficult or stressful life events (eg, fire, assault with a weapon, sexual assault or any other type of a highly distressing event or experience), and asks whether these events have ever happened to the participant, or if she/he witnessed, heard of them, or was exposed to them as part of her/his job. Furthermore, participants are also able to indicate if they are uncertain whether or not a

specific event happened to them, or that the event definitely did not happen to them. The extended self-report version additionally includes a non-specific item, item no. 17 (any other very stressful event or experience), which asks for a brief description of the worst event and when it happened. Participants are able to name their traumatic experience in their own words. Further, the extended self-report version asks for more details of the worst event (eg, whether someone's life was in danger or whether sexual violence was involved, and how many times similar events have occurred). In contrast to the other two versions, the LEC-5 interview version is conducted by an interviewer, who also asks for biographic information about the participant's family background, such as the participant's parents' educational principles and the parents' emotional handling of the participant. Afterwards, the 17 potential difficult or stressful events of the LEC-5 are assessed in an interview.

In the present study, the standard self-report version will be used to identify potentially traumatic events, because interviewers will be encouraged to explore the index trauma in more detail with the first question of the CAPS-5 interview and ensure that all interviewers are blind to the results of former interviews. However, all CAPS-5-interviewers will receive the LEC-5 for an overview of all experienced difficult or stressful events.

CAPS for DSM-5

Versions of the CAPS-5

The CAPS-5 is available in three versions: (1) last month version; (2) last week version and (3) worst month version.²¹ The last month version is the standard version that can be used to assess a current PTSD diagnosis in addition to PTSD symptom severity. The past week version assesses PTSD symptoms over a period of 1 week and can be used as a diagnostic tool in order to evaluate treatment progress. However, it cannot be used to establish a PTSD diagnosis. The worst month version can be used to establish a lifetime diagnosis of PTSD.²² In the present study, the last month version of the CAPS-5 will be used.

Administration of the interview

As a first step, the CAPS-5 assesses criterion A of the PTSD diagnosis according to the DSM-5. The traumatic events are then specified and categorised according to the kind of exposure (experienced, witnessed, learnt about, exposed to aversive details) in CAPS-5. Here, patients are asked to report their experiences in their own words. Subsequently, each DSM-5 symptom of PTSD is evaluated separately and in the order of the DSM-5 section. The interview is then continued using the exact words of the CAPS-5 instructions with the following exceptions: (1) the interviewer should use the exact words of the patient to describe/speak of the traumatic event, the same is true for the description of specific symptoms; (2) questions can be reformulated if specific information is already known; (3) if the questions do not produce sufficient information interviewers can use their own words to

specify the questions and (4) if necessary, examples can be obtained. Generally, there should be no comments made by interviewers that include suggestions for answers. In general, the interview should be conducted efficiently and smoothly using the minimal amount of questions in order to both keep the stress for the interviewee at a low level and to allow for a valid rating to be obtained. Also, note-taking should be reduced to obtaining only essential information in order to reduce delay, and the interviewer should remain respectful, yet purposeful, throughout the interview and focus on the traumatic events.

Scoring

As with the previous CAPS versions, the CAPS-5 scoring system is based on symptom frequency and intensity ratings. Intensity is rated on a four-point ordinal-scale. The rating scale should not be mentioned to the interviewee, as it is merely an orientation for the interviewer. For the frequency rating, either the frequency itself or the amount of time is used (eg, 'at least twice per month' or 'some of the time (20%–30%)'). For the CAPS-5, each item is rated with a single severity score in contrast to previous versions in which two ratings were combined into one severity rating.²² For the single severity rating, information about frequency and intensity is combined (eg, moderate severity rating='at least twice per month' and 'distress clearly present, less than 1 hour sleep loss'). The severity rating scale consists of five points: (1) absent; (2) mild/subthreshold; (3) moderate/threshold; (4) severely/significantly elevated and (5) extreme/incapacitating. There is a distinction between intensity and severity; nevertheless, both are related. Intensity is defined as the strength of a present symptom, whereas severity defines the total symptom load over a certain time period, and is a combination of intensity and frequency.²² Severity rating should only be made if both minimum frequency (eg, 'at least twice per month' or 'some of the time (20%–30%)') and intensity (eg, 'distress clearly present...' or 'reduction of positive emotional experience clearly present...') criteria are fulfilled. In cases of deviations in frequency and intensity, raters should use their clinical judgement. Each symptom needs to be related to the traumatic index event. For the remaining items, a trauma-related inquiry and rating scale is used, consisting of three ratings: (1) definite; (2) probable; (3) or unlikely, whereby a trauma-related rating of unlikely should not be used for the severity score or a PTSD diagnosis.²² For the CAPS-5 total symptom severity score, all item severity scores are summed after excluding the two dissociation severity scores. For the CAPS-5 symptom cluster severity scores, the severity score of each criterion is constructed by summing the corresponding items. Finally, the PTSD diagnostic status is defined by checking each DSM-5 criterion and dichotomising the CAPS-5 criteria according to the scoring rules (absence: <2; presence: ≥2).

After the interview section, interviewers have to make three global ratings regarding (1) the global validity, (2) the global severity and (3) the global improvement. Each

global rating is evaluated on a specific dimensional five-point Likert scale.

Psychometric properties

In the first evaluation of CAPS-5 psychometrics in a military veteran sample, the CAPS-5 diagnosis showed strong inter-rater reliability (Cohen's kappa (K)=0.78–1.00, depending on the scoring rule), strong test-retest reliability (K =0.83) and high concordance with CAPS-IV (K =0.84).²² The CAPS-5 total severity score showed high internal consistency (α =0.88) and inter-rater reliability (intraclass correlation (ICC)=0.91), good test-retest reliability (ICC=0.78), and high agreement with CAPS-IV severity score (r =0.83).²² Versions in other languages have also shown comparable indicators for very good psychometric properties.^{28–29} In conclusion, CAPS-5 appears to be a solid measure for PTSD diagnosis and symptom severity.²²

Translation–backtranslation procedure

The CAPS-5 interview was translated into German following a translation–backtranslation procedure. One author translated the interview into German and the other author translated this German version back into English. This back-translated English version was checked by members of the National Centre for PTSD. Potential translational inconsistencies or contradictions were then clarified via email and adaptations were done if required. The finalised German version was then authorised by the National Centre for PTSD (July 23, 2015).

Diagnostisches Kurzinterview bei psychischen Störungen open access

The Mini-DIPS-OA^{30–33} is a short version of the Diagnostisches Interview bei psychischen Störungen OA.³⁴ As a short semistructured clinician-administered diagnostic interview, the Mini-DIPS-OA can be used to assess the most common disorders in daily clinical practice, such as anxiety disorders, depressive disorders, bipolar disorders, obsessive–compulsive disorders and related disorders. Furthermore, trauma-related and stress-related disorders, eating disorders, somatic symptom and related disorders, substance-related and addictive disorders, and sleep-wake disorders can be diagnosed according to DSM-5, in addition to impulse-control problems, psychotic and sexual dysfunction symptoms, and suicidal tendencies.

Investigations of the psychometric properties of the DSM-IV version of the interview revealed good congruities between raters (on upper class levels of disorders (such as anxiety disorders): 91%–100%; on the disorder level: 88%–100%), with moderate to perfect Cohen's kappa coefficients (0.76–1.0), and Yule's Y coefficient of colligation (–0.84–1.0) in upper class levels within disorders and on disorders level (Cohen's kappa: 0.67–1.0; Yule's Y: 0.73–1.0).³⁰ Additionally, the validity of the mini-DIPS has been confirmed by cross-validation with the validated long version of the interview.^{30–33–36}

Primary outcome measure

PTSD checklist for DSM-5

The German version of the PTSD checklist (PCL-5)²⁵ will be used³⁷ as a primary outcome measure for the purpose of examining convergent validity. The PCL-5 is a self-report questionnaire to assess post-traumatic symptoms within the last month with respect to a traumatic event. It contains 20 item and each item reflects 1 of the 20 DSM-5 PTSD criteria. Participants report the intensity of each symptom on a five-point Likert-scale ranging from 0 (= 'not at all') to 4 (= 'extremely'). An overall PTSD symptom severity (ranging from 0 to 80) and a severity of each symptom cluster (ranges: intrusions=0–20, avoidance=0–8, negative alterations in cognition and mood=0–28, hyperarousal=0–24) can be generated. Scorings of ≥ 2 (= 'moderately') indicate that a symptom is present according to the levels specified in the DSM-5. Studies investigating the psychometric properties of the English PCL-5 version have revealed strong convergent and discriminant validity, very good sensitivity and specificity, and high test-retest reliability ($r_t \geq 0.82$),^{25–38} an excellent internal consistency for PCL-5 total score ($\alpha \geq 0.91$),^{25–38–40} and acceptable to excellent internal consistencies for the symptom cluster scores (intrusions: $\alpha \geq 0.80$; avoidance: $\alpha \geq 0.81$, negative alterations in cognition and mood: $\alpha \geq 0.82$, hyperarousal: $\alpha \geq 0.75$).^{39–40} Furthermore, confirmatory factor analyses have revealed a good model fit for the assumed DSM-5 model of PTSD, although a previously reported 7-factor hybrid model⁴¹ revealed the best model fit.^{25–38–40} Equivalently, good to excellent psychometric properties have also been reported for the French,⁴⁰ German,³⁷ Swedish⁴² and Shona⁴³ versions. The overall PTSD symptom severity can be used to administer a cut-off score. The cut-off score for indicating clinically relevant PTSD symptom severity has consistently ranged between 31 and 33 for the PCL-5 scale,^{25–37–40–42–43} whereas for the German version a cut-off score of ≥ 33 is recommended.³⁷

Hypotheses

We expect significant positive associations between PCL-5 and both CAPS-5 total and symptom cluster scores. Additionally, we expect that patients who receive a PTSD diagnosis according to CAPS-5 will also exceed the cut-off score of ≥ 33 on the PCL-5 (area under the curve (AUC) ≥ 0.80). Furthermore, we expect significant differences in the mean PCL-5 sum score and subscales between the three groups (PCL-5 sum scores: multiple-traumatisation >monotraumatisation >traumatised without PTSD).

Secondary outcome measures for assessing concurrent validity

Impact of Event Scale-Revised

The German version of the Impact of Event Scale-Revised (IES-R) self-rating questionnaire^{44–45} will be used⁴⁶ as a secondary outcome measure to examine convergent validity. The IES-R is a well-established measure in the context of trauma that has been repeatedly used as an

external criterion measure for PTSD-related measures. The IES-R is based on the DSM-IV PTSD criteria and consists of three subscales with a total of 2 items: (1) intrusions (eight items); (2) avoidance (eight items); and hyperarousal (six items), which show relatively high inter-correlations.⁴⁷ The IES-R assesses the patient's distress per item within the past week on a five-point Likert-scale (0 = 'not at all' to 4 = 'extremely'). Studies with the English version revealed good scale construct validity,^{47 48} adequate to excellent internal consistencies (total score: $\alpha \geq 0.90$; intrusions: $\alpha \geq 0.87$; avoidance: $\alpha \geq 0.84$; hyperarousal: $\alpha \geq 0.79$),^{44 47-49} and high test-retest reliability after 6 months ($r_{tt} \geq 0.89$).⁴⁴ As a cut-off score with good diagnostic accuracy, a score of 1.5 is recommended.⁴⁷ However, some methodological issues also diminish the quality of the IES-R. Specificity and sensitivity are not within an optimal range for a diagnostic tool.^{47 48} Furthermore, data concerning the factorial structure are inconsistent. On the one hand, mostly adequate model fit indices for a three-factor solution have been reported,⁴⁸ but on the other hand, some studies report a questionable factorial structure of the three-factor solution and suggest one-factor or four-factor solutions (with an additional sleep factor).^{47 49} Another limitation is related to the invariance of the IES-R. Whereas its configural invariance has mostly been demonstrated, the metric invariance caused by variances in the intrusions factor over time has not been fully supported.⁴⁹

However, considering the comparability with results of other studies, we decided to include the IES-R as a secondary outcome measure. The German version of the IES-R has also revealed good construct validity and adequate to excellent internal consistencies with respect to two different samples (intrusions: $\alpha \geq 0.71$; avoidance: $\alpha \geq 0.79$; hyperarousal: $\alpha = 0.90$).⁴⁶ In addition to the original answering format related to distress, there is another format for the German version, asking about the frequency of symptoms in the past week (0 = 'not at all', 1 = 'infrequently', 3 = 'sometimes', 5 = 'frequently'). By using an evaluation formula ($X = [-0.02 * \text{intrusions}] + [0.07 * \text{avoidance}] + [0.15 * \text{hyperarousal}] - 4.36$), a cut-off score of >0 indicates a potential suspicion of PTSD.⁵⁰ In the current study, both versions of the IES-R will be included.

Hypotheses

We expect significant positive associations between total scores from the IES-R and CAPS-5. Again, we expect that a significant proportion of patients diagnosed with PTSD according to CAPS-5 will also exceed the cut-off score of ≥ 1.5 (in the distress version) with the corresponding >0 (in the frequency version) for the IES-R (AUC ≥ 0.80). Additionally, we expect significant differences of mean IES-R sum scores between groups (IES-R sum score: multiple-traumatisation $>$ monotraumatisation $>$ traumatised without PTSD). Due to the questionable factor structure, the IES-R subscale scores will not be included in the analyses.

Beck depression inventory-Revised

In order to assess the severity of depressive symptoms, the German version of the Beck depression inventory (BDI-II) will be used (61.62). The BDI-II consists of 21 items assessing depressive symptoms during the past 2 weeks. Anchors differ per items, however, for each item, patients are able to use a four-point Likert scale (0–3). BDI-II sum scores can be used to differ between no (0–8), minimal (9–13), mild (14–19), medium (20–28) and severe levels of depression (29–63). The English version of the BDI-II has shown good concurrent validity and diagnostic accuracy^{51 52} in addition to excellent internal consistency ($\alpha = 0.91$).⁵³ The German version of the BDI-II has demonstrated good to excellent internal consistencies ($\alpha \geq 0.84$) in clinical and non-clinical samples and adequate concurrent validity.^{54 55}

Hypotheses

Due to the inclusion of the new PTSD cluster 'negative alterations in cognition and mood' and considering high comorbidity between PTSD and depression,^{56–60} we expect high correlations between the CAPS-5 total and the BDI-II sum score. In particular, we expect high correlations between the CAPS-5 score of the cluster 'negative alterations in cognition and mood' and the BDI-II. Furthermore, significant differences in mean BDI-II sum scores between groups are expected (BDI-II sum score: multiple-traumatisation $>$ mono-traumatisation $>$ traumatised without PTSD).

State-trait anxiety inventory

The state-trait anxiety inventory (STAI)⁶¹ is a self-report questionnaire that consists of two distinct scales measuring state and trait anxiety (STAI-S and STAI-T, respectively), with 20-items per scale. For both scales, participants are asked to rate each item on a four-point Likert-scale (STAI-S: 1 = 'not at all' to 4 = 'very much so'; STAI-T: 1 = 'almost never' to 4 = 'almost always'). The STAI-S asks for currently experienced anxiety, and the STAI-T asks how patients feel in general. The sum score of each scale lies between 20 and 80. For the present study, the German version of STAI-S and STAI-T will be used.⁶² The STAI-S shows excellent ($\alpha \geq 0.90$) and the STAI-T good to excellent ($\alpha \geq 0.88$) internal consistencies. Furthermore, both scales have shown good convergent and divergent validity, and norms for both sexes for different age classes based on German data have been reported.

Hypotheses

For the present study, we expect high correlations between the total CAPS-5 sum score and the STAI-S and STAI-T sum scores as well as significantly higher STAI-S and STAI-T sum scores for patients with PTSD compared with traumatised persons without PTSD (STAI-S/T sum score: multiple-traumatisation = monotraumatisation $>$ traumatised without PTSD).

Dissociative Experiences Scale: 20 items short version

The Dissociative Experiences Scale (DES) is a 44-item questionnaire^{63 64} that was constructed to measure dissociative symptoms according to the DSM-IV and International Classification of Diseases-10th revision (ICD-10) scales. Each item measures how often a dissociative situation occurred on a scale from 0% (= 'never') to 100% (= 'always'). The DES has shown very good psychometric properties: high test-retest reliability ($r_{tt} \geq 0.79$) and split half reliability ($r \geq 0.83$), excellent internal consistency ($\alpha \geq 0.93$) and good construct validity.⁶⁵ However, the factor structure and temporal stability of the DES have been reported as problematic. Thus, a short version with 20 items (DES-20) was constructed to address these problems. In the present study, the German version of the DES-20⁶⁵ will be used. This version presents excellent internal consistency ($\alpha = 0.93$), relatively high temporal stability and good divergent and convergent validity. Additionally, the DES-20 significantly discriminates between patients with borderline personality disorder and patients with PTSD from patients with other mental disorders.⁶⁵ However, it cannot be used to differentiate patients with borderline personality disorder from patients with PTSD.⁶⁵

Hypotheses

For the present study, we expect significantly higher DES-20 sum scores for patients with PTSD as compared with traumatised persons without PTSD (DES-20 sum score: multiple-traumatisation=monotraumatisation>traumatised without PTSD). Additionally, we expect higher DES-20 sum scores among patients with PTSD with a dissociative subtype according to DSM-5 compared with patients with PTSD without a dissociative subtype (DES-20 sum score: PTSD with dissociation > PTSD without dissociation).

Questionnaire of thoughts and feelings

The questionnaire of thoughts and feelings (QTF) is a German questionnaire 'Fragebogen zu Gefühlen und Gedanken' that was developed on the basis of the cognitive concepts of personality disorders, and especially the biosocial model of borderline personality disorder.^{66 67} The QTF is a self-rating questionnaire and consists of 34 items that describe typical borderline-specific statements regarding thoughts and feelings. Items are rated on a five-point Likert-scale (1 = 'do not agree at all' to 5 = 'I agree completely'). Item analysis has revealed overall satisfactory to good values, and analyses of psychometric properties have revealed good sensitivity to change, good internal consistency ($\alpha = 0.89$), and good test-retest reliability ($r_{tt} = 0.81$) in addition to good construct validity measured via self-rating questionnaires and a semistructured diagnostic interview.^{66 67} In the present CAPS-5 study, the short version of the QTF (14 items) will be used, which also has good to excellent psychometric properties.⁶⁷

Hypotheses

With respect to results of the present study, we expect significantly higher sum scores on the QTF for participants with multiple-traumatisation compared with both other groups, but no significant differences between monotraumatised PTSD patients and traumatised participants without PTSD (QTF sum score: multiple-traumatisation>monotraumatisation=traumatised without PTSD).

Screening zur komplexen Posttraumatischen Belastungsstörung

The 'screening zur komplexen Posttraumatischen Belastungsstörung' (SkPTBS)⁶⁸ is a German self-rating questionnaire for the screening of complex PTSD (CPTSD) according to the eleventh edition of the ICD (ICD-11).^{69 70} The questionnaire consists of three parts.⁷¹ In part 1, patients are asked if they have ever personally experienced or witnessed one of the 11 listed potentially traumatising events. The list is divided into two parts; on the left side, six non-interpersonal traumatic events are listed (such as a serious accident, natural disaster or a work-related traumatic event), whereas the right side lists five interpersonal traumatic events (such as emotional abuse, sexual assault or torture). Additionally, patients are able to name another event in case it is not listed, or to state that they have not experienced any of the listed events. Subsequently, patients who have experienced more than one event are instructed to circle the most distressing event. Part 2 assesses risk and resilience factors regarding the selected event. First, the age of patients when experiencing the event, the frequency of the experience (once, twice, three times, 4–5 times or more often), the duration of the event (not prolonged, over _ months, over _ years) and who caused the event (a stranger, a family member or better known person, force majeure or illness) are asked. Next, patients are asked to rate on a seven-point Likert-scale (0 = 'not at all' to 6 'fully correct') several sets of symptoms: whether (1) they were afraid for their life; (2) they received post-traumatic social support and (3) the event now seems unreal, such as in a dream or a movie. Finally, part 3 uses 14 items on a seven-point Likert-scale (0 = 'not at all' to 6 'fully correct') to assess the rate of personal experiences in everyday life as a consequence of the traumatic event (such as difficulty in trusting people, uncomfortable feelings when physically touched or feeling ashamed of the events that happened).

The psychometric properties of the SkPTBS are good to excellent with item difficulties and discriminatory power within the requested range, good convergent, divergent and predictive validity, and an excellent internal consistency ($\alpha = 0.91$).^{71 72} However, the questionnaire can only be used to assess the risk of having a CPTSD diagnosis, not to diagnose the disorder.⁷²

The evaluation of the SkPTBS will be carried out in four steps.^{71 72} For the first step, for each interpersonal traumatic event (right side of the first part) a score of 10 (except for sexual violence, which is scored with 50) is given, regardless of whether the event happened to the patient or the patient witnessed the event. All scores are

then summed up to a sum score A, which can range from 0 to 90. Next, the sum score B is determined by summing up the answers of items 4–14 of part 3 of the questionnaire, in which personal experiences in everyday life are assessed. The sum score B is calculated independently of missing values within these items. Thus, a possible sum score B could range from 0 to 66. Finally, the SkPTBS sum score is calculated using the following formula: sum score SkPTBS = sum score A + 2 * sum score B.

For the evaluation and interpretation of the questionnaire, the authors will provide a sheet that automatically evaluates the raw data, supports the interpretation and provides norms.⁷³ A SkPTBS score can be interpreted: (1) very low risk of CPTSD: SkPTBS score = 0–2.20; (2) low risk of CPTSD: SkPTBS score = 2.21–5.07; (3) high risk of CPTSD: 5.08–28.18 and (4) very high risk of CPTSD: SkPTBS score \geq 28.19.

Hypotheses

For the present study, we expect a significantly higher SkPTBS score for PTSD patients with multiple-traumatisation compared with monotraumatized PTSD patients and traumatized controls and a significantly higher SkPTBS score for PTSD patients with monotraumatisation compared with traumatized controls (SkPTBS score: multiple-traumatisation > monotraumatisation > traumatized without PTSD).

Symptom-Checklist-90-Revised-Global Severity Index

The Symptom-Checklist-90-Revised (SCL-90-R)-Global Severity Index (GSI)⁷⁴ is a self-report questionnaire assessing various psychological symptoms. In this study, the German version⁷⁵ will be used. In order to determine the concurrent validity of the CAPS-5, the GSI score will be used. The GSI reflects the overall distress caused by psychological symptoms during the past week and is calculated by summing up the scores of all 90 items of the SCL-90-R and dividing them by 90. Items are coded on a five-point Likert scale (0 = 'not at all' to 4 = 'extremely'). The internal consistency of the GSI is good to excellent ($\alpha \geq 0.89$) for the original version^{76 77} and excellent ($\alpha \geq 0.96$) for the German version of the SCL-90-R.⁷⁵

Hypotheses

Based on study results regarding associations between the GSI and trauma-associated symptoms, which show that complex traumatized patients present the highest GSI scores,⁷² we expect high correlations between the total CAPS-5 sum score and the GSI and significant differences of mean GSI scores between groups (SCL-90-R GSI: multiple-traumatisation > monotraumatisation > traumatized without PTSD).

Secondary outcome measures for assessing for testing divergent validity

SCL-90-R-subcales

Six of the nine subscales of the SCL-90-R^{74 75} will be used to determine the divergent validity: (1) somatisation (SOM); (2) obsession-compulsion (OBS); (3) interpersonal

sensitivity (INT); (4) phobic anxiety (PHO); (5) paranoid ideation (PAR) and (6) psychoticism (PSY). Due to former study results regarding the identification of PTSD patients by using SCL-90-R items, several items will be excluded for the calculation of the sum scores of the subscales and the analyses for the divergent validity of CAPS-5: (1) item 3 (OBS); (2) item 43 (PAR); (3) item 50 (PHO); (4) item 55 (OBS); (5) item 70 (PHOB) and (6) item 90 (PSY). However, the number of excluded items lies under the threshold of the maximally tolerated missing items for each subscale.⁷⁵ Internal consistencies for the subscales range from questionable for the PSY subscale ($\alpha \geq 0.69$), acceptable for the PAR subscale ($\alpha \geq 0.72$) and good for the subscales SOM, OBS, INT and PHO ($\alpha \geq 0.81$).^{76 77} The test-retest reliability of the SCL-90-R is high after 1 week of therapy ($r_{tt} \geq 0.80$), and the construct validity is good.⁷⁷ The German version of the SCL-90-R has shown acceptable to good internal consistencies with respect to the six selected subscales in clinical samples ($\alpha \geq 0.75$) and also good parameters regarding its construct validity.⁷⁵

Hypotheses

We expect no or at most small correlations between the CAPS-5 total sum score and the six subscales of the SCL-90-R, and no significant differences in mean scores of the six subscales between groups (SCL-90-R sum scores: multiple-traumatisation = monotraumatisation = traumatized without PTSD).

Patient Acceptance Questionnaire

The Patient Acceptance Questionnaire (PAQ) is a questionnaire that was developed to assess the satisfaction of patients when undergoing a structured diagnostic interview.⁷⁸ PAQ items were constructed based on theories and research concerning factors that are of high importance for the therapeutic relationship.⁷⁸ The questionnaire consists of two factors measuring the mental effort of patients during the interview, and emotional reactions during the interview. For the present study, the PAQ was slightly modified. The overall satisfaction rating scale was dropped, and patients will be able to make their response to the ten items see⁷⁸ on a five-point Likert scale (0 = 'completely disagree' to 4 = 'completely agree'). Furthermore, three items were added to assess patients' evaluation of the therapeutic relationship during prior therapeutic settings (1) 'I did not feel taken seriously', (2) 'I had the feeling of receiving too little attention' and (3) 'I had the feeling that no doctor/therapist is interested in me as a human being'). For the present study, the PAQ will be used to investigate the feasibility of the CAPS-5.

Hypotheses

We expect a medium to high satisfaction of patients with the CAPS-5 interview, yet no significant differences in satisfaction between groups are expected (PAQ sum scores: multiple-traumatisation = monotraumatisation = traumatized without PTSD).

Interviewer Acceptance Questionnaire

The Interviewer Acceptance Questionnaire (IAQ) was developed to assess the interviewers' satisfaction with the structured diagnostic interview in the interview situation itself. The IAQ consist of four items and assess whether interviewers (1) rate the interview as helpful, (2) the interview can be useful to organise patient's problems, (3) the questions of the interview are useful to justice to the patient's problems and (4) the interview pushed the patient to the limits of her/his resilience (reversed coded). Each item is rated on a five-point Likert scale (0 = 'completely disagree' to 4 = 'completely agree').

Hypotheses

We expect a medium to high satisfaction of the interviewers with the CAPS-5 interview. However, we expect a significantly higher IAQ score when assessing mono-traumatised PTSD patients and traumatised controls, compared with PTSD patients with multiple-traumatisations (IAQ score: mono-traumatisation=traumatised without PTSD>multiple-traumatisation)."

Balanced Inventory of Desirable Responding

The Balanced Inventory of Desirable Responding (BIDR) is a questionnaire for assessing socially desirable responses.^{79 80} It assesses two factors of socially desirable responses: (1) self-deceptive enhancement and (2) impression management. For the present study, the German version of the BIDR, which has shown to be a valid questionnaire with acceptable psychometric properties, will be used.⁸¹

Hypotheses

For the present study, we assume that there is no significantly increased tendency for socially desirable response for none of the groups (BIDR sum scores: multiple-traumatisation=mono-traumatisation=traumatised without PTSD).

Training of interviewers

Similar to the training procedure for the DIPS interviews,^{30 33 34} all interviewers will have to undergo a standardised training procedure. This training procedure consists of six phases: (1) The diagnostic-phase: in this phase, interviewers are trained in the correct administration of the DSM-5 criteria of the PTSD and PTSD subtypes. This step includes a theoretical introduction and illustrations based on case reports; (2) The introduction-phase: interviewers are trained in the correct administration and evaluation of the CAPS-5 interview. Additionally, they are trained how to deal with typical problematic situations that can arise during the CAPS-5 interview; (3) The monitoring-phase: In this phase, interviewers take part in at least two interviews that are conducted by a trained interviewer; (4) The exercise phase: interviewers train to conduct the interview with familiar, non-diagnosed persons; (5) The coding phase: At least two audiotaped CAPS-5 interviews, which were conducted by a trained interviewer, are rated by the trainees. The conclusions of

these ratings have to be in accordance with the rating of the trained interviewer. Ratings are in accordance if they agree with respect to the primary diagnosis (including subtypes), and if the symptom ratings do not significantly differ (± 1 on the rating scale) and (6) The certification phase: In this last phase, interviewers have to conduct at least two interviews with patients and their ratings have to be in accordance with a trained interviewer as defined in the coding phase. Additionally, the interviewers have to conduct interviews without making severe administrative errors. Administrative errors consist of the several violations: (1) inadequately introducing the interview to the patient; (2) marked aberration from the questions of the interview; (3) inadequate requests to clarify the diagnosis; (4) describing the numeric severity ratings for describing symptoms to the patients and (5) excessive comments or questions unrelated to the interview.

Interviewers applying the CAPS should be graduated clinical psychologists, or should have completed their training as a psychiatrist, or should be in the last phase of their master's degree in clinical psychology.

Patient and public involvement

No patient involvement.

Statistical analyses

Characteristics of the study sample will be described separately for all three groups using frequencies, means and SD. Potential group differences will be tested by using χ^2 tests for categorical variables, Mann-Whitney U tests for ordinally and not normally distributed data, and t-tests for normally distributed data.

The aim of the present study is to investigate the psychometric properties of the German CAPS-5 and its diagnostic accuracy in daily clinical routine. First of all, an item analyses (M, SD, skewness, kurtosis, range) for all CAPS-5 items will be applied to all three groups. The test-retest reliability will be assessed by conducting ICCs analyses between the CAPS-5 interviews at T1 and T2 (1 week after T1). The inter-rater reliability will be analysed by calculating Cohen's kappa coefficients between the actual assessment of CAPS-5 at T1 and the independent rating of the audio record of the interview at T1. In order to test the internal consistency of the CAPS-5 and its subscales, SEM-based internal consistency (McDonald's omega: 93–95) will be calculated. Construct validity, consisting of divergent and convergent validity, will be measured with Spearman's rank correlation analyses and a multivariate analyses of variance (MANOVA) with Holm-Bonferroni corrected post hoc ANOVAs with a priori contrasts. In order to test the diagnostic accuracy (true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN)) of the German CAPS-5, the PCL-5 score will be used as external criterion. Based on these results of the diagnostic accuracy, the sensitivity and specificity in addition to the receiver operating characteristics and the sensitivity, specificity, AUC, positive likelihood ratio (the ratio of TP vs FP test results), and negative likelihood ratio (the

ratio of TN to FN test results) will also be calculated. The CAPS-5 was constructed as a structured diagnostic interview to assess whether patients fulfil the DSM-5 criteria of PTSD. Accordingly, each item represents a specific PTSD criterion. Thus, by using a confirmatory factor analysis with Satorra-Bentler correction,^{82 83} the convergent validity of the German CAPS-5 can and will be analysed. Finally, network analyses will be used to investigate the associative strength between the different PTSD symptoms as assessed with the CAPS-5.

Power calculation

G*Power (V.3.1.9.3 for MacOS)^{84–86} was used to determine a sufficient sample size for a MANOVA with three groups: (1) PTSD patients with monotraumatization; (2) PTSD patients with multiple-traumatization and (3) traumatised controls without PTSD and 23 dependent variables (primary and secondary outcome measures) by using an α of 0.05, a power of 0.80 and a small effect size ($f=0.10$).⁸⁴ This calculation resulted in a total sample size of at least $n=168$. However, due to drop-outs (eg, due to incomplete interviews) and missing data, a rate of 30% of uncompleted data sets is expected, implying that an additional $n=51$ participants will be tested in order to ensure a minimal sample size of $n=73$ participants per group with a total sample size of $n=219$ participants.

DISCUSSION

The present study protocol describes the methods and the design of a non-interventional, multitrait–multimethods design multicentre study that aims to investigate the diagnostic accuracy, psychometric properties and feasibility of the German version of the CAPS-5 interview. The study will be carried out under routine clinical conditions and will include participants with different types and numbers of traumatic experiences to examine the feasibility of the CAPS-5 for diagnosing PTSD and PTSD symptom severity.⁸⁷ In clinical research and daily practice, the use of structured diagnostic interviews is highly recommended for the assessment of mental disorders, because they ensure a systematic assessment of symptoms combined with clinicians' expert knowledge.^{88–95} For the assessment of PTSD, the CAPS⁹⁶ is considered the diagnostic gold standard,^{25 43 97 98} and its use is recommended by PTSD guidelines.⁹⁹ The CAPS was designed to represent a standard in terms of an interview-based diagnostic measure of PTSD.^{23 96} It was created as a structured diagnostic interview based on the DSM criteria of PTSD, and given its excellent psychometric properties it became universally accepted for PTSD diagnosis in both research and practice.¹² Further, CAPS users are able to evaluate the interview according to different scoring rules, enabling more liberal evaluations for daily clinical practice or more conservative evaluations for research.²⁶ The DSM-IV version of the CAPS consistently exhibited good to excellent psychometric properties,^{18 23 96 100 101} with good diagnostic utility and good sensitivity to clinical

change.²³ Given the importance of the CAPS-5 for research and treatment, it is essential to scrutinise its diagnostic accuracy, psychometric properties and clinical utility under clinical routine conditions using a broad and heterogeneous sample. Based on these results, the German CAPS-5 can contribute to individualised planning of treatments and adequate evaluation of therapy efficacy. However, studies such as the current one have several strengths and limitations.

Strengths of the study

One strength of the study is that all interviews will be conducted as face-to-face interviews with patients in their actual treatment settings. Additionally, the detailed training and ongoing supervision of interviewers in this study should be highlighted. Thus, high quality of the interviews will be ensured. Another important strength of this study is that all interviewers will be blind to the results of former interviews, the results of the other assessments, and to a previously assessed trauma index. This helps to strengthen the results of this study. Furthermore, the study covers a broad range of psychometric properties such as test–retest and inter-rater reliability, convergent and divergent validity. Additionally, a sample encompassing both civilian and military participants will be included to ensure interpretability of the results in relation to both sample types and to allow generalisation of the study's results. Patients' acceptance of the CAPS-5 interview will be evaluated. Inclusion of participants who experienced monotraumatization and multiple-traumatization in addition to participants who experienced different kinds of traumatic events (such as motor accident, sexual abuse or combat actions) is another strength of the study and allows us to compare our samples with other samples and to generalise the results. Former studies have mostly focused on homogeneous study samples (eg, veterans or motor accident victims), and thus, it is questionable whether the results of these studies can be generalised to other samples (eg, sexual assault victims). Finally, this study will also investigate aspects such as socially desirable response patterns in the form of self-deceptive enhancement and impression management, which are rarely investigated in relation to diagnostic interviews.

Challenges and limitations

However, in addition to its strengths, this study also has several potential limitations and challenges. First, defining a trauma index as required by the CAPS-5 instruction can be very difficult for participants who have undergone multiple traumatic experiences. Especially when these events occurred several years or decades ago, the answers' validity could be affected by a variety of influencing factors (such as former treatment, avoidance behaviour or time factors). Furthermore, results of previous studies have shown that summarising several traumatic events as one trauma index leads to significantly higher symptom severity ratings.¹⁰² This, in turn could potentially bias the results in relation to the validity of the interview.

Additionally, the fact that participants will have to report stressful, stigmatising and shameful events to strangers could potentially lead to an increased dropout rate. This could be worsened by the fact that three different interviewers will interview each participant, and thus, events have to be reported three times. Another challenge might be the different arrangements of interviewer teams with different genders. It could potentially be difficult for some participants to talk to either male interviewers (such as women after a sexual assault) or female interviewers (such as men with abasement experiences). These difficulties could potentially have an influence on the answers' validity. Furthermore, given the naturalistic design of the study, participants will show different types of comorbidity. Thus, it might be possible that the results regarding psychometric properties may be affected by symptoms of other disorders (such as personality disorders or severe major depressive disorders). Finally, it may be possible that the final composition of the sample could lead to issues regarding some of the planned statistical analyses (such as network analyses) due to high heterogeneity or high rates of drop-outs.

ETHICS AND DISSEMINATION

Ethical considerations

The study received ethical approval by the Ethics Committees of the Faculty of Psychology at the Ruhr-Universität Bochum (reference numbers: 331 and 358). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Declaration of Helsinki 1975, as revised in 2008. Furthermore, this study will follow the principles of good scientific practice. The study participation is entirely voluntary. All participants will be informed about the aims, the procedure, the data collection and the dissemination plans of the study and will be asked to give their written informed consent. Participants may withdraw the participation at any moment without any negative consequences.

Dissemination plan

The results of the study will be presented nationally and internationally at scientific conferences and will be published in scientific journals.

Author affiliations

¹Clinical Psychology and Psychotherapy, Universität zu Köln, Köln, Germany

²Faculty of Psychology, Mental Health Research and Treatment Center, Ruhr-Universität Bochum, Bochum, Germany

³Department of Psychosomatic Medicine and Psychotherapy, LWL University Hospital, Ruhr-Universität Bochum, Bochum, Germany

⁴German Armed Forces Center for Military Mental Health, Berlin, Germany

⁵VA Boston Healthcare System, Massachusetts and Boston University School of Medicine, Boston, Massachusetts, USA

Twitter Simon E Blackwell @Blackwell_SE and Jan Christopher Cwik @Cwik_JC

Acknowledgements We gratefully thank Frank W Weathers and Juliette M Harik for providing us the original version of the CAPS-5 and their helpful support.

Contributors JCC, J-PS, SEB and MLW wrote the manuscript. JCC, MLW, HK, GDW, SH, HR, KK and J-PS designed the study. The translation-backtranslation process was accompanied by MB and BPM. All authors have read and approved the manuscript. Planning and conceptualisation of this study started in March 2015, and the recruitment of participants began in December 2016. Shortly before the planned enrolment of the first civil study patient, the authors decided to additionally involve a military subsample. Due to the already implemented recruitment procedures in the inpatient and outpatient units and to avoid a delayed recruitment, the study started with civil participants before the ethics agreement of the military units was given. This resulted in a retrospective study registration and delayed formulation of this study protocol.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The study received ethical approval by the Ethics Committees of the Faculty of Psychology at the Ruhr-Universität Bochum (reference numbers: 331 and 358).

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Marcella Lydia Woud <http://orcid.org/0000-0002-4974-505X>

Simon E Blackwell <http://orcid.org/0000-0002-3313-7084>

Jan Christopher Cwik <http://orcid.org/0000-0002-2290-353X>

REFERENCES

- 1 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 3. ed. Washington, DC: American Psychiatric Association, 1980.
- 2 Gersons BP, Carlier IV. Post-Traumatic stress disorder: the history of a recent concept. *Br J Psychiatry* 1992;161:742–8.
- 3 Wilson JP. The historical evolution of PTSD diagnostic criteria: from Freud to DSM-IV. *J Trauma Stress* 1994;7:681–98.
- 4 Kinzie JD, Goetz RR. A century of controversy surrounding posttraumatic stress-spectrum syndromes: the impact on DSM-III and DSM-IV. *J Trauma Stress* 1996;9:159–79.
- 5 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4 - TR edn. Washington, DC: American Psychiatric Association, 2000.
- 6 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th edn. Washington, DC: American Psychiatric Association, 2013.
- 7 Stein JY, Wilmet DV, Solomon Z. Does one size fit all? Nosological, clinical, and scientific implications of variations in PTSD criterion a. *J Anxiety Disord* 2016;43:106–17.
- 8 Brewin CR, Lanius RA, Novac A, et al. Reformulating PTSD for DSM-V: life after criterion a. *J Trauma Stress* 2009;22:366–73.
- 9 Kilpatrick DG, Resnick HS, Acierno R. Should PTSD criterion a be retained? *J Trauma Stress* 2009;22:374–83.
- 10 Weathers FW, Keane TM. The criterion a problem revisited: controversies and challenges in defining and measuring psychological trauma. *J Trauma Stress* 2007;20:107–21.
- 11 Karam EG, Andrews G, Bromet E, et al. The role of criterion A2 in the DSM-IV diagnosis of posttraumatic stress disorder. *Biol Psychiatry* 2010;68:465–73.
- 12 Weathers FW, Marx BP, Friedman MJ, et al. Posttraumatic stress disorder in DSM-5: new criteria, new measures, and implications for assessment. *Psychol Inj Law* 2014;7:93–107.
- 13 Ehlers A, Clark DM. A cognitive model of posttraumatic stress disorder. *Behav Res Ther* 2000;38:319–45.
- 14 Foa EB, Steketee G, Rothbaum BO. Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behav Ther* 1989;20:155–76.

- 15 Brewin CR, Holmes EA. Psychological theories of posttraumatic stress disorder. *Clin Psychol Rev* 2003;23:339–76.
- 16 Yufik T, Simms LJ. A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms. *J Abnorm Psychol* 2010;119:764–76.
- 17 Elhai JD, Palmieri PA. The factor structure of posttraumatic stress disorder: a literature update, critique of methodology, and agenda for future research. *J Anxiety Disord* 2011;25:849–54.
- 18 King DW, Leskin GA, King LA, et al. Confirmatory factor analysis of the clinician-administered PTSD scale: evidence for the dimensionality of posttraumatic stress disorder. *Psychol Assess* 1998;10:90–6.
- 19 Simms LJ, Watson D, Doebbellings BN. Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *J Abnorm Psychol* 2002;111:637–47.
- 20 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th edn. Washington, DC: American Psychiatric Association, 1994.
- 21 Weathers FW, Blake DD, Schnurr PP, et al. *Clinician-administered PTSD scale for DSM-5 (CAPS-5)*. National Center for Posttraumatic Stress Disorder, 2013.
- 22 Weathers FW, Bovin MJ, Lee DJ, et al. The Clinician-Administered PTSD scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans. *Psychol Assess* 2018;30:383–95.
- 23 Weathers FW, Keane TM, Davidson JRT. Clinician-administered PTSD scale: a review of the first ten years of research. *Depress Anxiety* 2001;13:132–56.
- 24 Cwik JC, Woud ML. *Deutsche Übersetzung Der CLINICIAN-ADMINISTERED PTSD scale for DSM-5 (CAPS-5)*. Bochum: Mental Health Research and Treatment Center, Ruhr-Universität Bochum, 2015.
- 25 Bovin MJ, Marx BP, Weathers FW, et al. Psychometric properties of the PTSD checklist for diagnostic and statistical manual of mental Disorders—Fifth edition (PCL-5) in veterans. *Psychol Assess* 2016;28:1379–91.
- 26 Weathers FW, Ruscio AM, Keane TM. Psychometric properties of nine scoring rules for the Clinician-Administered posttraumatic stress disorder scale. *Psychol Assess* 1999;11:124–33.
- 27 Schnyder U, Moergeli H. German version of clinician-administered PTSD scale. *J Trauma Stress* 2002;15:487–92.
- 28 Boeschoten MA, Van der Aa N, Bakker A, et al. Development and evaluation of the Dutch Clinician-Administered PTSD scale for DSM-5 (CAPS-5). *Eur J Psychotraumatol* 2018;9:1546085.
- 29 Boysan M, Guzel Ozdemir P, Yilmaz E, et al. Psychometric properties of the Turkish version of the Clinician-Administered PTSD scale for diagnostic and statistical manual of mental disorders, fifth edition (Turkish CAPS-5). *Psychiatry and Clinical Psychopharmacology* 2017;27:173–84.
- 30 Margraf J, Cwik JC. *Mini-DIPS open access: Diagnostisches Kurzinterview bei psychischen Störungen*. 2nd edn. Bochum: Forschungs- und Behandlungszentrum für psychische Gesundheit, Ruhr-Universität Bochum, 2017.
- 31 Weathers FW, Blake DD, Schnurr PP, et al. *Life events checklist 5 (LEC-5)*. Boston, Massachusetts: National Center for Posttraumatic Stress Disorder, 2013.
- 32 Cwik JC, Woud ML. *Deutsche Übersetzung Der life events CHECKLIST-5 (LEC-5)*. Bochum: Ruhr-Universität Bochum, 2015.
- 33 Margraf J, Cwik JC, Pflug V, et al. Strukturierte klinische Interviews Zur Erfassung psychischer Störungen über die Lebensspanne: Gütekriterien und Weiterentwicklungen Der DIPS-Verfahren. *Zeitschrift für Klinische Psychologie und Psychotherapie* 2017;46:176–86.
- 34 Margraf J, Cwik JC, Suppiger A, et al. *Dips open access: Diagnostisches interview bei psychischen Störungen*. 5th edn. Bochum: Forschungs- und Behandlungszentrum für psychische Gesundheit, Ruhr-Universität Bochum, 2017.
- 35 Margraf J. *Diagnostisches Kurz-Interview psychischer Störungen (Mini-DIPS)*. Weinheim: Beltz, 1994.
- 36 In-Albon T, Suppiger A, Schlup B, et al. Validität des Diagnostischen interviews bei psychischen Störungen (dips für DSM-IV-TR). *Zeitschrift für Klinische Psychologie und Psychotherapie* 2008;37:33–42.
- 37 Krüger-Gottschalk A, Knaevelsrud C, Rau H, et al. The German version of the posttraumatic stress disorder checklist for DSM-5 (PCL-5): psychometric properties and diagnostic utility. *BMC Psychiatry* 2017;17:379.
- 38 Blevins CA, Weathers FW, Davis MT, et al. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. *J Trauma Stress* 2015;28:489–98.
- 39 Wortmann JH, Jordan AH, Weathers FW, et al. Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychol Assess* 2016;28:1392–403.
- 40 Ashbaugh AR, Houle-Johnson S, Herbert C, et al. Psychometric validation of the English and French versions of the posttraumatic stress disorder checklist for DSM-5 (PCL-5). *PLoS One* 2016;11:e0161645–e45.
- 41 Armour C, Tsai J, Durham TA, et al. Dimensional structure of DSM-5 posttraumatic stress symptoms: support for a hybrid anhedonia and Externalizing behaviors model. *J Psychiatr Res* 2015;61:106–13.
- 42 Sween J, Bondjers K, Willebrand M. Psychometric properties of the PTSD checklist for DSM-5: a pilot study. *Eur J Psychotraumatol* 2016;7:30165.
- 43 Verhey R, Chibanda D, Gibson L, et al. Validation of the posttraumatic stress disorder checklist – 5 (PCL-5) in a primary care population with high HIV prevalence in Zimbabwe. *BMC Psychiatry* 2018;18:1–8.
- 44 Weiss DS, Marmar CR. The Impact of Event Scale - Revised. In: Wilson J, Keane TM, eds. New York: Guilford, 1997: 399–411.
- 45 Weiss DS. The Impact of Event Scale: Revised. In: WJ P, TC S, eds. *Cross-Cultural assessment of psychological trauma and PTSD international and cultural psychology series*. Boston, MA: Springer, 2007: 219–38.
- 46 Maercker A, Schützwohl M. Erfassung von psychischen Belastungsfolgen: die impact of event Skala-revidierte version (IES-R). *Diagnostica* 1998;44:130–41.
- 47 Creamer M, Bell R, Failla S. Psychometric properties of the impact of event Scale—Revised. *Behav Res Ther* 2003;41:1489–96.
- 48 Beck JG, Grant DM, Read JP, et al. The impact of event Scale-Revised: psychometric properties in a sample of motor vehicle accident survivors. *J Anxiety Disord* 2008;22:187–98.
- 49 King DW, Orazem RJ, Lauterbach D, et al. Factor structure of posttraumatic stress disorder as measured by the impact of event Scale-Revised: stability across cultures and time. *Psychological Trauma: Theory, Research, Practice, and Policy* 2009;1:173–87.
- 50 Maercker A, Schützwohl M. *Testbeschreibung: IES-R*. Zürich: Universität Zürich, 1998.
- 51 Dozois DJA, Dobson KS, Ahnberg JL. A psychometric evaluation of the Beck depression Inventory-II. *Psychol Assess* 1998;10:83–9.
- 52 Arnau RC, Meagher MW, Norris MP, et al. Psychometric evaluation of the Beck depression Inventory-II with primary care medical patients. *Health Psychology* 2001;20:112–9.
- 53 Beck AT, Steer RA, Ball R, et al. Comparison of Beck depression Inventories-IA and-II in psychiatric outpatients. *J Pers Assess* 1996;67:588–97.
- 54 Hautzinger M, Keller F, Kühner C, et al. *BDI-II. revision*. Frankfurt/Main: Harcourt Test Services, 2006.
- 55 Kühner C, Bürger C, Keller F, et al. Reliabilität und Validität des revidierten Beck-Depressionsinventars (BDI-II): Befunde AUS deutschsprachigen Stichproben. *Nervenarzt* 2007;78:651–6.
- 56 Breslau N, Davis GC, Peterson EL, et al. A second look at comorbidity in victims of trauma: the posttraumatic stress disorder-major depression connection. *Biol Psychiatry* 2000;48:902–9.
- 57 Contractor AA, Durham TA, Brennan JA, et al. DSM-5 PTSD's symptom dimensions and relations with major depression's symptom dimensions in a primary care sample. *Psychiatry Res* 2014;215:146–53.
- 58 Gonzalez OI, Novaco RW, Reger MA, et al. Anger intensification with combat-related PTSD and depression comorbidity. *Psychological Trauma: Theory, Research, Practice, and Policy* 2016;8:9–16.
- 59 Maes M, Mylle J, Delmeire L, et al. Psychiatric morbidity and comorbidity following accidental man-made traumatic events: incidence and risk factors. *Eur Arch Psychiatry Clin Neurosci* 2000;250:156–62.
- 60 Perkonig A, Kessler RC, Storz S, et al. Traumatic events and post-traumatic stress disorder in the community: prevalence, risk factors and comorbidity. *Acta Psychiatr Scand* 2000;101:46–59.
- 61 Spielberger CD, Gorsuch RL, Lushene R, et al. *Manual for the State-Trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press, 1983.
- 62 Laux L, Glanzmann P, Spielberger CD. *Das State-Trait-Angstinventar, Theoretische Grundlagen und Handanweisungen*. Weinheim: Beltz-PVU, 1981.
- 63 Bernstein EM, Putnam FW. Development, reliability, and validity of a dissociation scale. *J Nerv Ment Dis* 1986;174:727–35.
- 64 Frischholz EJ, Braun BG, Sachs RG, et al. The dissociative experiences scale: further replication and validation. *Dissociation: Progress in the Dissociative Disorders* 1990;3:151–3.
- 65 Spitzer C, Stieglitz R-D, Freyberger HJ. *Fragebogen zu Dissoziativen Symptomen (fds): ein Selbstbeurteilungsverfahren Zur*

- syndromalen Diagnostik dissoziativer Phänomene. 3 ed. Göttingen: Hogrefe, 2014.
- 66 Renneberg B, Schmidt-Rathjens C, Hippin R, et al. Cognitive characteristics of patients with borderline personality disorder: development and validation of a self-report inventory. *J Behav Ther Exp Psychiatry* 2005;36:173–82.
 - 67 Renneberg B, Seehausen A. Fragebogen zu gedanken und Gefühlen (FGG): ein screening instrument für Borderline-spezifisches Denken. *Zeitschrift für Klinische Psychologie und Psychotherapie* 2010;39:170–8.
 - 68 Dorr F, Bengel J. Screening Zur komplexen Posttraumatischen Belastungsstörung (SkPTBS). Universität Freiburg: Freiburg, 2017.
 - 69 World Health Organization. ICD-11 beta draft, 2017. Available: <https://icd.who.int/dev11/l-m/en> [Accessed 13.02.2019].
 - 70 World Health Organization. ICD-11 for mortality and morbidity statistics, 2018. Available: <https://icd.who.int/browse11/l-m/en/#/http%3a%2f%2fid.who.int%2fid%2fentfity%2f437815624>
 - 71 Dorr F, Firus C, Kramer R, et al. Entwicklung und Prüfung eines screenings Zur komplexen Posttraumatischen Belastungsstörung (SkPTBS). *Psychotherapie, Psychosomatik, Medizinische Psychologie* 2016;66:441–8.
 - 72 Dorr F, Sack M, Med PP, et al. Validierung des screenings Zur komplexen Posttraumatischen Belastungsstörung (SkPTBS) – revision. *Psychotherapie, Psychosomatik, Medizinische Psychologie* 2018.
 - 73 Dorr F, Bengel J. Auswertung zum SkPTBS [Excel spreadsheet]. Freiburg: Universität Freiburg, 2017. <http://portal.uni-freiburg.de/psychologie/abteilungen/Rehabilitationspsychologie/downloads/auswertung-zum-skptbs/view>
 - 74 Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the scl-90: a study in construct validation. *J Clin Psychol* 1977;33:981–9.
 - 75 Franke GH. Symptom Checkliste von L. In: R. Derogatis - Deutsche Version (SCL-90-R. 2nd edn. Göttingen: Beltz Test, 2000.
 - 76 Paap MCS, Meijer RR, Van Bebber J, et al. A study of the dimensionality and measurement precision of the SCL-90-R using item response theory. *Int J Methods Psychiatr Res* 2011;45:39–55.
 - 77 Derogatis LR. SCL-90-R: administration, scoring and procedures manual. 3rd edn. Minneapolis, MN: National Computer Systems, 1994.
 - 78 Suppiger A, In-Albon T, Hendriksen S, et al. Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. *Behav Ther* 2009;40:272–9.
 - 79 Paulhus DL. The balanced inventory of desirable responding. Toronto/Buffalo: Multi-Health Systems, 1998.
 - 80 Paulhus DL. Balanced inventory of desirable responding: reference manual for BIDR version 6. Vancouver, Canada: University of British Columbia, 1994.
 - 81 Musch J, Brockhaus R, Bröder A. Ein Inventar Zur Erfassung von zwei Faktoren sozialer Erwünschtheit. *Diagnostica* 2002;48:121–9.
 - 82 Satorra A, Bentler PM. Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika* 2010;75:243–8.
 - 83 Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 2001;66:507–14.
 - 84 G*Power [program]. 3.1.9.3 version: Uiversität Kiel, Germany 2013.
 - 85 Faul F, Erdfelder E, Buchner A, et al. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;41:1149–60.
 - 86 Faul F, Erdfelder E, Lang A-G, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;39:175–91.
 - 87 Pickover AM, Lipinski AJ, Dodson TS, et al. Demand/withdraw communication in the context of intimate partner violence: implications for psychological outcomes. *Journal of Anxiety Disorders* 2016;2017:95–102.
 - 88 Joiner TE, Walker RL, Pettit JW, et al. Evidence-Based assessment of depression in adults. *Psychol Assess* 2005;17:267–77.
 - 89 Silverman WK, Ollendick TH. Evidence-Based assessment of anxiety and its disorders in children and adolescents. *J Clin Child Adolesc Psychol* 2005;34:380–411.
 - 90 Ehler U. Eine Psychotherapie ist immer nur so gut wie ihre Diagnostik. *Verhaltenstherapie* 2007;17:81–2.
 - 91 Rettew DC, Lynch AD, Achenbach TM, et al. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *Int J Methods Psychiatr Res* 2009;18:169–84.
 - 92 Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009;374:609–19.
 - 93 Cwik JC, Margraf J. Information order effects in clinical psychological diagnoses. *Clin Psychol Psychother* 2017;24:1142–54.
 - 94 Cwik JC, Papen F, Lemke J-E, et al. An investigation of diagnostic accuracy and confidence associated with diagnostic checklists as well as gender biases in relation to mental disorders. *Front Psychol* 2016;7:7.
 - 95 Merten EC, Cwik JC, Margraf J, et al. Overdiagnosis of mental disorders in children and adolescents (in developed countries). *Child Adolesc Psychiatry Ment Health* 2017;11:5.
 - 96 Blake DD, Weathers FW, Nagy LM, et al. The development of a clinician-administered PTSD scale. *J Trauma Stress* 1995;8:75–90.
 - 97 Armour C, Müllerová J, Elhai JD. A systematic literature review of PTSD's latent structure in the diagnostic and statistical manual of mental disorders: DSM-IV to DSM-5. *Clin Psychol Rev* 2016;44:60–74.
 - 98 Pielmaier L, Maercker A. Diagnostik in Der Traumatherapie. *Psychotherapie im Dialog* 2012;13:30–5.
 - 99 Flatten G, Gast U, Hofmann A, et al. S3 - Leitlinie Posttraumatische Belastungsstörung. *Trauma & Gewalt* 2011;3:202–10.
 - 100 Weathers FW, Nagy LM, Danny G, et al. The development of a Clinician-Administered PTSD scale 1995;8.
 - 101 Palmieri PA, Weathers FW, Difede J, et al. Confirmatory factor analysis of the PTSD checklist and the Clinician-Administered PTSD scale in disaster workers exposed to the world Trade center ground zero. *J Abnorm Psychol* 2007;116:329–41.
 - 102 Priebe K, Kleindienst N, Schropp A, et al. Defining the index trauma in post-traumatic stress disorder patients with multiple trauma exposure: impact on severity scores and treatment effects of using worst single incident versus multiple traumatic events. *Eur J Psychotraumatol* 2018;9:1486124–24.