Meta-Psychology, 2024, vol 8, MP.2023.3958
https://doi.org/10.15626/MP.2023.3958
Article type: Original Article

Open data: Yes
Open materials: Yes
Open and reproducible analysis: Yes
Open reviews and editorial process: Yes
Preregistration: No

Edited by: Rickard Carlsson
Reviewed by: Williams M.N., Winter, A., Eisenbarth, H.
Analysis reproduced by: Lucija Batinović
All supplementary files can be accessed on OSF:
https://doi.org/10.17605/OSF.IO/A47ZK

The many faces of early life adversity - Content overlap in validated assessment instruments as well as in fear and reward learning research

Alina Koppold[1], Julia Ruge[1], Tobias Hecker[2 3], and Tina B. Lonsdorf[1 4]
[1]University Medical Center Hamburg-Eppendorf, Department of Systems Neuroscience, Hamburg, Germany
[2]University of Bielefeld, Clinical Psychology & Violence Research, Bielefeld, Germany
[3]University of Bielefeld, Institute of Interdisciplinary Conflict and Violence Research, Bielefeld, Germany
[4]University of Bielefeld, Biological Psychology and Cognitive Neuroscience, Bielefeld, Germany

The precise assessment of childhood adversity is crucial for understanding the impact of aversive events on mental and physical development. However, the plethora of assessment tools currently used in the literature with unknown overlap in childhood adversity types covered hamper comparability and cumulative knowledge generation. In this study, we conducted two separate item-level content analyses of in total 35 questionnaires aiming to assess childhood adversity. These include 13 questionnaires that were recently recommended based on strong psychometric properties as well as additional 25 questionnaires that were identified through a systematic literature search. The latter provides important insights into the actual use of childhood adversity questionnaires in a specific, exemplary research field (i.e., the association between childhood adversity and threat and reward learning). Of note, only 3 of the recommended questionnaires were employed in this research field. Both item-wise content analysis illustrate substantial heterogeneity in the adversity types assessed across these questionnaires and hence highlight limited overlap in content (i.e., adversity types) covered by different questionnaires. Furthermore, we observed considerable differences in structural properties across all included questionnaires such as the number of items, age ranges assessed as well as the specific response formats (e.g., binary vs. continuous assessments, self vs. caregiver). We discuss implications for the interpretation, comparability and integration of the results from the existing literature and derive specific recommendations for future research. In sum, the substantial heterogeneity in the assessment and operationalization of childhood adversity emphasizes the urgent need for theoretical and methodological solutions to promote comparability, replicability of childhood adversity assessment and foster cumulative knowledge generation in research on the association of childhood adversity and physical as well as psychological health.

Keywords: measurement heterogeneity, content analysis, early childhood adversity, childhood maltreatment, questionnaires

Introduction

Exposure to childhood adversity has been established as a strong risk factor and predictor for the development of psychopathology during the last decades (Anda et al., 2006; Danese & Widom, 2021; Felitti et al., 1998; Gilbert et al., 2009; C. Heim & Nemeroff, 2001; C. M. Heim et al., 2019; McLaughlin et al., 2015; Teicher et al., 2021). In particular as exposure to childhood adversity is rather common - approximately 60% of all children and adolescents are exposed to at least one adverse event (Madigan et al., 2023) - it is linked to substantial individual suffering and societal costs (Hughes et al., 2021). Childhood adversity has been defined in the literature as experiences that necessitate substantial adaptation by the child and deviate from the expected environment based on the assumption that appropriate input is fundamental for the normative (neuro-) developmental maturation (Cicchetti & Lynch, 1995; McLaughlin et al., 2021). In the literature, childhood adversity is assessed through a variety of approaches including official records, interviews or questionnaires (as either self-report or caregiver/parent report versions). Research on childhood adversity and in particular cumulative knowledge generation in the field is, however, hampered by measurement and operationalizational challenges (Elson et al., 2023; Flake & Fried, 2019; Flake et al., 2017). As for general trauma, there

is a lack of an agreed-up definition as well as substantial heterogeneity in assessment instruments. More precisely, it has recently been demonstrated that a multitude of different measures exist for the assessment of general trauma that differ in the number and type of specific trauma types assessed (Karstoft & Armour, 2022). As a consequence, individuals classified as exposed in one study, may not be classified as exposed in another study. This assessment heterogeneity may in part originate from definitions evolving over time. More precisely, the criteria to define "trauma", a necessary precondition for the diagnosis of post-traumatic stress disorder (PTSD), differ between editions of diagnostic criteria (e.g., DSM-III, DSM-IV, ICD-11, Frueh et al., 2004; Karstoft & Armour, 2022; Slep et al., 2015). As a consequence, different definitions of (presumably) the same construct (i.e traumatic experiences), are at the base of different questionnaires used for the assessment of general trauma experience. Similarly, standardized definitions and assessment instruments of childhood adversity are lacking (Slep et al., 2015). For instance, the DSM-5 and DSM-IV place childhood adversity categories (e.g., physical or sexual abuse) under V codes (i.e., non-mental disorder conditions). The ICD-10 includes child maltreatment exposures in the Z codes (similar to DSM's V codes), T codes (injury), and Y codes (external causes of injury and death). However, none of these codes (T, V, Y, Z) are operationalized by clear criteria for defining exposure (Slep et al., 2015), cover predominantly physical events while generally neglecting socio-emotional adverse events and do not discriminate between exposure and experience (McLaughlin et al., 2021). Relatedly, childhood adversity questionnaires developed early often focused on physical and sexual violence (e.g., Castelda et al., 2007; Felitti et al., 1998; Lynch & Cicchetti, 1998), while more recently developed assessment tools also cover additional facets such as emotional violence, neglect, peer, and/or witnessed violence (Teicher & Parigger, 2015). Both meta-scientific research and clinically-oriented research highlighted that such heterogeneity in measurement and assessment tools is a key threat for comparability, clinical translation and cumulative science alike (Elson et al., 2023; Flake & Fried, 2019; Flake et al., 2017). For instance, prospective and retrospective assessment instruments of childhood adversity did not identify the same individuals as "exposed" (Baldwin et al., 2019). Yet, the extent to which differences in content (i.e., adversity types) or differences in structural properties (e.g., response format, age ranges, valence of the experience) between different questionnaires contribute to heterogeneity in the literature remains unclear to date. To this end, a systematic investigation and illustration of measurement heterogeneity is an important first step to "ensure that we understand what we talk about when we talk about"childhood adversity" (cf. Karstoft & Armour, 2022) and hence putting this important research field on a solid foundation. In light of the co-existence of a plethora of distinct questionnaires for the assessment of general childhood adversity, it is a challenge for researchers, particularly those new to the field, to select the optimal assessment tool for their purposes from this rich buffet. Hence, the overarching aim of this work is to provide guidance by establishing a detailed overview on the currently used measures. To this end, we aim to inform decision-making, facilitate comparisons, and contribute to the cumulative generation and integration of knowledge from various findings.

More precisely, this study comprises three specific aims. The first aim involves examining the extent of item-based content overlap (i.e., childhood adversity types covered) between childhood adversity questionnaires that have been recommended based on strong psychometric properties. The second aim is to provide insight into item-based content overlap between childhood adversity questionnaires used in an exemplary research field (i.e., associations between childhood adversity and experimental threat and reward learning (Ruge et al., 2023). Finally, the third aim is to provide a comparative and comprehensive description and overview of the structural characteristics of the identified childhood adversity assessment questionnaires. Together, these results will allow us to distill clear considerations and recommendations for future research and provide much needed guidance in interpreting and integrating the current literature.

Method

1. Selection of Questionnaires

Only self-report and caregiver-report measures focusing on exposure to adverse events during childhood and adolescence were included. Official reports, interviews, and measures focusing on stress rather than adverse events were not considered in line with the approach of previously published content analyses (e.g., Karstoft and Armour, 2022 on General Trauma; Fried, 2017 on depression). A detailed list of all questionnaires is provided in Supplementary Table 1.

a. General Content Analysis.
First, we aimed to investigate the item-based overlap between questionnaires, which i) capture adverse childhood exposures, ii) are the most recurrently validated questionnaires from 2010 to 2020 (as reported in Georgieva et al., 2022) and iii) show strong psychometric properties in a recent systematic review (inter-

nal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, or responsiveness, Saini et al., 2019). We included all questionnaires that met at least 3 strong to moderate criteria, as determined by the Consensus-based Standards for the Selection of Health Status Measurement Instruments (COSMIN) checklist in both publications (Georgieva et al., 2022; Saini et al., 2019). In the study by Saini et al. (2019), a total of 10 questionnaires (CTF-long form, CTQ-short form, MACE, CATS; ETI-SR, AE-III, CCMI, BCAS, ETI, CEVQ) fulfilled these criteria. In the study by Georgieva et al. (2022), 5 questionnaires (CTQ-short form, MACE, CAPI, IPARAN, PAT2.0) met the specified criteria. The overlap of two questionnaires led to the inclusion of a total of 13 questionnaires. Both studies (Georgieva et al., 2022; Saini et al., 2019) relied on the childhood adversity definition of the WHO Consultation on Child Abuse Prevention states: 'Child abuse or maltreatment constitutes all forms of physical and/or emotional ill-treatment, sexual abuse or negligent treatment or commercial or other exploitation, resulting in actual or potential harm to the child's health, survival, development or dignity in the context of a relationship of responsibility, trust or power (World Health Organization, 2024)', which we hence also adopt here. Ten questionnaires were openly available online in the form of articles or dissertations; an additional two were obtained directly from the original authors, and one more was downloaded from a specific platform (ePROVIDE, https://eprovide.mapi-trust.org).

b. Field Specific Content Analysis.
Second, we aimed to investigate the item-based overlap between questionnaires which capture childhood adversity in "the field" using this specific research field as a case example. As learning is a central mechanism through which environmental inputs shape emotional and cognitive processes and ultimately behavior, learning mechanisms are key candidates potentially underlying the biological embedding of exposure to childhood adversity and its impact on development and risk for psychopathology (McLaughlin & Sheridan, 2016). Thus, we chose research on the association between childhood adversity and threat as well as reward learning (Ruge et al., 2023) as a case example. We extracted the questionnaires used to assess childhood adversity from the studies identified through a recent systematic review (Ruge et al., 2023). In brief, 25 different questionnaires were extracted from a total of 73 studies (threat: N=36, reward: N=37). Fifteen of the identified 25 questionnaires were included in the Jaccard overlap analyses that provide a metric about the item-based content overlap between pairs of measures (details in the section 'Statistical Analyses' below). We did not consider assessment instruments that focus on clinical diagnoses, low family socioeconomic status based on income or postcode and parental substance abuse only, questionnaires focussing on specific types of exposure only, such as threatening experiences only ["Child Abuse Potential Inventory" (CAPI, Milner, 1994), the Violence Exposure Scale Revised child version (VEX-R, Fox & Leavitt, 1995), the Violence Exposure Scale Revised parent version (VEX-R_parent, Fox & Leavitt, 1995), Exposure to Violence (ETV, Selner-O'Hagan et al., 1998), Conflict Tactics Scale (CTS, Straus et al., 1998)], deprivation only [Home Screening Questionnaire (HSQ, Frankenburg & Coons, 1986), Multidimensional Neglectful Behavior Scale (MNBS, Kantor et al., 2004), MacArthur SSS Scale of Subjective Social Status (MacArthur SSS Scale, Adler et al., 2000)] or bullying in school [Generalized Harassment Questionnaire (GHQ, Radoman et al., 2019), Peer Experiences Scale (PEQ, Casement et al., 2014)]. Additionally, the Traumatic Events Inventory (TEI) was excluded, as in the studies that employed the TEI, only a subset of three out of a total of 15 items was used (Morrison et al., 2022; Rowland et al., 2022). We did not include assessment instruments from which only a subset of questions was used in a specific study while no information was provided on which questions were included (Casement et al., 2014). Sixteen of the included questionnaires were openly available online in form of an article or dissertation, five were obtained from the original authors, and one was downloaded from specific platforms (ePROVIDE, https://eprovide.mapi-trust.org). Some questionnaires have undergone revisions or updates, e.g., the Life events checklist (LEC, Gray et al., 2004) has been modified to align with DSM-5 standards, resulting in the LEC-5. For our analyses, we used the exact version that was employed in the study of the example field.

2. Content Analyses

First, we extracted all items across questionnaires which resulted in 833 items for the general content analysis and 759 for the field-specific content analysis. Second, an item-based content overlap analysis (Fried, 2017; Karstoft & Armour, 2022) was conducted. In brief, based on the conservative approach used by Fried (2017) and Karstoft and Armour (2022), we identified matching adversity types across questionnaires. Two raters (AK, JR) independently determined the type of adversity of each item. First, based on two questionnaires, the two coders independently generated a pool of exposure types, guided by the subscales of the questionnaires. Subsequently, we compared our exposure type pools and reached an agreement. The established

pool of exposure types was then applied to all subsequent questionnaire items. Afterwards the coders were unblinded with respect to the results of the second coder and resolved mismatching assignments of items to exposure types collaboratively. Difficult decisions were discussed with the team members and decisions were made collaboratively. More precisely, if items were relatively similar and refer to the same adversity type, such as the item "There was someone to take me to the doctor if I needed it" (CTQ-SF's) and the item "You were not able to make sure your child got to a doctor or hospital when he/she needed it?" (PC-CTS), they were coded as measuring the same adversity type (i.e., physical neglect). Furthermore, categories were labeled as specific as possible to differentiate between different albeit related events. For instance, some questionnaires assess specific adversity types using a single broadly formulated item (e.g., sexual abuse in the LEC), while other questionnaires provide a more fine-grained assessment through multiple items referring to the same adversity type but different perpetrators (e.g., sexual abuse by household person or sexual abuse by a peer in the MACE) or adversity subtypes (e.g., sexual abuse nonspecific and sexual harassment in the MACE). As suggested by Follingstad et al. (2005), "emotional abuse" was operationalized here as including verbal abuse, and hence verbal abuse is not listed as a separate adversity category. Items regarding a child's behavior (e.g., sleep difficulties) as well as items about parental educational attitudes (e.g., "ignoring your child can be beneficial at times") were categorized as 'not event related' (see Figure 5,6) as these cannot be directly linked to potential maltreatment. A significant number of items did not cover a single specific content but were formulated more broadly covering a broader range of contents (e.g events). In these cases, a single item was scored more than once (for details see Figure Description 5 and 6). For instance, the item "Seeing someone die suddenly or get badly hurt or killed" (THS) was scored twice, once as content for 'witnessed murder' and once for 'witnessed violence'.

3. Statistical Analyses

The agreement in the classification of item-based content between both raters was assessed using Cohen's Kappa (unweighted, Watson and Petrie, 2010) in R Studio, utilizing the 'irr' package (Gamer et al., 2019). Jaccard-index based Item-wise Content Analysis .

The procedure as described in the following was identical for both the general content analysis and the field-specific analysis. In brief, the Jaccard Index, a widely used similarity coefficient for binary data, was calculated to estimate content overlap, which can range from 0 (no overlap) to 1 (complete overlap). The Jaccard Index, also known as the Jaccard similarity coefficient, is computed using the formula,

$$\frac{s}{u_1 + u_2 + s}$$

where s denotes the number of shared items between two questionnaires, and u_1 and u_2 represent the items that are exclusive to questionnaire1 and questionnaire2, respectively. As there is no widely accepted criterion for defining a weak or strong Jaccard similarity coefficient, the correlation coefficient rule proposed by Fried (2017) based on Evans (1996) is used here as benchmarks: a coefficient of 0.00-0.19 is considered very weak, 0.20-0.39, weak, 0.40-0.59 moderate, 0.60-0.79 strong, and 0.80-1.0 is considered very strong. The questionnaires used in the field-specific analysis are based on original studies from the systematic review (Ruge et al., 2023). In these original studies, the questionnaires were sometimes modified by the authors (e.g., by adding or removing items). In these cases (CLES, LEC, MNBS-CR, HSQ, GHQ, THQ, ETV, TEI, see Supplementary Table 1), the number of items included in the field-specific Jaccard overlap analysis corresponds to the number of items used in the respective studies, not the number of items in the original questionnaire. For the general content analysis, four questionnaires were excluded from Jaccard analyses due to their exclusive focus on a particular type of childhood adversity: the "Psychosocial Assessment Tool" (PAT2.0, Pai et al., 2008), tailored specifically for parents of children diagnosed with cancer; the "Identification of Parents at Risk for Child Abuse and Neglect" (IPARAN, Bouwmeester-Landweer, 2016), designed to assess parenting stress risk in parents with newborns; the "Child Abuse Potential Inventory" (CAPI, Milner, 1994); and the "Childhood Experiences of Violence Questionnaire" (CEVQ, Walsh et al., 2008), both designed to evaluate threat-related childhood adversity.

4. Comparative Overview: Descriptive Structural Characteristics of Childhood Adversity Questionnaires

To provide an overview of structural and format differences between questionnaires, we included all questionnaires from both content analyses, resulting in a total of 35, with 3 questionnaires overlapping (Figure 1). For each questionnaire, we extracted detailed information on the versions available and used in the content analyses described above. This included details on the response format, target population, the age range for which the original questionnaire was developed, valence and frequency of exposure, calculation of sum scores (quantification of exposure based

on responses), and available cut-off recommendations. Regarding questionnaires from the field-specific section, we recorded information from both the original questionnaire and any deviations found in the respective studies (e.g., number of items). For data analyses and visualizations as well as for the creation of the manuscript, we used the following R packages[1]. Data and code are available on Zenodo (doi: https:/doi.org/10.5281/zenodo.10695485).

Results

Inter-rater Agreement on Item-based Content

High agreement between rater 1 and rater 2 with respect to the item-based content were observed for both the general content analysis (Cohen's Kappa unweighted, $\kappa = 0.81$) and the field-specific content analysis ($\kappa = 0.86$).

Overview of Descriptive Structural Characteristics

The descriptives of the structural properties of the questionnaires revealed major differences between the instruments. First, the questionnaires differed significantly in their number of items. The smallest number of items was observed for the MacArthur SSS Scale, a deprivation-specific scale with 2 items, while the Assessing Environments had a maximum of 170 items (see Figure 2A). Second, the modality of the questionnaires varied from self-report, to caregiver report or self and caregiver report combined, semi-structured interview to age-specific self-report for adolescents, with self-reports generally being the most common modality (see Figure 2B). Notably, self-reports included both child self-reports prospectively and adult self-reports retrospectively (Figure 4B). Third, the purpose and target group varied from screening tools to diagnostic measures for both children and adults (see Figure 4B). Relatedly, some questionnaires assessed stressful or traumatic experiences (CTQ), others risk of exposure (HSQ, CLES), or mood and behavioral tendencies of parents (CAPI, IPARAN) vs. other specifically deprivation (MNBS), violence (VEX-R) or the stress level of children diagnosed with cancer (PAT2). Fourth, the specific age range was from 0-19 or lifetime, however, the age range from 0-17 was most common (Figure 3A). Interestingly, the end of childhood varied between an age of 17 and 19. Fifth, the responses varied from verbal options, i.e., dichotomous, open field questions, to different Likert scale versions (3-8 pt). Different Likert scale response options were often designed to capture the frequency of the event (see Figure 4A) and resulted in various scoring minima and maxima (Figure 3B). Moreover, some questionnaires also included nonverbal options (e.g., pictures in the MNBS-CR and VEX-R). Finally, the assessment of specific characteristics of the exposure to adversity, such as age, frequency and duration of exposure is captured in at least one characteristic by 26 out of all 35 questionnaires. More precisely, the age of exposure is assessed by 17 questionnaires, whereas 2 questionnaires focus specifically on the last 12 months (PC-CTQ and GHQ, for details on corresponding questionnaires see Supplementary Table 1). The duration or frequency is assessed by 18 questionnaires, and the experienced valence of the event by 8 questionnaires, moreover 2 questionnaires, the LEC and UCLA PTSD-RI, capture the role of the respondent in the event (victim, witnessed, learned about, Figure 3C).

General Content Analysis

Thirteen questionnaires assessing a total of 40 different adversity types were identified. Individual questionnaires captured between 7 (17.5%, CAPI) and 22 (55%, CCMI) of these adversity types (see Figure 5). The average item-based content overlap aiming to assess general childhood adversity (as calculated by the Jaccard Index for 9 out of the 13 questionnaires) ranged from 19% to 36%, indicating rather low similarity of adversity types across questionnaires (see Supplementary Table 2 for details). Overlap across questionnaires revealed an overall mean Jaccard Index of .29. At a descriptive

[1]R (Version 4.2.2; Team, 2022) and the R-packages ade4 (Bougeard and Dray, 2018; Chessel et al., 2004; Dray and Dufour, 2007, Version 1.7.20; Dray et al., 2007), bookdown (Version 0.30; Xie, 2016), corrplot2021 (Wei & Simko, 2021), data.table (Version 1.14.6; Dowle and Srinivasan, 2022), dplyr (Version 1.1.2; Wickham et al., 2023), effectsize (Version 0.8.2; Ben-Shachar et al., 2020), flextable (Version 0.9.4; Gohel and Skintzos, 2023), forcats (Version 0.5.2; Wickham, 2022a, ggalluvial-article (Brunson, 2020), ggplot2 (Version 3.4.0; Wickham, 2016), ggsankey (Version 0.0.99999; Sjoberg, 2023), gridGraphics (Version 0.5.1; Murrell and Wen, 2020), gt (Version 0.8.0; Iannone et al., 2022), irr (Version 0.84.1; Gamer et al., 2019), kableExtra (Version 1.3.4; Zhu, 2021), knitr (Version 1.41; Xie, 2015), lpSolve (Version 5.6.17; Berkelaar et al., 2022), officer (Version 0.6.3; Gohel, 2023), papaja (Version 0.1.1; Aust and Barth, 2022), patchwork (Version 1.1.2; Pedersen, 2022), purrr (Version 1.0.1; Wickham and Henry, 2023), rcartocolor (Version 2.0.0; Nowosad, 2018), RColorBrewer (Version 1.1.3; Neuwirth, 2022), readr (Version 2.1.3; Wickham et al., 2022), readxl (Version 1.4.1; Wickham and Bryan, 2022), stringr (Version 1.5.1; Wickham, 2022b), tibble (Version 3.2.1; Müller and Wickham, 2023), tidyr (Version 1.3.0; Wickham and Girlich, 2022), tidyverse (Version 1.3.2; Wickham et al., 2019), tinylabels (Version 0.2.3; Barth, 2022), viridis (Version 0.6.2; Garnier et al., 2021, 2022), and viridisLite (Version 0.4.1; Garnier et al., 2022Garnier et al., 2022)

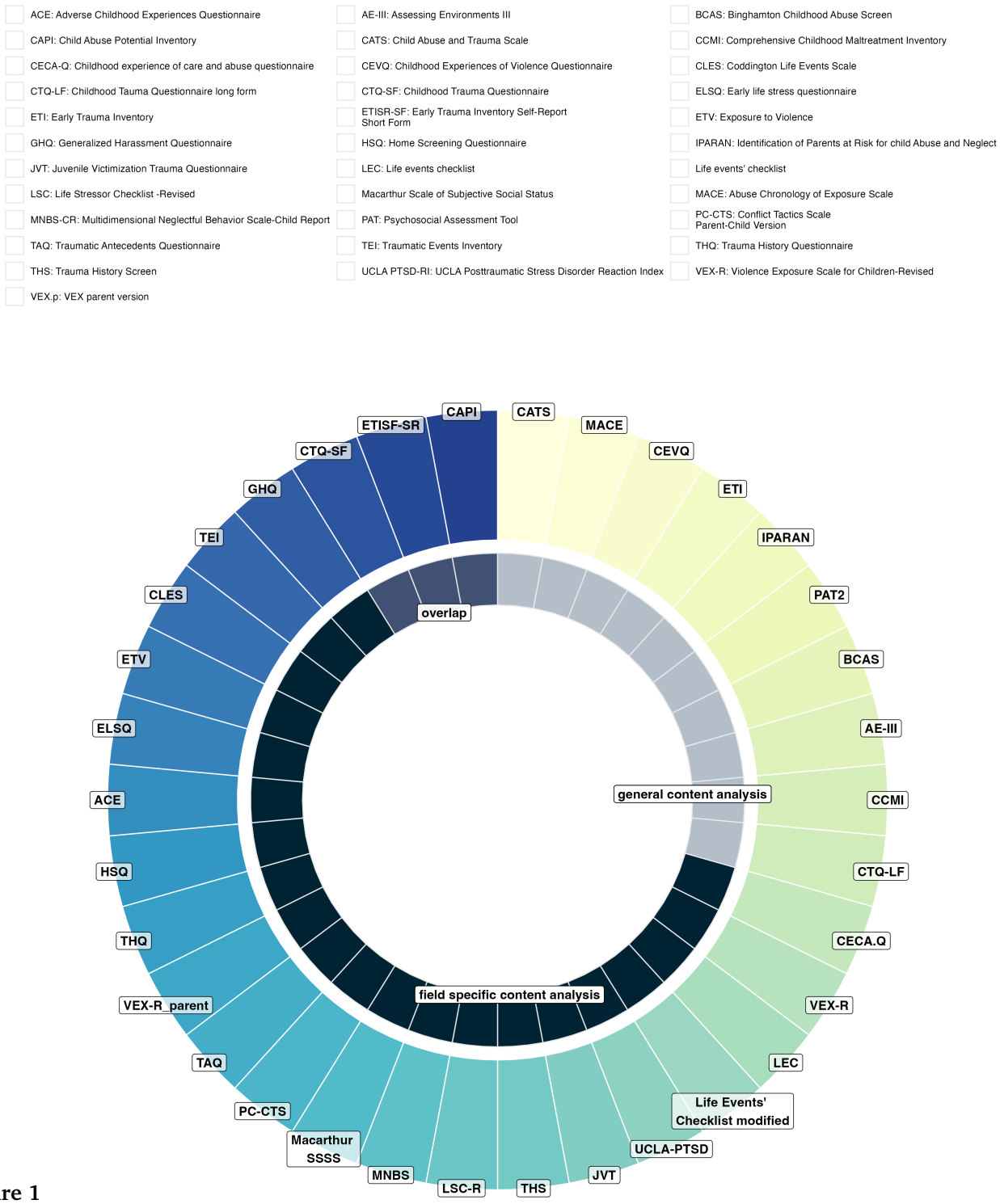


Figure 1

The illustration highlights which questionnaires (for details see Supplementary Table 1) were recommended based on strong psychometric properties in the literature (general content analysis, Georgieva et al., 2022; Saini et al., 2019) and observed in the field in studies included in a systematic literature search on the association between childhood adversity and threat as well as reward learning processes (field specific content analysis) or both (overlap, Figure 1). A color within the color scheme represents a specific questionnaire, each of which is indicated above the circle by its abbreviation in a box.

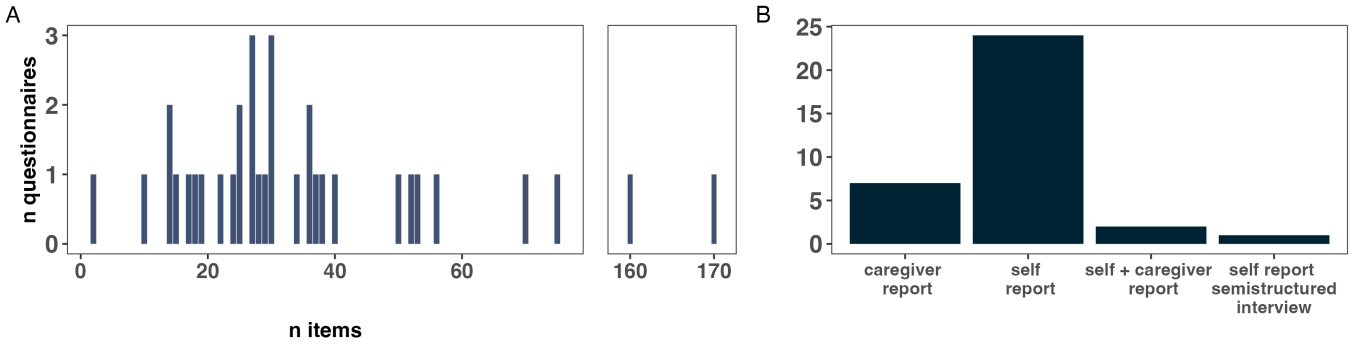


Figure 2

Histogram A illustrates the distribution of the number of items in each questionnaire.Individual questionnaires differed markedly in the number of items included (2 - 170 items, Figure 2A), inherently resulting in lower Jaccard Indices. Some studies in the field-specific content analysis, only used a subset of items of a respective questionnaire (for details see Supplementary Table 1), however, the number of items depicted here reflects the original number of items, whereas the field specific content analysis is based on the items used in the respective study in "the field". In total, 8 questionnaires (CLES, LEC, MNBS-CR, HSQ, GHQ, THQ, ETV and the TEI) were used in a modified version in the studies included in this review and hence included in the content analysis in this modified version. More specifically, for the CLES only 40 items out of originally 53 were used, for the modified version of the LEC (Scharfenort et al., 2016), items 1-3, 16, 17,19, 21, 24 were not used, for the MNBS-CR (Machlin et al., 2019; Milojevich et al., 2020), items 41-51 were not used, for the HSQ, items 6, 22, 29, 30, 34 were not used, for the GHQ, a modified version of the Generalized Workplace Harassment Questionnaire with 29 items, only 21 were used, for the THQ, only the last 6 items were used, for the ETV, only 13 out of 25 items were used, and for the TEI, only items 11, 12, 13 were used (Morrison et al., 2022; Rowland et al., 2022). For generation of the content analysis, we employed the modified versions of the questionnaire as reported in the publications included here. The extent of heterogeneity in assessment modalities, such as self-report and caregiver self-report (as shown in Figure 2B), was found to be quite diverse across questionnaires.

level, the BCAS had the lowest overlap with other questionnaires (.19) and CCMI had the largest overlap with all other questionnaires (.36; see Supplementary Table 2). Pairwise comparisons between the questionnaires, (Table 3), revealed - at a descriptive level, the lowest degree of overlap was between the CATS and the CTQ-SF (.14) as well as BCAS and CCMI, and the highest degree of overlap between the CTQ-LF and the AE-III (.58).

Field-specific Content Analysis

Twenty-two questionnaires assessing a total of 50 different adversity types were identified. Individual questionnaires captured between 1 (2%, MacArthur SSS Scale, PEQ) and 25 (50%, LSC-R) of these adversity types (see Figure 6). The average item-based content overlap aiming to assess general childhood adversity (as calculated by the Jaccard Index for 15 out of the 25 questionnaires) ranged from 1% to 34%, indicating rather low similarity of adversity types across questionnaires (see Supplementary Table 3 for details). Overlap across questionnaires revealed an overall mean Jaccard Index of .25. At a descriptive level, the PEQ had the lowest overlap with other questionnaires (.01), followed by the CECA-Q (.15). The TAQ had the largest overlap with all other questionnaires (.34; see Supplementary Table 3). At a descriptive level, pairwise comparisons between the questionnaires revealed that the lowest degree of overlap was between the PEQ and all other questionnaires (0.00), except the ACE (0.05), while the highest degree of overlap was observed between the UCLA-PTSD and the THQ (0.67). Interestingly, also in the general content analysis the CTQ (short form) revealed the lowest overlap with other questionnaires. Moreover, a study focusing on general adversity questionnaires (Karstoft & Armour, 2022) also identified the

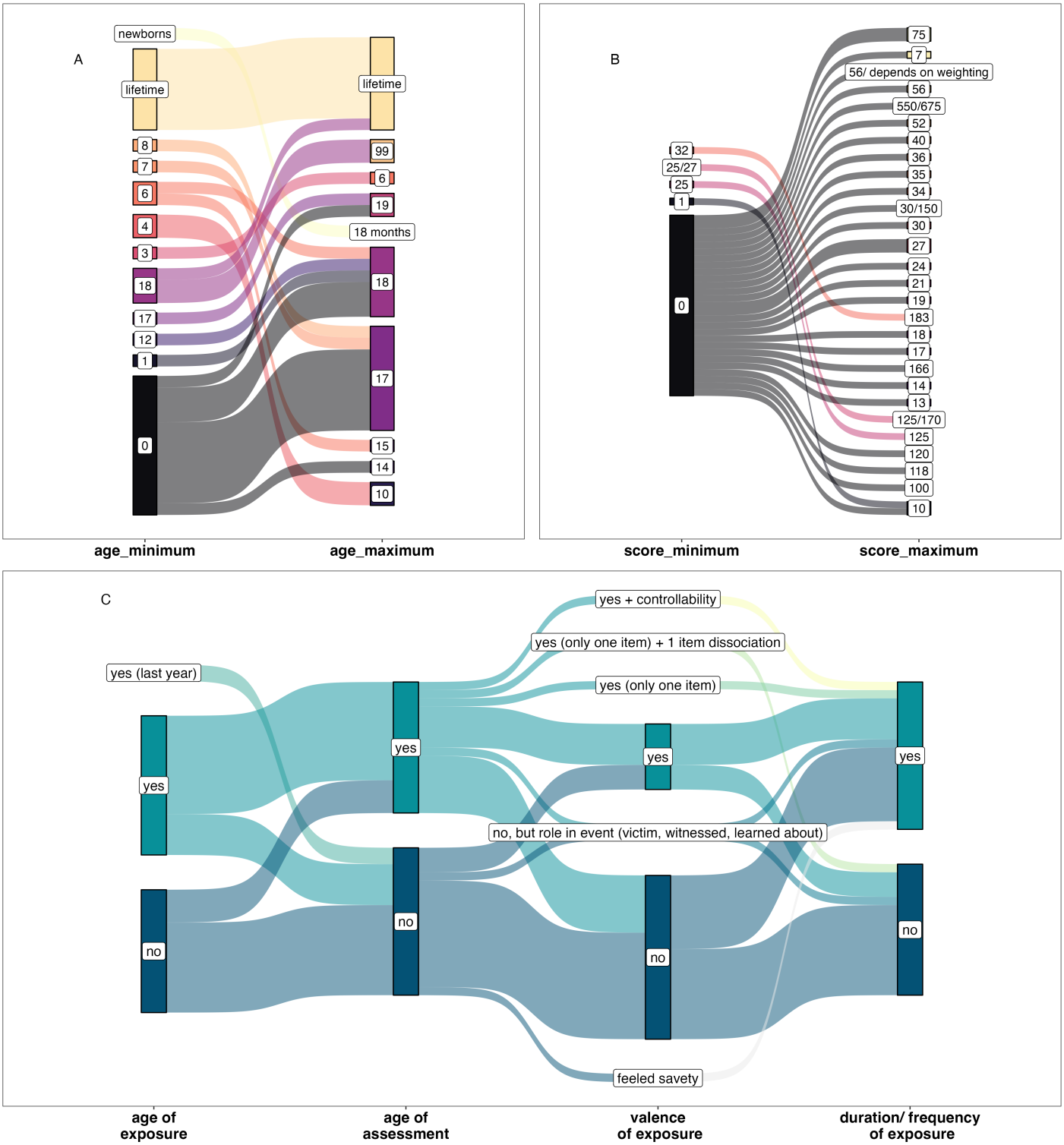


Figure 3

Sankeys illustrates the frequency of different age ranges based on the original questionnaire. Most refer to 0-17 years, beyond that a large part of the questionnaires can be applied spanning lifetime (A). The score ranges of the questionnaires are usually obtained by adding up responses of the items (in some questionnaires these are additionally assigned to subcategories and weighted, e.g., MACE). Notably, no single score range occurs more than once. Two values in a row indicate that for this questionnaire exists more than one scoring procedure (B). Specific characteristics of exposure to adversity (C) revealed that for 54.3 % (19/35) questionnaires the age of exposure was assessed, 45.7 % (16/35) age of assessment, 22.8 % (8/35) the valence of exposure, and 51.4 % (18/35) the frequency of exposure. For the HSQ that asks about the child's environment in depth (e.g., 'How often does someone get a chance to read stories to your child?') where these specific exposure questions were not relevant, they were coded as 'no'. Notably, in two questionnaires the age of exposure refers only to specific time points (JVT, ETV) or the previous year (PC-CTQ, GHQ). For detailed corresponding questionnaire information see Supplementary Table 1.

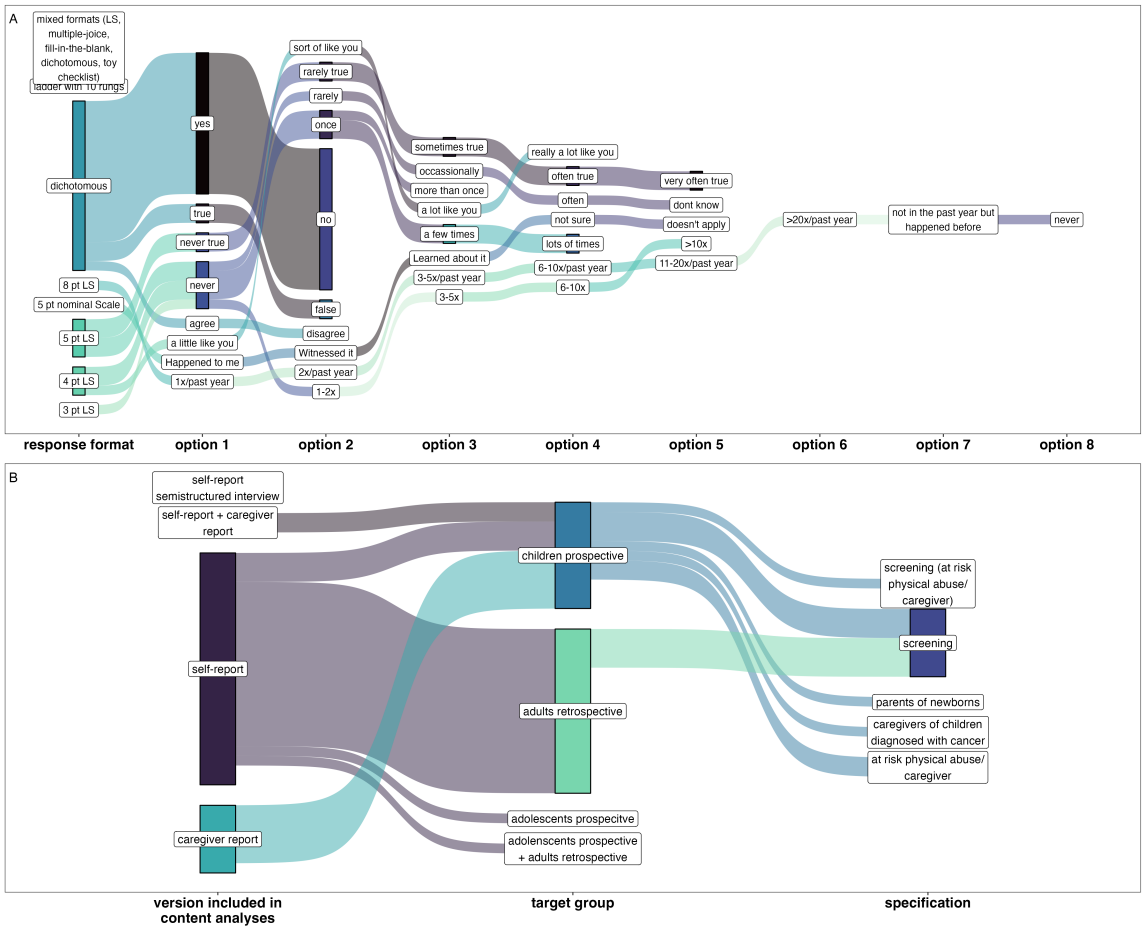


Figure 4

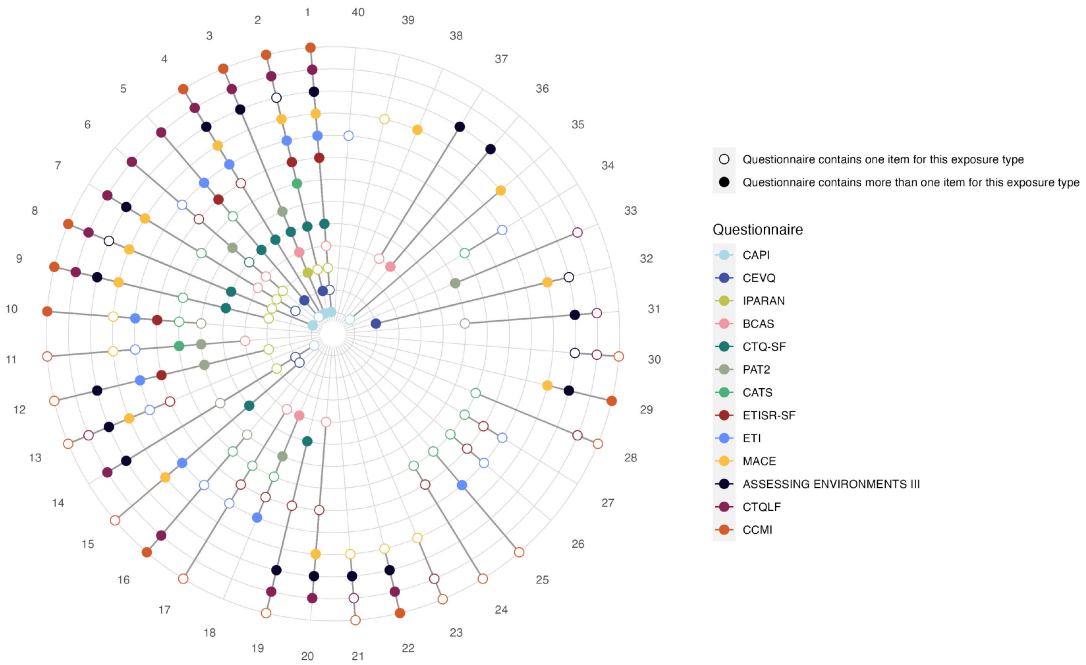
Sankey illustrates responses variations from verbal options, i.e., dichotomous, open field questions, to different Likert scale versions whereas some capture frequency some not and/ or specify the time frame of exposure (A). Note that option refers to a predefined set of answer choices. The majority of questionnaires were retrospective self-reports for adults (B). Questionnaires aiming to prospectively assess childhood varied substantially in modality (self vs. caregiver report) and purpose (e.g., at risk, maltreatment screening). Note: pt LS = point Likert Scale.

THQ as showing the highest overlap (.59) in content overlap analyses.

Table 1

Jaccard similarity index matrix across questionnaires. Values range from 0 (no overlap) to 1 (total overlap). Mean Jaccard index across questionnaires was .29, indicating weak overlap. Minimum mean overlap was .19 (BCAS) and maximum mean overlap .36 (CCMI; for details see Supplementary Table 2).

Questionnaire	CATS	CTQ-SF	ETI	ETISR-SF	MACE	BCAS	CCMI	CTQ-LF	AE-III
CATS		0.14	0.50	0.36	0.17	0.18	0.32	0.20	0.09
CTQ-SF	0.14		0.27	0.32	0.26	0.17	0.33	0.41	0.30
ETI	0.50	0.27		0.57	0.23	0.21	0.38	0.22	0.15
ETISR-SF	0.36	0.32	0.57		0.21	0.24	0.32	0.29	0.25
MACE	0.17	0.26	0.23	0.21		0.15	0.46	0.38	0.44
BCAS	0.18	0.17	0.21	0.24	0.15		0.14	0.19	0.24
CCMI	0.32	0.33	0.38	0.32	0.46	0.14		0.48	0.45
CTQ-LF	0.20	0.41	0.22	0.29	0.38	0.19	0.48		0.58
AE-III	0.09	0.30	0.15	0.25	0.44	0.24	0.45	0.58	



1 emotional abuse
2 sexual abuse non-specific
3 not event related
4 emotional neglect
5 physical abuse non-specific
6 substance abuse in household
7 witnessed domestic violence
8 physical neglect
9 physical abuse by household person
10 loss of close person
11 any other stressful or adverse event
12 witnessed illness/ injury
13 separation/ divorce parents
14 arguments
15 sexual harassment
16 threat of death or harm
17 witnessed violence
18 serious illness/ injury
19 feeling unwanted
20 feeling insufficient
21 self-reported socioeconomic status
22 feeling close
23 institutionalization
24 witnessed murder
25 war zone/ riots/ combat
26 serious accident
27 natural disaster
28 community violence
29 sexual abuse by household person
30 separation from family
31 family crime
32 bullying
33 financial problems
34 robbery
35 disinterest of parents
36 intimate partner violence in adolescents
37 crime committed
38 sexual abuse peer
39 care work
40 abortion / miscarriage (mother)

Figure 5

Co-occurrence of the 40 different childhood adversity types identified in the 13 questionnaires (with a total of 833 items) as recommended in (Georgieva et al., 2021; Saini et al., 2019). Three (7.5%) adversity types were only included in one individual questionnaire. Numbers refer to the event types listed below the circle. Colored dots in this figure represent the questionnaires listed in the legend on the right. Fully colored dots for an event type indicate that the respective questionnaire included more than a single item for this event type. Dots that are not filled indicate that only a single item captures this event in the respective questionnaire. Items highlighted in blue font indicate that this item was not filled out by the child but a caretaker (e.g., parent). In case a questionnaire item referred to two or more possible events (e.g., "witnessing severe illness OR death of a close person") and the exposure to of one of these events was sufficient to indicate the experience of childhood adversity , the item was coded as present twice in order to represent all possible content in the analysis. This was the case for the CCMI 7 items and AE-III 3 items.

Table 2

Jaccard similarity index matrix across questionnaires. Values range from 0 (no overlap) to 1 (total overlap). Mean Jaccard index across questionnaires was 0.25, indicating weak overlap. Minimum mean overlap was 0.01 for PEQ and maximum mean 0.34 for TAQ (for details see Supplementary Table 3).

Questionnaire	CECA.Q	CTQ-SF	ETISF-SR	JVQ	LEC	LSC	Life events' checklist	THS	UCLA-PTSD	ACE	CLES	ELSQ	TAQ	THQ	PEQ
CECA.Q		0.50	0.28	0.16	0.05	0.18	0.05	0.05	0.10	0.17	0.10	0.14	0.25	0.09	0.00
CTQ-SF	0.50		0.32	0.24	0.10	0.25	0.04	0.10	0.09	0.26	0.14	0.18	0.33	0.13	0.00
ETISF-SR	0.28	0.32		0.20	0.27	0.38	0.21	0.42	0.38	0.36	0.30	0.35	0.61	0.41	0.00
JVQ	0.16	0.24	0.20		0.21	0.31	0.09	0.27	0.35	0.38	0.12	0.32	0.26	0.37	0.05
LEC	0.05	0.10	0.27	0.21		0.27	0.12	0.47	0.42	0.10	0.17	0.26	0.25	0.53	0.00
LSC	0.18	0.25	0.38	0.31	0.27		0.39	0.28	0.34	0.37	0.29	0.37	0.47	0.37	0.00
Life events' checklist	0.05	0.04	0.21	0.09	0.12	0.39		0.24	0.22	0.22	0.32	0.30	0.20	0.25	0.00
THS	0.05	0.10	0.42	0.27	0.47	0.28	0.24		0.53	0.15	0.29	0.33	0.36	0.56	0.00
UCLA-PTSD	0.10	0.09	0.38	0.35	0.42	0.34	0.22	0.53		0.22	0.26	0.36	0.38	0.67	0.00
ACE	0.17	0.26	0.36	0.38	0.10	0.37	0.22	0.15	0.22		0.31	0.46	0.46	0.17	0.05
CLES	0.10	0.14	0.30	0.12	0.17	0.29	0.32	0.29	0.26	0.31		0.35	0.42	0.24	0.00
ELSQ	0.14	0.18	0.35	0.32	0.26	0.37	0.30	0.33	0.36	0.46	0.35		0.46	0.33	0.06
TAQ	0.25	0.33	0.61	0.26	0.25	0.47	0.20	0.36	0.38	0.46	0.42	0.46		0.36	0.00
THQ	0.09	0.13	0.41	0.37	0.53	0.37	0.25	0.56	0.67	0.17	0.24	0.33	0.36		0.00
PEQ	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.06	0.00	0.00	

Exploratory Analyses

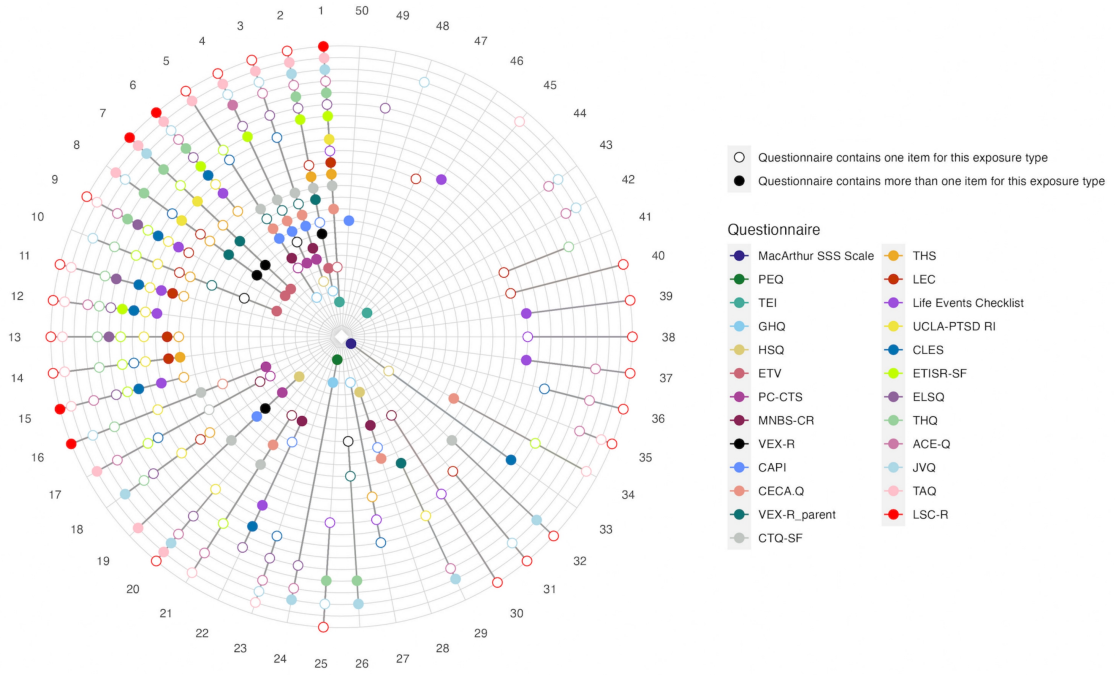
Based on a reviewer's comment, we also explored the content overlap within caregiver and self-report questionnaires by conducting a further content analysis. For caregiver questionnaires, a mean Jaccard Index of 0.22 across these questionnaires (CAPI, PC-CTS, VEX-R Caregiver Version, HSQ, CLES, IPARAN, PAT) was observed. Supplementary Table 4 lists the indices specific to each pair of caregiver questionnaire comparisons. We refer to Supplementary Figure 2 for an illustration of adversity types included in questionnaires separated for questionnaires using a) retrospective self report b) self report in pediatric samples as well as c) caregiver report in pediatric samples. The Jaccard indices for self-report questionnaires are illustrated in Table 2.

Discussion

Heterogeneity in measurement as well as assessment tools has been highlighted as a key threat for comparability between studies and cumulative science alike (Elson et al., 2023; Flake & Fried, 2019; Flake et al., 2017). Here we investigated item-based content overlap (i.e., adversity types) between different questionnaires used to assess childhood adversity. This will provide a first step for establishing a solid and common ground for future research on this important predictive risk factor for the development of both mental and physical disorders. More precise, we focus on item-based content overlap between i) validated questionnaires that have been recommended based on strong psychometric properties (n=13, Georgieva et al., 2021; Saini et al., 2019) as well as on item-based content overlap between questionnaires used in one translationally highly relevant exemplary research field (e.g., associations between childhood adversity and threat as well as reward learning, n=24 questionnaires). It is noteworthy that only 3 of the recommended questionnaires, namely CAPI, ETI-SR, and CTQ-SF, were in fact employed in the exemplary research field of associations between threat and reward learning and childhood adversity. The ETI-SR and CTQ-SF capture a broad range of different adversity types during childhood (CTQ-SF: 10 types, ETI-SR: 15 types; see Figure 6) through retrospective self-reports in adults while the CAPI is an established questionnaire for the assessment of parents' tendency towards physical abuse of children. These specific three questionnaires may be in fact used by scientists in practice, because they are fast to complete (according to the authors <15 min), provide clear scoring procedures for the evaluation, are available in multiple languages, and are freely available - at least the CTQ-SF and ETI-SR (MacDonald et al., 2016; Thombs et al., 2007). Limited Item-based Content Overlap across Childhood Adversity Questionnaires Both the general and field-specific item-based content analyses illustrate a limited overlap and hence substantial heterogeneity in content (i.e., adversity types) and number of adversities between the different questionnaires. None of the pairwise comparisons between two questionnaires yielded a content overlap of 0.60 or higher (mean content overlap .29 and .25 in the general and field-specific item-wise content analyses, respectively). This rather weak overlap among questionnaires indicates that different childhood adversity questionnaires should be compared with caution and may not be used interchangeably although this is often done in practice (see Ruge et al., 2023). In fact, different questionnaires may not even measure the same underlying "experience". These results from childhood adversity questionnaires are in line with recent reports of low content overlap in assessment instruments of general trauma (Karstoft & Armour, 2022) as well as previous reports on limited agreement in the classification of "exposed" and "unexposed" individuals across different childhood adversity questionnaires - for instance between the CTQ and the KERF (german version of the MACE questionnaire, kappa .39, Koppold et al., 2023).

Substantial Heterogeneity in Childhood Adversity Types across Questionnaires

Our results across both content analyses also provide a valuable overview on the different adversity types considered as childhood adversity across questionnaires (i.e., general content analysis: n=40, field-specific content analysis: n=50, see Figure 5 and 6), which provides valuable conceptual information for the field. In addition, our results allow us to extract how frequently individual childhood adversity types (i.e., item-based content) are included across questionnaires. More precisely, the the most consistently included content across questionnaires were sexual abuse non-specific' (20 questionnaires) and 'physical abuse non-specific' (17 questionnaires), while 'sexual abuse peer' (2 questionnaires), 'financial problems' (in household, 4 questionnaires), or 'family crime' (6 questionnaires) were only included in a minority of questionnaires. In providing a comprehensive overview on the item-based content of childhood adversity questionnaires, we aim to facilitate the identification of questionnaires particularly suitable for specific research questions by making decision-relevant information accessible to the community (e.g., specificity and broadness of items, not covered content). It is particularly striking that childhood adversity types listed as the most common worldwide (United Nations Population Fund, 2024), such as child



1	sexual abuse non-specific	11	any other stressful or adverse event	21	feeling unwanted	31	kidnapping	41	exposure to toxics/ chemicals
2	physical abuse non-specific	12	witnessed illness/ injury	22	arguments	32	sexual harassment	42	intimate partner violence in adolescents
3	physical neglect	13	natural disaster	23	separation from family	33	self-reported socioeconomic status	43	hate crime
4	emotional abuse	14	serious accident	24	bullying	34	feeling insufficient	44	sexual abuse by household person
5	emotional neglect	15	separation/ divorce parents	25	robbery	35	institutionalization	45	feeling close
6	loss of close person	16	physical abuse by household person	26	burglary/ theft/ vandalism	36	family crime	46	work event (e.g. job loss)
7	witnessed violence	17	substance abuse in household	27	loss of important possession	37	crime committed	47	causing death or harm
8	threat of death or harm	18	war zone/ riots/ combat	28	disinterest of parents	38	financial problems	48	sexual abuse peer
9	serious illness/ injury	19	not event related	29	community violence	39	abortion / miscarriage (mother)	49	adoption
10	witnessed murder	20	witnessed domestic violence	30	care work	40	witnessed serious accident	50	educational attitude

Figure 6

Co-occurrence of the 50 different childhood adversity types identified in the 24 questionnaires (with a total of 750 items) derived from the publications included in the systematic literature search (Ruge et al., 2023). Six (12%) adversity types were only included in one individual questionnaire. Numbers refer to the event types listed below the circle. Colored dots in this figure represent the questionnaires listed in the legend on the right. Fully colored dots for an event type indicate that the respective questionnaire included more than a single item for this event type. Dots that are not filled indicate that only a single item captures this event in the respective questionnaire. In case a study did not use all items of a questionnaire (see Supplementary Table 1), these are not included here. Items highlighted in blue font indicate that this item was not filled out by the child but a caretaker (e.g., parent). In case a questionnaire item referred to two or more possible events (e.g., "witnessing severe illness OR death of a close person") and the exposure to one of these events was sufficient to indicate the experience of childhood adversity, the item was coded as present twice in order to represent all possible content in the analysis. This was the case for the ETISR-SF 3 items, LSC-R 2 items, Life events checklist 1 item, CLES 1 item, ACE-Q 1 item, HSQ 2 items, TAQ 1 item, THQ 1 item, UCLA 3 items, THS 1 item. In total, 8 questionnaires (CLES, LEC, MNBS-CR, HSQ, GHQ, THQ, ETV and the TEI) were used in a modified version in the studies included in this review and hence included in the content analysis in this modified version. More specifically, for the CLES only 40 items out of originally 53 (Smith & Pollak, 2021) were used, for the modified version of the LEC (Scharfenort et al., 2016), items 1-3, 16, 17,19, 21, 24 were not used, for the MNBS-CR (Machlin et al., 2019; Milojevich et al., 2020), items 41-51 were not used, for the HSQ (Milojevich et al., 2020), items 6, 22, 29, 30, 34 were not used, for the GHQ, a modified version of the Generalized Workplace Harassment Questionnaire (Radoman et al., 2019) with 29 items, only 21 were used, for the THQ (Young et al., 2019), only the last 6 items were used, for the ETV (Estrada et al., 2020), only 13 out of 25 items were used, and for the TEI, only items 11, 12, 13 were used (Morrison et al., 2022; Rowland et al., 2022). For generation of this figure as well as the content analysis, we employed the modified versions of the questionnaire as reported in the publications included here.

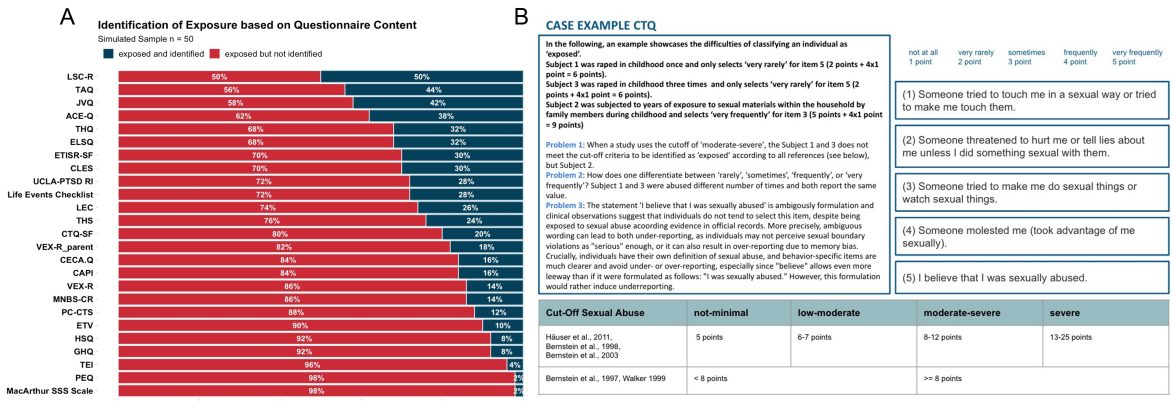


Figure 7

The barplot (A) illustrates, in a simulated sample of 50 subjects, each of whom experienced one of the 50 identified exposure types from the field-specific analysis (Figure 6), how many of the 50 individuals would be identified by each questionnaire. The "most diverse" exposure list would therefore be the LSC-R, which would only identify half of the exposed subjects. Exposure-specific questionnaires such as the MacArthur SSS Scale or the PEQ naturally perform even worse. Since some questionnaires weight responses to create sum-scores that determine whether certain cutoffs are exceeded, in B we depict the problem of forming sum scores for (1) the identification of exposed individuals and furthermore the fallacy of deriving severity from sum-scores.

marriage and child prostitution are not covered in any of 35 questionnaires included here. This impressively but also depressingly highlights a cultural gap and a potentially systematic bias. Furthermore, institutionalization, estimated to impact approximately 3.18 million to 9.42 million children worldwide (Desmond et al., 2020), is addressed in merely 6 out of the 35 included questionnaires, covered with only a single item each. This is particularly problematic as deprivation experiences due to institutionalization are highlighted prominently as a key adversity type in a central theoretical framework on the consequences of childhood adversity (i.e., DMAP, Sheridan and McLaughlin, 2014). In fact, our results show that in field specific research, deprivation is often assessed as single event types (such as official records of adoption or institutionalization, low family SES) without further characterization of individuals' exposure to or experience of other childhood adversity types (Ruge et al., 2023). This is particularly problematic because this approach lacks specificity to conclude derivation-specific effects, even though it's often done, and in addition hampers comparability of studies. In addition, it can be questioned if conceptualizing institutionalization or adoption as an adverse event and more specifically a deprivation experience is a valid approach. For instance, it can be questioned whether a person that has been adopted (i.e., exposure) has ever in fact experienced deprivation experiences. Similarly, conceptualizing the exposure to institutionalization as deprivation experience neglects the increased risk of institutionalized children to be exposed to threat-related or other adverse experiences and hence the exposure to institutionalization may not be linked specifically and exclusively to deviation experiences (see below 'fuzzy categories,' Smith and Pollak, 2022).

Potential Reasons for limited Item-based Content Overlap

Several considerations on potential reasons for the overserved generally low item-based content overlap between different questionnaires used to assess childhood adversity need to be discussed. First, despite all aiming to assess childhood adversity generally, some of the questionnaires included were - in part - designed for different purposes (e.g., screening, prospective risk assessment, retrospective report), different target groups (e.g., caregiver, exposed individual, pediatric sample vs. adult sample), or different target contexts (e.g., high

risk samples or general population samples in institutions, schools, refugee centers). Even though this may explain why only some questionnaires include items that are potentially traumatic (according to the DSM-5) such as natural disaster or exposure to war (Weems et al., 2021) while others do not, the substantial heterogeneity still is a challenge for cumulative knowledge generation in the field. Second, individual questionnaires differed markedly in the number of items included (2-170 items). Simply because more childhood adversity types are assessed, questionnaires with a high number of items theoretically will result in higher Jaccard indices (Karstoft & Armour, 2022) and prevalence rates for childhood adversity types (Breslau & Kessler, 2001). Yet, our data provide little support for generally higher overlap indices for questionnaires with higher numbers of items. Third, in some of the studies included in the field-specific analyses, specific items of questionnaires were either added and dropped ad hoc (Lambert et al., 2017; Milojevich et al., 2020) or the wording of items was changed (i.e., GHQ changed workplace harassment to bullying content). For instance, if all items that do not measure sexual abuse were excluded, it follows that this modified questionnaire has limited content overlap with other questionnaires. The practice of ad-hoc modification of (validated) questionnaires has been criticized previously (Flake et al., 2017) as it hampers the interpretation of the total sum score, replicability, and the reliability as well as validity of the measurement (Flake & Fried, 2019) and may also contribute to lower content-based overlap metrics.

Heterogeneity in Descriptive Structural Characteristics

In addition to heterogeneous content covered across childhood adversity questionnaires, differences in potentially relevant structural characteristics may impact on the interpretation and results in the literature. More precisely, a range of response formats was employed including verbal report, open field questions, Likert-type, pictures, and binary (yes/no; true/false) response types (for details see Supplementary Table 1). Different response formats are well known to impact the answer for a number of reasons and biases (for details see Menold and Bogner, 2014). In addition to response formats, it is noteworthy that only one quarter of the included questionnaires provided scoring or cut-off recommendations (see Supplementary Table 1). Furthermore, these procedures vary even for a single questionnaire, for instance with respect to scoring recommendations (Life Stressor Checklist-Revised, Wolfe et al., 2012) or cut-off criteria (Childhood Trauma Questionnaire, Bernstein et al., 1997; "Childhood trauma questionnaire: A retrospective self-report: manual," 1998). As a consequence, the classification in "exposed" vs. "unexposed" may vary even within users of the same questionnaire (see Figure 7). Moreover, the procedure of summing items to form a total or sum scores and deriving severity of exposure from this score has been criticized (Carlson et al., 2011; Lacey & Minnis, 2020). However, it has been shown that the number of different traumatic event types experienced improves the prediction of PTSD risk, whereas the additional assessment of event frequencies did not substantially enhance prediction rates further (Wilker et al., 2015). Future comparative work is necessary to shed further light on this topic. Furthermore, it might be a promising avenue for future work to take the experienced valence, controllability, or predictability (i.e., the experience) more strongly into account than the exposure itself. Together, we call for more attention to such descriptive structural properties of childhood adversity questionnaires and attempts to reduce ambiguity and hope that our systematic illustration serves as a facilitator along this path (see also BOX).

Heterogeneity in Target Age Groups and the Assessment of Developmental Timing

Despite all questionnaires aiming to assess childhood adversity, the specific target age group differs widely with respect to the minimum age, maximum age as well as age range. This applies to both age at exposure and age at assessment. While some questionnaires exclude late adolescence (e.g., the well established CTQ-CF, target age: 0-17 years) other questionnaires specifically focus on late adolescence and early adulthood - a developmental time well known to be characterized by substantial brain maturation processes (Giedd et al., 1999; Houston et al., 2013; Sowell et al., 1999, 2001). Relatedly, across both content analyses, more than half of all 35 questionnaires assess information on the developmental timing of adverse events and the duration or frequency of exposure (see Figure 3). This seems indeed informative as positive correlations between childhood adversity chronicity and the severity of Posttraumatic Stress Disorder (PTSD) has been reported (English et al., 2005; Jonson-Reid et al., 2012).

Valence and Controllability

Furthermore, although predictability has been highlighted to play a crucial role in the mechanisms linked to psychopathology according to a prominent theory on the consequences of childhood adversity (DMAP, McLaughlin et al., 2021 none of the questionnaires included here included assessments of (un)predictability. Relatedly, a single questionnaire assesses controllability of individual adversity and 8 of 35 questionnaires

included assessments of the valence of adversity - although in some questionnaires (MACE, TAQ or ETI-SR) not for all items. These experienced characteristics of the adverse event are assumed to be associated with neurobiological changes and should be assessed in the future (Smith & Pollak, 2021). In sum, as the universally applicable childhood adversity assessment instrument yet has to be invented, our systematic and comparative overview on different childhood adversity questionnaires, their specific content and key structural characteristics (target age group, assessment of predictability, valence or controllability) will aid researchers select the most appropriate instrument for their specific purposes. In addition, we anticipate that this overview will facilitate cumulative knowledge generation and critical comparisons across results in the literature by considering the respective childhood adversity questionnaire specific advantages and disadvantages. To this end, it might be more important to synchronize screening instruments and subsequently apply in-depth instruments, to balance different demands and needs at different stages of the research process.

Limitations

Several limitations of our work deserve mentioning. First, the list of childhood adversity questionnaires used is not exhaustive, as additional questionnaires may include exposure types not covered here. This would, however, most likely increase rather than reduce heterogeneity. Second, while focusing exclusively on questionnaires aligns with previously published item-based content analyses (e.g., Karstoft and Armour, 2022 on General Trauma; Fried, 2017 on depression), content agreement between questionnaires and interview-based assessment of childhood adversity may be a valuable avenue for the future. Third, despite high inter-rater agreement in coding of items (K = .86 and K = .81), this approach cannot be fully standardized. In the future, AI tools may be a promising avenue for helping to disambiguate this task. For maximal transparency, data files and analysis code of our work are publicly available. Relatedly, assigning childhood adversity categories to items proved to be quite challenging for some items. For instance, the item "Tied up or locked in a closet." (ETI), cannot be unambiguously categorized as either emotional or physical abuse - both of which are commonly used childhood adversity types in the literature - as it clearly contains elements of both. This example showcases a deeper challenge: the general challenge of assigning complex experiences into distinct categories. As illustrated by the example above, these are often overlapping and have hence been referred to as 'fuzzy categories' (Smith & Pollak, 2021, 2022). An additional criticism with respect to such subcategories is that there is in fact little evidence in the literature that specific subcategories do indeed map onto distinct, specific (neurobiological) alterations (Smith & Pollak, 2022; Young et al., 2019). In fact, it has been shown that stress-response systems are not sensitive to specific exposures, but more to individual differences and specific characteristics of experiences (Korte et al., 2005; Smith & Pollak, 2021). Furthermore, based on the available evidence (McLaughlin et al., 2021; Pollak & Smith, 2021; Smith & Pollak, 2021) it seems more promising for future research to consider the subjective experience and evaluation as well as chronicity and developmental timing of childhood adversity rather than focusing exclusively on such 'fuzzy' categories (Danese & Widom, 2021; Smith & Pollak, 2022).

Summary

In sum, our systematic overview of item-based content (i.e., adversity types) of questionnaires assessing childhood adversity identified substantial heterogeneity in assessment tools and operationalization as a general challenge that may in part originate from a lack of agreed upon definition of childhood adversity. Our work may serve as a starting point for moving towards "an empirical basis for classifying adversity" (Pollak & Smith, 2021). The comprehensive overview on the item-based content as well as structural characteristics of different childhood adversity questionnaires will facilitate selection of tools tailored to the specific aims of a study. In addition, it allows researchers to easily identify differences and commonalities across assessment tools which is particularly valuable when aiming to bring together results across different studies. It is important for the facilitation of cumulative knowledge generation to discuss where to go from here and how to meet the challenges and opportunities in future work and in interpreting the existing literature. To this end, we provide a detailed list of actionables (see BOX) for guidance.

We anticipate that our work will aid the improvement of comparability, replicability and cumulative knowledge generation in this societal and clinically highly relevant research field.

Box: Methods-focused Considerations and Future Directions to advance research in measuring and reporting Childhood Adversity

these considerations include general considerations as well as considerations that can be distilled directly from the present work

Selection of an childhood adversity questionnaire for a research project:

- If available, questionnaires with at least satisfactory psychometric properties according to all COSMIN criteria (Mokkink et al., 2010) should be prioritizedeven though we acknowledge that classical psychometric criteria may not be applicable to questionnaires assessing childhood adversity (i.e., no underlying latent construct that is to be assessessed) and may not even be desirable in questionnaires aiming to assess a broad range of exposure types (e.g., Cronbach's alpha, Flake et al., 2017).

- Fit between content (i.e, adversity types) covered by a questionnaire and the research question and sample should be critically considered. For instance, the intended use of the questionnaire (such as screening or diagnosis) as well as the specific group being targeted (retrospective self-reports in adults vs. pediatric samples).

- (Additional) Assessment tools that enable a fine-grained evaluation of potentially relevant childhood adversity characteristics (i.e., onset and duration of exposure, controllability, (un)predictability) should be considered and reported (also descriptively when not of key interest to the studies aims).

- For the assessment of deprivation, it's crucial to utilize questionnaires that enable a detailed characterization of the construct rather than relying on a single item. However, it's noteworthy that most of the questionnaires examined in this analysis did not meet this criterion. Even though a study may focus on a specific type of childhood adversity, it is highly recommended to provide a nuanced screening and overview also on other childhood adversity types. This disambiguates the integration of results across studies and facilitates cumulative knowledge generation (Röseler et al., 2024).

Reporting Standards:

- Questionnaires used should be described in sufficient detail (i.e., content, target age group, subscales, types of maltreatment are assessed) and psychometric properties should be provided if applicable (however see above first bullet point).

- To facilitate cumulative knowledge generation, prevalence of adversity types (e.g., based on subscale scores and, if available, severities) in the respective sample should be characterized and reported even though this may not be the focus of a specific study (e.g., supplementary material).

Fostering reproducibility:

- Ad-hoc modifications of questionnaires should be avoided as these (may) compromise construct validity and reproducibility (Flake & Fried, 2019).

- If modifications are unavoidable, these must be reported with sufficient detail, ensuring that construct validity, and sound scientific reasoning are maintained (see Flake et al., 2017 for guidelines).

- Validated cut-offs should be adhered to and detailed cut-off information provided. Simply referring to prior publications may often lead to dead-ended reference chains. We highlight that cut-off scores may be more meaningful with respect to the experienced valence, controllability, and predictability of experiences, rather than for the number of adverse experiences.

- To facilitate cumulative knowledge generation, materials used in the assessment, such as questionnaires and interviews, should be made openly available, although copyright restrictions need to be considered.

- Generally, we call for more sharing of materials as only 26 out of 35 questionnaires were openly available. Publicly available material will facilitate rapid testing of new and existing hypotheses or psychometric validations (e.g., test-retest reliability), thus a mutual benefit of the research community, however, copyright should be considered.

Integration into the body of literature:

- To compare the results between different studies focusing on childhood adversity, the challenges of different questionnaires used and their (item-wise) content overlap as well as structural characteristics (e.g., developmental timing, experience vs. exposure, prospective vs. retrospective assessment Danese and Widom, 2020) needs to be considered even though no data may be available to allow for a direct comparison.

- Prevalences of trauma types in the respective sample, specifically the trauma load of the control group (if available in methods) should be considered.

- Exposure and experience should be distinguished in assessment and interpretation (McLaughlin et al., 2021) e.g., by using questionnaires which assess the experienced valence, controllability, and predictability of the event. While exposure directly relates to the fact of being present at a specific adverse event, experience involves subjectively perceiving it as aversive.

Author Contact

Correspondence concerning this article should be addressed to Alina Koppold, Postal address. E-mail: al.koppold@gmail.com

Alina Koppold - ORCID https://orcid.org/0000-0002-3164-3389 Julia Ruge - ORCID https://orcid.org/0000-0001-5818-5683 Tobias Hecker - ORCID https://orcid.org/0000-0001-9272-0512 Tina B. Lonsdorf - ORCID https://orcid.org/0000-0003-1501-4846

Acknowledgements

The authors thank Eiko Fried for making his R code available (Fried, 2017).

Conflict of Interest and Funding

The authors declare no competing financial interests.

Financial Support This work was funded by grants by the German Research Foundation (Deutsche Forschungsgemeinschaft) to TBL: DFG LO1980/4-1 and GRK 2753/1 – Project Number 449640848, subproject A01).

Author Contributions

The authors made the following contributions. Alina Koppold: Conceptualization, Methodology, Investigation, Validation, Formal analysis, Visualization, Data Curation, Writing - Original Draft; Julia Ruge: Investigation, Data Curation; Tobias Hecker: Writing – review & editing; Tina B. Lonsdorf: Conceptualization, Writing - Original Draft, Supervision, Funding acquisition. Author names order connects to the contribution (AK first author, JR & TH co-authors, TBL senior author).

Open Science Practices

This article earned the Open Data, Open Materials, and Open Code badge for making the data, materials, and code openly available. It has been verified that the analysis reproduced the results presented in the article. The entire editorial process, including the open reviews, is published in the online supplement.

References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology*, *19*(6), 586–592. https://doi.org/10.1037/0278-6133.19.6.586

Anda, R. F., Felitti, V. J., Bremner, J. D., Walker, J. D., Whitfield, C., Perry, B. D., Dube, S. R., & Giles, W. H. (2006). The enduring effects of abuse and related adverse experiences in childhood: A convergence of evidence from neurobiology and epidemiology. *European Archives of Psychiatry and Clinical Neuroscience*, *256*(3), 174–186. https://doi.org/10.1007/s00406-005-0624-4

Aust, F., & Barth, M. (2022). Papaja: Prepare reproducible apa journal articles with r markdown [https://github.com/crsh/papaja].

Baldwin, J. R., Reuben, A., Newbury, J. B., & Danese, A. (2019). Agreement between prospective and retrospective measures of childhood maltreatment: A systematic review and meta-analysis. *JAMA Psychiatry*, *76*(6), 584. https://doi.org/10.1001/jamapsychiatry.2019.0097

Barth, M. (2022). Tinylabels: Lightweight variable labels [https://cran.r-project.org/package=tinylabels].

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). Effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815

Berkelaar, M., et al. (2022). Lpsolve: Interface to 'lp$_s$*olve*' *v*.5.5*to*s*olvelinear*/*integerprograms* [https://CRAN.R-project.org/package=lpSolve].

Bernstein, D. P., Ahluvalia, T., Pogge, D., & Handelsman, L. (1997). Validity of the childhood trauma questionnaire in an adolescent psychiatric population. *Journal of the American Academy of Child Adolescent Psychiatry*, *36*(3), 340–348. https : / / doi . org / 10 . 1097 / 00004583 - 199703000-00012

Bougeard, S., & Dray, S. (2018). Supervised multiblock analysis in r with the ade4 package. *Journal of Statistical Software*, *86*(1), 1–17. https://doi.org/10.18637/jss.v086.i01

Breslau, N., & Kessler, R. C. (2001). The stressor criterion in dsm-iv posttraumatic stress disorder: An empirical investigation. *Biological Psychia-*

try, *50*(9), 699–704. https://doi.org/10.1016/S0006-3223(01)01167-2

Brunson, J. C. (2020). Ggalluvial: Layered grammar for alluvial plots. *Journal of Open Source Software*, *5*(49), 2017. https://doi.org/10.21105/joss.02017

Carlson, E. B., Smith, S. R., Palmieri, P. A., Dalenberg, C., Ruzek, J. I., Kimerling, R., Burling, T. A., & Spain, D. A. (2011). Development and validation of a brief self-report measure of trauma exposure: The trauma history screen. *Psychological Assessment*, *23*(2), 463–477. https://doi.org/10.1037/a0022294

Casement, M. D., Guyer, A. E., Hipwell, A. E., McAloon, R. L., Hoffmann, A. M., Keenan, K. E., & Forbes, E. E. (2014). Girls' challenging social experiences in early adolescence predict neural response to rewards and depressive symptoms. *Developmental Cognitive Neuroscience*, *8*, 18–27. https://doi.org/10.1016/j.dcn.2013.12.003

Castelda, B. A., Levis, D. J., Rourke, P. A., & Coleman, S. L. (2007). Extension of the sexual abuse questionnaire to other abuse categories: The initial psychometric validation of the binghamton childhood abuse screen. *Journal of Child Sexual Abuse*, *16*(1), 107–125. https://doi.org/10.1300/J070v16n01_06

Chessel, D., Dufour, A.-B., & Thioulouse, J. (2004). The ade4 package – i: One-table methods [https://cran.r-project.org/doc/Rnews/]. *R News*, *4*(1), 5–10.

Childhood trauma questionnaire: A retrospective self-report: Manual [[Computer software]. NCS Pearson]. (1998).

Cicchetti, D., & Lynch, M. (1995). Failures in the expectable environment and their impact on individual development: The case of child maltreatment. In *Developmental psychopathology, vol. 2: Risk, disorder, and adaptation* (pp. 32–71). John Wiley Sons.

Danese, A., & Widom, C. S. (2020). Objective and subjective experiences of child maltreatment and their relationships with psychopathology. *Nature Human Behaviour*, *4*(8), 811–818. https://doi.org/10.1038/s41562-020-0880-3

Danese, A., & Widom, C. S. (2021). The subjective experience of childhood maltreatment in psychopathology. *JAMA Psychiatry*, *78*(12), 1307. https://doi.org/10.1001/jamapsychiatry.2021.2874

Desmond, C., Watt, K., Saha, A., Huang, J., & Lu, C. (2020). Prevalence and number of children living in institutional care: Global, regional, and country estimates. *The Lancet Child Adolescent Health*, *4*(5), 370–377. https://doi.org/10.1016/S2352-4642(20)30022-5

Dowle, M., & Srinivasan, A. (2022). Data.table: Extension of 'data.frame' [https://CRAN.R-project.org/package=data.table].

Dray, S., & Dufour, A. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20. https://doi.org/10.18637/jss.v022.i04

Dray, S., Dufour, A.-B., & Chessel, D. (2007). The ade4 package – ii: Two-table and k-table methods [https://cran.r-project.org/doc/Rnews/]. *R News*, *7*(2), 47–52.

Elson, M., Hussey, I., Alsalti, T., & Arslan, R. C. (2023). Psychological measures aren't toothbrushes. *Communications Psychology*, *1*(1), 25. https://doi.org/10.1038/s44271-023-00026-9

English, D. J., Upadhyaya, M. P., Litrownik, A. J., Marshall, J. M., Runyan, D. K., Graham, J. C., & Dubowitz, H. (2005). Maltreatment's wake: The relationship of maltreatment dimensions to child outcomes. *Child Abuse Neglect*, *29*(5), 597–619. https://doi.org/10.1016/j.chiabu.2004.12.008

Estrada, S., Richards, C., Gee, D. G., & Baskin-Sommers, A. (2020). Exposure to violence and nonassociative learning capability confer risk for violent behavior. *Journal of Abnormal Psychology*, *129*(7), 748–759. https://doi.org/10.1037/abn0000579

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Brooks/Cole Pub. Co.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. *American Journal of Preventive Medicine*, *14*(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Flake, J. K., & Fried, E. I. (2019). Measurement schmeasurement: Questionable measurement practices and how to avoid them [preprint] [PsyArXiv. https://doi.org/10.31234/osf.io/hs7wm].

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Follingstad, D. R., Coyne, S., & Gambone, L. (2005). A representative measure of psychological ag-

gression and its severity. *Violence and Victims*, *20*(1), 25–38. https://doi.org/10.1891/vivi.2005.20.1.25

Fox, N., & Leavitt, L. (1995). Violence exposure scale revised (vex-r).

Frankenburg, W. K., & Coons, C. E. (1986). Home screening questionnaire: Its validity in assessing home environment. *The Journal of Pediatrics*, *108*(4), 624–626. https://doi.org/10.1016/S0022-3476(86)80853-8

Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, *208*, 191–197. https://doi.org/10.1016/j.jad.2016.10.019

Frueh, B. C., Elhai, J. D., & Kaloupek, D. G. (2004). Unresolved issues in the assessment of trauma exposure and posttraumatic reactions. In G. M. Rosen (Ed.), *Posttraumatic stress disorder* (pp. 63–84). John Wiley Sons Ltd. https://doi.org/10.1002/9780470713570.ch4

Gamer, M., Lemon, J., & S., I. F. P. (2019). Irr: Various coefficients of interrater reliability and agreement [https://CRAN.R-project.org/package=irr].

Garnier, S., Ross, N., Rudis, R., Camargo, P. A., Sciaini, M., & Scherer, C. (2021). Viridis - colorblind-friendly color maps for r. https://doi.org/10.5281/zenodo.4679424

Garnier, S., Ross, N., Rudis, R., Camargo, P. A., Sciaini, M., & Scherer, C. (2022). Viridis - colorblind-friendly color maps for r. https://doi.org/10.5281/zenodo.4679424

Georgieva, S., Tomas, J. M., & Navarro-Pérez, J. J. (2021). Systematic review and critical appraisal of childhood trauma questionnaire short form (ctq-sf). *Child Abuse Neglect*, *120*, 105223. https://doi.org/10.1016/j.chiabu.2021.105223

Georgieva, S., Tomás, J. M., Navarro-Pérez, J. J., & Samper-García, P. (2022). Systematic review and critical appraisal of five of the most recurrently validated child maltreatment assessment instruments from 2010 to 2020. *Trauma, Violence, Abuse*, 152483802210976. https://doi.org/10.1177/15248380221097694

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal mri study. *Nature Neuroscience*, *2*(10), 861–863. https://doi.org/10.1038/13158

Gilbert, R., Widom, C. S., Browne, K., Fergusson, D., Webb, E., & Janson, S. (2009). Burden and consequences of child maltreatment in high-income countries. *The Lancet*, *373*(9657), 68–81. https://doi.org/10.1016/S0140-6736(08)61706-7

Gohel, D. (2023). *Officer: Manipulation of microsoft word and powerpoint documents*.

Gohel, D., & Skintzos, P. (2023). *Flextable: Functions for tabular reporting*.

Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric properties of the life events checklist. *Assessment*, *11*(4), 330–341. https://doi.org/10.1177/1073191104269954

Heim, C., & Nemeroff, C. B. (2001). The role of childhood trauma in the neurobiology of mood and anxiety disorders: Preclinical and clinical studies. *Biological Psychiatry*, *49*(12), 1023–1039. https://doi.org/10.1016/S0006-3223(01)01157-X

Heim, C. M., Entringer, S., & Buss, C. (2019). Translating basic research knowledge on the biological embedding of early-life stress into novel approaches for the developmental programming of lifelong health. *Psychoneuroendocrinology*, *105*, 123–137. https://doi.org/10.1016/j.psyneuen.2018.12.011

Houston, S. M., Herting, M. M., & Sowell, E. R. (2013). The neurobiology of childhood structural brain development: Conception through adulthood. *The Neurobiology of Childhood*, *16*, 3–17. https://doi.org/10.1007/978-3-662-45758-0_265

Hughes, K., Ford, K., Bellis, M. A., Glendinning, F., Harrison, E., & Passmore, J. (2021). Health and financial costs of adverse childhood experiences in 28 european countries: A systematic review and meta-analysis. *The Lancet Public Health*, *6*(11), e848–e857. https://doi.org/10.1016/S2468-2667(21)00232-2

Iannone, R., Cheng, J., Schloerke, B., Hughes, E., & Seo, J. (2022). Gt: Easily create presentation-ready display tables.

Jonson-Reid, M., Kohl, P. L., & Drake, B. (2012). Child and adult outcomes of chronic child maltreatment. *Pediatrics*, *129*(5), 839–845. https://doi.org/10.1542/peds.2011-2529

Kantor, G. K., Holt, M. K., Mebert, C. J., Straus, M. A., Drach, K. M., Ricci, L. R., MacAllum, C. A., & Brown, W. (2004). Development and preliminary psychometric properties of the multidimensional neglectful behavior scale-child report. *Child Maltreatment*, *9*(4), 409–428. https://doi.org/10.1177/1077559504269530

Karstoft, K.-I., & Armour, C. (2022). What we talk about when we talk about trauma: Content overlap and heterogeneity in the assessment of trauma exposure. *Journal of Traumatic Stress*, jts.22880. https://doi.org/10.1002/jts.22880

Koppold, A., Kastrinogiannis, A., Kuhn, M., & Lonsdorf, T. B. (2023). Watching with argus eyes: Characterization of emotional and physiological responding in adults exposed to childhood maltreatment and/or recent adversity. *Psychophysiology*. https://doi.org/10.1111/psyp.14253

Korte, S. M., Koolhaas, J. M., Wingfield, J. C., & McEwen, B. S. (2005). The darwinian concept of stress: Benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neuroscience Biobehavioral Reviews*, *29*(1), 3–38. https://doi.org/10.1016/j.neubiorev.2004.08.009

Lacey, R. E., & Minnis, H. (2020). Practitioner review: Twenty years of research with adverse childhood experience scores–advantages, disadvantages and applications to practice. *Journal of Child Psychology and Psychiatry*, *61*(2), 116–130. https://doi.org/10.1111/jcpp.13135

Lambert, H. K., King, K. M., Monahan, K. C., & McLaughlin, K. A. (2017). Differential associations of threat and deprivation with emotion regulation and cognitive control in adolescence. *Development and Psychopathology*, *29*(3), 929–940. https://doi.org/10.1017/S0954579416000584

Lynch, M., & Cicchetti, D. (1998). An ecological-transactional analysis of children and contexts: The longitudinal interplay among child maltreatment, community violence, and children's symptomatology. *Development and Psychopathology*, *10*(2), 235–257. https://doi.org/10.1017/S095457949800159X

MacDonald, K., Thomas, M. L., Sciolla, A. F., Schneider, B., Pappas, K., Bleijenberg, G., Bohus, M., Bekh, B., Carpenter, L., Carr, A., et al. (2016). Minimization of childhood maltreatment is common and consequential: Results from a large multinational sample using the childhood trauma questionnaire. *PLOS ONE*, *11*(1), e0146058. https://doi.org/10.1371/journal.pone.0146058

Machlin, L., Miller, A. B., Snyder, J., McLaughlin, K. A., & Sheridan, M. A. (2019). Differential associations of deprivation and threat with cognitive control and fear conditioning in early childhood. *Frontiers in Behavioral Neuroscience*, *13*, 80. https://doi.org/10.3389/fnbeh.2019.00080

Madigan, S., Deneault, A.-A., Racine, N., Park, J., Thiemann, R., Zhu, J., Dimitropoulos, G., Williamson, T., Fearon, P., Cénat, J. M., McDonald, S., Devereux, C., & Neville, R. D. (2023). Adverse childhood experiences: A meta-analysis of prevalence and moderators among half a million adults in 206 studies. *World Psychiatry*, *22*(3), 463–471. https://doi.org/10.1002/wps.21122

McLaughlin, K. A., Peverill, M., Gold, A. L., Alves, S., & Sheridan, M. A. (2015). Child maltreatment and neural systems underlying emotion regulation. *Journal of the American Academy of Child Adolescent Psychiatry*, *54*(9), 753–762. https://doi.org/10.1016/j.jaac.2015.06.010

McLaughlin, K. A., & Sheridan, M. A. (2016). Beyond cumulative risk: A dimensional approach to childhood adversity. *Current Directions in Psychological Science*, *25*(4), 239–245. https://doi.org/10.1177/0963721416655883

McLaughlin, K. A., Sheridan, M. A., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2021). The value of dimensional models of early experience: Thinking clearly about concepts and categories. *Perspectives on Psychological Science*, *16*(6), 1463–1472. https://doi.org/10.1177/1745691621992346

Menold, N., & Bogner, K. (2014). Gestaltung von ratingskalen in fragebögen. sdm survey guidelines. *SDM Survey Guidelines*. https://doi.org/10.15465/SDM-SG_015

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

Milojevich, H. M., Machlin, L., & Sheridan, M. A. (2020). Early adversity and children's emotion regulation: Differential roles of parent emotion regulation and adversity exposure. *Development and Psychopathology*, *32*(5), 1788–1798. https://doi.org/10.1017/S0954579420001273

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., Bouter, L. M., & De Vet, H. C. (2010). The cosmin checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*, *10*(1), 22. https://doi.org/10.1186/1471-2288-10-22

Morrison, K. E., Stenson, A. F., Marx-Rattner, R., Carter, S., Michopoulos, V., Gillespie, C. F., Powers, A.,

Huang, W., Kane, M. A., Jovanovic, T., & Bale, T. L. (2022). Developmental timing of trauma in women predicts unique extracellular vesicle proteome signatures. *Biological Psychiatry*, *91*(3), 273–282. https://doi.org/10.1016/j.biopsych.2021.08.003

Müller, K., & Wickham, H. (2023). Tibble: Simple data frames. *CRAN.R-project.org*. https://CRAN.R-project.org/package=tibble

Murrell, P., & Wen, Z. (2020). Gridgraphics: Redraw base graphics using 'grid' graphics. *CRAN.R-project.org*. https://CRAN.R-project.org/package=gridGraphics

Neuwirth, E. (2022). Rcolorbrewer: Colorbrewer palettes. *CRAN.R-project.org*. https://CRAN.R-project.org/package=RColorBrewer

Nowosad, J. (2018). 'cartocolors' palettes. *nowosad.github.io*. https://nowosad.github.io/rcartocolor

Pai, A. L. H., Patiño-Fernández, A. M., McSherry, M., Beele, D., Alderfer, M. A., Reilly, A. T., Hwang, W.-T., & Kazak, A. E. (2008). The psychosocial assessment tool (pat2.0): Psychometric properties of a screener for psychosocial distress in families of children newly diagnosed with cancer. *Journal of Pediatric Psychology*, *33*(1), 50–62. https://doi.org/10.1093/jpepsy/jsm053

Pedersen, T. L. (2022). Patchwork: The composer of plots. *CRAN.R-project.org*. https://CRAN.R-project.org/package=patchwork

Pollak, S. D., & Smith, K. E. (2021). Thinking clearly about biology and childhood adversity: Next steps for continued progress. *Perspectives on Psychological Science*, *16*(6), 1473–1477. https://doi.org/10.1177/17456916211031539

Radoman, M., Akinbo, F. D., Rospenda, K. M., & Gorka, S. M. (2019). The impact of startle reactivity to unpredictable threat on the relation between bullying victimization and internalizing psychopathology. *Journal of Psychiatric Research*, *119*, 7–13. https://doi.org/10.1016/j.jpsychires.2019.09.004

Röseler, L., Kaiser, L., Doetsch, C. A., Klett, N., Seida, C., Schütz, A., Aczel, B., Adelina, N., Agostini, V., Alarie, S., Albayarak-Aydemir, N., Aldoh, A., Al-Hoorie, A. H., Azevedo, F., Baker, B. J., Barth, C. L., Beitner, J., Brick, C., Brohmer, H., … Zhang, Y. (2024). The Replication Database: Documenting the Replicability of Psychological Science. https://doi.org/10.31222/osf.io/me2ub

Rowland, G. E., Mekawi, Y., Michopoulos, V., Powers, A., Fani, N., Bradley, B., Ressler, K. J., Jovanovic, T., & Stevens, J. S. (2022). Distinctive impacts of sexual trauma versus non-sexual trauma on ptsd profiles in highly trauma-exposed black women. *Journal of Affective Disorders*, *317*, 329–338. https://doi.org/10.1016/j.jad.2022.08.099

Ruge, J., Ehlers, M. R., Kastrinogiannis, A., Klingelhöfer-Jens, M., Koppold, A., & Lonsdorf, T. B. (2023). How adverse childhood experiences get under the skin: A systematic review integration and methodological discussion on threat and reward learning mechanisms. *PsyArXiv Preprint*. https://doi.org/10.31234/osf.io/nfpbj

Saini, S. M., Hoffmann, C. R., Pantelis, C., Everall, I. P., & Bousman, C. A. (2019). Systematic review and critical appraisal of child abuse measurement instruments. *Psychiatry Research*, *272*, 106–113. https://doi.org/10.1016/j.psychres.2018.12.068

Scharfenort, R., Menz, M., & Lonsdorf, T. B. (2016). Adversity-induced relapse of fear: Neural mechanisms and implications for relapse prevention from a study on experimentally induced return-of-fear following fear conditioning and extinction. *Translational Psychiatry*, *6*(7), e858–e858. https://doi.org/10.1038/tp.2016.126

Selner-O'Hagan, M. B., Kindlon, D. J., Buka, S. L., Raudenbush, S. W., & Earls, F. J. (1998). Assessing exposure to violence in urban youth. *Journal of Child Psychology and Psychiatry*, *39*(2), 215–224. https://doi.org/10.1111/1469-7610.00315

Sheridan, M. A., & McLaughlin, K. A. (2014). Dimensions of early experience and neural development: Deprivation and threat. *Trends in Cognitive Sciences*, *18*(11), 580–585. https://doi.org/10.1016/j.tics.2014.09.001

Sjoberg, D. (2023). Ggsankey: Sankey alluvial and sankey bump plots. *CRAN.R-project.org*.

Slep, A. M. S., Heyman, R. E., & Foran, H. M. (2015). Child maltreatment in dsm-5 and icd-11. *Family Process*, *54*(1), 17–32. https://doi.org/10.1111/famp.12131

Smith, K. E., & Pollak, S. D. (2021). Rethinking concepts and categories for understanding the neurodevelopmental effects of childhood adversity. *Perspectives on Psychological Science*, *16*(1), 67–93. https://doi.org/10.1177/1745691620920725

Smith, K. E., & Pollak, S. D. (2022). Early life stress and perceived social isolation influence how children use value information to guide behavior. *Child Development*, *93*(3), 804–814. https://doi.org/10.1111/cdev.13727

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, *2*(10), 859–861. https://doi.org/10.1038/13154

Sowell, E. R., Thompson, P. M., Tessner, K. D., & Toga, A. W. (2001). Mapping continued brain growth and gray matter density reduction in dorsal frontal cortex: Inverse relationships during postadolescent brain maturation. *The Journal of Neuroscience*, *21*(22), 8819–8829. https://doi.org/10.1523/JNEUROSCI.21-22-08819.2001

Straus, M. A., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child maltreatment with the parent-child conflict tactics scales: Development and psychometric data for a national sample of american parents. *Child Abuse Neglect*, *22*(4), 249–270. https://doi.org/10.1016/S0145-2134(97)00174-9

Team, R. C. (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https://www.R-project.org/

Teicher, M. H., Gordon, J. B., & Nemeroff, C. B. (2021). Recognizing the importance of childhood maltreatment as a critical factor in psychiatric diagnoses, treatment, research, prevention and education. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-021-01367-9

Teicher, M. H., & Parigger, A. (2015). The "maltreatment and abuse chronology of exposure" (mace) scale for the retrospective assessment of abuse and neglect during development. *PLOS ONE*, *10*(2), e0117423. https://doi.org/10.1371/journal.pone.0117423

Thombs, B. D., Lewis, C., Bernstein, D. P., Medrano, M. A., & Hatch, J. P. (2007). An evaluation of the measurement equivalence of the childhood trauma questionnaire-short form across gender and race in a sample of drug-abusing adults. *Journal of Psychosomatic Research*, *63*(4), 391–398. https://doi.org/10.1016/j.jpsychores.2007.04.010

United Nations Population Fund. (2024). Child marriage [Accessed: 2024-05-29]. https://www.unfpa.org/child-marriage#readmore-expand

Walsh, C. A., MacMillan, H. L., Trocmé, N., Jamieson, E., & Boyle, M. H. (2008). Measurement of victimization in adolescence: Development and validation of the childhood experiences of violence questionnaire. *Child Abuse Neglect*, *32*(11), 1037–1057. https://doi.org/10.1016/j.chiabu.2008.05.003

Watson, P. F., & Petrie, A. (2010). Method agreement analysis: A review of correct methodology. *Theriogenology*, *73*(9), 1167–1179. https://doi.org/10.1016/j.theriogenology.2010.01.003

Weems, C. F., Russell, J. D., Herringa, R. J., & Carrion, V. G. (2021). Translating the neuroscience of adverse childhood experiences to inform policy and foster population-level resilience. *American Psychologist*, *76*(2), 188–202. https://doi.org/10.1037/amp0000780

Wei, T., & Simko, V. (2021). R package 'corrplot': Visualization of a correlation matrix. *github.com*. https://github.com/taiyun/corrplot

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2022a). Forcats: Tools for working with categorical variables (factors). *CRAN.R-project.org*. https://CRAN.R-project.org/package=forcats

Wickham, H. (2022b). Stringr: Simple consistent wrappers for common string operations. *CRAN.R-project.org*. https://CRAN.R-project.org/package=stringr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2022). Readxl: Read excel files. *CRAN.R-project.org*. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). Dplyr: A grammar of data manipulation. *CRAN.R-project.org*. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Girlich, M. (2022). Tidyr: Tidy messy data. *CRAN.R-project.org*. https://CRAN.R-project.org/package=tidyr

Wickham, H., & Henry, L. (2023). Purrr: Functional programming tools. *CRAN.R-project.org*. https://CRAN.R-project.org/package=purrr

Wickham, H., Hester, J., & Bryan, J. (2022). Readr: Read rectangular text data. *CRAN.R-project.org*. https://CRAN.R-project.org/package=readr

Wilker, S., Pfeiffer, A., Kolassa, S., Koslowski, D., Elbert, T., & Kolassa, I.-T. (2015). How to quantify exposure to traumatic stress? reliability and

predictive validity of measures for cumulative trauma exposure in a post-conflict population. *European Journal of Psychotraumatology*, *6*(1), 28306. https://doi.org/10.3402/ejpt.v6.28306

Wolfe, J., Kimerling, R., Brown, P. J., Chrestman, K. R., & Levin, K. (2012). Life stressor checklist–revised. *American Psychological Association*. https://doi.org/10.1037/t04534-000

World Health Organization. (2024). Child maltreatment [Accessed: 2024-05-29]. https://apps.who.int/violence-info/child-maltreatment/

Xie, Y. (2015). *Dynamic documents with r and knitr (2nd ed.)* Chapman; Hall/CRC. https://yihui.org/knitr/

Xie, Y. (2016). *Bookdown: Authoring books and technical documents with r markdown*. Chapman; Hall/CRC. https://bookdown.org/yihui/bookdown

Young, D. A., Neylan, T. C., Chao, L. L., O'Donovan, A., Metzler, T. J., & Inslicht, S. S. (2019). Child abuse interacts with hippocampal and corpus callosum volume on psychophysiological response to startling auditory stimuli in a sample of veterans. *Journal of Psychiatric Research*, *111*, 16–23. https://doi.org/10.1016/j.jpsychires.2019.01.011

Zhu, H. (2021). Kableextra: Construct complex table with 'kable' and pipe syntax. *CRAN.R-project.org*. https://CRAN.R-project.org/package=kableExtra