Method

Research Staff

The research staff consisted of the co-authors and research assistants. Three of the co-authors are faculty members at a public university. One co-author (CK) is a researcher with a background in clinical and developmental psychology, a second one (SMM) is a social epidemiologist in a school of public health, and a third co-author (AB) has a background in developmental-behavioral pediatrics and is a practicing physician. One co-author (BG) is a Ph.D. student in Developmental Psychology, and another (FC) was a post-baccalaureate researcher at the time of the study and started a Ph.D. program in Developmental Psychology in Fall 2021. Three co-authors are librarians from health and social sciences with experience in systematic and scoping reviews (AR, CB, and SH) in medicine, public health, and psychology.

Ten articles included in the initial database and seven articles included in the list of eligible articles were co-authored by one of the two co-authors (CK and AB); neither co-author was involved in decisions regarding screening or data extraction for their studies.

The search strategies were created by the three librarians (AR, CB, and SH). Screening was conducted by three co-authors (CK, SMM, and AB); a graduate student in public health; and seven undergraduates and recent graduates from Psychology and Child Psychology.

Data extraction started with an iterative process involving up to five of the co-authors (CK, SMM, AB, BG, FC) to clarify and improve the coding instructions. For Part 1, data extraction was completed by CK and 26 undergraduates and recent graduates. For Part 2, data extraction was completed by three co-authors (CK, BG, FC), assisted by 20 undergraduate research assistants in psychology and child psychology. For Part 3, data extraction for protective factors was completed by CK and seven undergraduate research assistants. Data extraction for

researcher recommendations was completed by three co-authors (CK, AB, and BG), and a research assistant.

Identifying and Screening the Articles

Identifying Relevant Articles

We identified relevant articles following guidelines for scoping reviews (Arksey & O'Malley, 2005). Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines were followed to conduct the review and analysis (Tricco et al., 2018). The librarian co-authors (AR, CB, and SH) created the search strategies for nine databases: PsycINFO, Ovid MEDLINE, ERIC via Ebsco, Education Source, CINAHL, Embase, Social Services Abstracts, PAIS Index, and Scopus. These databases represent literature from the fields of medicine, public health, psychology, education, nursing, social science, and public policy. The nine selected databases represented a breadth of literature and best served the research question. The searches were run between November 26 and December 2, 2019. The time frame was from 1/1/1999 to the search period. We excluded articles prior to 1999 as we were explicitly interested in research responding to the 1998 article (Felitti et al., 1998) that coined the term ACEs.

The full electronic search strategies for each database are detailed below.

Embase
("adverse child* experience*" or "adverse child* event*").tw,kw.
Limit 1 to yr= "1999-Current"

Education Source
TI ("Adverse Child* Experience*" OR "Adverse Child* Event*") OR AB ("Adverse Child* Experience*" OR "Adverse Child* Event*") OR KW ("Adverse Child* Experience*" OR "Adverse Child* Event*")
Limit 1 to "1999-current"

ERIC via EBSCO

TI ("Adverse Child* Experience*" OR "Adverse Child* Event*") OR AB ("Adverse Child* Experience*" OR "Adverse Child* Event*") OR KW ("Adverse Child* Experience*" OR "Adverse Child* Event*")
Limit 1 to "1999-current"

MEDLINE via Ovid

("adverse child* experience*" or "adverse child* event*").tw,kw.
exp Adverse Childhood Experiences/
or/1-2
limit 4 to yr="1999-Current"

PsycINFO

("Adverse Child* Experience*" OR "Adverse Child* Event*").ti,ab,id
Limit 1 to yr="1999-current"

Scopus

TITLE-ABS-KEY ("adverse child* experience*" OR "adverse child* event*") AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015) OR LIMIT-TO (PUBYEAR , 2014) OR LIMIT-TO (PUBYEAR , 2013) OR LIMIT-TO (PUBYEAR , 2012) OR LIMIT-TO (PUBYEAR , 2011) OR LIMIT-TO (PUBYEAR , 2010) OR LIMIT-TO (PUBYEAR , 2009) OR LIMIT-TO (PUBYEAR , 2008) OR LIMIT-TO (PUBYEAR , 2007) OR LIMIT-TO (PUBYEAR , 2006) OR LIMIT-TO (PUBYEAR , 2005) OR LIMIT-TO (PUBYEAR , 2004) OR LIMIT-TO (PUBYEAR , 2003) OR LIMIT-TO (PUBYEAR , 2002) OR LIMIT-TO (PUBYEAR , 2001) OR LIMIT-TO (PUBYEAR , 2000) OR LIMIT-TO (PUBYEAR , 1999))

Social Services Abstracts

ti(("Adverse Child* Experience*" OR "Adverse Child* Event*")) OR ab(("Adverse Child* Experience*" OR "Adverse Child* Event*")) OR if(("Adverse Child* Experience*" OR "Adverse Child* Event*"))
Limited to 1999-present

PAIS Index

"adverse childhood experience" OR "adverse childhood experiences" OR "adverse child experience" OR "adverse child experiences" OR "adverse childhood events" OR "adverse childhood event" OR "adverse child events" OR "adverse child event"
Limits applied year: 1999; 2000; 2001; 2002; 2003; 2004; 2005; 2006; 2007; 2008; 2009; 2010; 2011; 2012; 2013; 2014; 2015; 2016; 2017; 2018; 2019

CINAHL
TI (("adverse childhood experience" or "adverse childhood experiences" or "adverse child experience" or "adverse child experiences" or "adverse childhood events" or "adverse childhood event" or "adverse child events" or "adverse child event")) OR AB (("adverse childhood experience" or "adverse childhood experiences" or "adverse child experience" or "adverse child experiences" or "adverse childhood events" or "adverse childhood event" or "adverse child events" or "adverse child event")) OR MW (("adverse childhood experience" or "adverse childhood experiences" or "adverse child experience" or "adverse child experiences" or "adverse childhood events" or "adverse childhood event" or "adverse child events" or "adverse child event")) OR MH (("adverse childhood experience" or "adverse childhood experiences" or "adverse child experience" or "adverse child experiences" or "adverse childhood events" or "adverse childhood event" or "adverse child events" or "adverse child event")) OR SU (("adverse childhood experience" or "adverse childhood experiences" or "adverse child experience" or "adverse child experiences" or "adverse childhood events" or "adverse childhood event" or "adverse child events" or "adverse child event")) OR DH (("adverse childhood experience" or "adverse childhood experiences" or "adverse child experience" or "adverse child experiences" or "adverse childhood events" or "adverse childhood event" or "adverse child events" or "adverse child event"))
Limiters - Published Date: 19990101-20201231

Because the research question was to document how various disciplines define ACEs, the initial search strategy was designed to capture literature that mentioned ACEs or similar terms. The use of subject headings was tested, and in most cases led to a search that was over-sensitive based on our needs; thus, search strategies in most databases did not utilize subject headings. We did not use ACEs as a search term, as it is also used as an acronym for terms that were not relevant to the current study (e.g., ACE inhibitors). Duplicate references were removed to the extent possible, and items were uploaded for screening to Rayyan, a database for conducting systematic reviews (Ouzzani et al., 2016).

Inclusionary and Exclusionary Criteria for Title/Abstract Screening

After reaching adequate inter-rater reliability (described below), pairs of team members independently reviewed the titles and abstracts of the selected articles based on the following a

priori inclusionary criteria: (1) the abstract must describe a study based on original analysis of quantitative data from humans, and (2) the abstract must include a measure of risk factors referred to with the term ACEs or mention ACEs or a similarly broad term encompassing multiple risk factors similar to ACEs (e.g., "adverse family experiences," "childhood adversities," "exposure to violence and other early adversities," "childhood stressors"). At this stage, we included articles using broad terms similar to ACEs to check the full text in the next stage. We also included articles where there was uncertainty about the measure, as long as there was a suggestion that some measure of risk factors similar to ACEs was used (e.g., "a measure of early adverse experiences," "childhood exposure to violence and other adverse events").

As our goal was to examine methodological decisions made by researchers in original, empirical studies, we excluded reviews, meta-analyses, case studies, dissertations, solely qualitative studies, all gray literature, and other items that did not provide enough information for data extraction (e.g., conference proceedings, commentaries, letters to the editor) whenever we could identify them through the titles and abstracts. If no abstract was available, we excluded the study only if it was clear from the title that the study met exclusionary criteria (e.g., inclusion of the words "review," or "meta-analysis"). We based our decisions on the titles and abstracts with one exception: we based exclusion on journal or publication type if it referred to "Dissertation Abstracts." Although inclusion in peer-reviewed journals could not be used as a search criterion in all the databases, other criteria (listed in the next section) made it likely that almost all of the identified articles would be in peer-reviewed journals.

Inclusionary and Exclusionary Criteria for Full-Text Screening

The librarians used all available means (including interlibrary loans and directly contacting the authors) to obtain the full texts of the articles included after the title/abstract

screening. We limited our search to publicly available studies. Given the large number of articles identified through the librarians' rigorous search, we did not conduct further searches through the reference sections of screened articles, or manual searches of relevant journals or forward citations, nor did we contact content experts.

We had initially planned to focus the review on only the term "adverse childhood experiences." However, during the abstract screening phase, we encountered many terms that overlapped with ACEs that included combinations of these and other similar words: childhood adverse experiences/conditions, (cumulative) childhood/developmental adversity, adverse caregiving/family experiences, psychosocial/early life adversity, childhood (psychosocial) stress, childhood misfortune/difficulties/victimization/poly-victimization, childhood/early life stress/stressors/stressful life events, toxic stress, stressful family environment, lifetime stress exposure, early/cumulative trauma/risk factors/victimization, (early) childhood trauma/traumatic events/experiences, pathological/negative/unstable childhood experiences, potentially traumatic events/experiences, complex trauma, negative family events, potential family life difficulties, and interpersonal violence. Thus, to both capture the broad range of research in this area but to limit the review to articles directly relevant to the field of ACEs, we defined the inclusionary criteria for full-text screening as articles in which the authors used a term in the title, Introduction, or Method that included the words "adverse" or "adversity" to refer to what they were studying (regardless of the name of the measure and other terms used interchangeably with it) or used the words "adverse" or "adversity" to describe the measure (regardless of the term used to refer to it). For example, if the authors stated that they studied "cumulative childhood adversities" using the "Risky Families" measure, or if they wrote that they studied "early life stress" using a measure of "developmental adversity," we included these articles. It should be

noted, however, that these criteria meant that some articles using measures that overlapped with Felitti and colleagues' ACEs measure were excluded if they did not use a term that included "adverse" or "adversity" to refer to it.

To capture the literature generated by the original ACEs study (Felitti et al., 1998), included articles also had to have a measure of cumulative risk, with at least two distinct types of adversity risk factors (e.g., maltreatment and parental psychopathology) between ages 0 and 18. Because a key innovation of the original ACE study was combining information on maltreatment and household dysfunction, we counted different types of maltreatment as one type of risk factor and excluded articles that defined ACEs only in terms of maltreatment (i.e., physical, sexual, and emotional abuse, and physical and emotional neglect). These are included under "wrong measure" in Figure 1 in the paper. In addition, we excluded studies in which the goal was to train practitioners or staff on ACEs or to ask them if they were screening for ACEs. We also excluded duplicates and articles in languages other English. The other exclusionary criteria (publication type, collection of data from humans) were the same as for the title/abstract screening. The screening process is copied below.

Process of Screening and Establishing Inter-Rater Reliability for Screening and Data Extraction

Title/Abstract Screening. The large number of screeners required three rounds of training on three sets of 40 to 50 articles. Agreement was defined as the number of articles yielding the same decision divided by the number of articles screened. Screeners reviewed discrepancies and clarified the inclusionary and exclusionary criteria in an iterative manner (Arksey & O'Malley, 2005; Levac et al., 2010) in between rounds. After the third round, 60% agreement was reached among all eleven screeners, but 85% agreement among the three co-

authors (CK, SSM, AB). In cases where the three co-authors agreed on an article, the other screeners agreed with them on 87% to 100% of the articles. Although we had indicated in our protocol that we would proceed after reaching 80% agreement across the whole set of screeners, we decided that this level of agreement might have been unrealistically high for 11 screeners. Thus, we deviated from the protocol in this respect.

Next, pairs of team members independently reviewed each title/abstract record, assigned on the basis of the last name of the first author. Decisions about whether to include the article in the full-text review were made using a decision tree that incorporated the inclusionary and exclusionary criteria. Decisions were entered as Include, Exclude, or Maybe. The pairs assigned to each record met to resolve uncertainties and discrepancies after all articles were reviewed and unblinded. If they could not agree, the record was included in the next step. We did not tabulate reasons for exclusion or calculate agreement at this stage, as the goal was to reach 100% agreement.

Full-Text Screening. The same process was repeated in Rayyan (Ouzzani et al., 2016) for full-text screening; however, we also tabulated reasons for exclusion. Only the first reason was included in the analyses, though we attempted to use the same exclusionary criteria in the same order. If a resolution could not be reached, screeners consulted a co-author not involved in the decision (CK, SSM, or AB).

We created a data extraction form and revised the variable definitions in an iterative manner based on pilot testing, as recommended by studies of scoping methodology (Arksey & O'Malley, 2005; Levac et al., 2010). The codes were divided among team members, with each code extracted from each article independently by two screeners, referring as necessary to published supplementary materials or other articles cited as sources.

Inter-Rater Reliability During Data Extraction. We were not able to use kappa as a measure of inter-rater reliability, as we had indicated in our protocol. Many of the codes involved very uneven proportions of responses across response categories, and codes were not independent of each other within articles (e.g., multiple ACEs or purposes could be assigned to each article), both of which yield unreliable kappa estimates (Eugenio & Glass, 2004; Sim & Wright, 2005). Therefore, we used percent agreement as an index of inter-rater reliability across all codes.

For codes that were contingent on other codes, we calculated reliability only on articles on which there was agreement on the initial code (e.g., reliability of whether there was a reported conflict of interest was calculated only on articles where the coders agreed on the presence of a conflict interest statement).

We created a data extraction form and initially coded articles in several batches of 10 to 50 in an iterative manner, as recommended by studies of scoping methodology (Arksey & O'Malley, 2005; Levac et al., 2010). We first coded the articles independently, discussed the discrepancies, clarified the definitions, and came to a consensus decision. After clarifying the coding instructions for Parts 1 and 3, we conducted reliability checks between screeners on different sets of articles until we reached 80% agreement on all codes before coding additional articles. This final step was not feasible for Part 2 due to the number and nature of the codes.

New coders were trained using articles we had already reached consensus decisions on. New coders were provided with 10 articles that had already been coded and were then asked to code 20 more on their own. Another coder reviewed their discrepancies with the consensus and met to discuss the coding instructions. The process was repeated for another 20 articles. We then asked coders to code up to 50 articles at a time (depending on the number of already coded articles) until they reached at least 80% agreement with the consensus, after which they coded articles that had not been coded previously. There were no new coders for coding dominant narratives who came on board after the codes were established.

Disagreements were resolved by consensus or referred to one of the co-authors who was not involved in the coding of that article if consensus could not be reached. Consensus meetings were conducted regularly (after about every 10-200 articles, depending on the code) to maximize inter-rater reliability.

Percent agreement ranged from 89% to 99% for the definition of ACEs and from 94% to 98% for study purpose. We achieved 92% agreement for whether there was any information on funding, 84% to 95% agreement for type of funding, 94% agreement on whether there was a conflict of interest statement, and 97% agreement on whether a conflict of interest was declared. Agreement was 97% for the country of the sample and 88% for whether the study was affiliated with another study.

Articles were coded for dominant narratives primarily by three of the co-authors (CK, BG, or FC) and teams of two to three research assistants. The research assistants initially met with each other to arrive at a consensus decision for their team and then met with one of the co-authors to arrive at final consensus decisions. In cases where a coder coded the article as unclear, we did not calculate reliability for the other categories. Complete agreement for all roles

assigned to each category was 74% for Non-biological Measures of Health, 77% for Psychosocial Circumstances, 80% for ACEs, and ranged from 86% to 100% for all other categories.

For protective factors, we agreed on the presence/absence of a protective factor on 90% of the articles. For type of protective factor on articles for which we agreed on the presence of a protective factor, we reached 87% agreement for individual factors, 78% for family psychosocial factors, 97% for family socioeconomic factors, 80% for proximate community factors, 96% for societal factors, and 97% for whether it was unclear. Overall agreement for recommendations was 79%. We achieved 97% agreement on whether there was a recommendation for primordial/primary prevention, 83% agreement on whether there was a recommendation for secondary/tertiary prevention, 90% agreement on whether the researchers made both types of recommendations, and 88% agreement on whether the recommendation was too vague or nonexistent.

Data Extraction (Operational Definitions of the Codes)

Study purpose in relation to ACEs was relevant to the whole scoping review. The remainder of this section lists the other codes relevant to other parts of the review.

Study Purpose in Relation to ACEs

We coded the purpose of the study in relation to ACEs, as described by the authors in the Abstract and at the end of the Introduction as (1) descriptive (e.g., describing the prevalence of ACEs in a population) or observational/relational (examining associations between ACEs and other variables), (2) focused on issues related to screening for ACEs (e.g., feasibility and utility of screening at a clinical setting, screening for ACEs in large administrative data sets), (3) focused on interventions related to ACEs (reducing ACEs, mitigating their outcomes, or

examining whether ACEs moderated intervention effects), (4) focused on methodological issues (e.g., reliability and validity of assessment of ACEs, inter-relatedness of the items, validity of prospective versus retrospective recall of ACEs), (5) secondary (e.g., using ACEs as covariates, or as criteria for establishing the validity of another measure, or in an exploratory study where ACEs were not mentioned in the Introduction or the main objectives but were referred to in the Method section). A study could be coded as having more than one purpose; however, if it was coded as having ACEs as a secondary purpose, we did not code additional purposes.

The main goal for this code was to separate out articles in which ACEs played a primary role from articles in which they were incidental. For Parts 1 and 3, we analyzed all variables other than year and journal of publication (included in Part 1) only for articles in which ACEs were included in a primary purpose (94.1% of all coded articles). For Part 2, we further restricted analyses to studies where a descriptive/relational purpose was coded (89.5% of all coded articles).

Year and Journal of the Article

We coded the year of publication from the full text. If an article was only available electronically, we used the date it was electronically published. We also extracted the journal in which it was published.

Operationalization of ACEs

We categorized the types of adversities (e.g., physical abuse, household substance use) included in the measures in studies of "adversity" as defined above (see Table 1 in the paper). This categorization proved quite challenging due to variability across articles. For example, the 1998 article includes items on "household members" being mentally ill, using substances, or going to prison. However, across the articles we coded, "household member" sometimes

changed to "mother," "parents," or "caregivers." The operationalization of legal involvement was especially diverse, referring to a broad range of behaviors and constructs, from going to "jail" or "prison" to "antisocial behavior" and "criminal behavior" and any combination of these and related terms. In the end, we compromised between "lumping" (e.g., categorizing death or divorce of parents as one type of ACE) and "splitting" (e.g., counting parental death under one type and divorce under another) and coded moderately broad categories (e.g., coding as one type of ACE any item that referred to legal involvement or antisocial tendencies in anyone in the household).

Next, we coded each article according to which types of ACEs it included. When multiple types of adversity were included spanning different ages (e.g., childhood abuse prior to age 18 and adult intimate partner violence after age 18), we only included ACEs occurring prior to age 18. When it was not clear what an item was measuring (e.g., not having enough to eat), we tried to code it based on how the authors defined the item (e.g., as physical neglect or financial hardship). If the item did not match the authors' description (e.g., when items defined as physical neglect in the original ACEs questionnaire were referred to as "neglect" by the authors), we went with the authors' description. If we could not determine how to code an item, we coded it under "other adversities." If we could not identify all the ACEs included in a measure, or if it was unclear which items were based on ages 0 to 18, we coded ACEs as "unclear."

Funding Sources

We first coded dichotomously whether the article included a section on funding sources for the study. We then coded each source as (1) a government agency or a government-associated organization representing the public interest, (2) an academic institution or housed in an academic institution (e.g., an academic center), (3) a private organization, (4) a professional

society, (5) other or could not be specified. We coded sources within the same organization separately [e.g., if a study received funding from both the National Institute of Mental Health (NIMH), and the National Institute on Alcohol Abuse and Alcoholism (NIAAA), both of which are part of the National Institutes of Health (NIH), we coded these as two funders]. Many of the articles were based on other studies, and authors were not consistent in reporting funding sources for the current study versus the parent study. Therefore, we decided to code all funding sources included in their statement, whether these referred to the current or the parent study. However, we did not code funding sources for the authors (e.g., endowed chairs, fellowships, young investigator awards) if these were not labeled as funding sources for the study.

Conflicts of Interest

We coded whether there was any statement on conflict of interest and whether the authors declared any conflict of interest. Both were coded dichotomously as yes or no.

Country of the Sample

We coded the country or countries from which the sample was recruited.

Affiliation of the Study

We coded dichotomously whether the study was affiliated with another research or public surveillance study (i.e., whether the participants were recruited from another study or measures were drawn from a regularly administered survey such as the Behavioral Risk Factor Surveillance Survey (BRFSS) and specified its name if available. This information was typically available in the Participants or other sections of the Method or the end of the Introduction. This code was not included in the registered protocol.

Dominant Narratives

We coded research goals mentioned in the first part of the Abstract and the Introduction (which were almost always in the last one to two paragraphs). Coders read relevant sections of the rest of the articles only to gain context and to clarify the research goals (e.g., to decide whether researchers framed a variable as a health risk vs. a health outcome, understand the operationalization of the constructs, determine which variables were entered as predictors vs. outcomes in data analyses). We checked supplementary materials and original sources cited by the researchers when it was necessary to elucidate how they operationalized a construct. We did not code research questions mentioned in the Method or Results sections if they were not included in the first part of the Abstract or in the Introduction.

We deviated from the registered protocol in that we had originally planned to code study design in relation to ACEs and the role of socioeconomic status (SES) in the design. However, as reported in the first part, we found that the vast majority of the studies in this literature had descriptive/relational designs; thus, there was limited variation to code study designs. In addition, we decided to expand our focus from SES to additional social determinants of health inequities using the WHO framework (Solar & Irwin, 2010). Thus, we coded the roles played by different elements in the research design in relation to ACEs within the WHO framework and coded SES within this framework.

The modified WHO framework we used for coding is depicted in the paper. For each research goal, we coded (a) which role each element directly related to ACEs played in the research design (population of interest, grouping, predictor, outcome, bidirectional, moderator, mediator, or part of an SES composite); and (b) where each element fell into the categories in the WHO scheme, which we modified slightly for the purpose of this review.

We coded each goal separately (with an exception for mediational hypotheses, as explained later) but entered the information together in the relevant categories. For example, if one goal was to compare rates of ACEs between genders and a second goal was to determine whether ACEs predicted anxiety, we coded both outcome (for the first goal) and predictor (for the second goal) under ACEs.

If an element fit under multiple categories and each category was considered in the Introduction, we coded it under all relevant categories. For example, if the researchers studied Latino farmworkers and mentioned issues in the Introduction related to both Latinos and farmworkers, we coded under both Race/Ethnicity and Occupation. However, if they were studying outpatients with schizophrenia, but only focused on schizophrenia in the Introduction without mentioning the health system, we did not code anything under Health System. In cases where researchers referred to a broad construct but had a specific measure to define it, we coded the specific measure. For example, if "low income" was operationalized only as being eligible for Medicaid, we coded that construct under Health System, not Income.

We only coded elements related to ACEs, so no elements studied in parallel with or independent of ACEs were coded. For example, if a goal included both Education and ACEs as independent predictors of obesity, we did not code Education as predictor. However, if the goal included ACEs as a predictor and Education as a moderator of obesity or tested the joint or interactive effects of ACEs and Education, we did code Education.

We did not code experimental manipulations (e.g., where treatments were administered as part of a research study). We also did not code covariates unless the researchers examined them to test mediation and not just to rule out their effects.

WHO Framework.

Socioeconomic and Political Context.

(1) *Governance, Laws and Policies.* This category was coded when a research goal referred broadly to governance, laws and policies (i.e., macroeconomic, social and public).

(2) *Culture and Societal Values.* This category was coded when a goal referred to a cultural or societal factor. For example, if a study's sample was entirely rural, and the authors discussed unique cultural values in rural areas (e.g., stigma around mental health) in the Introduction, we coded population under Culture and Societal Values. It is not easy to differentiate a group from its cultural/societal values. In these cases, we examined how the researchers framed the question. For example, if they focused on the stigma, we coded it under Culture and Societal Values. If they focused on the fact that the sample was rural, we coded it under Geographic Location.

Structural Determinants of Health Inequities.

(1) *Social Class.* This category was coded when a research goal referred to social class (upper class, middle class, etc.) without further operationalizing the measure.

(2) *Wealth and Assets.* This category was coded when a research goal referred to measures of wealth and assets, such as home ownership.

(3) *Income.* This category was coded when a goal referred to income and related measures (e.g., income-to-poverty ratio).

(4) *Education.* This category was coded when a goal referred to education, including level of education attained, academic performance, attendance, and school

engagement. However, we coded expectations and attitudes about education under Psychosocial Circumstances.

(5) ***Occupation and Employment Status.*** This category was coded when a research goal referred to occupation or employment (including job performance and attendance), with clearly stated implications in the Introduction for income, prestige, or intrinsic exposure to health risks and health care. For example, if all participants were veterans, the researchers discussed trauma related to military service, and examined the relation between ACEs and PTSD among veterans, we coded population under Occupation.

(6) ***Gender and LGBTQ Status.*** This category was coded when a research goal referred to biological sex, gender identity, or LGBTQ status.

(7) ***Race/Ethnicity/Culture/Religion.*** This category was coded when a research goal referred to identity related to race, ethnicity, nationality, religion, or cultural group.

(8) ***Geographic Location.*** This category was coded when a research goal referred to a geographic location broader than the neighborhood and the researchers discussed distinguishing characteristics of that location in the Introduction. For example, if a goal was to examine the prevalence of ACEs in Argentina and the researchers discussed disparities in access to healthcare in Argentina in the Introduction, we coded population under Geographic Location. We did not code this category when researchers mentioned the location without mentioning unique characteristics of that location.

(9) *Composite Measures of SES.* Given the variability in how SES was defined, we coded Composite Measure of SES when a research goal referred to SES or a related term as a broad construct and used a combination of variables (e.g., income and occupation) to measure it, even if these variables were entered separately in the analyses. Composite measures were either coded as Well-differentiated, when all components clearly fell under one of our categories, or as Undifferentiated, when not all components were clearly operationalized or fit within our framework. For example, if a goal included SES as a moderator, and SES was operationalized as Income and Education, we coded moderator under Well-differentiated Composite and entered composite under Income and Education. If a goal included SES as a moderator, but it was not operationalized clearly or its operationalization included a measure that did not fit into our scheme (e.g., Education and "poverty"), we coded moderator under Undifferentiated Composite and entered composite only under Education.

We also examined the extent to which Intermediary Determinants of Health (e.g., eligibility for free school lunches) were used to define SES, because combining factors that determine SES (e.g., Income) with factors that result from SES (e.g., Receipt of Needs-based Social Welfare Services) may muddle understanding of determinants of health versus health inequities.

(10) *Other Social Determinants of Health Inequities.* This category was coded when a research goal referred to an additional social determinant of health inequities that did not fit within the categories outlined above. Examples included immigration status and general discrimination. Although age is not a social

determinant of health inequities, if a cohort was studied because of its collective experiences related to health inequities (e.g., exposure to war), we coded this under Other Social Determinants.

Intermediary determinants of health: Material circumstances.

(1) ***Receipt of Needs-Based Social Welfare Services.*** This category was coded when a research goal referred to any welfare program providing needs-based assistance.

(2) ***Housing Instability.*** This category was coded when a research goal referred to housing instability, homelessness and/or material characteristics of the home environment.

(3) ***Hunger and Lack of Access to Adequate Food.*** This category was coded when a research goal referred to hunger and/or lack of access to adequate food.

(4) ***Neighborhood Conditions.*** This category was coded when a research goal referred to conditions in the immediate neighborhood, including poverty and disadvantage, as well as neighborhood resources. Neighborhood included the immediate neighborhood and relatively small geographic units (e.g., small towns) where people typically live, go to school, and work. Larger geographic units, such as large metropolitan areas, states or countries were coded under Geographic Location.

(5) ***Other Material Determinants.*** This category was coded when a research goal referred to additional material circumstances that did not fit within the categories outlined above.

Intermediary Determinants of Health: Psychosocial Factors.

(1) ***Composition of the Household.*** This category was coded when a research goal referred to the composition of the household, with implied or explicitly stated consequences for SES (e.g., marital status, family size).

(2) ***Family Psychosocial Characteristics.*** This category was coded when a research goal referred to psychosocial characteristics of family members in relation to measures collected on their offspring. For example, if a goal included parental history of ACEs as a predictor of children's ACEs, we coded predictor under Family Psychosocial Characteristics and outcome under ACEs. This category included parental illness, family history of genetic illnesses, parents' coping styles, parent-child relationship, and breastfeeding. However, it did not include characteristics of family members related to the Social Determinants of Health, such as parental education or immigration status.

(3) ***Parenthood.*** This category was coded when a research goal included measures collected on pregnant women or caregivers, and researchers discussed the unique characteristics of parenthood in the Introduction. For example, when a goal consisted of examining the relationship between ACEs and depression in pregnant women, we coded population under Parenthood. However, if researchers studied the relationship between parental ACEs and children's depression, we coded predictor under Family Psychosocial Characteristics and outcome under Non-biological Measures of Health.

(4) ***ACEs.*** This category was coded when the research goal referred to the variables operationalized as "adversity" in the study, including specific types of ACEs as well as cumulative scores. Any characteristic of ACEs as defined by the

researchers fell under this category, including the timing or clustering of ACEs. If another category in our coding scheme was included in the researchers' own definition of ACEs (e.g., homelessness), we did not code that category separately. Operationalization of ACEs was coded in the first part of the review, and these codes were available to the current coders.

(5) ***Psychosocial Circumstances.*** This category was coded when a research goal referred to psychosocial circumstances related to stress and coping, such as stressors and traumatic experiences, life stress, perpetration or being a victim of interpersonal violence, social support, coping styles, resilience, health locus of control and health literacy, attitudes towards health-care utilization, and psychological or personality characteristics (e.g., shame, temperament). However, if a goal referred to family relationships as a source of stress or support, we coded it under Family Psychosocial Characteristics.

(6) ***Other Psychosocial Factors.*** This category was coded when a research goal referred to additional intermediary determinants that did not fit within the categories outlined above.

Intermediary Determinants of Health: Behaviors and Biological Factors.

(1) ***Health-Risk and Health-Promotive Behaviors.*** This category was coded when a research goal referred to health-risk (e.g., substance use, disordered eating, sleep difficulties, cognitive endophenotypes of disorders) or health-promotive behaviors (e.g., physical activity, healthy diet), and researchers justified their purpose of studying the variable because of its risk for, or protection against, a future condition (e.g., cardiovascular disease). For example, if a goal included a measure

of substance use or suicidality, we coded it under this category if the researchers framed the measure as a risk for future health problems or mortality, but under Non-biological Measures of Health if they framed it as an end in itself (e.g., suicidal thinking as a negative outcome in itself, not as a risk for mortality).

(2) *Genetic Factors.* This category was coded when a research goal referred to static genetic factors, such as variants of certain alleles and genetic risk for disorders. However, we coded family history of genetic disorders under Family Psychosocial Characteristics.

(3) *Biomarkers of Risk for Physical Health Conditions.* This category was coded when a research goal referred to biomarkers framed by the researchers as indices of risk for future physical health conditions, such as allostatic load, blood pressure, neurological functioning, functioning of the HPA axis, insulin resistance, pubertal status, cortisol levels, telomere length, methylation, and BMI.

Intermediary Determinants of Health: Systems.

(1) *Health System.* This category was coded when a research goal referred to characteristics of the health system, access to the health system, receipt of interventions and prescription medications, and other measures of utilization or inadequate utilization of the health system. This did not include experimental interventions implemented by the researchers as a part of the study.

(2) *Child Welfare System.* This category was coded when a research goal referred to the child welfare system, including all types of contact with the child welfare system (e.g., CPS referrals, foster care, adoption). However, when the outcome

included placement in the child welfare system, we coded this under Non-

biological Measures of Health.

Measures of Health and Well-Being.

(1) ***Biological Measures of Health Conditions and Physical Health.*** This category

was coded when a research goal referred to biological measures of health

conditions and physical health, and researchers framed these measures as

outcomes in and of themselves, not as risk factors. This category included (a)

measures of general physical health, quality of life related to physical health,

diagnoses or self-reports of specific physical conditions, that is, conditions not

classified as "mental or behavioral disorders" in ICD-10 (World Health

Organization, 1992); (b) symptom trajectory and measures of disease treatment or

management of physical health conditions (e.g., viral load in individuals with

HIV); (c) biological measures of mental or physical health that were not framed

as a risk factor for future problems but instead as a correlate of existing conditions

(e.g., hippocampal volume as a correlate of PTSD, preterm birth as a physical

health outcome).

(2) ***Non-Biological Measures of Health, Well-Being, and Social Functioning.*** This

category was coded when a research goal referred to non-biological measures of

health and well-being and framed these measures as outcomes and not as risk

factors. This category included (a) diagnoses and self-reports of mental health

conditions included in ICD-10 under Mental and Behavioral Disorders (World

Health Organization, 1992); (b) internalizing and externalizing symptomatology

and behavior problems; (c) symptom trajectory or other measures of disease

treatment or management of mental health conditions; (d) suicidal ideation and attempts, as well as non-suicidal self-injury if these were framed as outcomes and not as risks; (e) developmental delays in mental or cognitive functioning; (f) teenage pregnancy, and induced or elective abortions, if these were framed in the context of mental health; (g) involvement in the child welfare system as an outcome; (h) self-reports of delinquency or crimes with no mention of actual involvement with the justice system; (i) any self-reported measures of general health, well-being, functional limitations, and quality of life that did not differentiate between mental and physical health.

(3) ***Mortality.*** This category was coded when a goal referred to mortality, including general or premature mortality and completed suicides.

Involvement with the Justice System. Although the justice system is not included in the WHO framework, we included it in the modified framework because it is relevant to research on ACEs. Unlike the Health and Child Welfare Systems, it is not designed to improve health and well-being outcomes. Thus, we included it as a separate category along with Measures of Health. We coded research goals referring to involvement with the justice system, excluding cases when this was included in the researchers' definition of ACEs. This category did not include Antisocial Personality Disorder or measures of aggression or engagement in illegal or delinquent behavior without a clear reference to involvement with the justice system. For example, if researchers examined self-reports of engaging in theft but it was not clear whether the participants engaged with the justice system as a result, we coded this under Non-biological Measures of Health.

***Unclear or Inconsistent Research Goals*.** This category was coded when research goals stated in the Abstract and Introduction were unclear or inconsistent with one another or with the Methods or data analyses. In these instances, no other variables were coded.

Variable Roles. Table 1 below lists the definitions of population of interest, predictor, grouping, moderator, mediator, and outcome variables, examples of each, proposed roles in a research narrative, implications of these variable roles, and title of a sample article for each that exemplifies the role. Table 2 lists the potential plots we examined in the literature.

Table 1

Variable Codes, Examples, Proposed Roles in the Narrative, Implications, and Sample Articles

What was coded?	Operational definition	Example	Proposed roles as narrative elements	Why does the code matter?	Title and authors of sample article for the code

| Predictor | When a variable in the research design had at least two levels and (a) was entered as a predictor in statistical analyses; or (b) was described by the authors as a "predictor", or "cause of" or "risk for" another variable, or as one variable having an "effect" on another; or (c) when one variable clearly preceded another (e.g., ACEs vs. employment rates in adulthood) | A research goal is to test whether ACEs predict depression

(ACEs = *predictor*) | Beginning of the story; potential setting, victims, heroes, or villains | What is viewed as the root of the problem? To what extent does the story start with Social Determinants? | "Cumulative exposure to childhood stressors and subsequent psychological distress. An analysis of US panel data" (Björkenstam et al., 2015)

(ACEs = *predictor)* |

Population	When researchers expressed interest in a population in the Introduction and the sample for a research goal was drawn only from that population. The researchers must have provided a rationale for choosing their sample or have discussed the sample's relevant characteristics in the Introduction.	A research goal is to ascertain the prevalence of ACEs in Bolivia, the sample consists of individuals from Bolivia, and the researchers discuss at least some distinctive characteristics of Bolivia in the Introduction	Whose ACEs are being examined; potential setting, victims	Whose adversities get counted? Implications for interventions tailored for specific target populations?	"Adverse Childhood Experiences Among Direct Support Professionals" (Keesler, 2018)
	Population was also coded if a goal was to compare the prevalence of ACEs in the sample to previously published rates in the general population.				(Occupation = *population*)
		(Geographic Location = *population*)			
	Population was *not* coded if a sample was drawn from a population but the authors did not mention the population in the Introduction other than stating that the sample was drawn from that population.				

| Grouping | When a research goal included examining the prevalence of ACEs across different groups included in the study. | A research goal is to compare the prevalence of ACEs between men and women (Gender/LGBTQ status = *grouping*) or between patients with fibromyalgia and controls (Biological Measures of Health = *grouping*).

A sample consists solely of individuals in the military, but a research goal was to compare ACEs between those exposed to combat and those who were not exposed (Occupation = *grouping*). | Whose ACEs are being compared; victims | Are there disparities in ACEs between groups? To what extent are Social Determinants considered as influencers of the unequal distribution of ACEs (i.e., an Intermediary Determinant of Health)? Implications for interventions tailored for specific target populations? | "Comparing the rates of early childhood victimization across sexual orientations: Heterosexual, lesbian, gay, bisexual, and mostly heterosexual" (Zou & Andersen, 2015)

(Gender/LGBTQ status = g*rouping*) |

Moderator	When researchers conducted formal moderation analyses, or when a research goal was to examine differences in predictor-outcome relationships across subgroups	A research goal is to test whether gender moderates the relationship between ACEs and depression (Gender/LGBTQ status = *moderator)*	Agent or factor affecting chain of events in the story; potential hero or villain	Which factors may make a difference in vulnerability to negative outcomes? Potential entry points for interventions?	"The Racial and Gender Differences in the Impact of Adverse Childhood Experiences on Juvenile Residential Placement" (Zettler et al., 2018) (Gender/LGBTQ status and Race/Ethnicity/Culture/Religion = *moderator*)

Mediator	When researchers conducted formal mediation analyses, or when a research goal included how a variable might explain the nature of the relationship between other variables.	A research goal is to test whether coping skills mediate the relationship between ACEs and depression	Mechanism of action in the story	How do negative outcomes come about? To what extent are Intermediary Determinants treated as *mediators* of the relationship between Social Determinants and Health? Potential entry points for interventions?	"Relationship Distress as a Mediator of Adverse Childhood Experiences and Health: Implications for Clinical Practice with Economically Vulnerable Racial and Ethnic Minorities" (Wheeler et al., 2019)
	If the researchers' main goal was a mediation analysis (e.g., X mediates the relationship between A and B), but they stated the steps as separate goals (e.g., one goal was to test whether A was related to X and another was whether X was related to B), we based the coding on the whole analysis (A as *predictor*, B as *outcome*, and X as the *mediator*) and did not code the subgoals separately.	(Psychosocial Circumstances = *mediator*)			(Psychosocial Circumstances = *mediator*)
	In a small number of cases, mediational hypotheses were based on testing the significance of covariates; we coded these covariates as *mediators* if the authors made it clear that they were testing these covariates as potential mediators.				

| Outcome | Coded similarly to *predictor*; but considered the variable that succeeds the *predictor*.

Outcome was also coded under ACEs when a research goal was to examine the prevalence of ACEs, the effects of developmental timing of ACEs, or associations among ACEs (e.g., using Latent Class Analysis).

However, when clusters of ACEs were used to predict other outcomes, we coded ACEs as *predictor*. | A research goal is to test whether ACEs predict depression (Non-Biological Measures of Health = *outcome)* or to examine the prevalence of ACEs in South Carolina (ACEs = *outcome*) | Ending of the story; potential victims | How does the story end? To what extent are Social Determinants included in the ending? Implications for social mobility and intergenerational transmission of determinants of health inequities? | "Children in foster care: adverse childhood experiences and psychiatric diagnoses" (Jamora et al., 2009)

(Non-biological Measures of Health = *outcome*) |

| Bidirectional | Coded when a research goal entailed correlational analyses or chi-square tests, and it was impossible to discern which variables the researchers viewed as the predictor versus the outcome or which variable preceded the other (e.g., ACEs and mental disorders in children). | A research goal is to test whether ACEs and ADHD symptoms are correlated with each other in children

(ACEs and Non-Biological Measures of Health = *bidirectional*) | Unclear | To account for cases when it was not clear what was a *predictors* and what was an *outcome* | "The associations of cumulative adverse childhood experiences and irritability with mental disorders in detained male adolescent offenders" (Bielas et al., 2016)

(ACEs and Psychosocial Circumstances = *bidirectional*) |

Table 2

Plots Examined Through Analysis of the Research Goals

Type of Plot	Proposed Plot Elements	Title and Authors of A Sample Article and Elements That Fit into the Plot
Story of Decline	predictor→ outcome (AND no moderators or mediators)	"The differential impact of adverse childhood experiences in the development of pre-diabetes in a longitudinal cohort of US adults" (Campbell et al., 2018)
		(ACEs = predictor; Biomarkers of Risk for Physical Health Conditions = outcome)

Story of Helplessness and Control	predictor AND moderator or mediator → outcome	"Pathways from Childhood Adversity to Problem Behaviors in Young Adulthood: The Mediating Role of Adolescents' Future Expectations" (Brumley et al., 2017)
		(ACEs = predictor; Psychosocial Circumstances = mediator; Non-biological Measures of Health = outcome
The WHO Story (Social Determinants of Health Inequities → Intermediary Determinants of Health→ health inequities)	(a) at least one Social Determinant as a predictor, moderator, grouping variable, or population, AND (b) at least one Intermediary Determinant (including ACEs) as a mediator or outcome.	"Adverse childhood experiences among Latina women and its association with mastery of stress and health" (Alvarez et al., 2019) (Gender/LGBTQ status and Race/Ethnicity/Culture/Religion = population; ACEs = predictor; Psychosocial Circumstances = mediator; Biological and Non-Biological Measures of Health = outcome)

The Intermediary Determinant Story (Intermediary Determinants → health)	(a) at least one Intermediary Determinant (including ACEs) as a predictor, moderator, grouping variable, or population, AND (b) at least one Intermediary Determinant or Measure of Health (Biological or Non-Biological Measures or Mortality) as an outcome	"Adverse Childhood Experiences Are Associated with Unmet Healthcare Needs among Children with Autism Spectrum Disorder" (Berg et al., 2018) (ACEs = predictor; Health System = outcome; Non-Biological Measures of Health = population)
The Intergenerational Transmission Story (Intermediary Determinants → Social Determinants)	(a) any Intermediary Determinant (including ACEs and Family Psychosocial Characteristics) as a predictor AND (b) any Social Determinant as an outcome.	"Chronic School Absenteeism and the Role of Adverse Childhood Experiences" (Stempel et al., 2017) (ACEs = predictor; Education = outcome)

Protective Factors

We coded as "yes," "no," or "unclear" whether any protective factors for ACEs, defined as such by the authors, were intentionally incorporated into the study design, regardless of the results. We determined how variables were defined and treated in the design through reading the Title, the Abstract and the research questions at the end of the Introduction and checking the Method and Results sections and the rest of the Introduction as necessary. We considered a variable a protective factor if it was explicitly treated in the design as a factor that prevents the emergence of ACEs or works as a buffer against the harmful effects of ACEs. These variables were commonly treated as moderators in the research designs. We also coded protective factors in studies in which researchers investigated the potentially beneficial effect of a variable on an outcome alongside ACEs, although the variable was not treated as a moderator in the analyses. When in doubt (e.g., for gene variants that could be either risk or protective factors), we examined how the variable was framed by the authors in the Introduction.

We did not code interventions conducted by the researchers as protective factors (these were coded under Study Purpose as Intervention studies and are summarized in part 1). We also did not code a protective factor for studies in which researchers examined a potentially protective factor as a negative outcome of ACEs (e.g., to test whether ACEs were associated with worse scores on a resilience measure), as a mediator of the relationship between ACEs and poor outcomes (e.g., to examine whether lack of social support mediated the relationship between ACEs and poor health outcomes) or as a covariate whose effects were meant to be ruled out (e.g., maternal education as a covariate in examining the relationship between ACEs and health). We did not code a variable as a protective factor if the researchers did not refer to it as such in the Introduction but engaged in post-hoc speculation about its protective effects in the Discussion.

Finally, we did not code protective factors that were alluded to in the text of the articles but were not explicitly incorporated into the research design.

Once protective factors were identified, we coded whether each protective factor included in the research design was at the level of the (1) individual (e.g., genes, gender, personality traits, resilience, coping skills), (2) psychosocial characteristics of the family (e.g., attachment; quality of relationships with the family of origin or with the family formed as an adult), (3) socioeconomic characteristics of the family (e.g., assets, income, parental education), (4) proximate community (e.g., social support from peers, schools, community-based organizations, sense of community, neighborhood resources, support from health care providers), (5) broader society (e.g., policies and laws). When in doubt, we examined how the authors framed the variable (e.g., whether connection to nature was framed as an individual preference or as a resource that could be provided by the community). A study could be coded as having protective factors at more than one level (e.g., if protective factors were defined as resilience and peer support).

Researchers' Recommendations for Translating Findings into Action

We examined the end of the Abstract and the Discussion sections of the articles to identify researchers' recommendations regarding ACEs or for populations who were found in the study to have high rates of ACEs, even when the researchers only referred to other articles in which there were specific recommendations. When in doubt, we read through the rest of the Discussion and the article to determine whether a sentence referred to any actionable recommendation. The codes were based on how the authors framed the issues and defined ACEs. For example, if the authors treated peer victimization as an outcome of ACEs (whose definition in that study did not include peer victimization) and made recommendations to prevent peer

victimization, we coded this as secondary/tertiary in relation to ACEs. However, if they included peer victimization in their list of ACEs and made recommendations to prevent peer victimization, we coded this under primordial/primary prevention.

We coded recommendations couched in tentative terms (e.g., parenting programs may be beneficial; parenting programs may be beneficial, but more research is needed to determine whether they may be more beneficial for mothers or fathers) as long as the researchers did not seem to indicate that the effectiveness of recommendation was so doubtful that more research was needed before any action could be taken (e.g., parenting programs may be beneficial, but more research is needed to determine if they are). We did not code recommendations for future research because the goal of the current study was to focus on recommendations for translation of research to action, as opposed to more research.

Although our initial goal was to create relatively fine-grained categories of recommendations (e.g., screening; increasing public awareness; training providers; population-wide surveillance; reducing disparities; categorizing recommendations based on who the target population was or who the individuals were who were tasked with carrying out the recommendations; recommendations for translation to action that were mentioned but were too vague), we were unable to make these differentiations with adequate inter-rater reliability. Recommendations were instead categorized based on type of prevention (primordial, primary, secondary, tertiary). However, it was difficult to reliably determine clear boundaries between these. The distinction between primordial and primary prevention can depend on researchers' (often implicit or unexamined) causal theories about the origins of ACEs. The line between primary and secondary prevention is similarly blurry, as there could be cases where secondary prevention strategies could lead to primary prevention of ACEs (e.g., a recommendation to treat

mental disorders stemming from ACEs could also be viewed implicitly as a primary prevention

strategy for the next generation). It is also not always clear when a recommendation is for

secondary prevention to avoid downstream effects as opposed to tertiary prevention to mitigate

effects that have already occurred. Thus, we combined across these categories and grouped

recommendations into the following four categories:

(1) Primary/Primordial Interventions. For this category, we coded whether the

researchers recommended an action that would prevent or reduce the occurrence of ACEs (as

they defined them), whether through changing structural conditions that lead to ACEs (e.g.,

socioeconomic conditions, laws, policies, cultural norms) or by reducing the occurrence of ACEs

without necessarily changing structural conditions (e.g., home visiting programs focused on

high-risk groups). We included in this category recommendations for parents or parents-to-be

that explicitly mentioned preventing or minimizing ACEs in their current or future children.

(2) Secondary/Tertiary Interventions. For this category, we coded whether the

researchers recommended interventions aimed at preventing negative outcomes after exposure to

ACEs (e.g., screening for ACEs; intervening with mechanisms that lead from ACE exposure to

poor outcomes, such as resilience-based interventions) or preventing worsening of outcomes

after exposure to ACEs (e.g., implementing trauma-informed policies and educating service

providers about ACEs; improving practice to address the needs of ACE-exposed individuals who

have developed disorders). We also included in this category large-scale surveillance, such as

state-wide or country-wide surveys of ACEs.

(3) Both (1) and (2).

(4) Recommendations That Were Too Broad or Vague To Be Coded Under the

Previous Categories, Or No Recommendations Regarding ACEs. We coded a recommendation

as too vague if (a) it was not clear whether the researchers were recommending primordial/primary or secondary/tertiary prevention (e.g., minimizing the burden of ACEs; "multifaceted solutions" to address ACEs), (b) the statement did not provide any specificity or point practitioners or law-/policy-makers in one direction over another regarding next steps (e.g., therapy vs. therapy to target feelings of helplessness; early intervention vs. early intervention for emotion regulation; reduce parental stress levels vs. parenting skills programs), (c) the researchers pointed out problems without proposing an actionable recommendation for what to do about them (e.g., children living in stressful environments are more prone to ACEs; hospitals not screening for ACEs), (d) the researchers indicated that an action might be beneficial, but added that more research was needed to determine whether it was actually beneficial or the extent of its benefits.

To code an article as including a recommendation, we further required that researchers specify either the target of the intervention (e.g., children suffering from trauma) or whose responsibility it was to act on the recommendation (e.g., health care providers, therapists). Thus, recommendations such as preventing ACEs in society in general, "psychosocial interventions to reduce early adversities," or stating that "family bonding is a protective factor," or "advocacy for children" were coded as vague.

References

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Eugenio, B. D., & Glass, M. (2004). The Kappa statistic: A second look. *Computational Linguistics*, *30*(1), 95–101. https://doi.org/10.1162/089120104773633402

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: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine*, *14*(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1), 69. https://doi.org/10.1186/1748-5908-5-69

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 1–10.

Sim, J., & Wright, C. C. (2005). The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257–268. https://doi.org/10.1093/ptj/85.3.257

Solar, O., & Irwin, A. (2010). *A conceptual framework for action on the social determinants of health* [(Discussion Paper Series on Social Determinants of Health, 2)]. World Health Organization.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., & Weeks, L. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines* (Vol. 1). World Health Organization.