



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

School-based child sexual abuse interventions

Citation for published version:

Lu, M, Barlow, J, Meinck, F, Walsh, K & Wu, Y 2022, 'School-based child sexual abuse interventions: A systematic review and meta-analysis', *Research on Social Work Practice*.
<https://doi.org/10.1177/10497315221111393>

Digital Object Identifier (DOI):

[10.1177/10497315221111393](https://doi.org/10.1177/10497315221111393)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Research on Social Work Practice

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



School-based Child Sexual Abuse Interventions: A Systematic Review and Meta-analysis

Research on Social Work Practice
2022, Vol. 0(0) 1–23
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10497315221111393
journals.sagepub.com/home/rsw



Mengyao Lu¹ , Jane Barlow¹, Franziska Meinck^{2,3,4}, Kerryann Walsh⁵, and Yumeng Wu⁶

Abstract

Purpose: The purposes of this systematic review were to systematically summarize components in existing school-based child sexual abuse (CSA) prevention programs and identify predictors for program effectiveness.

Method: Building upon the most comprehensive systematic review on this topic, we conducted systematic searches in both English-language from September 2014 to October 2020 and Chinese-language from inception to October, 2020. Meta-regressions were performed to identify predictors for program effectiveness.

Results: Thirty-one studies were included with a total sample size of 9049 participants. Results from meta-analyses suggested that interventions are effective in increasing participants' CSA knowledge as assessed via questionnaires ($g = 0.72$, 95% CI [0.52–0.93]) and vignette-based measures ($g = 0.55$, 95% CI [0.35–0.74]). Results from meta-regression suggested that interventions with more than three sessions are more effective than interventions with fewer sessions. Interventions appear to be more effective with children who are 8 years and older than younger children.

Discussion: CSA is a global issue that has significant negative effects on victims' physical, psychological, and sexual well-being. Our findings also provide recommendations for future research, particularly in terms of optimizing the effectiveness of school-based CSA prevention programs, and the better reporting of intervention components as well as participant characteristics.

Keywords

child sexual abuse, child sexual abuse prevention, school-based interventions

Introduction

Child sexual abuse (CSA) refers to “the involvement of a child in sexual activity that he or she does not fully comprehend, is unable to give informed consent to, or for which the child is not developmentally prepared and cannot give consent, or that violates the laws or social taboos of society” (WHO, 2006, p.62). CSA is a serious problem for children worldwide. The most recent meta-analysis of 55 studies from 24 countries showed that the prevalence of CSA (defined as non-contact abuse, contact abuse, forced intercourse, and mixed sexual abuse) ranged from 8% to 31% for girls and 3% to 17% for boys (Barth, Bermetz, Heim, Trelle, & Tonia, 2013). CSA is associated with a range of adverse psychological, sexual, and economic consequences. For example, one meta-analysis found a significant relationship ($d = .32$ to $d = .67$) between CSA and psychological distress (e.g., anger, anxiety, and depression) as well as dysfunction in adult women (e.g., revictimization, self-mutilation, sexual problems, substance abuse, and suicidality; Manglio, 2009). The global costs

related to physical, psychological, and sexual violence against children are estimated to be between 3% and 8% of the global gross domestic product (GDP) (Perezniето, Montes, Routier, & Langston, 2014). As home to the world's second-largest child population (UNICEF, 2014), China experiences a significant burden in terms of CSA. Findings from East Asian and Pacific regions indicate that the economic costs of CSA are substantial, with one study estimating the value of Disability-Adjusted Life Years (DALYs) lost due to CSA at \$18,378.3

¹University of Oxford, UK

²University of Edinburgh, UK

³North-West University, South Africa

⁴University of the Witwatersrand, South Africa

⁵Queensland University of Technology, Australia

⁶Memorial Sloan Kettering Cancer Center, USA

Corresponding Author:

Mengyao Lu, Department of Social Policy and Intervention, University of Oxford, 32 Wellington Square, Oxford, OX1 2ER, UK.

Email: mengyao.lu@spi.ox.ac.uk

million which accounts for 0.39% of the Gross Domestic Product (GDP) in China alone (Fang et al., 2015).

School-based universal interventions for preventing CSA are variously designed to provide students with skills that can be used to recognize, react, and report abuse (Finkelhor, 2009) and thereby reduce the occurrence and re-occurrence of sexual abuse in children and adolescents (National Institute of Justice, 2012). Most school-based personal safety training programs share common goals, including teaching children about body integrity, body safety, appropriate and inappropriate touching, gender roles, and private body parts (Bronson, 2019). In acknowledgment of the limited power that children have themselves to prevent experiencing CSA, these programs were developed to *empower* children by enhancing their *knowledge* about CSA and personal safety and developing their *competence* in self-protective behaviors, thereby helping them adopt strategies to avoid potentially unsafe situations and disclose their experiences to protective adults (Kim & Kang, 2017). In general, these interventions can be classified into two categories (Kopp & Miltenberger, 2009): (a) information-based trainings (IBT) in which information is delivered to children by an instructor in class using materials such as video tapes, plays, or activity books; and (b) behavioral skills trainings (BST) in which similar information is delivered but is taught using modeling, active rehearsals, and application of self-protection knowledge and skills to problem-based (i.e., “what if...”) scenarios.

Reviews of the literature over almost three decades, consistently show that children who have received school-based CSA interventions show greater gains in knowledge about CSA prevention concepts than children who have not received such programs (e.g., Kenny, Capri, Ryan, & Runyon, 2008; MacMillan, MacMillan, Offord, Griffith, & MacMillan, 1994; Walsh, Zwi, Woolfenden, & Shlonsky, 2015). One gold standard systematic review of 24 studies with 5802 participants in primary (elementary) and secondary (high) schools in Canada, China, Germany, Spain, Taiwan, Turkey, and the United States found that school-based CSA interventions were effective in increasing participants’ knowledge of CSA and self-protective skills and did so without increasing or decreasing their fear or anxiety (Walsh et al., 2015).

To develop and deliver effective school-based CSA interventions, program developers, practitioners, school systems, and policymakers need to understand what works, for whom, and how intervention effects can be optimized. By identifying and integrating components associated with greater effectiveness, interventions can be strengthened and resource wastage associated with implementation of sub-optimal programs can be reduced (Kaminski, Valle, Filene, & Boyle, 2008). This is particularly important when delivering such interventions in low and middle-income countries (LMICs) facing economic hardship, severe resource restrictions, and service gaps. To date, however, despite a growing body of research on school-based CSA interventions, there has been little attempt to systematically identify the content and

structural intervention components which may contribute to the effectiveness of these programs.

To address this research gap, this systematic review, for the first time in this field, will use Intervention Component Analysis (ICA; Sutcliffe, Thomas, Stokes, Hinds, & Bangpan, 2015) to systematically identify intervention components in existing school-based CSA interventions and establish which component(s) are associated with improved outcomes for children who participated. To this end, advancing on the substantial work undertaken by Walsh et al. (2015), the aims of this review are to: (a) map the intervention components of school-based CSA interventions that are aimed at improving children’s knowledge and behaviors in relation to CSA prevention; and (b) identify which intervention component(s) appear(s) to contribute to the most significant improvements in knowledge and behavior outcomes for participating children.

Method

To meet the above aims, the following research methods were employed: (a) a systematic review and meta-analysis of overall program effectiveness to update the most recent and comprehensive review of school-based CSA intervention programs (Walsh et al., 2015) by identifying studies published between September 2014 and October 2020, and widening the inclusion criteria to include studies published in the Chinese-language databases from inception to October 2020; (b) an intervention component analysis to classify and synthesize intervention components in existing school-based CSA interventions; and (c) a meta-regression to identify the intervention components contributing to program effectiveness.

Data Sources, Inclusion Criteria, and Study Identification

New evidence should be added to existing systematic reviews to ensure their currency, and updating existing systematic reviews is more efficient than developing a new protocol to address the same research question (Garner et al., 2016). For example, the Cochrane Review policy recommends that systematic reviews should be updated within 2 years to ensure the up-to-dateness of reviews. This also avoids potential resource wastage or misleading conclusions (Garner, et al., 2016). Elliott et al. (2017) introduced the concept of “living systematic reviews, which are a novel approach to continually updating the reviews and incorporating new evidence when it becomes available” (p.24). To this end, given the consistently evolving evidence relating to school-based CSA prevention programs, in order to keep the evidence as up-to-date as possible, we updated and advanced on the substantial work undertaken by Walsh et al. (2015), and replicated the literature search strategy and screening methods used.

Following Walsh et al.’s (2015) inclusion and exclusion criteria, this review included only studies that had utilized the

most rigorous evaluation designs, randomized controlled trials (RCTs), cluster-RCTs, or quasi-RCTs in which participants were allocated to an intervention or control group using explicit methods such as using day of the week, alphabetical order, or other sequential allocation such as by class or school (Walsh et al., 2015). The review included only studies of school-based CSA intervention programs that focused on improving knowledge of CSA and CSA prevention concepts, or skill acquisition in self-protective behaviors. Included studies were undertaken only with students aged five to 18 years, who were enrolled in primary or secondary schools.

A clear limitation of the Walsh et al.'s (2015) review was that it included mostly papers published in English for the period up to and including September 2014. The search strategy for the current review attempted to also capture studies from some of the world's most populous regions by searching for both English-language and Chinese-language papers because there is a developing body of research on CSA prevention in China and the language skills were available to the review team. English-language databases were searched for the period September 2014 to October 2020. Chinese-language databases were searched from inception to October 2020. In all, twenty-three electronic databases were searched to identify studies meeting the pre-defined inclusion criteria.

Eighteen English-language databases were searched: Applied Social Sciences Index and Abstracts, British Education Index, Campbell Collaboration Library, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Library, CINAHL, ClinicalTrials.gov, DARE, ERIC, EMBASE, Global Health, ICTRP: apps.who.int/trialsearch/, Medline, ProQuest Dissertation, PsycINFO, Social Sciences Citation Index, Social Services Abstracts (ProQuest), and Sociological Abstracts. Five Chinese-language databases were searched: CNKI Academic Journals Full-text Database, China Doctoral Dissertations Full-text Database, China Masters' Theses Full-text Database, Taiwan Electronic Periodical Services, and Wanfang Database. Customized search strategies were developed for each database as search terminology varied across databases. The search strategies are available upon request. We also consulted with a subject librarian in the searching process. Reference lists of all included studies and previous systematic reviews were hand-searched for eligible studies not identified in database searches. Thirteen authors of included studies were contacted via email to identify further ongoing or unpublished trials, of which nine replied with information resulting in the identification of one additional study.

To reduce the potential for missed studies, it is recommended that two or more reviewers undertake the screening process (Stoll et al., 2019). As such, the titles and abstracts of records identified in the database searches were screened against inclusion and exclusion criteria by the first author (ML) and double checked by a second reviewer (YW), both of whom are native Mandarin Chinese speakers and fluent in English. Those studies considered to be not relevant were excluded. Disagreements in judgments between the two

reviewers were resolved by discussion, aiming for consensus. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009) was used in all stages of the review process. The systematic review protocol was registered with PROSPERO (ID: CRD 42019140564).

Data Extraction

Data were extracted to align with the three study methods as noted above: (a) meta-analysis to update evidence on overall program effectiveness; (b) intervention component analysis to classify intervention components; and (c) meta-regression to identify intervention components contributing to program effectiveness.

Two reviewers (ML and YW), working independently, extracted all data by coding each paper line-by-line to ensure that all available information was captured. Where more than one publication referred to the same study or dataset, we used the relevant paper with the most comprehensive and clear description of methods and outcome reports from a peer-reviewed journal as the primary reference in accordance with systematic review guidelines (Higgins et al., 2011).

We extracted characteristics of the included studies including authors, study design, country, intervention outcome measures, and participant age. We extracted intervention outcome data for the following outcomes: knowledge of CSA and/or CSA prevention concepts typically measured via questionnaires or vignettes. Consistent with Walsh et al. (2015) review, we extracted questionnaire and vignette measures as separate knowledge outcomes because these required different methods of administration and measured different domains of children's knowledge: factual knowledge and applied knowledge (Walsh et al., 2015).

We extracted two categories of intervention components: structural components (e.g., mode of delivery, location, and provider) and content components (e.g., the information delivered during sessions). For structural components, we adopted the Template for Intervention Description and Replication (TIDieR) checklist (Hoffmann et al., 2014). The purpose of the TIDieR checklist is to ensure that interventions are described in sufficient detail to enable replication (Hoffmann et al., 2014). The TIDieR checklist is an official extension of the CONSORT (Item 5; Moher et al., 2010) and SPIRIT statements (Item 11; Moher & Chan, 2014). We extracted the following TIDieR checklist intervention characteristics from study reports: why (rationale of the intervention), what (materials used in the intervention), who (intervention providers), how (modes of delivery), where (location of intervention), when and how much (duration and number of sessions), tailoring (the adaptation of the intervention during the study), modification (the modification of the intervention during the study), and how well (fidelity assessment of the intervention during the study). For content components we identified and inductively coded the themes or

concepts addressed in the programs (Sutcliffe et al., 2015) and then summarized and categorized these components into ten dimensions. We contacted one study author for further clarification of the key information (Bustamante et al., 2019). We entered the extracted data into an Excel spreadsheet and imported this to R Studio (RStudio Team, 2015).

Data Analysis

We followed a three-stage data analysis process. To update evidence of overall program effectiveness, the first stage comprised an effectiveness synthesis, which aimed to identify statistically significant differences in the outcomes reported in the study results and to understand the differences between interventions in terms of the effectiveness of the component parts (Sutcliffe et al., 2015). To calculate the overall effect size for the primary outcome of interest—CSA knowledge gains—a random effect model was used for the purpose of meta-analysis due to the anticipated variability in the true effect as a result of differences in interventions, teaching methods, and sample characteristics between studies. A random effect model is useful for providing an overall effect estimate and characterizing the heterogeneity of effects across studies (DerSimonian & Laird, 2015).

We estimated the amount of residual heterogeneity and ratio of true to total variance I^2 . The I^2 is interpreted as the proportion of the observed variability in a set of effect sizes that reflects real differences among true effects (Borenstein, Cooper, Hedges, & Valentine, 2009). A value of 0% indicates no observed heterogeneity, and I^2 values of 25%, 50%, and 75% are considered as indicating low, moderate, and high heterogeneity, respectively (Higgins, Thompson, Deeks, & Altman, 2003).

Using methods not previously used in examining the effectiveness of school-based CSA interventions, to identify intervention components, and to understand variation in participant outcomes that may be attributed to characteristics of interventions, in the second stage we deductively coded, categorized, and described intervention *structural* components using the TIDieR checklist items as detailed above, and further explained below. We inductively coded, categorized, and described intervention *content* components that appeared in narrative text of the study reports. Once content components had been identified, we then mapped each study report against these newly-devised categories.

As an advance on previous reviews, using meta-regression we also explored intervention components as moderator variables, an approach that has not previously been possible owing to insufficient information provided in study reports (Davis & Gidycz, 2000; Walsh et al., 2015). Meta-regression allows for the evaluation of one or more covariates simultaneously (Baker et al., 2009). To identify the components that contribute to program effectiveness, in the third stage, with structural and content components clearly identified through the TIDieR checklist and ICA coding, we ran meta-regression

to explore sources of heterogeneity among studies by testing whether effect size estimates varied significantly by intervention structural and content components. These components included: (a) modes of delivery; (b) providers; (c) the level of adaptation of the intervention (tailored or untailored); (d) the level of modification of the intervention; (e) fidelity assessment; (f) training for program providers prior to the program implementation; (g) intervention dosage; and (h) intervention intensity. We also examined whether participants' characteristics contributed to program effectiveness. To further examine whether combinations of intervention components and study characteristics contributed to program effectiveness, we explored the association between effect sizes and intervention content components as well as study characteristics using multiple meta-regression (Harrer, Cuijpers, Furukawa, & Ebert, 2019).

Calculation of Effect Sizes

We calculated standardized effect sizes for all outcomes of interest. When possible, Hedges' g was calculated for between-group effects for all continuous outcome measures. Where studies reported outcome data at mid-intervention, post-intervention, and/or at follow-up periods, the immediate post-intervention timepoint was used. Where multiple studies reported the outcomes for one trial, the study that reported the outcome of most relevance to the current review was used, namely, CSA knowledge gains or self-protective skills. In addition, some trials had multiple treatment arms (i.e., these studies did not conform to the standard two-arm "intervention group" vs. "control group" design and included two or more intervention groups). For the purpose of including the results of these studies in the meta-analysis, we followed the formulae for combining groups recommended by Higgins and Green (2011) and synthesized the results of the intervention arms to obtain one single comparison with the control group.

Assessment of Risk of Bias in Included Studies

The methodological quality of the included studies was assessed using the Cochrane Risk of Bias Tool (v1.0; Higgins et al., 2011), which includes criteria to assess the risk of bias across six different domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, and selective reporting. Each domain was assigned a judgment of "low," "unclear," or "high" risk of bias.

Sensitivity Analysis

Where appropriate, we performed sensitivity analyses to examine if outliers were associated with the magnitude and significance of effect sizes. Outliers were defined as studies in

which confidence intervals did not overlap with confidence intervals of the pooled effect (Harrer, et al., 2019).

Results

Results of the Search

The study flow diagram (PRISMA flowchart) is presented in Figure 1. Database searches yielded 5,423 records which were then imported into the reference management software, Mendeley (version 1.19.3; Elsevier, 2018). One thousand six hundred and sixty records remained for title and abstract screening after duplicates were removed. One additional record was identified through contacting authors. Two papers (Holloway & Pulido, 2018; Pulido et al., 2015) reported the

same trial, and we selected the first (Pulido et al., 2015) because it reported both pre- and post-test data.

After title and abstract screening against the inclusion and exclusion criteria, twenty-eight records progressed to full-text review alongside 29 records from the Walsh et al. (2015) review, which used identical inclusion and exclusion criteria to assess eligibility. Full-text reports of the 28 newly-identified records were assessed including four studies published in Chinese-language journals. Twenty-one papers were excluded: five papers were excluded as these were already accounted for in the Walsh et al. (2015) review and sixteen papers were excluded for reasons documented in the PRISMA flowchart, including the four Chinese-language studies—these were excluded on the study design criteria. One newly-identified study (Jin, Chen, Jiang, & Yu, 2017) was

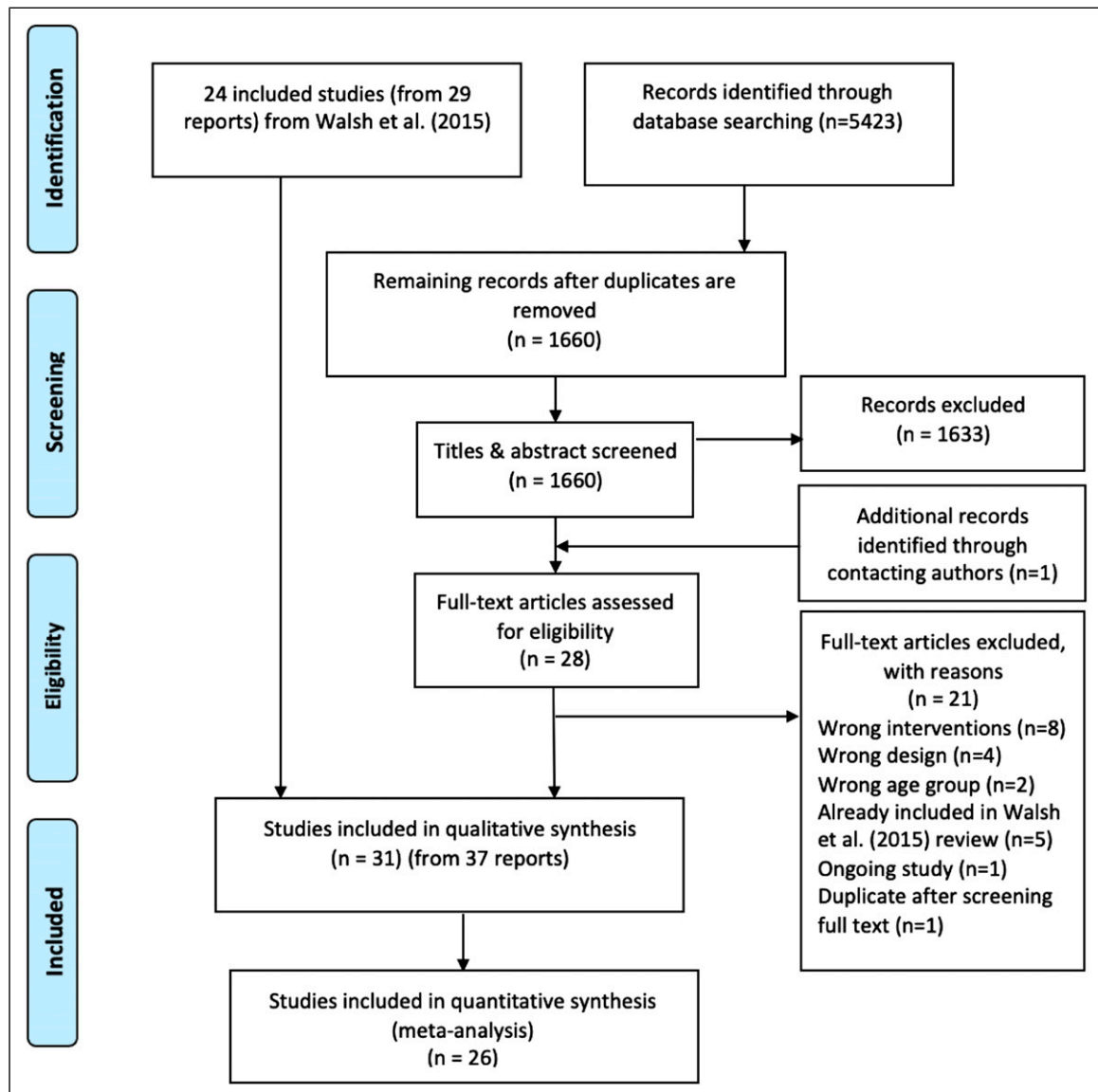


Figure 1. PRISMA flowchart.

Table 1. Characteristics of Included Studies.

Authors	Design	Country	Intervention	Outcome Measures	Participant Mean Age (years)
Chen et al. (2012)	RCT	Taiwan	Red Flag/Green Flag People and Red Flag/Green Flag People II	Children's Sexual Knowledge Questionnaire (CSKQ); Children's Awareness of Scary Secrets (CASSQ); What If Situation Test (WIST) (Chinese version)	9.02
Fryer et al. (1987a)	RCT	USA	Children Need to Know Personal Safety Training Program	Children Need to Know Knowledge/Attitude Test (CNTKKAT)	Not reported
Harvey et al. (1988)	RCT	USA	Good Touch-Bad Touch	Custom-made measure	5.80
Jin et al. (2017)	RCT	China	Body Safety Training	What If Situation Test (WIST)	Mean age not reported; grade level: 1 st –5 th grades 13.44
Lee and Tang (1998)	RCT	China	Behavioral Safety Training	Personal Safety Questionnaire (PSQ); WIST	Mean age not reported; grade level 1 st –6 th grades
Tutty (1997)	RCT	Canada	Who do you tell?	Children's Knowledge of Abuse Questionnaire (CKAQ)-Revised PSQ; WIST	6.10; 11.10
Wurtele et al. (1986)	RCT	USA	Film: Touch Body Safety Training Combined of the above two intervention	Good Touch Bad Touch Curriculum Test	Mean age not reported; grade level: 4 th grade 7.20
Çeçen-Eroĝu & Kaf Hasircı (2013)	RCT	Turkey	Good Touches/Bad Touches	Touch Discrimination Task (TDT); custom-made measure	9.3 (range 6–14)
Blumberg (1991)	Cluster-RCT	USA	STOP (stop, tell someone, own your body, protect yourself); Multimedia program: child abuse primary prevention program	Custom-made measure (questionnaire-based measure)	8.70
Bustamante et al. (2019)	Cluster-RCT	Ecuador	I have the right to feel safe at all times	Custom-made measure	10.60
Dake et al. (2003)	Cluster-RCT	USA	Not specified	Personal Safety Issues Test	6.90
Dawson (1987)	Cluster-RCT	USA	Not specified	PSQ; WIST	Mean age not reported; grade level: 3 rd –4 th grades
Grendel (1991)	Cluster-RCT	USA	CSA program developed by Women Helping Women	What Would You Do: What I Know About Touching scale	9.7 (range 8–10)
Hazzard et al. (1991)	Cluster-RCT	USA	Feeling Yes, Feeling No	CKAQ-R-III (questionnaire-based)	8.30 (experimental group); 8.50 (control group)
Jones et al. (2020)	Cluster-RCT	Australia	Name	Custom-made measure	
Kolko et al. (1989)	Cluster-RCT	USA	Red Flag/Green Flag	Custom-made measure	

(continued)

Table 1. (continued)

Authors	Design	Country	Intervention	Outcome Measures	Participant Mean Age (years)
Krahé and Knappert (2009)	Cluster-RCT	Germany	Live performance of a theater play titled "No Child's Play" DVD (recorded version of the live performance)	Custom-made measure	7.55
Kraizer (1991)	Cluster-RCT	USA	The Safe Child Program	CNTKKAT	Mean age not reported; age range 3–10 (k-3 rd grade)
Morris et al. (2017)	Cluster-RCT	USA	Safe @ Last	Custom-made measure	Mean age not reported; grade level: 1 st –6 th grade
Müller et al. (2014)	Cluster-RCT	Germany	Cool and Safe	CKAQ	9.00
Oldfield et al. (1996)	Cluster-RCT	USA	Project TRUST	CKAQ-Revised	Mean age not reported; grade level: 1 st –6 th grade
Poche et al. (1988)	Cluster-RCT	USA	Videotape (name not specified)	N/A	Not reported
Pulido et al. (2015)	Cluster-RCT	USA	Safe Touches	CKAQ	8.34
Urbann et al. (2020)	Cluster-RCT	Germany	Safe with Sam	Custom-made measure (questionnaire-based measure); WIST (vignette-based measure)	9.66 (range 8–12)
Wolfe et al. (1986)	Cluster-RCT	USA	Play Intervention	Custom-made measure	10.20
Crowley (1989)	Quasi-RCT	USA	Good Touches/Bad Touches: A Program to Prevent Child Sexual Abuse	PSQ; Sexual Abuse Knowledge Inventory (SAKI)	Mean age not reported; grade level: 4 th –5 th grade
Daigneault et al. (2012)	Quasi-RCT	Canada	ESCAPE: a French adaptation of the American child assault prevention program	WIST (French version); CKAQ	7.75
del Campo Sánchez & Sánchez (2006)	Quasi-RCT	Spain	Not specified	PSQ; custom-made measure	Not reported
Hébert et al. (2001)	Quasi-RCT	Canada	ESCAPE: an adaptation of the American child assault prevention program	CKAQ; PSQ; Unnamed measure	7 years and 2 months
Saslowsky and Wurtele (1986)	Quasi-RCT	USA	Film called "Touch"	PSQ; WIST	6.20; 11.10
Snyder (1986)	Quasi-RCT	USA	Good Secrets Bad Secrets	Good Secrets, Bad Secrets Quiz	9.60

conducted in China and published in an English-language journal. A total number of seven new studies (from eight reports) were identified. Together, the new searches (seven studies from eight reports) and the Walsh et al. (2015) searches (24 studies from 29 reports) resulted in 31 included studies (from 37 reports).

Study Characteristics

Characteristics of included studies are presented in Table 1.

Designs. Of the 31 studies, seven studies were RCTs, eighteen were cluster-RCTs, and six were quasi-RCTs. The unit of randomization in eighteen studies were groups of children (e.g., classes, schools, or administrative collectives such as districts). Of these, eighteen studies were cluster-RCTs (as above) and three were quasi-RCTs. The unit of randomization in the remaining ten trials was individual school students. Of these, seven studies were RCTs and three were quasi-RCTs.

Participants. A total of 9049 participants were included across the 31 studies (5802 participants from Walsh et al.'s review and 3247 from the newly-included studies). Sample sizes in cluster-RCTs ranged from 74 (Poche et al., 1988) to 1269 (Oldfield et al., 1996). For the standard RCTs, the sample sizes ranged from 46 (Chen et al., 2012) to 565 (Jin et al., 2017).

The studies were conducted in a number of different countries, a majority ($n = 18$) being from the United States, three were conducted in Canada and Germany, two studies were conducted in China, and one apiece in Ecuador, Spain, Taiwan, and Turkey.

The mean age of participants could be identified in 21 studies. Authors of 10 studies did not report the mean age of participants. Where identified, participant mean age ranged from 5.8 years (Harvey et al., 1988) to 13.44 years old (Lee & Tang, 1998). The proportion of females in the included studies ranged from 44.20% (Pulido et al., 2015) to 56.10% (Jones et al., 2020). As with Walsh et al.'s (2015) review, only one study included female only participants (Lee & Tang, 1998). Gender data for participants were not reported in five studies. Of note, gender data were reported in all newly-included studies.

Outcome Measures. Questionnaire-based measures for studies included in the meta-analysis included the Personal Safety Questionnaire (PSQ), Children's Knowledge of Abuse Questionnaire and versions thereof (CKAQ, CKAQ-R, and CKAQ-III-R), Children Need to Know Knowledge/Attitude Test, What I Know About Touching Scale; Good Touch Bad Touch Curriculum Test; Children's Awareness of Scary Secrets, Children's Sexual Knowledge Questionnaire (CSKQ); Good Secrets Bad Secrets Quiz; Personal Safety Issues Test; Sexual Abuse Knowledge Inventory (SAKI), other custom-designed and unnamed measures were also used. Vignette-based measures included the What If Situations Test (WIST),

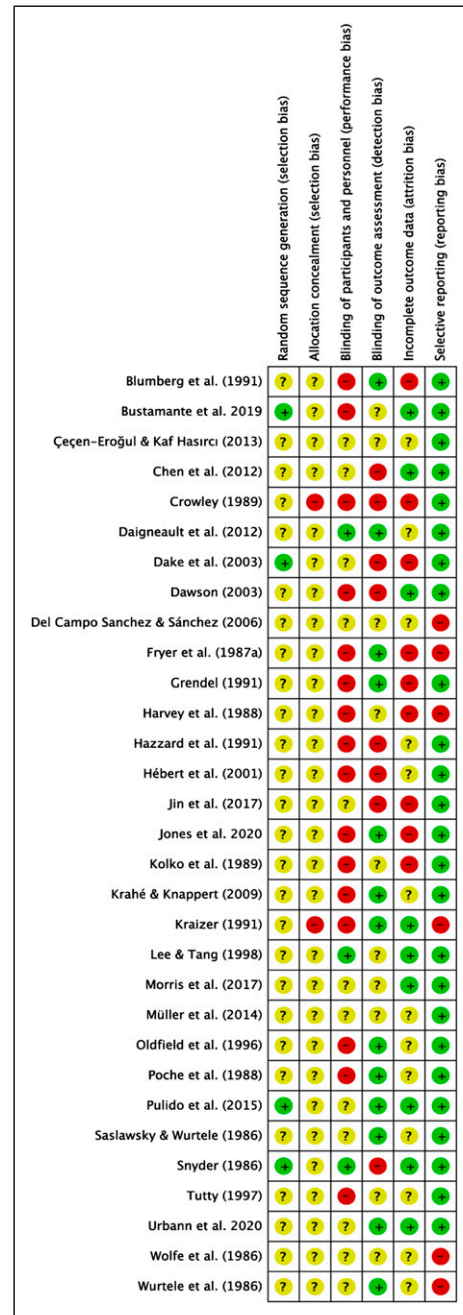


Figure 2. Risk of bias assessment of included studies.

Touch Discrimination Task (TDT), What Would You Do? (WWYD), and other unnamed measures were also used.

Risk of Bias in Included Studies

Random Sequence Generation (Selection Bias). Figure 2 (summary for individual studies) and Figure 3 (summary across all studies) provide details on the quality of the included studies. Notably, only four studies reported successful randomization, two of which were identified in the new search. The remaining

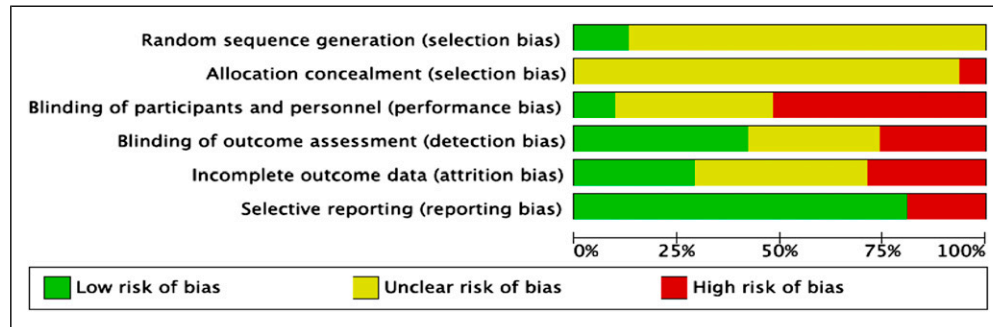


Figure 3. Review authors' judgments about each risk of bias item across all included studies.

27 studies did not explicitly state the randomization process used, therefore were rated as unclear risk.

Allocation Concealment (Allocation Bias). Twenty-nine of the included studies did not report the strategies used for allocation concealment. We therefore classified these studies as unclear risk of bias. Further to this, two studies were rated as high risk of bias by Walsh et al. (2015) due to failure of randomization and the involvement of school staff in the randomization process. None of the seven studies identified in the new search reported allocation concealment procedures, and these were also rated as unclear risk of bias.

Blinding of Participants and Personnel (Performance Bias). Only three trials reported blinding of participants and personnel, we therefore classified these studies as low risk of bias. Sixteen studies (two of which were newly-included studies) were rated as high risk of bias due to the potential bias resulting from knowledge about the study group. Eleven studies did not report on the blinding of participants and personnel, as a result, we rated them as unclear risk of bias. Five of the seven newly-included studies did not report clearly regarding blinding of participants and personnel. Therefore, we classified these studies as unclear risk of bias.

Blinding of Outcome Assessment (Detection Bias). Blinding of outcome assessment was not reported in ten studies; therefore, we rated them as unclear risk of bias. Thirteen studies (including six studies identified in our new search) were classified as low risk of bias because authors reported using strategies to minimize outcome assessment bias. Eight studies were rated as high risk of bias as no strategies were in place to blind outcome assessors to group membership or to ensure children completed the assessment independently.

Incomplete Outcome Data (Attrition Bias). Consistent with Walsh et al.'s (2015) criteria for attrition bias, we classified studies as low risk of bias when attrition rates were less than 10% and high risk of bias when attrition rates were more than 10%. Following this classification criteria, nine studies were rated as low risk of bias because they reported no attrition or

loss to follow-up or reported attrition rates lower than 10%. Nine trials reported attrition rates higher than 10%, we therefore rated them as high risk of attrition bias. Thirteen studies did not report attrition rates, we classified these studies as unclear risk of bias.

Selective Reporting (Reporting Bias). Six studies were rated as high risk of reporting bias because data for outcomes were not reported completely in the results section. The remaining 25 studies (including seven studies identified in our new search) were rated as low risk of bias.

Meta-analysis of the Overall Effectiveness of School-based CSA Interventions. Twenty-five studies were included in the quantitative synthesis (see Figure 1). Two meta-analyses were conducted. A total of 23 studies with 7191 participants were included in the meta-analysis of participant knowledge outcomes assessed using questionnaire-based measures (i.e., assessment of factual knowledge), and a total of 13 studies (2264 participants) were included in the meta-analysis of knowledge outcomes assessed using vignette-based measures (i.e., assessment of applied knowledge).

The results of the two meta-analyses are presented in Figures 4 and 5. The random effects meta-analysis showed a significant medium to large effect size (Hedges' g) of 0.72 (95% CI [0.52, 0.93]) and 0.55 (95% CI [0.35, 0.74]) for factual knowledge with questionnaire-based measures and applied knowledge with vignette-based measures, respectively. This indicates that students who have participated in CSA interventions showed higher levels of factual and applied CSA knowledge than participants who did not receive a CSA intervention program. As expected, there was significant heterogeneity in effect sizes across studies using both measures. These results suggest the importance of examining moderating variables and the value of conducting of meta-regression.

Intervention Components Analysis

Structural Components. An overview of intervention structural components is presented in Table 2. The detailed descriptions

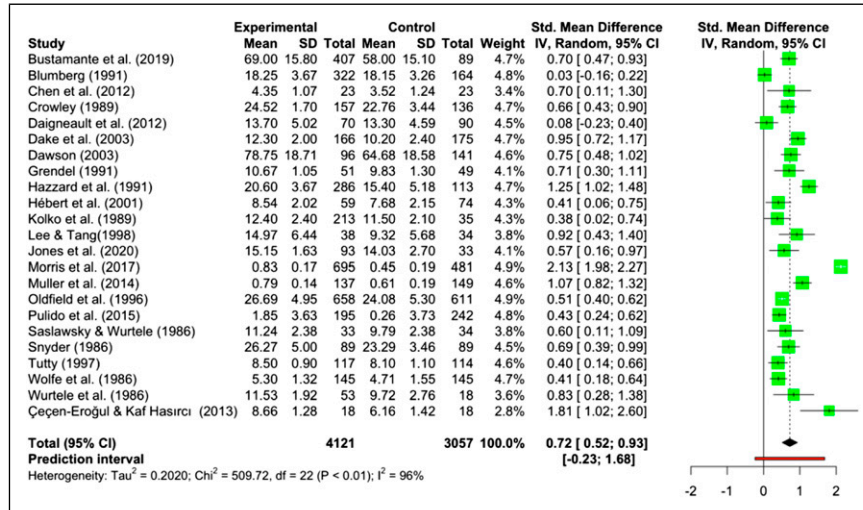


Figure 4. Questionnaire-based CSA knowledge gains.

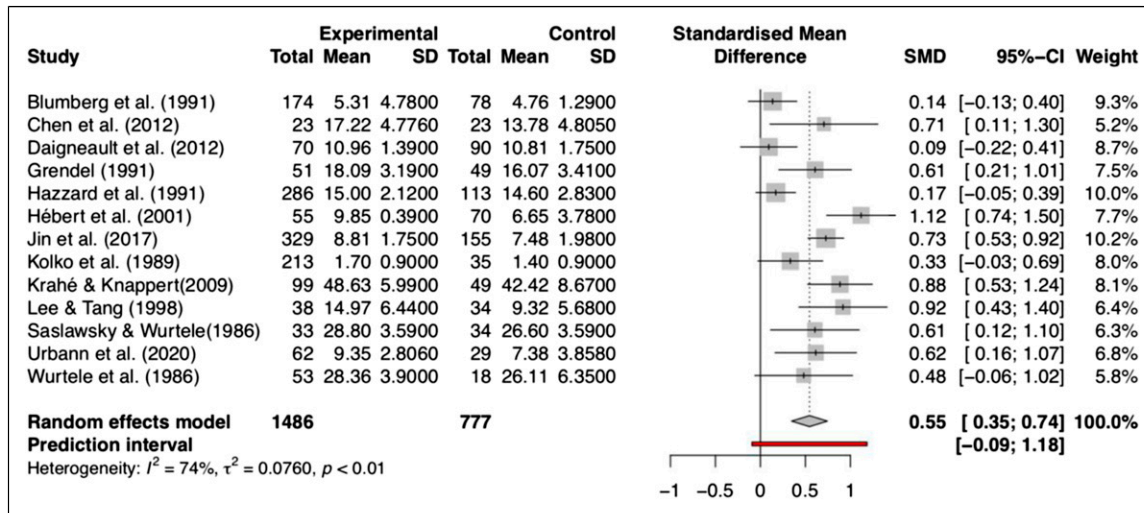


Figure 5. Vignette-based CSA knowledge gains.

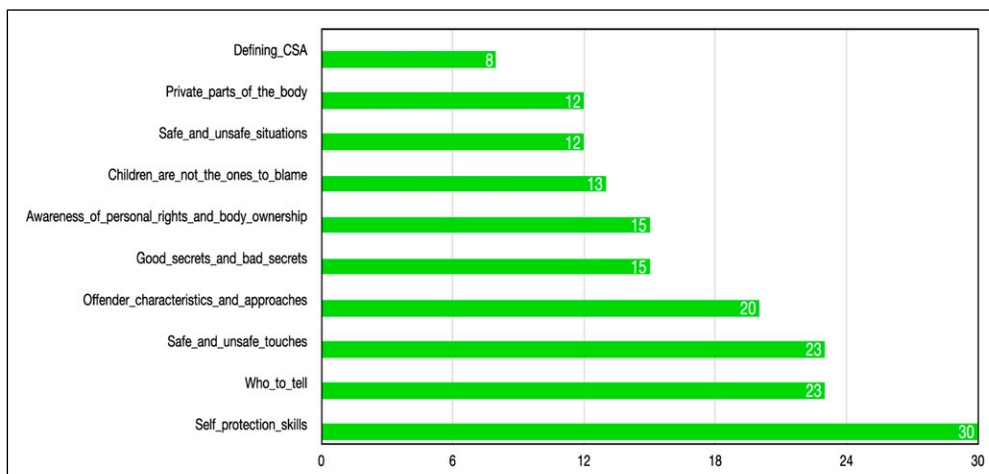


Figure 6. Content components in interventions evaluated in the included studies (n = 31).

Table 2. Description of Structural Components of Interventions.

Structural Component	Description
Mode of delivery	Coded as free text: Home activities (homework, parental involvement); live performance (performance, theater play); instructional (music, games, stories, films, pictures, presentations, instructions; review); behavioral (role-play, participant modeling, practice, shaping); feedback (social reinforcement, praise, discussion)
Providers	Coded as free text: Professionals (teachers, social workers, counselors, nurses, or community workers; police); parents; volunteers
Tailored or untailored	Coded as present (1) or absent (0)
Modification	Coded as present (1) or absent (0)
Fidelity assessment	Coded as present (1) or absent (0)
Training for program providers prior to program implementation	Coded as present (1) or absent (0)
Intervention dosage	Coded as 60 min or less (1); 61 min or over (0)
Intervention intensity	Coded as 1–3 sessions (1); 4–6 sessions (0)

of intervention structural components and their coding descriptions are shown in [Table 3](#).

Modes of Delivery. Twenty-one studies incorporated role-play teaching strategies to achieve the intervention goals. Eighteen studies included discussion as a teaching strategy in the program to enhance participants' understanding of CSA. Seventeen studies employed the use of video recordings or films and six studies incorporated parental engagement or home activities outside the classroom as modalities. Two studies used live performance or theater plays to achieve intervention goals.

Intervention Dosage and Intensity. The duration of interventions ranged from one single 50-minute session ([Pulido et al., 2015](#)) to 10 one-hour sessions delivered over 10 weeks ([Bustamante et al., 2019](#)). Of the seven newly-included studies, the duration of intervention programs ranged from one single 50-minute session ([Pulido et al., 2015](#)) to ten one-hour sessions delivered over a 10-week period ([Bustamante et al., 2019](#)). One study ([Daigneault et al., 2012](#)) included three types of additional booster sessions beyond the initial program sessions, including a complete version of the original program (*ESPACE*), or a brief version of the original program (for third and fourth graders), or a program called “*Confidence, Solidarity, Respect*” (90 minutes) which revisited *ESPACE* concepts and addressed gender-based violence and respect (for fifth and sixth graders). These boosters were delivered 2 years after the initial intervention.

Provider of the Intervention. The intervention providers fell into three categories: professionals, parents, and volunteers (including student volunteers). Twenty-three studies were delivered by professionals such as teachers, guidance counselors, social workers, mental health professionals, and police. One study was delivered by both parents and professionals ([Jin et al., 2017](#)). Five studies were delivered by volunteers

including students. Three studies did not provide details about the intervention provider.

Training Prior to Program Implementation. Twenty studies reported providing training sessions for parents and/or teachers prior to program implementation.

Program Tailoring. Twenty-two studies did not report whether or not the program was tailored or adapted. The remaining nine studies reported tailored components. Tailored components included using different teaching strategies to accommodate children's developmental levels and interests (e.g., teaching younger children songs about safe and unsafe touches and using less complex role plays with younger children), and using culturally adapted names for characters in practical scenarios and examples.

Program Modification. Five studies reported modifying the program during the implementation process. Modifications included using a shorter version of the program in order to keep the length of the entire program comparable with school conditions, modifying program length for students in different grade or year levels, and adjusting the volume of content presented to the students in a given session.

Fidelity Assessment. Four included studies reported fidelity assessment. All were newly included studies. [Bustamante et al. \(2019\)](#) included volunteers to assist program implementation to maintain a degree of standardization between classrooms. [Pulido et al. \(2015\)](#) and [Urbann et al. \(2020\)](#) utilized a fidelity checklist/questionnaire to track whether the program was delivered according to the study protocol. In [Jin et al.'s \(2017\)](#) study, each intervention session was monitored by one of the researchers and fidelity checklists were also used to ensure that program providers covered key concepts.

Table 3. Overview of Intervention Structural Components.

Study ID	How (Mode of Delivery)	What (Intervention Materials)	Who (Providers)	Training prior to the program implementation			Tailored or Untailored	Modification	How much (Intervention dosage and intensity)	Where (Study Location)	How well (Fidelity Assessment)
				Yes	No	Yes					
Blumberg et al. (1991)	Behavioral and instructional	Presentation; puppets	Professionals and parents	Yes	No	Yes	No	One*1-hour session	Classroom	No	
Bustamante et al. (2019)	Behavioral and instructional	Workbook	Professionals and volunteers	Yes	No	Yes	No	Ten*1-hour sessions over 10 weeks	School	Yes	
del Campo Sánchez & Sánchez (2006)	Not specified	Unclear	Not specified	Yes	Yes	Yes	Yes	Two*50-minute sessions delivered at the beginning of the school day before children begin their regular class.	Classroom	No	
Chen et al. (2012)	Behavioral and instructional	Activity books	Professionals	Yes	Yes	Yes	Yes	Two*50-minute sessions delivered at the beginning of the school day before children begin their regular class.	Classroom	No	
Crowley (1989)	Behavioral and instructional	Film	Professionals	Yes	Yes	Yes	Yes	Two*45-minute sessions; once per week for 2 weeks	Classroom	No	
Daigneault et al. (2012)	Behavioral	Unclear	Professionals	Yes	No	Yes	No	One*90-minute workshop	Classroom	No	
Dake et al. (2003)	Behavioral and instructional	Video tapes	Professionals	Yes	No	Yes	No	Two*1-hour sessions	Classroom	No	
Dawson (1987)	Behavioral and instructional	Film	Professionals	No	No	No	No	One*1-hour session	Classroom	No	
Fryer et al. (1987a)	Behavioral	Presentations	Not specified	No	No	No	No	One*20-minute presentation for 8 days	Classroom	No	
Grendel (1991)	Instructional and feedback	Film	Professionals	Yes	No	No	No	One*50-minute lecture and film presentation	Classroom	No	
Harvey et al. (1988)	Behavioral and instructional	Storybook; puppets	Professionals	Yes	Yes	Yes	Yes	Two*45-minute sessions; once per week for 2 weeks	Classroom	No	
Hazzard et al. (1991)	Instructional and feedback	Video tapes comic books, home activities	Professionals	Yes	No	No	Yes	Three*1-hour sessions	Classroom	No	
Hébert et al. (2001)	Behavioral	Unclear	Professionals	Yes	No	No	No	One*60–75 minute workshop	Classroom	No	
Jin et al. (2017)	Behavioral and instructional	Workbook	Professionals and parents	Yes	No	No	No	Three sessions were finished within 15 days. Each session is about 30 minutes.	Classroom	Yes	
Jones et al. (2020)	Behavioral; instructional	Online materials; activities	Professionals	Yes	No	No	No	Typically 1 hour a week/fortnight for 5–10 weeks. Minimum play time of 5 hours	School	No	
Kolko et al. (1989)	Instructional	Workbook; film	Professionals and volunteers	Yes	No	No	No	Two*45-minute sessions on consecutive days	Classroom	No	
Krahé and Knappert (2009)	Live performance and instructional	Theater play	Professionals	No	Yes	Yes	No	One*60-minute session	Classroom	No	
Kraizer (1991)	Behavioral and instructional	Video tapes	Professionals	Yes	No	Yes	No	Preschool and kindergartners: 1 session per day for 10 days. 1–3rd graders: 1 session per day for 5 days.	Classroom	No	

(continued)

Table 3. (continued)

Study ID	How (Mode of Delivery)	What (Intervention Materials)	Who (Providers)	Training prior to the program implementation	Tailored or Untailored	Modification	How much (Intervention dosage and intensity)	Where (Study Location)	How well (Fidelity Assessment)
Lee and Tang (1998)	Behavioral and feedback	Workbook	Professionals	No	No	No	Two*45-minute sessions	Classroom	No
Morris et al., (2017)	Behavioral	Workbook	Professionals	Yes	Yes	No	Four-sessions each lasting 35 minutes, delivered in a weekly basis over 1 month	Classroom	No
Müller et al. (2014)	Instructional and feedback	Online materials	Professionals	Yes	No	No	Two hours during normal lesson time	Classroom	No
Oldfield et al. (1996)	Live performance and instructional	Theater play	Professionals and volunteers	Yes	No	No	The presentation lasts for 30 minutes and is followed by a 15 minutes student Q & A period.	Classroom	No
Poche et al. (1988)	Behavioral, feedback, and instructional	Video tapes	Professionals	No	No	No	Video only (25 minutes) videotape and rehearsal 45 minutes standard program 60 minutes	Classroom	No
Pulido et al. (2015)	Home, instructional, and behavioral	Workbook; puppets	Professionals	Yes	No	No	One*50-minute workshop	Classroom	Yes
Urbann et al. (2020)	Behavioral; instructional	Puppets; film; workbook; activities	Professionals	No	No	No	Over 2 weeks, 3 unites taught per week, each lasting 180 minutes	School	Yes
Saslawsky and Wurtele (1986)	Instructional	Film	Volunteers	No	No	No	Thirty-five minutes film viewing and 15 minutes discussion following the film	Classroom	No
Snyder (1986)	Behavioral	Presentation	Professionals	Yes	No	No	Three sessions were finished within 15 days.	Classroom	No
Tutty (1997)	Home, behavioral, and instructional	Workbook	Professionals	No	No	No	Each session is about 30 minutes.	Classroom	No
Wolfe et al. (1986)	Live performance; behavioral	Theater play	Volunteers	Yes	No	No	Two*45-60 minute sessions on consecutive days	Classroom	No
Wurtele et al. (1986)	Instructional, behavioral, and feedback	Workbook; film	Volunteers	No	No	Yes	Two*5-minute plays written and performed by volunteer medical students who consulted with child abuse specialists. One*70-minute session	Classroom	No
Çeçen-Eroĝu & Kaf Hasirci (2013)	Not specified	Video tapes	Unclear	No	Yes	No	Intervention 1 and 2: 50-minute session; Intervention 3: one 60-minute session	Classroom	No
						No	One-hour sessions were carried out in 4 consecutive days	Classroom	No

Content Components. The first author and the last author (ML and YW) inductively coded, categorized, and systematically described intervention *content* components that appeared in the narrative text of the study reports. Any discrepancies between the two reviewers were resolved by discussion or consultation with a third reviewer. Cohen's *kappa* was calculated to measure the inter-rater agreement between the two reviewers; the result $k=0.72$ indicated substantial agreement between reviewers regarding the coded components (McHugh, 2012). The details of the coding process are available upon request.

In total, the inductive coding process generated 10 content components. An overview of intervention content components is presented in Table 4. Interventions were heterogeneous, incorporating two (*Who Do You Tell?* evaluated by Tutty, 1997) to eleven content intervention components (a CSA Prevention Program developed by Women Helping Women, evaluated by Grendel, 1991). Table 5 and Figure 6 present an analysis of the frequencies of interventions incorporating each content component. "self-protection skills" appeared to be the most frequently taught component across the interventions, which was included in 30 of the 31 studies. In contrast, "defining CSA" was noted as the least commonly taught component, with only eight interventions including this.

Results from Meta-regression Analyses

Meta-regression of Intervention Structural Components. The results of the meta-regression for structural components showed that only one of the eight components was associated with significant program effects as shown in Table 6. In order to explore how many intervention sessions are associated with greater program effectiveness, we categorized the total number of intervention sessions into three major groups: (a) up to 3 sessions (1) versus 3 sessions and more (0); (b) up to 4 sessions (1) versus 4 sessions or more (0); and (c) up to 5 sessions (1) versus 5 sessions or more (0). Results from the meta-regression showed that although interventions with more than three sessions and more than four sessions are both found to be negatively associated with effect size, we observed a statistically significant relationship only between interventions with three or more sessions ($\beta = -0.6637$, 95% CI $[-1.1554, -0.1720]$, $p = 0.0106$), indicating that interventions with more than three sessions tended to be more effective than interventions with three or fewer sessions. We did not observe a statistically significant relationship between other structural components and effect size.

Meta-regression of Intervention Content Components

We conducted univariate meta-regression for questionnaire-based knowledge outcomes to examine the association between study effect size and intervention content components. The results are presented in Table 7. Interventions including the specific content component "Safe situations and unsafe situations" tended to show larger effects ($\beta = 0.2439$, 95% CI

$[-0.1766, 0.6643]$, $p = 0.2412$). We planned to conduct exploratory multiple meta-regression to investigate the association between different combinations of intervention content components and effect size, but the data provided in study reports did not allow for meaningful comparisons as I^2 values for all content components were strongly heterogeneous across studies. We also planned to conduct meta-regression analyses for vignette-based knowledge measures, however, there were insufficient data.

Meta-regression Results of Participant Characteristics

Demographic characteristics such as participant age, gender, ethnicity, socioeconomic position, and ability level are potential effect moderators and might explain existing heterogeneity (Walsh et al., 2015). However, participant demographic information was poorly reported in the included studies, and we were able to extract data only for participant age and gender.

We conducted meta-regression to explore the effects of participants' age on knowledge outcomes measured using different instruments (questionnaire-based measures vs. vignette-based measures). We treated eight years of age as the cut-off for the meta-regression based on the cut-off used in previous studies/reviews (Davis & Gidycz, 2000; Tutty, 1997) or used corresponding school grades as the cut-off when ages were not provided (e.g., kindergarten to third grade and fourth grade and above). The selection of this cut-off is educationally relevant as it is consistent with the distinction between the early years of school and middle school. For studies using questionnaire-based measures, this categorization resulted in two groups of studies: (a) four studies with participants younger than eight years old; (b) sixteen studies including participants aged eight years and older. The results of meta-regression suggest that when testing factual knowledge utilizing questionnaire-based measures immediately after receiving the program, older children (eight years and older) show greater knowledge gains than children who are younger than eight years old ($\beta = -0.4574$, 95% CI $[-0.8203, -0.0946]$, $p = 0.0162$).

For studies using vignette-based measures, four studies included participants younger than eight years old. Eight studies included participants aged eight years and older. The results of the meta-regression show that when testing applied knowledge using vignette-based measures immediately after the program, we did not observe any association between participants' age and program effectiveness ($\beta = 0.0323$, 95% CI $[-0.4295, 0.4941]$, $p = 0.8792$).

In the next set of meta-regressions, we explored whether program effectiveness varied with participants' gender. The meta-regression results are presented in Table 8. Overall, we did not observe any association between participants' gender and program effectiveness ($\beta = -0.2552$, 95% CI $[-1.1506, 0.6402]$, $p = 0.5542$).

Sensitivity Analysis and Publication Bias

We conducted sensitivity analyses to examine whether study quality was associated with the magnitude and significance of effect sizes. Five studies (Blumberg et al., 1991; Daigneault et al., 2012; Hazzard et al., 1991; Morris et al., 2017; Çeçen-Eroğlu & Kaf Hasırcı, 2013) were identified in which effect size confidence intervals did not overlap with confidence interval of the pooled effect (Harrer et al., 2019). The results of the sensitivity analysis in which these five studies were removed showed an overall effect size of 0.6360 (95% CI [0.5302, 0.7417], $p < 0.0001$), which was lower than the original effect size of 0.7234 (95% CI [0.5159, 0.9310], $p < 0.0001$) but was still in the range for a medium effect. I^2 value was reduced from 95.7% to 58.8% indicating that even with studies at higher risk of bias removed, this remains a heterogeneous set of studies.

Finally, we assessed for potential publication bias for the 23 studies included in the meta-analysis. According to Egger's test, there was no evidence of publication bias (*intercept*: -1.286 ; 95% CI [$-6.03, 3.46$]; $p = 0.600785$). Although this result did not reach statistical significance, we cannot fully rule out the existence of publication bias because the subjective visual assessment of the funnel plot (Figure 7) suggested that there were fewer studies at bottom left and bottom right on both sides of the center, indicating the potential for missing studies.

Discussion

This systematic review and meta-analysis advances on the substantive work undertaken in previous systematic reviews in several ways. First, by replicating searches using identical search and screening strategies and searching Chinese-language databases, this review also captured studies from one of the worlds' most populous regions. Second, this systematic review and meta-analysis is the first to catalogue CSA prevention programs' structural and content components in a way that enables systematic assessment of factors associated with program effectiveness. By systematically mapping out intervention components using inductive thematic analysis, the findings from this review advance the field's understanding of intervention components in existing school-based CSA prevention programs. This innovation can, in turn, serve several purposes including providing criteria for program development, and the potential to develop an auditing tool for program quality assessment. Third, we provided a very robust analysis of the predictors for greater program effectiveness, which are of particular importance for contexts such as LMICs where resources are limited and the impetus to avoid resource wastage and provide only effective interventions is high.

The majority of included studies ($n = 26$) were conducted in high-income countries (HICs), only five studies were conducted in LMICs and upper-middle income countries. It

should be noted that programs used in these evaluation studies (e.g., *Red Flag/Green Flag People*, *Body Safety Training*, *Behavioral Safety Training*, and *Good Touches/Bad Touches*) were all originally developed in HICs and the majority of the studies did not report any cultural adaptation undertaken during the implementation process. Therefore, findings from our systematic review indicate an urgent need for developing context-specific school-based CSA interventions that have been adapted to address different sociocultural contexts and gender norms, as well as evaluate the effectiveness of these culturally adapted programs. Further to this, having screened all the studies from database searches, we excluded four Chinese studies that did not meet our study design criterion (i.e., RCT), suggesting the need for rigorous evaluation studies such as RCTs to be conducted in China, and more broadly, in LMICs, to deepen our understanding of factors that influence the effectiveness of programs in these settings.

Structural Components and Participants Characteristics

Findings from our analyses added weight to the argument that once-off single session personal safety programs undertaken in large school assemblies are clearly unlikely to be effective. Furthermore, previous literature suggested that dividing programs into several sessions allows participants to maintain their attention for the entire period of the intervention and increases the opportunity for repeated presentations of key concepts, which reinforces earlier learning (Sanderson, 2004). Davis and Gidycz (2000) proposed that prevention education has a cumulative effect, with children's knowledge and skills continuing to improve with further exposure to program content, and although repeated exposure is known to be an effective instructional strategy (Hattie & Yates, 2013), it is yet to be tested empirically with this specific category of educational interventions.

As for participants' characteristics, results from the meta-regression are consistent with Walsh et al.'s (2015) findings, which indicated that when using questionnaire-based measures, older children (8 years and older) demonstrate greater knowledge gains than younger children (less than 8 years) exposed to the same programs. Several reasons might explain the results. First, younger children's cognitive, social and emotional, and behavioral development may account for some of their challenge in learning CSA prevention concepts. Second, it may also be that programs evaluated in studies in this review were not ideal in terms of developmental appropriateness for children in the early childhood years. The ways in which these concepts are taught may benefit from greater tailoring to suit children in this age range, for example, in the use of interactive toys, games, and multimedia.

Third, as the findings from the meta-regression showed, the way in which program outcomes are measured appears to matter. Further to this, the measures used to assess factual and

Table 4. Description of Content Components of Interventions.

Content Component	Description
Defining CSA	Teaching children about what CSA means
Private parts of the body	Teaching children to identify the private parts of their bodies, or distinguish public and private parts
Safe and unsafe touches	Teaching children to recognize appropriate and inappropriate touches and distinguish between these touches
Self-protection skills	Teaching children self-protection and a range of skills for responding in potentially abusive situations
Children are not the ones to blame	Teaching children that sexual abuse of a child is never the child's fault
Safe and unsafe situations	Teaching children to recognize safe and unsafe situations
Awareness of personal rights/body ownership	Teaching children that their body belongs to them and no one can touch it without their permission
Offender characteristics and approaches	Teaching children that the perpetrators can be someone known or unknown to them and providing information about grooming strategies including gifts and favors
Good and bad secrets	Teaching children the difference between good and bad secrets, or to distinguish secrets and surprises
Who to tell	Teaching children about the importance of disclosure and pointing them to appropriate and effective sources of help

applied CSA knowledge have typically been independently completed self-report paper and pencil tests. These may favor children with higher reading and comprehension skills and, indeed, more experience with test taking. With the ubiquitous presence and use of digital technologies in schools, study assessors may now begin to capture the opportunities for diverse and more inclusive methods for data collection with young children, meaning that in future, there will be less possibility for testing methods to interfere with the assessment of program effects.

Finally, we did not observe a significant association between program providers and effect size, which is consistent with the findings of a previous meta-analysis (Davis & Gidycz, 2000). Fidelity assessments that record which components were used in the actual program implementation process would be useful in examining the degree to which the intervention was delivered as intended. We were not able to conduct such analysis due to limited reporting with only four studies reporting on fidelity assessment (Bustamante et al., 2019; Jin et al., 2017; Pulido et al., 2015; Urbann et al., 2020). It is noteworthy that these were more recent studies, and this may indicate the quality of study reporting on this criteria is starting to improve.

Intervention Content Components

Although we did not observe significant relationships between any of the individual program content components and the overall effectiveness of the intervention, information about “safe and unsafe situations” tended to be relatively more effective than other components in improving participants’ understanding of CSA knowledge and concepts. The reason for the lack of effect for other content components remains unclear. One possible explanation being that the focus on program components often fails to consider other features of interventions which might contribute to differences in effect

sizes, such as component sequencing, coordination, and guiding supervision infrastructure (Chorpita & Daleiden, 2007). Alternatively, the number of included trials was small, and data were strongly heterogenous, which limited the ability of statistical analysis to examine associations.

Limitations

This review has a number of limitations. First, we are aware that many CSA interventions could show changes in children’s CSA knowledge and behaviors in a longer time frame. However, we included only immediate post-test outcome assessments in this review, which does not allow for the identification of whether these results are maintained over time, or for the possibility of sleeper effects that might have been identified with longer term follow-up. It should be highlighted that improvement in self-protection skills and CSA knowledge do not imply that children will implement the skills and knowledge obtained from these programs in real-life situations, and there is no evidence that these school-based CSA prevention interventions reduce actual instances of sexual abuse. Second, the intervention components were coded based on a relatively small number of studies, which might have resulted in spurious associations or masked true differences in effectiveness related to other intervention components that were not examined in our meta-regression. Moreover, our content components analysis relied solely on the narrative text of study reports. We did not seek out the actual programs to determine whether or not the study reports included a comprehensive description of program contents. The publication of reporting guidelines such as the TIDieR checklist (Hoffmann et al., 2014) make it incumbent upon study authors to ensure that they have fully explained the programs they are evaluating.

In terms of the limitations of the included studies, we were not able to conduct moderator analyses with several variables

Table 5. Overview of Intervention Content Components.

Study ID	Defining CSA	Private parts of the body	Safe and unsafe touches	Self-protection skills (say no, stay away, disclosure methods)	Children are not the ones to blame	Safe and unsafe situations	Awareness of personal rights/body ownership	Offender characteristics and approaches	Good and bad secrets	Who to tell	Sum of components
Bustamante et al. (2019)	0	0	1	1	1	1	0	1	1	1	7
Blumberg et al. (1991)	1	1	1	1	1	1	1	1	1	1	10
Chen et al. (2012)	0	1	1	1	0	1	1	1	1	1	8
Crowley (1989)	0	1	1	1	1	1	1	0	1	1	8
Daignault et al. (2012)	0	0	0	1	0	0	1	0	0	1	3
Dake et al. (2003)	1	0	1	1	1	1	0	0	1	1	7
Dawson (1987)	1	0	1	1	1	0	1	1	1	1	8
del Campo Sánchez & Sánchez (2006)	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Fryer et al. (1987a)	0	0	0	1	0	1	0	1	0	0	3
Grendel (1991)	0	1	1	1	1	1	1	1	1	1	9
Harvey et al. (1991)	1	0	1	1	1	1	1	1	0	1	8
Hazzard et al. (1991)	1	1	1	1	1	1	0	1	0	1	8
Hébert et al. (2001)	0	0	0	1	0	0	1	0	0	1	3
Jin et al. (2017)	0	1	1	1	1	0	0	0	1	0	5
Jones et al. (2020)	1	1	1	1	1	1	0	1	0	1	8
Kolko et al. (1989)	1	0	1	1	0	0	0	0	0	1	4
Krahé and Knappert (2009)	0	0	0	1	0	1	0	0	1	0	3
Kraizer (1991)	0	0	1	1	0	0	1	1	1	1	6
Lee and Tang (1998)	0	1	1	1	1	0	1	0	0	0	5

(continued)

Table 5. (continued)

Study ID	Defining CSA	Self-protection					Awareness of personal rights/body ownership	Offender characteristics and approaches	Good and bad secrets	Who to tell	Sum of components
		Private parts of the body	Safe and unsafe touches	Self-protection skills (say no, stay away, disclosure methods)	Children are not the ones to blame	Safe and unsafe situations					
Morris et al. (2017)	0	0	0	1	0	1	0	0	1	4	
Müller et al. (2014)	0	1	1	1	0	1	1	1	0	6	
Oldfield et al. (1996)	0	0	1	1	0	1	0	0	0	3	
Poche et al. (1988)	0	0	0	1	0	1	0	0	1	3	
Pulido et al. (2015)	0	1	1	1	1	1	0	1	1	7	
Saslowsky and Wurtele (1986)	0	0	1	1	0	1	1	0	1	5	
Snyder (1986)	0	0	1	1	0	1	0	1	1	5	
Tutty (1997)	0	0	1	1	0	0	0	0	0	2	
Urbann et al. (2020)	0	1	1	1	0	1	1	1	1	7	
Wolfe et al. (1986)	1	0	0	1	0	1	0	0	1	4	
Wurtele et al. (1986)	0	1	1	1	0	1	1	0	1	6	
Çeçen-Eroğlu & Kaf Hasircı (2013)	0	0	1	1	1	1	1	1	1	6	
Total	8	12	23	30	19	24	15	23	15	23	

Table 6. Results from Meta-regression for Structural Components.

Structure Component	β (95% CI, <i>p</i> -value)	R ²	I ² , %
Delivery modes		0	92.29
behavioral* (intercept)-reference	0.8574 ([0.3201, 1.3948], <i>p</i> = 0.0043)		
behavioral and feedback	0.0598 ([−1.2193, 1.3388], <i>p</i> = 0.9211)		
behavioral and instructional	−0.2387 ([−0.9150, 0.4377], <i>p</i> = 0.4595)		
home, behavioral, and instructional	−0.4428 ([−1.3639, 0.4782], <i>p</i> = 0.3179)		
instructional	−0.3722 ([−1.3357, 0.5912], <i>p</i> = 0.4190)		
instructional and feedback	0.1622 ([−0.6611, 0.9855], <i>p</i> = 0.6774)		
instructional and live performance	−0.3481 ([−1.5142, 0.8179], <i>p</i> = 0.5301)		
instructional, behavioral, and feedback	−0.0276 ([−1.3390, 1.2837], <i>p</i> = 0.9644)		
live performance	−0.4486 ([−1.6364, 0.7392], <i>p</i> = 0.4292)		
Providers		3.61%	90.27
Professionals* (intercept)-reference	0.7932 ([0.5534, 1.0331], <i>p</i> < 0.0001)		
Volunteers	−0.2011 ([−0.8153, 0.4131], <i>p</i> = 0.5004)		
Professionals and volunteers	−0.2588 ([−0.8335, 0.3158], <i>p</i> = 0.3565)		
Professionals and parents	−0.7650 ([−1.6841, 0.1541], <i>p</i> = 0.0974)		
Intervention dosage		1.46%	92.45
Up to 60 min (1) versus 60 min and more (0)	0.2738 ([−0.2233, 0.7708], <i>p</i> = 0.2649)		
Intervention intensity		24.11%	89.92
Up to 3 sessions (1) versus more than 3 sessions (0)	−0.6637 ([−1.1554, −0.1720], <i>p</i> = 0.0106)	9.85%	91.36
Up to 4 sessions (1) versus more than 4 sessions (0)	−0.5206 ([−1.0929, 0.0516], <i>p</i> = 0.0724)	0	92.86
Up to 5 sessions (1) versus more than 5 sessions (0)	0.0946 ([−0.6554, 0.8447], <i>p</i> = 0.7956)	0	92.40
Tailored (1) versus untailored (0)	0.1559 ([−0.3050, 0.6167], <i>p</i> = 0.4895)	0	92.50
Modification (1) versus no modification (0)	0.2787 ([−0.2825, 0.8400], <i>p</i> = 0.3134)	0	92.84
Training (1) versus no training (0)	−0.1321 ([−0.6323, 0.3682], <i>p</i> = 0.6179)	0	92.62
Fidelity assessment (1) versus no fidelity assessment (0)	−0.1761 ([−0.9003, 0.5481], <i>p</i> = 0.6184)	0	

Table 7. Results from Meta-regression for Content Components.

Content component	β -value (95% CI, <i>p</i> -value)	R ²	I ² , %
Defining CSA	−0.0922 ([−0.5694, 0.3850], <i>p</i> = 0.6919)	0	92.68
Private parts of the body	0.0126 ([−0.4193, 0.4445], <i>p</i> = 0.9522)	0	92.74
Safe and unsafe touches	−0.0801 ([−0.6286, 0.4683], <i>p</i> = 0.7642)	0	92.49
Self-protection skills	0.7238 ([0.5161, 0.9314], <i>p</i> = <.0001)	0	92.58
Children are not the ones to blame	0.1019 ([−0.3226, 0.5264], <i>p</i> = 0.6227)	0	92.59
Safe and unsafe situations	0.2439 ([−0.1766, 0.6643], <i>p</i> = 0.2412)	1.80%	92.10
Awareness of body rights	−0.0876 ([−0.5119, 0.3367], <i>p</i> = 0.6722)	0	92.67
Offender characteristics and approaches	0.1161 ([−0.3312, 0.5634], <i>p</i> = 0.5949)	0	92.71
Good and bad secrets	0.0010 ([−0.4259, 0.4280], <i>p</i> = 0.9960)	0	92.66
Who to tell	0.0124 ([−0.5414, 0.5662], <i>p</i> = 0.9633)	0	92.21

Table 8. Results from Meta-regression for Participant Characteristics.

Participants Characteristics	β (95% CI, <i>p</i> -value)	R ²	I ² , %
Up to 8 years old (1); 8 years and older (0)	Questionnaire-based measures: −0.4574 ([−0.8203, −0.0946], <i>p</i> = 0.0162) Vignette-based measures: 0.0323 ([−0.4295, 0.4941], <i>p</i> = 0.8792)	21.65%	79.63
No males (<5%) (1); 50–60% male (2); Mostly males (>60%) (3); All males (>95%) (4)	−0.2552 ([−1.1506, 0.6402], <i>p</i> = 0.5542)	0	88.72

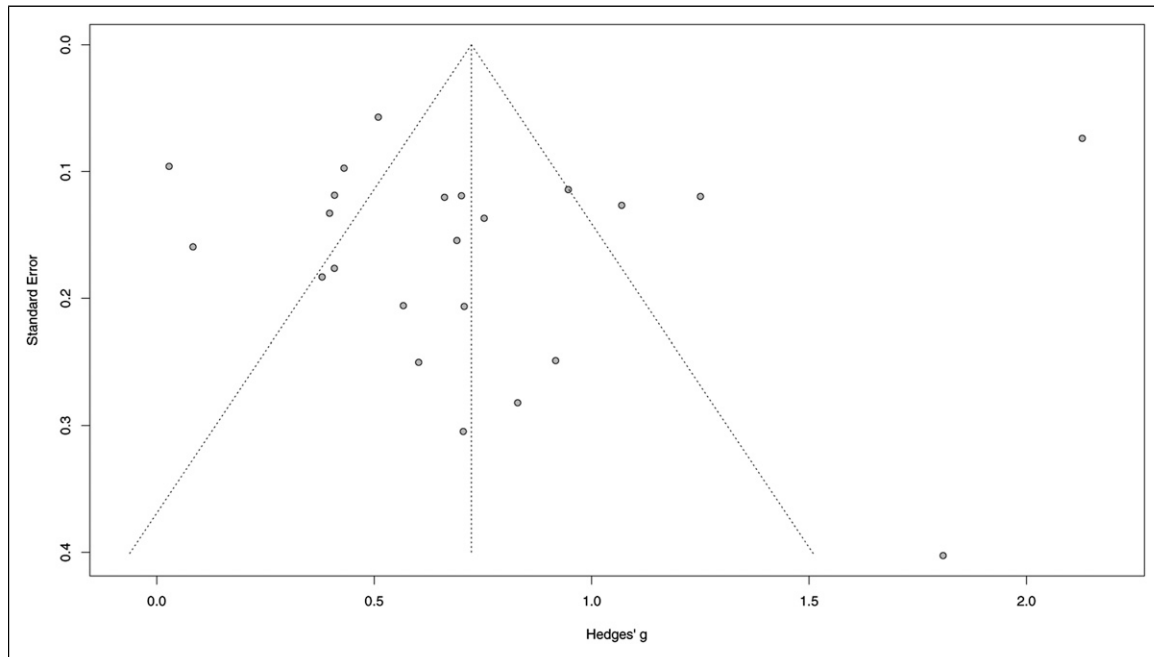


Figure 7. Funnel plot used to explore publication bias.

such as fidelity assessment and the program's theory of change due to poor reporting of the characteristics of interventions in many study reports. We were also not able to conduct exploratory multiple meta-regression to investigate the association between different combinations of intervention content components due to the high heterogeneity across studies. Furthermore, the results of the meta-regression should be seen as exploratory because intervention components were coded based on limited information regarding the theoretical basis of the CSA intervention components in study reports. In addition, the relationship explained by meta-regression analysis is an observational association across trials and it does not indicate causation (Thompson & Higgins, 2002). Finally, it should also be noted that meta-regression will always be subject to the risk of ecological fallacy because it attempts to make inferences about individuals using study-level information (Higgins & Green, 2011).

Implications for Research and Practice Implications

Despite the limitations, our findings have important research, practice, and policy implications. First, clearer reporting of implementation fidelity and intervention components is needed in order to rigorously control for potential sources of confounding. Second, it is important to investigate the program's theory of change as a moderator of effect size. Future studies should clearly report program theory and further explore the heterogeneity across interventions. Third, researchers need to identify additional moderators and intervention components/moderators that could explain the heterogeneity that was identified across studies, particularly in terms of identifying the

factors associated with greater effectiveness in terms of school-based CSA prevention programs. It is also important to include follow-up assessments in a longer-term assessments in future studies, as immediate post-test outcomes might potentially over/underestimate intervention effectiveness. Finally, the majority of included studies in this review were conducted in Western countries and only a limited numbers of studies were conducted in LMICs and upper-middle income countries such as China. In total, we found seven new studies (from eight reports, one of which was conducted in China). Specifically, in terms of the two studies conducted in China (Jin et al., 2017; Lee & Tang, 1998), both had imported programs from the United States (*Body Safety Training* and *Behavioral Safety Training*), indicating the need for the development of programs that are appropriately designed for rather than adapted to Chinese cultural contexts. Furthermore, more evaluation studies in LMICs and upper-middle income countries such as China are needed in future to strengthen context-specific evidence of school-based CSA intervention effectiveness.

In terms of optimizing the delivery of school-based CSA interventions, professionals and policy makers should focus on ensuring that programs being offered include components that appear to be associated with greater effectiveness such as *safe situations and unsafe situations*. Program developers could self-audit their programs against these components to ensure they are delivering programs that have the best chance of success.

Conclusion

CSA is a global issue that has significant negative effects on victims' physical, psychological, and sexual well-being. The

findings of the review suggest that school-based CSA interventions are effective in increasing participants' CSA knowledge as assessed by both questionnaire-based and vignette-based measures. This review provides the CSA prevention field with a comprehensive overview of the components in school-based CSA prevention programs that are associated with better outcomes. According to our findings, an effective CSA intervention should include at least three or more sessions that cover the content component “safe and unsafe situations.” Program developers should take participants' age into consideration when designing the interventions and use age-appropriate outcome measures to evaluate children's performance. Our findings also provide recommendations for future research, particularly in terms of optimizing the effectiveness of school-based CSA prevention programs, and the better reporting of intervention components as well as participant characteristics.

Author's Note

This review was registered with PROSPERO (ID: CRD 42019140564).

Declaration of Conflicting Interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: FM was supported by an Economic and Social Research Council (ESRC) Future Research Leader Award [ES/N017447/1] and the UKRI Global Challenges Research Fund (GCRF) [ES/S008101/1].

ORCID iD

Mengyao Lu  <https://orcid.org/0000-0002-4338-8418>

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Baker, W. L., Michael White, C., Cappelleri, J. C., Kluger, J., & Coleman, C. I., From the HealthOutcomes, Policy, and Economics (HOPE) Collaborative Group. (2009). Understanding heterogeneity in meta-analysis: the role of meta-regression. *International Journal of Clinical Practice*, 63(10), 1426–1434. <https://doi.org/10.1111/j.1742-1241.2009.02168.x>
- Barth, J., Bernmetz, L., Heim, E., Trelle, S., & Tonia, T. (2013). The current prevalence of child sexual abuse worldwide: A systematic review and meta-analysis. *International Journal of Public Health*, 58(3), 469–483. <https://doi.org/10.1007/s00038-012-0426-1>
- *Blumberg, E. J., Chadwick, M. W., Fogarty, L. A., Speth, T. W., & Chadwick, D. L. (1991). The touch discrimination component of sexual abuse prevention training: Unanticipated positive consequences. *Journal of Interpersonal Violence*, 6(1), 12–28. <https://doi.org/10.1177/088626091006001002>
- Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. *The Handbook of Research Synthesis and Meta-Analysis*, 2, 221–235. <https://doi.org/10.7758/9781610448864.14>
- Bronson, L. M. (2019). *Preventing child sexual abuse: A prevention program for elementary school students*. (Publication No. 13812317) [Master's thesis, California State University, Long Beach, United States]. ProQuest UMI Dissertations and Theses Global.
- *Bustamante, G., Andrade, M. S., Mikesell, C., Cullen, C., Endara, P., Burneo, V., & Grunauer, M. (2019). “I have the right to feel safe”: Evaluation of a school-based child sexual abuse prevention program in Ecuador. *Child Abuse & Neglect*, 91, 31–40. <https://doi.org/10.1016/j.chiabu.2019.02.009>
- *Çeçen-Eroğu, A. R., & Hasırcı, K. (2013). The effectiveness of psycho-educational school-based child sexual abuse prevention training program on Turkish elementary students. *Educational Sciences: Theory and Practice*, 13(2), 725–729. <https://files.eric.ed.gov/fulltext/EJ1017300.pdf>
- *Chen, Y. C., Fortson, B. L., & Tseng, K. W. (2012). Pilot evaluation of a sexual abuse prevention program for Taiwanese children. *Journal of Child Sexual Abuse*, 21(6), 621–645. <https://doi.org/10.1080/10538712.2012.726699>
- Chorpita, B. F., & Daleiden, E. L. (2007). *Biennial report: Effective psychosocial interventions for youth with behavioral and emotional needs*. Child and Adolescent Mental Health Division, Hawaii Department of Health. <https://doi.org/10.1037/e405592008-002>
- *Crowley, K. J. (1989). *Evaluation of good touches/bad touches: A program to prevent child sexual abuse in school-age children* (Publication No. 9011660) [Psy.D dissertation, University of California, Irvine, United States]. ProQuest UMI Dissertations and Theses Global.
- *Daigneault, L., Hébert, M., McDuff, P., & Frappier, J. Y. (2012). Evaluation of a sexual abuse prevention workshop in a multicultural, impoverished urban area. *Journal of Child Sexual Abuse*, 21(5), 521–542. <https://doi.org/10.1080/10538712.2012.703291>
- *Dake, J. A., Price, J. H., & Murnan, J. (2003). Evaluation of a child abuse prevention curriculum for third-grade students: assessment of knowledge and efficacy expectations. *Journal of School Health*, 73(2), 76–82. <https://doi.org/10.1111/j.1746-1561.2003.tb03576.x>
- Davis, M. K., & Gidycz, C. A. (2000). Child sexual abuse prevention programs: A meta-analysis. *Journal of Clinical Child Psychology*, 29(2), 257–265. https://doi.org/10.1207/s15374424jccp2902_11
- *Dawson, G. F. (1987). *An evaluation of cognitive and affective outcomes of a prevention program for childhood sexual abuse* [Unpublished Doctoral Thesis, Memphis State University, Tennessee, United States]. ProQuest UMI Dissertations and Theses Global.
- *del Campo Sánchez, A., & Sánchez, F. L. (2006). Evaluación de un programa de prevención de abusos sexuales a menores en Educación Primaria [Evaluation of school-based child sexual abuse prevention program]. *Psicothema*, 18(1), 1–8. <http://www.psicothema.com/english/psicothema.asp?id=3168>

- DerSimonian, R., & Laird, N. (2015). Meta-analysis in clinical trials revisited. *Contemporary Clinical Trials*, *45*, 139–145. <https://doi.org/10.1016/j.cct.2015.09.002>
- Elliott, J. H., Synnot, A., Turner, T., Simmonds, M., Akl, E. A., McDonald, S., & Pearson, L. (2017). Living systematic review: 1. Introduction—the why, what, when, and how. *Journal of Clinical Epidemiology*, *91*, 23–30. <https://doi.org/10.1016/j.jclinepi.2017.08.010>
- Fang, X., Fry, D. A., Ji, K., Finkelhor, D., Chen, J., Lannen, P., & Dunne, M. P. (2015). The burden of child maltreatment in China: a systematic review. *Bulletin of the World Health Organization*, *93*(3), 176–185C. <https://doi.org/10.2471/blt.14.140970>
- Finkelhor, D. (2009). The prevention of childhood sexual abuse. *The Future of Children*, *19*(2), 169–194. <https://doi.org/10.1353/foc.0.0035>
- *Fryer, G. E. Jr, Kraizer, S. K., & Miyoshi, T. (1987a). Measuring actual reduction of risk to child abuse: A new approach. *Child Abuse & Neglect*, *11*(2), 173–179. [https://doi.org/10.1016/0145-2134\(87\)90055-x](https://doi.org/10.1016/0145-2134(87)90055-x)
- Garner, P., Hopewell, S., Chandler, J., MacLehose, H., Akl, E. A., Beyene, J., & Schünemann, H. J. (2016). When and how to update systematic reviews: Consensus and checklist. *Bmj*, *354*. <https://doi.org/10.1136/bmj.i3507>
- *Grendel, M. A. (1991). *Cognitive and emotional effects of a brief child sexual abuse prevention program for first graders* (Publish No. 9205491) [Doctoral dissertation, University of Cincinnati, Ohio, United States]. ProQuest UMI Dissertations and Theses Global.
- Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). *Doing meta-analysis in R: A hands-on guide*. PROTECT Lab Erlangen.
- *Harvey, P., Forehand, R., Brown, C., & Holmes, T. (1988). The prevention of sexual abuse: Examination of the effectiveness of a program with kindergarten-age children. *Behavior Therapy*, *19*(3), 429–435. [https://doi.org/10.1016/s0005-7894\(88\)80014-5](https://doi.org/10.1016/s0005-7894(88)80014-5)
- Hattie, J., & Yates, G. C. (2013). *Visible learning and the science of how we learn*. Routledge. <https://doi.org/10.4324/9781315885025>
- *Hazzard, A., Webb, C., Kleemeier, C., Angert, L., & Pohl, J. (1991). Child sexual abuse prevention: Evaluation and one-year follow-up. *Child Abuse & Neglect*, *15*(1–2), 123–138. [https://doi.org/10.1016/0145-2134\(91\)90097-w](https://doi.org/10.1016/0145-2134(91)90097-w)
- *Hébert, M., Lavoie, F., Piché, C., & Poitras, M. (2001). Proximate effects of a child sexual abuse prevention program in elementary school children. *Child Abuse & Neglect*, *25*(4), 505–522. [https://doi.org/10.1016/s0145-2134\(01\)00223-x](https://doi.org/10.1016/s0145-2134(01)00223-x)
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, d5928. <https://doi.org/10.1136/bmj.d5928>
- Higgins, J., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. <http://www.handbook.cochrane.org>
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., & Lamb, S. E. (2014). Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, *348*, g1687. <https://doi.org/10.1136/bmj.g1687>
- Holloway, J. L., & Pulido, M. L. (2018). Sexual abuse prevention concept knowledge: Low income children are learning but still lagging. *Journal of Child Sexual Abuse*, *27*(6), 642–662. <https://doi.org/10.1080/10538712.2018.1496506>
- *Jin, Y., Chen, J., Jiang, Y., & Yu, B. (2017). Evaluation of a sexual abuse prevention education program for school-age children in China: a comparison of teachers and parents as instructors. *Health Education Research*, *32*(4), 364–373. <https://doi.org/10.1093/her/cyx047>
- *Jones, C., Scholes, L., Rolfe, B., & Stieler-Hunt, C. (2020). A serious-game for child sexual abuse prevention: An evaluation of orbit. *Child Abuse & Neglect*, *107*, Article 104569. <https://doi.org/10.1016/j.chiabu.2020.104569>
- Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*, *36*(4), 567–589. <https://doi.org/10.1007/s10802-007-9201-9>
- Kenny, M. C., Capri, V., Ryan, E. E., & Runyon, M. K. (2008). Child sexual abuse: From prevention to self-protection. *Child Abuse Review*, *17*(1), 36–54. <https://doi.org/10.1002/car.1012>
- Kim, S. J., & Kang, K. A. (2017). Effects of the child sexual abuse prevention education (C-SAPE) program on south Korean fifth-grade students' competence in terms of knowledge and self-protective behaviors. *The Journal of School Nursing*, *33*(2), 123–132. <https://doi.org/10.1177/1059840516664182>
- *Kolko, D. J., Moser, J. T., & Hughes, J. (1989). Classroom training in sexual victimization awareness and prevention skills: An extension of the Red Flag/Green Flag people program. *Journal of Family Violence*, *4*(1), 25–45. <https://doi.org/10.1007/bf00985655>
- Kopp, B., & Miltenberger, R. G. (2009). Evaluating the acceptability of four versions of a child sexual abuse prevention program. *Child & Family Behavior Therapy*, *31*(3), 192–202. <https://doi.org/10.1080/07317100903099183>
- *Krahé, B., & Knappert, L. (2009). A group-randomized evaluation of a theatre-based sexual abuse prevention programme for primary school children in Germany. *Journal of Community & Applied Social Psychology*, *19*(4), 321–329. <https://doi.org/10.1002/casp.1009>
- *Kraizer, S. (1991). *The Safe Child Program for the prevention of child abuse: Development and evaluation of a school-based curriculum*. [Unpublished doctoral dissertation, The Union Institute, Ohio, United States].
- *Lee, Y. K. S., & Tang, C. S. K. (1998). Evaluation of a sexual abuse prevention program for female Chinese adolescents with mild mental retardation. *American Journal on Mental Retardation*, *103*(2), 105–116. [https://doi.org/10.1016/s0145-2134\(98\)00124-0](https://doi.org/10.1016/s0145-2134(98)00124-0)
- MacMillan, H. L., MacMillan, J. H., Offord, D. R., Griffith, L., & MacMillan, A. (1994). Primary prevention of child physical

- abuse and neglect: A critical review. Part I. *Journal of Child Psychology and Psychiatry*, 35(5), 835–856. <https://doi.org/10.1111/j.1469-7610.1994.tb02298.x>
- Manglio, R. (2009). The impact of child sexual abuse on health: A systematic review or reviews. *Clinical Psychology Review*, 29(7), 647–657. <https://doi.org/10.1016/j.cpr.2009.08.003>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica: Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/bm.2012.031>
- Mendeley. (2018). *Elsevier. (2018). Mendeley* (Version 1.19.3) [Computer Software]. https://www.mendeley.com/?interaction_required=true
- Moher, D., & Chan, A. W. (2014). SPIRIT (standard protocol items: recommendations for interventional trials). *Guidelines for reporting health research: A user's manual*, pp. 56–67. <https://doi.org/10.1002/9781118715598.ch7>
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., & Altman, D. G. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ(Clinical Research Ed)*, 340, Article c332. <https://doi.org/10.1136/bmj.c869>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269. <https://doi.org/10.1371/journal.pmed.1000097>
- *Morris, M. C., Kouros, C. D., Janecek, K., Freeman, R., Mielock, A., & Garber, J. (2017). Community-level moderators of a school-based childhood sexual assault prevention program. *Child Abuse & Neglect*, 63, 295–306. <https://doi.org/10.1016/j.chiabu.2016.10.005>
- *Müller, A. R., Röder, M., & Fingerle, M. (2014). Child sexual abuse prevention goes online: Introducing “Cool and Safe” and its effects. *Computers & Education*, 78, 60–65. <https://doi.org/10.1016/j.compedu.2014.04.023>
- National Institute of Justice. (2012). *School-based child sexual abuse prevention programs*. <https://www.crimesolutions.gov/PracticeDetails.aspx?ID=28>
- *Oldfield, D., Hays, B. J., & Megel, M. E. (1996). Evaluation of the effectiveness of Project Trust: an elementary school-based victimization prevention strategy. *Child Abuse & Neglect*, 20(9), 821–832. [https://doi.org/10.1016/0145-2134\(96\)00070-1](https://doi.org/10.1016/0145-2134(96)00070-1)
- Pereznieto, P., Montes, A., Routier, S., & Langston, L. (2014). The costs and economic impact of violence against children. *Child-Fund International* (September), 8–10. <https://shorturl.at/kwDFX>
- *Poche, C., Yoder, P., & Miltenberger, R. (1988). Teaching self-protection to children using television techniques. *Journal of Applied Behavior Analysis*, 21(3), 253–261. <https://doi.org/10.1901/jaba.1988.21-253>
- *Pulido, M. L., Dauber, S., Tully, B. A., Hamilton, P., Smith, M. J., & Freeman, K. (2015). Knowledge gains following a child sexual abuse prevention program among urban students: A cluster-randomized evaluation. *American Journal of Public Health*, 105(7), 1344–1350. <https://doi.org/10.2105/ajph.2015.302594>
- RStudio Team. (2015). *RStudio, Inc* [Computer Software]. <http://www.rstudio.com/>
- Sanderson, J. (2004). Child-focused sexual abuse prevention programs. Crime and Misconduct Commission: Research and Issues Paper Series, 50, pp. 1–8. <http://shorturl.at/koTW3>
- *Saslowsky, D. A., & Wurtele, S. K. (1986). Educating children about sexual abuse: implications for pediatric intervention and possible prevention. *Journal of Pediatric Psychology*, 11(2), 235–245. <https://doi.org/10.1093/jpepsy/11.2.235>
- *Snyder, C. M. (1986). *An evaluation of the “good secrets, bad secrets” assault prevention program* (Publish No. 8705048) [PhD dissertation, International University, California, United States].
- Stoll, C. R., Izadi, S., Fowler, S., Green, P., Suls, J., & Colditz, G. A. (2019). The value of a second reviewer for study selection in systematic reviews. *Research Synthesis Methods*, 10(4), 539–545. <https://doi.org/10.1002/jrsm.1369>
- Sutcliffe, K., Thomas, J., Stokes, G., Hinds, K., & Bangpan, M. (2015). Intervention component analysis (ICA): a pragmatic approach for identifying the critical features of complex interventions. *Systematic Reviews*, 4(1), 140. <https://doi.org/10.1186/s13643-015-0126-z>
- Thompson, S. G., & Higgins, J. P. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21(11), 1559–1573. <https://doi.org/10.1002/sim.1187>
- *Tutty, L. M. (1997). Child sexual abuse prevention programs: Evaluating who do you tell. *Child Abuse & Neglect*, 21(9), 869–881. [https://doi.org/10.1016/s0145-2134\(97\)00048-3](https://doi.org/10.1016/s0145-2134(97)00048-3)
- *Urbann, K., Bienstein, P., & Kaul, T. (2020). The evidence-based sexual abuse prevention program: strong with Sam. *The Journal of Deaf Studies and Deaf Education*, 25(4), 421–429. <https://doi.org/10.1093/deafed/ena019>
- Walsh, K., Zwi, K., Woolfenden, S., & Shlonsky, A. (2015). *School-based education programmes for the prevention of child sexual abuse*. The Cochrane Library. <https://doi.org/10.1002/14651858.cd004380.pub3>
- *Wolfe, D. A., MacPherson, T., Blount, R., & Wolfe, V. V. (1986). Evaluation of a brief intervention for educating school children in awareness of physical and sexual abuse. *Child Abuse & Neglect*, 10(1), 85–92. [https://doi.org/10.1016/0145-2134\(86\)90036-0](https://doi.org/10.1016/0145-2134(86)90036-0)
- *Wurtele, S. K., Saslowsky, D. A., Miller, C. L., Marrs, S. R., & Britcher, J. C. (1986). Teaching personal safety skills for potential prevention of sexual abuse: A comparison of treatments. *Journal of Consulting and Clinical Psychology*, 54(5), 688. <https://doi.org/10.1037/0022-006x.54.5.688>