Check for updates

Original Research

Journal of Interpersonal Violence
1–26
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/08862605241233271
journals.sagepub.com/home/jiv
Sage

Caregiver Report of Children's Exposure to Adverse Life Events: Concordance Between Questionnaire and Interview Approaches

Austen McGuire[1], Yo Jackson[2], Damion J. Grasso[3], George M. Slavich[4] and Neal Kingston[5]

Abstract

Assessment practices for measuring adverse life events (ALEs) are often characterized by considerable variability, which is associated with inconsistency and reproducibility issues when conducting research on children with ALE exposure. One aspect of assessment variability for caregiver report of children's ALE history that has received minimal attention is assessment format. To address this issue, the current study evaluated concordance between two main ALE assessment formats: interviews and questionnaires. This involved examining overall endorsement of ALEs and concordance among multiple characteristics of ALE exposure, including type, polyvictimization, frequency, severity, and age of onset. Fifty-eight

---

[1] Medical University of South Carolina, Charleston, South Carolina, USA
[2] The Pennsylvania State University, University Park, Pennsylvania, USA
[3] University of Connecticut School of Medicine, Farmington, Connecticut, USA
[4] University of California, Los Angeles, California, USA
[5] University of Kansas, Lawrence, Kansas, USA

Corresponding Author:
Austen McGuire, Department of Psychiatry & Behavioral Sciences, National Crime Victims Research and Treatment Center, Medical University of South Carolina, 67 President Street, MSC861, Charleston, SC 29425, USA.
Email: mcguirau@musc.edu

caregivers (M_{age} = 33.72; 60% Black; 55% below the federal poverty line) of preschool and school-age children were administered an ALE assessment in both a questionnaire and interview format across two sessions. The sum scores and concordance rates between format responses were compared based on ALE type, polyvictimization, frequency, severity, and age of onset of exposure. Results indicated that most total or sum scores were similar between formats, with the exception of ALE severity scores. However, there was most often low-to-moderate concordance across the 50 types of ALEs examined in the current study, suggesting that a different constellation of events comprised each sum or total score. This was also the case across all characteristics of the ALEs and most notably for the severity of ALE. Based on these findings, the format of assessment may be associated with inconsistent reporting of children's ALE exposure across multiple characteristics of ALE. Researchers may need to utilize multiple types of ALE assessments when relying on caregiver report of a child's ALEs.

Keywords

childhood adversity, stress, children, assessment, methodology

Caregivers are key informants in the assessment of children's exposure to adverse life events (ALEs; e.g., maltreatment, violence exposure, serious illness, and natural disasters). This is especially true for the assessment of ALEs in young children (approximately <8 years) because these children lack crucial cognitive and language developmental skills needed to properly interpret and respond to questions about exposure (e.g., Oh et al., 2018; Tingskull et al., 2015). Numerous validated caregiver-report instruments exist (e.g., Bethell et al., 2017; Oh et al., 2018); however, they vary significantly regarding how and what questions are asked. This variability limits the comparability of findings across studies and impedes the application of this knowledge to clinical practice. Elucidating differences across these assessment methods will be critical for enhancing evidence-based assessment of ALEs in clinical and research settings. One fundamental difference in ALE assessment is the *format* or how responses to questions are solicited from informants. This is largely divided into two classes of assessment approaches: (a) questionnaires and (b) interviews. With questionnaires, respondents are often asked to report on an inventory of predetermined ALEs without direct guidance from an administrator using paper-pencil or computer-assisted programs. For ALE interview formats, respondents engage with an administrator who employs a structured, semistructured, or unstructured protocol to solicit responses (Harkness & Monroe, 2016).

Most research comparing interview and questionnaire approaches for ALE assessment comes from studies examining self-report among adolescents and adults (e.g., Harkness & Monroe, 2016). Overall, this evidence appears to suggest that considerable differences exist in information obtained from each method and low concordance regarding ALE prevalence or types of ALE exposure endorsed (e.g., Duggal et al., 2000; Harkness & Monroe, 2016). Relative to research on self-reported ALEs, research examining differences among caregiver report methods for a child is more limited. To date, only three studies have directly compared the use of caregiver-reported questionnaire and interview methods when assessing ALE exposure in preschool or school age children (Allen et al., 2012; Glackin et al., 2019; Wagner et al., 2006). Among these studies, two found strong, positive correlations between formats on the overall number of reported ALEs, suggesting reliability between formats (Allen et al., 2012; Wagner et al., 2006). Glackin et al. (2019) found agreement between questionnaire versus interview methods to range between 10.9% and 45.3% for certain ALE groupings, with Cohen's kappa estimates reaching a max of .36.

While the scant research presently available on caregiver-reported ALE exposure for preschool or school age children suggests some concordance, several important limitations complicate these findings and raise questions as to whether both formats produce similar reports. First, among the three studies examining caregiver report, no study was comprehensive regarding the array of ALE types inventoried. For example, Allen et al. (2012) exclusively inquired about short-term ALEs (less than 4 weeks) and excluded ALEs thought to be less common among socioeconomically healthy families (e.g., caregiver going to jail). Moreover, Wagner et al. (2006) and Glackin et al. (2019) did not assess for several types of interpersonal victimizations, such as injury after physical assault. The limited inventory of ALE types is problematic, as it may exclude certain ALEs known to be relatively common among children (e.g., family dysfunction related ALEs; Finkelhor et al., 2015) and, therefore, fail to identify a proportion of exposed children. This may be especially true among children from low-income and underrepresented populations for which these types of experiences tend to occur more often, compared to white or more socioeconomically privileged children (Slopen et al., 2016). Further, comparing methods using a small number of items may result in inflated concordance between methods, obscuring method variance.

Second, there was a lack of independence of caregiver-reported ALE endorsements between approaches. For example, Allen et al. (2012) and

Wagner et al. (2006) assessed ALEs in the interview contingent on whether they were first endorsed by caregivers on a prior questionnaire. As a result, it is unclear which events caregivers might have self-generated during the interview, as the interview was designed to confirm the information provided on the questionnaire rather than serve as an independent data source. Further, since the interview was used in a confirmatory manner, there was a lack of counterbalancing in the administration of measures such that questionnaires were always administered first. Counterbalancing administration is necessary when comparing methods to avoid priming effects (Peirce et al., 2009). Although Glackin et al. (2019) did not use a confirmation procedure, the authors included different sets and types of ALEs between formats, forcing them to rely on broadened categories of ALEs when examining concordance.

Third, no prior studies have examined differences in concordance rates among certain characteristics of ALEs (e.g., severity or age of exposure), and only one study examined ALE subtypes. Glackin et al. (2019) reported varied rates of agreement when examining any ALE exposure and among broad groupings of ALE types, including sexual violence, and direct or indirect violence. Allen et al. (2012) and Wagner et al. (2006) found good agreement on total or mean number of ALE types (i.e., polyvictimization); however; these methods may have elicited different constellations of ALE types that are obscured by simply tallying the number of types (e.g., Grasso et al., 2016; Monroe & Slavich, 2020; Slavich, 2019). It may be important to assess characteristics of ALE exposure beyond type, as previous research suggests their utility in predicting outcomes; however, measuring these characteristics brings their own set of difficulties. For example, measurement of ALE severity is challenging given its dependence on the subjective appraisal of the experience by the respondent (e.g., Litrownik et al., 2005; Trickey et al., 2012).

Evidence is limited regarding how questionnaires and interviews would compare when assessing these characteristics. For example, Harkness and Monroe (2016) argue that the interview method is the only approach capable of capturing these characteristics because of the potential to probe for follow-up details, such as its severity, frequency, and exposure timing. However, this argument appears to have emerged from a comparison of approaches that each assessed unique content rather than from a direct, empirical comparison of approaches assessing the same content. Therefore, conclusions about reporting differences between interview and questionnaire formats among the dimensions of ALE exposure are based on differences in assessment tool structure, not necessarily differences in modality. Research that examines the concordance and discordance between approaches that assess the same ALE characteristics is an essential endeavor.

Present Study

To address these gaps in understanding, this study sought to provide empirical evidence on differences in caregiver-reported ALEs of their child on a questionnaire format versus an interview format. Caregivers from diverse backgrounds were asked to complete independent but similar measures of ALE exposure in counterbalanced interview and questionnaire formats assessing a comprehensive set of ALE exposures and multiple dimensions of these ALEs, including type, frequency, severity, and age of onset. It was hypothesized that caregivers would demonstrate similar levels of exposure (i.e., mean ALE dimension values) and higher concordance rates (i.e., kappa values in the moderate to strong range) of ALE polyvictimization (i.e., number of different types of exposure), frequency, and age of onset compared to ALE severity between the questionnaire and interview formats.

Method

Participants

Participants were 58 caregivers ($M_{age}[SD]=33.72[6.91]$; 96.6% biological mothers) reporting on their preschool or school-age child ($M_{age}[SD]=6.13[1.60]$, range: 3–9 years; 43.1% female). Participants were from a large midwestern county and were recruited from local community agencies and organizations serving low-income families, such as local Head Starts and food pantries. To be eligible for participation, the caregiver needed to (a) have a child between ages 3 and 9 years and (b) be the legal guardian of the child. Caregivers were not permitted to participate if the caregiver (a) had a self-reported diagnosis of a developmental or autism spectrum disorder and/or (b) was not a native English speaker. The inclusion and exclusion criteria helped to ensure that caregivers were able to accurately and successfully complete study activities. Sixty caregivers completed data collection; however, one caregiver was excluded for not returning to data collection after the first data collection session and another caregiver was excluded for not properly completing the second form of the ALE assessment.

Measures

Demographics. Demographic information for the caregiver, child, and family was obtained using a demographic questionnaire. This questionnaire included questions on age, sex, socioeconomic status, living situation, race, and ethnicity for the caregiver and child.

Child's History of ALE Exposure. History of exposure to ALEs across the child's lifetime was obtained using the PAIR (Preschoolers' Adjustment and Intergenerational Risk; Griffith et al., 2020) Intergenerational Trauma Measure (PAIRIT). The PAIRIT is a questionnaire measure of ALE exposure designed to assess exposure to 50 different types of ALEs, including the type, severity, frequency, age of first exposure, and age of last exposure of each ALE. The full version of the measure was created by synthesizing multiple published tools for assessing ALE exposure in children, such as the Preschool Age Psychiatric Assessment (Egger & Angold, 2004) and the Juvenile Victimization Questionnaire (Finkelhor et al., 2011). Included are several categories of ALEs, such as those associated with exposure to violence (including both community and familial), direct and indirect victimization, natural disasters, maltreatment, household dysfunction, death or separation from close others, and physical difficulties for the child. The PAIRIT's design was influenced by previously validated measures, as well as several review articles on ALE assessment that describe various limitations of currently available ALE measurement tools (e.g., Milne & Collin-Vézina, 2015; Vanaelst et al., 2012).

There were two versions of the PAIRIT: the original questionnaire format and an interview format. For the ALE assessment questionnaire and interview formats, the same 50 types of ALE exposure were included in both methods, and each format used the same question structure and answer format. For each format, the measure first asked whether a child had experienced a certain type of event. For each endorsed event, caregivers answered a set of follow-up questions to assess frequency, age of onset or first exposure, and severity relevant to the ALE. Age of last exposure was also asked but not included in the study analysis due to low number of endorsements (e.g., need to experience an event and experience more than once). To measure frequency, caregivers provided the approximate number of times an ALE occurred on a five-point Likert scale from 1 ("1 time") to 5 ("More than 10 times"). For continuous ALEs without discrete start and stop times (e.g., "Has your child been exposed to domestic violence?"), a five-point Likert scale with qualitative time descriptions was used ranging from 1 ("Almost never") to 5 ("Always"). Caregivers reported on the age of onset of the event by reporting the first time the event had occurred using a six-point Likert scale ranging from 1 ("Within the last month") to 6 ("Six or more years ago"). Parents also reported on the severity or impact of the event(s) on a seven-point Likert scale ranging from 1 ("Very Bad"*)* to 7 ("Very Good"). For both assessment formats, a polyvictimization score was also calculated by summing together the number of different types of exposure that caregivers reported, which could range from 0 to 50.

Differences between the formats were related to administration of the questions. For the questionnaire format, all 50 types of exposure and follow-up questions were presented to the participant in the same set order via paper-and-pencil. Participants were permitted to ask questions to the research staff as needed but completed the questionnaire independently. In comparison to the questionnaire, the interview was administered by a member of the research team face-to-face. The interview also followed a semistructured discussion format at the beginning, commonly used in ALE exposure interviews (Harkness & Monroe, 2016). That is, at the start of the interview, caregivers were asked to provide a list of events the child had experienced. The interviewer then first began probing for additional information on the events reported by the caregiver from the initial list, with each event beginning with an open-ended prompt (e.g., "Please tell me as much information as you can about this event"). After going through the caregiver's initial list of events, the interviewer asked whether the child was exposed to any of the remaining types of events not initially discussed from the list of 50 events. If an event was endorsed, the interviewer proceeded with follow-up questions on frequency, age of onset, and severity. If a caregiver did not provide a response related to the follow-up questions, the interviewer would ask the follow-up questions directly to the caregiver and provide them with the answer response categories.

Procedures

All procedures for the current study were approved by the institutional review board at the primary author's institution. The study used repeated-measures within-subjects design, split between two data collection sessions. Interested and eligible participants were contacted by study personnel and scheduled for their first session at a community location convenient for the family (e.g., local community center, church). For counterbalancing, participants were randomly assigned to begin with either the interview or questionnaire format ALE assessment format. Additionally, all interviews were conducted with the same individual, who identified as a white, cisgender male and was a mental health professional with experience serving ALE-exposed populations.

Session 1. Following consent procedures, the participant began by completing the measure battery, which included the demographics form and other measures not included in the current study (e.g., measures of child behavioral and emotional functioning). These were administered on a laptop computer for the participant. After completing these measures, the research staff member administered the first type of ALE assessment, which varied in format depending on group assignment. At the end of the session, the participant met with a research assistant to conduct debriefing, which included answering participants' questions, handout of resources (e.g., a list of mental health services, self-care suggestions), and scheduling the next data collection session.

Session 2. Session 2 occurred at least 7 days and no more than 21 days after the first session. This is a time frame used in testing of ALE assessments that can reduce the chance of reporting fatigue or memory recall bias from one measure to the other, while also minimizing the potential for new ALEs to occur in between testing sessions (Thabrew et al., 2012). At the second session and following a second review of consent and study procedures, the participant first completed the second form of the ALE assessment. Lastly, the participant took part in a second debriefing session with the research staff and received financial compensation for participation.

Data Analysis

The first part of data analysis examined descriptive information for the main variables. Next, paired sample *t*-tests were conducted to test for statistically significant differences among the type/polyvictimization, frequency, severity, and age of onset scores between the interview and questionnaire methods and administration sequence (i.e., those administered the interview first vs. those administrated the questionnaire first). The *t*-statistics were examined for significance at a *p*-value level of $<.05$. A post-hoc power analysis for the paired sample *t*-tests determined that a sample size of 58 participants achieves 80% power to detect an effect size using Cohen's $d > .37$. Secondly, response concordances between the interview and questionnaire methods among the ALE characteristics (type, frequency, duration, and severity) were examined by calculating the overall Cohen's kappa coefficient. An unweighted Cohen's kappa coefficient was calculated for the each of the 50 ALE types given these items were binary, and a weighted Cohen's kappa coefficient (with linear weighting) was calculated for the frequency, age of onset, and severity items as these items were ordinal (Cohen, 1968; Tang et al., 2015). Kappa classifications cut-offs were: poor: $<.00$, slight: 00 to .20, fair: .21 to .40, moderate: .41 to .60, substantial: .61 to .80, and almost perfect: .81 to 1.00 (Landis & Koch, 1977). For frequency, participants received a score of 0 if they did not endorse an ALE type overall. Because of the concern with accuracy with small number of endorsements and need to endorse an event to answer follow-up questions regarding severity and age of onset for each ALE type, kappa values were not calculated for ALE characteristics when five or fewer participants endorsed an ALE type on both the questionnaire and interview formats.

Results

Descriptive information on caregivers and their child is shown in Table 1. Most caregivers identified as Black or African American (60.3%), followed by white (31.0%) and multiracial (5.2%). This was similar for caregiver report of their child's race, who were primarily Black or African American (58.6%), followed by white (25.9%) and then multiracial (12.1%). Based on yearly family income and number of individuals per household, most families (55.2%) were below the federal poverty line, while 20.7% of families were between 100% and 200% above the federal poverty line and 24.1% were at or above 200% of the federal poverty line, according to the 2021 federal poverty line standards (U.S. Department of Health and Human Services, 2021).

On average, participants completed each ALE assessment format 10.53 days apart (*SD* = 4.56; range: 7–21 days). The top five most endorsed events on the questionnaire included: A new child moving into the home, having a caregiver admitted to the hospital, primary caregivers separating, changing primary childcare providers, and moving to a new home. The top five most endorsed events on the interview were similar to the questionnaire, with the exception of the endorsement of caregiving having a mental health illness, as opposed to primary caregivers separating. Among the 50 different ALEs, there were eight types not endorsed on the questionnaire and seven types not endorsed on the interview by any participant, with five of these events not being endorsed on either the questionnaire or interview (see Table 2). Based on the level of endorsement, there were 27 types of ALEs that could not be evaluated for concordance of their characteristics due to fewer than five endorsements on either format of the ALE assessment.

Agreement for Type and Polyvictimization

There was no significant difference in mean polyvictimization (i.e., number of types of events) score between the questionnaire and interview methods, $t(57) = 1.63$, $p = .11$. Mean polyvictimization scores were also similar on both the questionnaire and interview versions between those who started with the questionnaire and those who started with the interview (all $t\text{s} < .57$, $p\text{s} > .57$). The number of endorsements for overall type, along with the Cohen's kappa coefficients, is provided in Table 2. There were 12 events for which kappa

Table 1. Participant Demographics.

Caregiver Demographics	Mean (*SD*) or %	Median	Range
Caregiver age	33.72 (6.91)	32.00	22.00–61.00
Caregiver relationship			
Biological mother	96.6%		
Biological father	1.7%		
Grandmother	1.7%		
Caregiver race			
NA/AN	0.0%		
Asian	3.4%		
Black/African American	60.3%		
NH/OPI	1.7%		
White	31.0%		
Mixed race/multiracial	3.4%		
Caregiver % Hispanic/Latino	5.2%		
Family SES			
50% below FPL	36.2%		
50% to 100% of FPL	19.0%		
100% to 150% of FPL	12.1%		
150% to 200% of FPL	8.6%		
Above 200% of FPL	24.1%		
Child Demographics	**Mean (*SD*) or %**	**Median**	**Range**
Child age (years)	6.13 (1.60)	6.44	3–9
Child sex (% female)	43.1%		
Child race			
NA/AN	0.0%		
Asian	5.4%		
Black/African American	58.6%		
NH/OPI	0.0%		
White	25.9%		
Mixed race/multiracial	12.1%		
Child % Hispanic/Latino	8.6%		
Child ALE characteristics			
Questionnaire polyvictimization	8.47 (5.69)	8.00	.00–26.00
Questionnaire mean frequency	.29 (.22)	.25	.00–.98
Questionnaire mean severity	4.05 (.84)	4.00	2.00–7.00
Questionnaire mean age of onset	3.87 (.85)	4.00	1.00–5.33
Interview polyvictimization	9.48 (4.47)	9.00	2.00–24.00
Interview mean frequency	.30 (.19)	.23	.04–.92
Interview mean severity	3.76 (.72)	4.00	2.00–5.00
Interview mean age of onset	4.02 (.72)	3.80	1.67–5.57

Note. ALE = Adverse life event; FPL = Federal Poverty Line; NA/AN = Native American/Alaska Native; NH/OPI = Native Hawaiian or Other Pacific Islander; SD = Standard deviation; SES = socioeconomic status.

Table 2. Concordance Between Interview and Questionnaire Assessments of Adverse Life Event Type.

Adverse Life Event Type	Total Endorsed *N*	Interview and Questionnaire Endorsed *n*	Only Interview Endorsed *n*	Only Questionnaire Endorsed *n*	Kappa [95% CI]
Changed childcare provider	51	29	21	1	.22 [0.04, 0.40]*
Moved to a new home	47	39	2	6	.64 [0.42, 0.87]*
Caregiver admitted to a hospital	40	22	12	6	.38 [0.15, 0.61]*
Caregivers separated	38	18	10	10	.31 [0.07, 0.56]*
Heard about something terrible	32	7	11	14	.04 [−0.22, 0.29]
Caregiver experiences mental illness	32	18	11	3	.52 [0.31, 0.73]*
New child moved into the home	29	23	1	5	.79 [0.64, 0.95]*
Close adult died	28	20	5	3	.72 [0.53, 0.90]*
Reduced standard of living	25	3	7	15	.01 [−0.23, 0.25]
Caregiver or close individual arrested	24	11	6	7	.47 [0.22, 0.71]*
Exposed to domestic/interpersonal violence	24	13	7	4	.57 [0.34, 0.79]*
Lost contact with close individual	23	7	7	9	.28 [0.01, 0.55]*
New caregiver moved into home	20	10	7	3	.55 [0.31, 0.79]*
Family/Close individual in jail or prison	20	12	4	4	.66 [0.44, 0.87]*
Seriously sick/diagnosed with acute illness	20	7	9	4	.38 [0.11, 0.65]*
Forced to leave home	19	6	4	9	.34 [0.07, 0.62]*
Seriously sick/diagnosed with chronic illness	19	13	5	1	.74 [0.55, 0.94]*
Serious, accidental fall	18	9	4	5	.57 [0.31, 0.82]*
Saw something terrible	16	3	7	6	.18 [−0.12, 0.49]
Admitted to a hospital	15	11	3	1	.80 [0.62, 0.99]*
Caregiver alcoholism or drug use	15	9	4	2	.69 [0.45, 0.92]*
Attended childcare in unsafe location	13	6	5	2	.56 [0.28, 0.85]*
Car accident	13	10	1	2	.84 [0.66, 1.00]*
Community violence	13	2	5	6	.16 [−0.16, 0.48]*
Pet died	11	10	0	1	.94 [0.83, 1.00]*
Close individual attempted or completed suicide	9	7	0	2	.86 [0.66, 1.00]*

(continued)

Table 2. (continued)

Adverse Life Event Type	Total Endorsed *N*	Interview and Questionnaire Endorsed *n*	Only Interview Endorsed *n*	Only Questionnaire Endorsed *n*	Kappa [95% CI]
Caregiver died	8	2	3	3	.34 [−0.06, 0.75]
Discrimination	7	4	2	1	.70 [0.38, 1.00]*
Broke a bone	6	4	2	0	.78 [0.49, 1.00]*
Removed from home due to neglect	6	1	5	0	.26 [−0.15, 0.68]
Caregivers divorced	5	3	1	1	.73 [0.38, 1.00]*
Saw another person hurt/injured	5	1	0	4	.31 [−0.16, 0.79]
Seriously burned	4	2	1	1	.65 [0.20, 1.00]*
Peer/friend died	4	0	3	1	—
Consistently ignored by peers	4	0	0	4	—
Physically attacked by another child	4	2	1	1	.65 [0.20, 1.00]*
Poisoned	3	3	0	0	1.00
Almost drowned	2	1	1	0	.66 [0.04, 1.00]*
Attacked by an animal	2	0	2	0	—
Sibling died	2	0	1	1	—
Natural disaster	2	2	0	0	1.00
Been in a fire	2	1	1	0	.66 [0.04, 1.00]*
Hit by a car	1	0	0	1	—
Mugged or robbed	1	0	1	0	—
Removed from home due to physical abuse	1	0	1	0	—
Involved in war/terrorism event	0	0	0	0	—
Kidnapped or taken hostage	0	0	0	0	—
Removed from home due to sexual abuse	0	0	0	0	—
Removed from home due to emotional abuse	0	0	0	0	—
Moved to new country	0	0	0	0	—

Note. Values are sorted from highest to lowest for total endorsement. Blank values (—) are types of adverse life events where a kappa value could not be calculated because there were no participants who endorsed the item on both the questionnaire and interview formats. CI = confident interval.

*$p < .05$.

values could not be calculated because no participant endorsed an event on both forms. Among those types of ALEs where kappa values could be calculated, there were 4 events in the poor to slight range (i.e., $\kappa = < .00–.20$), 15 events in the fair to moderate range (i.e., $\kappa = .21–.60$), 14 events in the substantial range (i.e., $\kappa = .61–.80$), and 5 events in the almost perfect to perfect range (i.e., $\kappa = .81–1.00$). However, it is important to note that for some events, there was an overall endorsement of that type of event on either the interview or questionnaire by less than five participants. Among events endorsed by five or more participants, there were five ALE types where the number of overall endorsed exposure to that type of ALE was greater for the questionnaire only (a) heard something terrible, (b) reduced standard of living, (c) lost contact with close individual, (d) experienced community violence, and (e) saw another person hurt/injured, compared to being endorsed on both the questionnaire and interview or only the interview. There were also three events where the number of overall endorsed exposure to that type of event was greater for the interview only (a) seriously sick/diagnosed with chronic illness, (b) saw something terrible, and (c) removed from home due to neglect, compared to being endorsed on both formats or only on the questionnaire.

Agreement for ALE Frequency

There was no significant difference in average frequency scores between the questionnaire and interview, $t(57) = .63$, $p = .53$. Overall mean frequency scores were also similar between the questionnaire and interview versions for those who started with the questionnaire and those who started with the interview (all $t\text{s} < .46$, $p\text{s} > .65$). The weighted Cohen's kappa coefficients for frequency scores are in Table 3. Among those types of ALEs where weighted Cohen's kappa values could be calculated, there was 1 event in the poor to slight range, 15 events in the fair to moderate range, 6 events in the substantial range, and 1 event in the almost perfect to perfect range.

Agreement for ALE Severity

A paired sample *t*-test found an overall significant difference in mean severity scores between the questionnaire and interview method, $t(57) = 2.50$, $p = .02$, indicating that the mean severity score for the questionnaire was higher than the mean for the interview format (Table 1). Overall mean severity scores were similar on both the questionnaire and interview versions between those who started with the questionnaire and those who started with the interview (all $t\text{s} < 1.74$, $p\text{s} > .09$). The weighted Cohen's kappa

Table 3. Concordance Between Interview and Questionnaire Assessments of Adverse Life Event Frequency, Severity, and Age of Onset.

Adverse Life Event Type	Interview and Questionnaire Endorsed *N*	Frequency Kappa [95% CI]	Severity Kappa [95% CI]	Age of Onset Kappa [95% CI]
Moved to a new home	39	.54 [0.37, 0.70]*	.25 [0.07, 0.43]*	.55 [0.37, 0.74]*
Changed childcare provider	29	.32 [0.16, 0.47]*	.27 [0.02, 0.53]*	.43 [0.18, 0.68]*
New child moved into the home	23	.71 [0.58, 0.84]*	.08 [−0.01, 0.25]	.61 [0.35, 0.86]*
Caregiver admitted to a hospital	22	.40 [0.20, 0.59]*	−.01 [−0.16, 0.15]	.32 [−0.02, 0.66]
Close adult died	20	.60 [0.39, 0.80]*	.67 [0.43, 0.91]	.76 [0.61, 0.91]*
Caregivers separated	18	.33 [0.12, 0.54]*	.26 [−0.02, 0.54]	.57 [0.30, 0.84]*
Caregiver experiences mental illness	18	.35 [0.19, 0.52]*	.29 [0.08, 0.51]*	.53 [0.24, 0.82]*
Exposed to domestic/interpersonal violence	13	.29 [0.11, 0.47]*	.40 [0.02, 0.79]	.33 [−0.04, 0.69]
Seriously sick/diagnosed with chronic illness	13	.66 [0.48, 0.84]*	.38 [0.10, 0.66]*	.50 [0.15, 0.85]
Family/Close individual in jail or prison	12	.57 [0.36, 0.78]*	.20 [−0.19, 0.59]	.67 [0.40, 0.94]*
Caregiver or close individual arrested	11	.35 [0.13, 0.57]*	.67 [0.26, 1.00]*	.41 [0.01, 0.82]*
Admitted to a hospital	11	.72 [0.55, 0.89]*	−.21 [−0.40, −.08]	.49 [0.13, 0.85]*
New caregiver moved into home	10	.42 [0.21, 0.62]*	.47 [0.09, 0.86]*	.29 [−0.09, 0.67]
Pet died	10	.88 [0.76, 1.00]*	.10 [−0.33, 0.53]	.50 [0.13, 0.87]*
Car accident	10	.66 [0.46, 0.87]*	.21 [−0.09, 0.51]	.84 [0.67, 1.00]*
Serious, accidental fall	9	.41 [0.17, 0.65]*	.73 [0.48, 0.98]	.39 [0.00, 0.79]
Caregiver alcoholism or drug use	9	.55 [0.33, 0.77]*	.27 [−0.02, 0.56]	.05 [−0.27, 0.36]
Lost contact with close individual	7	.36 [0.10, 0.62]*	.19 [−0.13, 0.50]	.56 [0.13, 0.99]*
Heard about something terrible	7	.11 [−0.11, 0.32]	.36 [−0.24, 0.97]	.13 [−0.20, 0.44]
Seriously sick/diagnosed with acute illness	7	.30 [0.09, 0.51]*	−.09 [−0.61, 0.43]	.76 [0.58, 0.94]*
Close individual attempted or completed suicide	7	.72 [0.48, 0.97]*	.22 [−0.06, 0.50]	.55 [0.11, .1.00]*
Attended childcare in unsafe location	6	.58 [0.34, 0.82]*	−.15 [0.27, 0.04]	.25 [0.−16, 0.66]
Forced to leave home	6	.44 [0.17, 0.70]*	−.20 [0.56, 0.16]	.50 [0.05, 0.95]*
Broke a bone	4	—	—	—

(continued)

Table 3. (continued)

Adverse Life Event Type	Interview and Questionnaire Endorsed *N*	Frequency Kappa [95% CI]	Severity Kappa [95% CI]	Age of Onset Kappa [95% CI]
Discrimination	4	—	—	—
Caregivers divorced	3	—	—	—
Reduced standard of living	3	—	—	—
Poisoned	3	—	—	—
Saw something terrible	3	—	—	—
Seriously burned	2	—	—	—
Caregiver died	2	—	—	—
Natural disaster	2	—	—	—
Physically attacked by another child	2	—	—	—
Community violence	2	—	—	—
Almost drowned	1	—	—	—
Been in a fire	1	—	—	—
Saw another person hurt/injured	1	—	—	—
Removed from home due to neglect	1	—	—	—
Hit by a car	0	—	—	—
Attacked by an animal	0	—	—	—
Sibling died	0	—	—	—
Peer/Friend died	0	—	—	—
Involved in war/terrorism event	0	—	—	—
Consistently ignored by peers	0	—	—	—
Mugged or robbed	0	—	—	—
Removed from home due to physical abuse	0	—	—	—
Kidnapped or taken hostage	0	—	—	—
Removed from home due to sexual abuse	0	—	—	—
Removed from home due to emotional abuse	0	—	—	—
Moved to new country	0	—	—	—

Note. Values are sorted from highest to lowest for total endorsement on both interview and questionnaire formats. Blank values (—) are types of adverse life events where a weighted kappa value could not be calculated because either (a) there were no participants who endorsed the item on both the questionnaire and interview formats, or (b) the item was not endorsed by both the interview and questionnaire methods more than five times. CI = confident interval.
*$p < .05$.

coefficients for severity scores are provided in Table 3. Among those types of ALEs where weighted Cohen's kappa values could be calculated, there were 8 events in the poor to slight range, 12 events in the slight to moderate range, 3 events in the substantial range, and 0 events in the almost perfect to perfect range.

Agreement for ALE Age of Onset

There were no significant differences in average age of onset scores between formats, $t(57)=1.84$, $p=.07$. Overall mean age of onset scores was similar on both the questionnaire and interview versions between those who started with the questionnaire and those who started with the interview (all $t\text{s} < 1.63$, $p\text{s} > .11$). Among those types of ALEs where weighted Cohen's kappa values could be calculated, there were 2 events in the poor to slight range, 16 events in the fair to moderate range, 4 events in the substantial range, and 1 event in the almost perfect to perfect range.

Discussion

Assessment practices of ALEs for children are often characterized by considerable variability, which can contribute to inconsistency and concerns with accuracy when conducting research on children with ALE exposure (e.g., Bethell et al., 2017). The current study examined concordance between two primary methods of caregiver-reported ALE assessment: interviews and questionnaires. This included a novel approach to examining overall endorsement of an ALE among 50 different event types, as well as multiple characteristics of the exposures, including polyvictimization, frequency, severity, and age of onset.

Based on the overall "scores" associated with ALE exposure obtained (e.g., sum polyvictimization or average frequency), there appeared to be strong agreement between the interview and questionnaire approaches insofar as there were no significant differences in total scores for polyvictimization, frequency, nor age of onset, regardless of which assessment method was used first. This finding is partially consistent with prior research suggesting good concordance between questionnaire and interview formats assessing caregiver-reported polyvictimization or number of ALE types (e.g., Allen et al., 2012). The only exception was the mean severity score, for which caregivers tended to report a more negative impact of the ALEs on the interview format relative to the questionnaire format. This finding partially supports the hypothesis that there would be less concordance between approaches when assessing ALE severity relative to other characteristics. This finding is also partially in line with Wagner et al. (2006), who found that caregivers tended to overreport severity of life events on a questionnaire compared to an interview, after interview responses were validated by a set of independent research coders.

One reason why severity scores may have differed across methods may relate to the greater subjectiveness of determining or defining severity or impact (e.g., Litrownik et al., 2005). That is, judgments about the severity or impact of an event may involve more nuance or factors related to the individual, such as what areas of life were positively versus negatively influenced, as compared to more objective characteristics of an event, such as the number of times an ALE occurred. It may have been the case that having participants discuss their perception of severity aloud during the interview allowed them to engage in more cognitive appraisal of the event, which introduced more variance into their ratings, compared to only reading and then answering the question on the questionnaire alone (e.g., Schlechter et al., 2021). Interestingly, however, the direction of impact, which was more negative for the interview (i.e., lower severity scores) relative to the questionnaire format, contrasts with some literature on reporting of stressful or difficult topics. Some evidence suggests that individuals will often try to minimize the negative impact of an event or situation when reporting face-to-face in an interview as compared to other data collection methods, perhaps to increase social desirability (e.g., Reddy et al., 2006; Tourangeau & Yan, 2007). These findings may suggest that caregivers are less likely to minimize severity when reporting about their child, as opposed to themselves.

Moving beyond the findings related to overall ALE scores, examining the concordance among the individual types of ALEs and the characteristics of these ALEs (i.e., frequency, severity, and age of onset) paints a different picture. While overall scores were largely similar between formats, the ALEs that contributed to those scores were markedly different for each ALE assessment format. For general endorsement of a specific type of ALE exposure (i.e., 1=yes, 0=no), concordance based on Cohen's kappa coefficients ranged from none to almost perfect. Interestingly, there would have been several types of ALE not observed or not believed to have occurred in the sample had a single assessment method been used. For example, there were two ALEs that were exclusively reported on the questionnaire (i.e., hit by a car, consistently ignored by peers), and three ALEs exclusively reported on the interview (i.e., attacked by an animal, mugged or robbed, removed from home due to physical abuse). There were also multiple ALE types for which the number of endorsements exclusively reported on one method was greater than endorsements on the alternative method, as well as endorsement on both methods (e.g., exposure to community violence was endorsed six times on the questionnaire format exclusively, whereas it was exclusively endorsed on the interview format five times and on both formats two times). These observations suggest that researchers studying childhood ALEs should consider using multiple assessment methods rather than relying on a single assessment method. This may be particularly important for research studies that seek to recruit or compare groups of children based on exposure to a specific ALE (e.g., exposed vs. nonexposed), for which a single method may lead to a nonrepresentative sample or in cases where researchers might need to consider risk of contamination when there are actually children with exposure to an ALE in a supposed "nonexposed" group (Shenk et al., 2016).

In a similar fashion, concordance rates for other ALE characteristics ranged from very poor to substantial across all types of ALEs. These findings demonstrating concordance values that fall predominantly in the fair-to-moderate range for both overall endorsement and endorsement of characteristics of specific ALE types align with some of the general grouping patterns observed in prior studies on concordance between interview and questionnaire formats. For example, Glackin et al. (2019) reported kappa values between interview and questionnaire formats for any type of ALE and indirect violence to be in the fair range, whereas exposure to noninterpersonal violence was in the poor range. The present data extend these prior findings by demonstrating this variability at the level of the individual type of ALE exposure. This study is also novel in its ability to simultaneously examine several characteristics of a certain type of ALE (i.e., frequency, severity, and age of onset) demonstrating that concordance also varies within a specific ALE type.

Among the various ALEs, there were several ALEs where the concordance estimates for each ALE characteristic were almost all in the moderate to almost perfect range, such as a close adult dying, being in a car accident, having a serious accidental fall, being diagnosed with a chronic illness, having a caregiver or close individual arrested, or a close individual attempted or died by suicide. Some of these trends appear to suggest that events involving death or serious physical health challenges to the child may tend to show more robust agreement between formats when relying on caregiver report. There were also various ALEs with all or almost all concordance estimates below the moderate range, such as a caregiver being admitted to the hospital, hearing about something terrible happening (e.g., the child hearing someone close to them was hurt), or exposure to domestic/interpersonal violence. Many of these events occurred to the caregiver themselves. Further examination across the different types of events showed that events occurring to a child's caregiver were often in the fair range at best (e.g., domestic violence, caregiver admitted to the hospital, and caregiver died). The only events in the substantial agreement or higher range for events involving caregivers were for concordance of type regarding caregivers' divorcing or a caregiver having an alcohol or drug use concern. Given some previous findings on reporting on the self as compared to close individuals (e.g., Lewinsohn et al., 2003), one might expect concordance to be higher among those events the child experienced that involved the caregiver since the caregiver is the reporter. However, this was not the case. Examination of endorsement of events on the questionnaire and interview related to caregivers suggests that one explanation for this might be that caregivers had a significant and chronic history of ALE exposure themselves. Thus, caregivers may have had more difficulty in reporting consistently for these events because they were having to also consider their own ALE history. This has been shown in previous research to influence consistency of reporting ALEs because of needing to think about a greater number of events, as well as possible mental health challenges that can interfere with recall (e.g., Glackin et al., 2019; Schweizer & Dalgleish, 2011).

There were also several ALEs where concordance estimates across the ALE characteristics fluctuated from poor-to-substantial or almost perfect. This included events such as attending childcare in an unsafe location, death of pet, and being seriously sick/diagnosed with acute illness. Further, in these situations, the kappa value was also often the lowest for ALE severity relative to frequency and age of onset. These findings demonstrate the importance of considering assessment format beyond simply whether a child has had an ALE, to also consider how format may influence how these events are described. This is especially necessary as the fields moves to a more multidimensional approach of ALE assessment and seeks to utilize multiple characteristics of ALE within the same model when evaluating how ALE may be predictive or associated with children's functioning (e.g., mental health, developmental abilities: Grasso et al., 2021).

Recommendations and Future Research for ALE Assessment

There are several important recommendations for research with young children that may have exposure to ALE. Considering the variability in endorsement, one recommendation is to employ two different assessment methods or sources of information that contain similar questions when assessing ALE history. This is not itself a novel idea, but the present findings extend work on this topic and underscore the importance of using this dual-assessment approach to help ensure a more accurate and comprehensive assessment of ALEs (Jackson et al., 2019; Shaffer et al., 2008). This can provide several advantages toward ensuring accurate research findings, such as helping to ensure that certain types of ALEs are not missed. This might involve administering two ALE assessment formats, such as a questionnaire and interview format. Other options might be to administer a questionnaire measure via paper-and-pencil format and a computer version of that measure, or to obtain reports on ALE history from multiple caregivers (e.g., Diaz et al., 2017). These types of approaches may be especially necessary when working with children and families from low resourced or racial and ethnic minority populations, who are often exposed to higher rates of ALEs compared to white or higher economic status populations, often due to systematic and structural inequities (Slopen et al., 2016). That is, these children may have exposure to a wider range of ALEs; thus, multiple assessment methods are needed to accurately document the ALEs.

Importantly, the authors recognize that this is not a simple task. Assessing ALEs can be challenging due to the nature of the questions being asked, as well as the resource and time burden on participants and researchers. However, it may be necessary to weigh these tradeoffs, such as determining whether the risks associated with administering two ALE assessments outweigh the risks of having inaccurate or incomplete assessment of children's ALEs. Future research on this topic may help identify both comprehensive and efficient methods for achieving both goals of obtaining accurate assessments and reducing both caregiver and researcher burden. For example, when administering two forms of an ALE assessment, it may be less burdensome and more efficient for caregivers to complete a questionnaire first independently, which could then help inform a subsequent interview with a researcher or professional. The findings of nonsignificant differences in administration order from the current study suggest that this may not significantly bias reports, but further research would be necessary to determine if this would be generalizable across other assessment methods (e.g., computer-administered questionnaire or online interview).

The results related to less method concordance when assessing ALE severity suggest additional recommendations. When using an impact-based approach to quantifying severity, one option may be to ask caregivers to report on ALE severity as it relates to more specific domains of wellbeing or functioning (e.g., psychological health, physical health, and/or development) as opposed evaluating overall severity. It is also worth noting that ALE severity as measured by perceived negative/positive impact on the child is one of many possible approaches to characterizing ALE severity (e.g., Trickey et al., 2012). There may exist other methods for quantifying severity that demonstrate more robustness across assessment formats, such as proximity, method of victimization/injury, or amount of physical harm. However, while there exist other potentially more objective measures of severity (e.g., physical injury), these methods do not always translate appropriately across all ALE types. Future research might examine method concordance across different operational definitions or ways of quantifying ALE severity.

Limitations and Conclusion

Several limitations of this study should be noted. First, the study was limited with respect to the number of analyses that could be conducted on the concordance of certain types of ALEs due to low-or-no endorsement of that type of ALE. With a larger sample, it may have been possible to capture more types of ALE exposure. Another limitation is that all interviews were completed the same interviewer. This did not allow for testing possible interviewer influences on caregiver reporting related to race, ethnicity, and other important individual differences that have been shown in previous literature to influence reporting of ALEs (e.g., Glackin et al., 2019). Further, as this study occurred during the COVID-19 pandemic, the use of facemasks during assessments could have affected ALE reports. Although there is no known research on reporting of ALEs while wearing facemasks, literature from other assessment fields has demonstrated that wearing face masks can influence adult's perceptions of an interviewer's perceived trust, safety, or emotional response (e.g., Biermann et al., 2021). These are factors that have previously been shown to influence reporting behaviors of sensitive topics, including ALE exposure (e.g., Decker et al., 2011). Finally, this study was limited in its assessment of some specific forms of maltreatment, including primary forms of abuse and neglect. Due to concerns about population participation and involvement of child protective services if reporting child maltreatment, questions regarding maltreatment exposure were not assessed. Although questions pertaining to removal from the household due to these concerns were included as a proxy, such questions are not fully adequate substitutes when assessing for exposure to these types of ALEs. Relatedly, the present study did not assess other features of events that could be relevant for health, such as the event's primary life domain (e.g., housing, intimate relationships) or social–psychological characteristics (e.g., interpersonal loss, physical danger). Additional research is thus needed to assess these characteristics, which have been shown to be relevant for a variety of outcomes in childhood (e.g., Slavich & Shields, 2018; Slavich et al., 2019).

Despite these limitations, this study provides several important insights into caregiver reporting behaviors of their child's ALE history. These data demonstrate that although overall or sum scores of a child's ALE history may be similar between formats, the types of events used to generate those scores are composed of different types of ALEs. This suggests that researchers may have inaccurate or incomplete assessments of children's ALE experiences when using a single format or assessment approach. Further, this study was the first to demonstrate that lack of concordance between ALE assessment methods extends beyond just overall endorsement of a type of event to also include concerns about concordance for the characteristics of these ALEs—namely frequency, severity, and age of onset of exposure. Researchers should ensure they are utilizing appropriate and accurate assessments of a child's ALE history and consider administering two versions of a ALE assessment to help ensure that a comprehensive ALE history is obtained.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This work was supported by National Institute of Mental Health grants 5R01MH079252-09 (YJ), 1F31MH124340-01A1 (AM), and manuscript preparation for Dr. McGuire was supported in part by T32MH018869. GMS was supported by grant #OPR21101 from the California Governor's Office of Planning and Research/California Initiative to Advance Precision Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of these organizations.

ORCID iDs

Austen McGuire https://orcid.org/0000-0002-7820-2927
Damion J. Grasso https://orcid.org/0000-0001-9577-2764

References

Allen, J. L., Rapee, R. M., & Sandberg, S. (2012). Assessment of maternally reported life events in children and adolescents: A comparison of interview and checklist methods. *Journal of Psychopathology and Behavioral Assessment*, *34*, 204–215. https://doi.org/10.1007/s10862-011-9270-5

Bethell, C. D., Carle, A., Hudziak, J., Gombojav, N., Powers, K., Wade, R., & Braveman, P. (2017). Methods to assess adverse childhood experiences of children and families: Toward approaches to promote child well-being in policy and practice. *Academic Pediatrics, 17*, S51–S69. https://doi.org/10.1016/j.acap.2017.04.161

Biermann, M., Schulze, A., Unterseher, F., Atanasova, K., Watermann, P., Krause-Utz, A., Stahlberg, D., Bohus, M., & Lis, S. (2021). Trustworthiness appraisals of faces wearing a surgical mask during the Covid-19 pandemic in Germany: An experimental study. *PLoS One*, *16*(5), e0251393. https://doi.org/10.1371/journal.pone.0251393

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagree-ment or partial credit. *Psychological Bulletin*, *70*, 213–220. https://doi.org/10.1037/h0026256

Decker, S. E., Naugle, A. E., Carter-Visscher, R., Bell, K., & Seifert, A. (2011). Ethical issues in research on sensitive topics: Participants' experiences of distress and benefit. *Journal of Empirical Research on Human Research Ethics*, *6*, 55–64. https://doi.org/10.1525/jer.2011.6.3.55

Diaz, A., Peake, K., Nucci-Sack, A., & Shankar, V. (2017). Comparison of modes of administration of screens to identify a history of childhood physical abuse in an adolescent and young adult population. *Annals of Global Health*, *83*, 726–734. https://doi.org/10.1016/j.aogh.2017.10.023

Duggal, S., Malkoff-Schwartz, S., Birmaher, B., Anderson, B. P., Matty, M. K., Houck, P. R., Bailey-Orr, M., Williamson, D. E., & Frank, E. (2000). Assessment of life stress in adolescents: Self-report versus interview methods. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 445–452. https://doi.org/10.1097/00004583-200004000-00013

Egger, H. L., & Angold, A. (2004). The Preschool Age Psychiatric Assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. In R. DelCarmen-Wiggins & A. Carter (Eds.), *Handbook of infant, toddler, and preschool mental health assessment* (pp. 223–243). Oxford University Press.

Finkelhor, D., Hamby, S., Turner, H., & Ormrod, R. (2011). *The Juvenile Victimization Questionnaire: 2nd revision (JVQ-R2)*. Crimes Against Children Research Center.

Finkelhor, D., Turner, H. A., Shattuck, A., & Hamby, S. L. (2015). Prevalence of childhood exposure to violence, crime, and abuse: Results from the national survey of children's exposure to violence. *JAMA Pediatrics*, *169*, 746–754. https://doi.org/10.1001/jamapediatrics.2015.0676

Glackin, E. B., Forbes, D., Heberle, A. E., Carter, A. S., & Gray, S. A. (2019). Caregiver self-reports and reporting of their preschoolers' trauma exposure: Discordance across assessment methods. *Traumatology*, *25*, 172–180. https://doi.org/10.1037/trm0000179

Grasso, D. J., Dierkhising, C. B., Branson, C. E., Ford, J. D., & Lee, R. (2016). Developmental patterns of adverse childhood experiences and current symptoms and impairment in youth referred for trauma-specific services. *Journal of Abnormal Child Psychology*, *44*, 871–886. https://doi.org/10.1007/s10802-015-0086-8

Grasso, D. J., DiVietro, S., Beebe, R., Clough, M., & Lapidus, G. (2021). Quantifying severity of maltreatment, adversity, and trauma from child protective services case record files. *Journal of Interpersonal Violence*, *36*(17–18), 8142–8163. https://doi.org/10.1177/0886260519847774

Griffith, R., Huffhines, L., Jackson, Y., & Nowalis, S. (2020). Strategies for longitudinal recruitment and retention with parents and preschoolers exposed to significant adversity: The PAIR project as an example of methods, obstacles, and recommendations. *Children and Youth Services Review*, *110*, 104770. https://doi.org/10.1016/j.childyouth.2020.104770

Harkness, K. L., & Monroe, S. M. (2016). The assessment and measurement of adult life stress: Basic premises, operational principles, and design requirements. *Journal of Abnormal Psychology*, *125*, 727–745. https://doi.org/10.1037/abn0000178

Jackson, Y., McGuire, A., Tunno, A. M., & Makanui, P. K. (2019). A reasonably large review of operationalization in child maltreatment research: Assessment approaches and sources of information in youth samples. *Child Abuse & Neglect*, *87*, 5–17. https://doi.org/10.1016/j.chiabu.2018.09.016

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

Lewinsohn, P. M., Rohde, P., & Gau, J. M. (2003). Comparability of self-report checklist and interview data in the assessment of stressful life events in young adults. *Psychological Reports*, *93*, 459–471. https://doi.org/10.2466/pr0.2003.93.2.459

Litrownik, A. J., Lau, A., English, D. J., Briggs, E., Newton, R. R., Romney, S., & Dubowitz, H. (2005). Measuring the severity of child maltreatment. *Child Abuse & Neglect*, *29*, 553–573. https://doi.org/10.1016/j.chiabu.2003.08.010

Milne, L., & Collin-Vézina, D. (2015). Assessment of children and youth in child protective services out-of-home care: An overview of trauma measures. *Psychology of Violence*, *5*, 122–132. https://doi.org/10.1037/a0037865

Monroe, S. M., & Slavich, G. M. (2020). Major life events: A review of conceptual, definitional, measurement issues, and practices. In K. L. Harkness & E. P. Hayden (Eds.), *The Oxford handbook of stress and mental health* (pp. 7–26). Oxford University Press.

Oh, D. L., Jerman, P., Boparai, S. K. P., Koita, K., Briner, S., Bucci, M., & Harris, N. B. (2018). Review of tools for measuring exposure to adversity in children and adolescents. *Journal of Pediatric Health Care*, *32*, 564–583. https://doi.org/10.1016/j.pedhc.2018.04.021

Peirce, J. M., Burke, C. K., Stoller, K. B., Neufeld, K. J., & Brooner, R. K. (2009). Assessing traumatic event exposure: Comparing the traumatic life events questionnaire to the structured clinical interview for DSM-IV. *Psychological Assessment*, *21*, 210–218. https://doi.org/10.1037/a0015578

Reddy, M. K., Fleming, M. T., Howells, N. L., Rabenhorst, M. M., Casselman, R., & Rosenbaum, A. (2006). Effects of method on participants and disclosure rates in research on sensitive topics. *Violence and Victims*, *21*, 499–506. https://doi.org/10.1891/0886-6708.21.4.499

Schlechter, P., Fritz, J., Cassels, M., Neufeld, S. A., & Wilkinson, P. O. (2021). The Youth and Childhood Adversity Scale: A step towards developing a new measure of adversity and its severity. *European Journal of Psychotraumatology*, *12*(1), 1981573. https://doi.org/10.1080/20008198.2021.1981573

Schweizer, S., & Dalgleish, T. (2011). Emotional working memory capacity in post-traumatic stress disorder (PTSD). *Behaviour Research and Therapy*, *49*(8), 498–504. https://doi.org/10.1016/j.brat.2011.05.007

Shaffer, A., Huston, L., & Egeland, B. (2008). Identification of child maltreatment using prospective and self-report methodologies: A comparison of maltreatment incidence and relation to later psychopathology. *Child Abuse & Neglect*, *32*(7), 682–692. https://doi.org/10.1016/j.chiabu.2007.09.010

Shenk, C. E., Noll, J. G., Peugh, J. L., Griffin, A. M., & Bensman, H. E. (2016). Contamination in the prospective study of child maltreatment and female adolescent health. *Journal of Pediatric Psychology*, *41*, 37–45. https://doi.org/10.1093/jpepsy/jsv017

Slavich, G. M. (2019). Stressnology: The primitive (and problematic) study of life stress exposure and pressing need for better measurement. *Brain, Behavior, and Immunity*, *75*, 3–5. https://doi.org/10.1016/j.bbi.2018.08.011

Slavich, G. M., & Shields, G. S. (2018). Assessing lifetime stress exposure using the Stress and Adversity Inventory for Adults (Adult STRAIN): An overview and initial validation. *Psychosomatic Medicine*, *80*, 17–27. https://doi.org/10.1097/PSY.0000000000000534

Slavich, G. M., Stewart, J. G., Esposito, E. C., Shields, G. S., & Auerbach, R. P. (2019). The Stress and Adversity Inventory for Adolescents (Adolescent STRAIN): Associations with mental and physical health, risky behaviors, and psychiatric diagnoses in youth seeking treatment. *Journal of Child Psychology and Psychiatry*, *60*, 998–1009. https://doi.org/10.1111/jcpp.13038

Slopen, N., Shonkoff, J. P., Albert, M. A., Yoshikawa, H., Jacobs, A., Stoltz, R., & Williams, D. R. (2016). Racial disparities in child adversity in the US: Interactions with family immigration history and income. *American Journal of Preventive Medicine*, *50*, 47–56. https://doi.org/10.1016/j.jpeds.2013.11.003

Tang, W., Hu, J., Zhang, H., Wu, P., & He, H. (2015). Kappa coefficient: A popular measure of rater agreement. *Shanghai Arch Psychiatry*, *27*, 62–67. https://doi.org/10.11919/j.issn.1002-0829.215010

Thabrew, H., de Sylva, S., & Romans, S. (2012). Evaluating childhood adversity. In G. Fava, N. Sonino, & T. Wise (Eds.), *The psychosomatic assessment* (pp. 35–57). Karger.

Tingskull, S., Svedin, C. G., Agnafors, S., Sydsjö, G., deKeyser, L., & Nilsson, D. (2015). Parent and child agreement on experience of potential traumatic events. *Child Abuse Review*, *24*, 170–181.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*, 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Trickey, D., Siddaway, A. P., Meiser-Stedman, R., Serpell, L., & Field, A. P. (2012). A meta-analysis of risk factors for post-traumatic stress disorder in children and adolescents. *Clinical Psychology Review*, *32*, 122–138. https://doi.org/10.1016/j.cpr.2011.12.001

U.S. Department of Health and Human Services (2021). 2021 poverty guidelines. Office of the Assistant Secretary for Planning and Evaluation (ASPE). https://aspe.hhs.gov/2021-poverty-guidelines.

Vanaelst, B., De Vriendt, T., Huybrechts, I., Rinaldi, S., & De Henauw, S. (2012). Epidemiological approaches to measure childhood stress. *Paediatric and Perinatal Epidemiology*, *26*, 280–297. https://doi.org/10.1111/j.1365-3016.2012.01258.x

Wagner, C., Abela, J. R., & Brozina, K. (2006). A comparison of stress measures in children and adolescents: A self-report checklist versus an objectively rated interview. *Journal of Psychopathology and Behavioral Assessment*, *28*, 250–260. https://doi.org/10.1007/s10862-005-9010-9

Author Biographies

Austen McGuire, PhD, is a postdoctoral fellow in the National Crime Victims Research and Treatment Center at the Medical University of South Carolina. His research focuses on evidence-based assessment of adverse life events, in addition to understanding risk and resilience factors for children exposed to maltreatment and evidence-based treatment dissemination and implementation.

Yo Jackson, PhD, is a clinical child psychologist and Associate Director of Penn State's Child Maltreatment Solutions Network. Her research focuses on the development of models of resilience for youth exposed to trauma and the intergenerational transmission of trauma. She also works on trauma assessment, as well as the assessment of emotion regulation and cognitive functioning for adversity exposed youth and families.

Damion Grasso, PhD, is a licensed clinical psychologist and Associate Professor in the Departments of Psychiatry and Pediatrics at the University of Connecticut School of Medicine. His research focuses on biobehavioral risk mechanisms linking childhood adversity with emergent psychopathology, as well as evidence-based assessment and treatment of trauma-related disorders.

George M. Slavich, PhD, is a clinical psychologist and Professor in the Department of Psychiatry and Biobehavioral Sciences at the University of California, Los Angeles, where he directs the UCLA Laboratory for Stress Assessment and Research. His research focuses on the conceptualization, assessment, and management of life stress, as well as psychological and biological mechanisms linking stress with poor mental and physical health.

Neal Kingston, PhD, is a University Distinguished Professor and Director of Graduate Studies in the Educational Psychology Department at the University of Kansas, and Director of the Achievement and Assessment Institute. His research focuses on how large-scale assessment can better support student learning through the use of learning maps and diagnostic classification models.