



# Effectiveness of social-emotional development programs applied in primary education: an umbrella systematic review

Eliana M. Moreno<sup>1,2\*</sup>, Ana Montero<sup>2</sup>, and José M. Armada-Crespo<sup>2</sup>

*1 Department of Psychology-University of Córdoba; Maimonides Institute for Research in Biomedicine of Córdoba (IMIBIC); Reina Sofía University Hospital (Spain).*

*2 Faculty of Educational Sciences and Psychology, University of Córdoba (Spain)*

**Título:** Eficacia de los programas de desarrollo socioemocional aplicados en educación primaria: Una revisión sistemática paraguas.

**Resumen:** En la actualidad ha aumentado el desarrollo de programas de entrenamiento en habilidades socioemocionales en la infancia, debido a que se los considera como una herramienta válida para la adaptación y afrontamiento de una gran variedad de situaciones, tanto académicas como personales. Sin embargo, son escasos los estudios que aporten una visión integral de las evidencias disponibles en el contexto de educación primaria. Se presenta una revisión sistemática de tipo paraguas basada en el método PRISMA, que incluye revisiones sobre programas de desarrollo socioemocional aplicados en educación primaria, con el objetivo de sintetizar sus características y recopilar los principales resultados reportados. Se utilizaron las bases de datos: ERIC, WOS, PSYCINFO, SCOPUS y COCHRANE. Tras un proceso por pares ciegos se seleccionaron y analizaron 15 revisiones. Utilizando las herramientas AMSTAR-2 y SANRA se encontró que el 60% de los estudios secundarios presenta una calidad críticamente baja o baja. Se identificaron 39 programas reportados en revisiones de buena calidad, un 51.2% presentaron evidencias moderadas o fuertes e informaron efectos significativos principalmente en ajuste del comportamiento, competencia social y emocional y habilidades académicas. Se discute el impacto de la calidad metodológica encontrada y las evidencias reportadas en la interpretación y generalización de los hallazgos.

**Palabras clave:** Habilidades socioemocionales. Competencias socioemocionales. Educación primaria. Escuela. Programas basados en la evidencia. Revisión sistemática.

**Abstract:** Currently, the development of training programs in socioemotional skills in childhood has increased because they are considered as a valid tool for adaptation and coping with a variety of situations, both academic and personal. However, there are few studies that show a comprehensive view of available evidences. This research presents an umbrella review based on PRISMA method guidelines. It includes reviews on socioemotional development programs applied in Primary Education with the aim of synthesizing their characteristics and compiling the main results on their effectiveness. The following databases were used: ERIC, WOS, PSYCINFO, SCOPUS and COCHRANE. After a blind peer process, 15 reviews that met the inclusion criteria were selected and analysed. Using the AMSTAR-2 and SANRA tools, it was found that 60% of secondary studies have critically low or low quality. Thirty nine programs reported in good quality reviews were identified, 51.2% presented moderate or strong evidence and reported significant effects mainly on behavioral adjustment, social and emotional competencies and academic skills. The impact of the methodological quality found and the evidences on the interpretation and generalization of the findings is discussed.

**Keywords:** Socioemotional skills. Socioemotional competencies. Primary education. School. Evidence based training programs. Systematic review.

## Introduction

Social-emotional skills refer to a set of learned behaviors that are put to work when we interact with others. These skills enable us to express feelings, attitudes, opinions, and to defend personal rights (Zins et al., 2000). Alzahrani et al. (2019) add that they refer not only to how one interacts with others, but also to the ability to manage own's own emotions and adequately respond to events that happen around us. For their part, Wu et al. (2018) indicate that development of social and emotional competence during early childhood is essential to ensuring preparedness for school, and is associated with many important aspects of life, including academic achievement, behavior, and mental and physical health. Improvement in these competencies is associated with better academic achievement and with positive social behaviors (Bear & Watkins, 2006).

Adequate development of these skills requires the individual's involvement in a long-term process of learning specific competencies and skills. Therefore, even though these skills can be worked on at any age, it is more productive to

introduce them in early childhood, and specifically in the school context (Björklund et al., 2014; Cornell et al., 2017; Diamond & Lee, 2011; Durlak et al., 2010). The stage of primary education is considered a critical period when boys and girls can have different opportunities to practice emotional control and deal with their emotions and social interactions, ensuring continued learning throughout later years of schooling. As one of the most powerful agents of socialization, school contributes not only to the transmission of knowledge but also to whole-person development (Grajales, 2003). Together with other socialization agents, such as family and society, school can play an important role in promoting the learning of social-emotional competencies by taking specific actions.

In recent years, many interventions have been proposed to promote these skills in the school environment. In this line, a wide range of programs have been integrated into the curriculum under the name of Social-Emotional Learning Programs (SEL) (Low et al., 2015). According to Durlak et al. (2011), these interventions can be considered to facilitate the process of developing a set of basic cognitive, affective and behavioral competencies, for managing and recognizing emotions, achieving one's purposes, acknowledging others' perspectives, establishing positive interpersonal relation-

**\* Correspondence address [Dirección para correspondencia]:**

Eliana M. Moreno. Calle San Alberto Magno, s/n, 14071 Córdoba (Spain).

E-mail: [eliana.moreno@uco.es](mailto:eliana.moreno@uco.es)

(Article received: 30/07/2021; revised: 19/04/2022, accepted: 20/05/2022)

ships, making responsible decisions, and managing relationships with others in a positive way.

International agencies that collect evidence on interventions, such as the Education Endowment Foundation (EEF) (2021), point out that in recent years more studies have been conducted in primary schools than in secondary schools. In England, several studies have posited a relationship between SEL interventions and academic outcomes. In the United States, the Every Student Succeeds Act (ESSA) (2021) reports on the effectiveness of certain social skills programs, citing results related to academic issues, disruptive behavior, interpersonal relationships, bullying, and so on. In this line, Fernández et al. (2014) report data on the effectiveness of 25 programs that target a population from early childhood education and primary education, finding that 45% of them showed significant differences in favor of improved social competence, and 30.30% showed significant values in reducing behavioral problems, as well as other effects.

Several previous studies mention difficulties in categorizing programs and analyzing their effects, due to a lack of agreement on which skills are being taught, variability in reported results, and a lack of consensus in the foundational theoretical models (Rubiales et al., 2018). The What Works Clearinghouse (WWC) (2021) establishes five categories of skills related to SEL, mentioning social awareness, self-knowledge, self-management, responsible decision-making, and relationship management. Fernández et al. (2014) and Pandey et al. (2018) take a different line, classifying programs according to their methodological or procedural characteristics. Elsewhere, McCallops et al. (2019) mention empathy, self-awareness, awareness of others, self-regulation, and motivation. A similar categorization system was established by Goldberg et al. (2019) and Sklad et al. (2012), who organize programs according to their results, grouping them into categories of social and emotional adjustment, behavioral adjustment, scholastic achievement, and internalizing symptoms. This system will be taken as reference in the present study, as it allows greater flexibility for integrating the diversity of information found in these studies.

There is a lack of consensus in categorizing the different programs and in the dimensions comprised, along with great diversity in objectives, designs, methodologies and presentation of results (Rubiales et al., 2018). Many authors have tried to synthesize the existing information in secondary studies that offer a systematic or literature review (Oros et al., 2011; Pérez-Escoda et al., 2012). However, there is still little information on program characteristics and effects organized by stage of education, and there is much variability in terms of the review quality and presentation of the information. The evidence reflected in secondary studies on the subject must be analyzed to obtain a broader and more integrated framework of understanding.

The aim of this study is to identify reviews and meta-analyses that include primary studies on programs to develop social-emotional skills within the school context, at the primary education level; to evaluate the methodological charac-

teristics of these reviews; and to provide a comprehensive synthesis of the programs and their results. The following review questions will be explored:

RQ1: What are the methodological characteristics and quality of the reviews and meta-analyses identified?

RQ2: What is the level of evidence for programs identified in good quality reviews?

RQ3: For programs with higher levels of evidence, what are the main effects reported?

RQ4: For programs with higher levels of evidence, what characteristics do they have (skills taught, components, teaching-learning methodology)?

## Method

An umbrella-type systematic review was carried out, including only other reviews or meta-analyses on one specific topic, to be studied and synthesized (Aromataris et al., 2015). The PRISMA methodology (Page et al., 2021) was followed, and we applied the specific recommendations of the Joanna Briggs Institute (JBI) and other relevant authors for conducting this type of review of reviews (Aromataris et al., 2015; Smith et al., 2011). The protocol meets the criteria required by the International Prospecting of Systematic Reviews (PROSPERO), and has been registered under the code: CRD42021239217. The entire process of identification, selection and evaluation of studies was carried out by two independent reviewers; after comparing their results, any doubtful cases were resolved by consensus or by appealing to a third reviewer.

### Search strategy

The following databases were used: ERIC, WOS, PSYCINFO, SCOPUS and COCHRANE. Key words in English and Spanish were used to identify the resources, making different combinations with Boolean operators: ("meta-analysis"; review; "systematic Review") AND ("emotional skills"; "emotional competences"; "socio emotional"; "social emotional"; "social skills") AND (school; "primary school"; "elementary school"; Child\*) AND (program\*; training; intervention, learning).

### Selection criteria

Inclusion criteria were established as follows: secondary studies of systematic, narrative or meta-analytic reviews, which reported on application of programs for the development of social-emotional skills carried out within the school environment, having a universal, curricular nature, with a primary school population, between the ages of 5 and 12 years. Scientific articles published in English or Spanish between the years 2000 and 2020 were included. The decision to include narrative or literature reviews was based on the fact that this type of study usually provides more detailed information on the characteristics of the interventions.

Our exclusion criteria eliminated reviews that included studies conducted in the clinical setting or outside the school setting, or which used any specifically-indicated population or students with specific characteristics (e.g., special educational needs or identified disorders). We eliminated primary studies and merely descriptive reviews that did not offer results from program application, or that included studies applied in stages other than primary education, or with subjects under 5 years of age or over 12. We also eliminated studies that did not address social-emotional skills as their main element.

Reviews were selected if they met all inclusion criteria and excluded if they met at least one of the exclusion criteria.

### Data extraction and analysis

Following the criteria and recommendations of Aromataris et al. (2015) and Smith et al. (2011), a matrix was designed for data categorization, extraction and analysis. This process was carried out in three phases. In the first phase, data on the general characteristics of all the studies were extracted and analyzed, such as: type of review, type of primary studies included, country and language, period of years covered, population.

In the second phase, the quality of all the selected studies was evaluated. The systematic reviews and meta-analyses were evaluated using the AMSTAR-2 tool (Ciapponi, 2018; Shea et al., 2017), which contains 16 domains that explore general methodological quality, seven of which are considered critical, making it possible to rate confidence in the results of a scale as either High, Medium, Low or Critically Low. The critical domains are: existence and record of a previous protocol (item 2); exhaustive search strategy (item 4); presence of a justified list of excluded studies (item 7); evaluation and discussion of risks for bias (items 9 and 13); statistical analyses and analyses of publication biases (items 11 and 15, which are only applied in meta-analyses). The items are completed with ratings of Yes, No, Partial Yes, Does not apply. Non-systematic literature reviews or narratives were evaluated with the SANRA (Scale for the Assessment of Narrative Review Articles) (Baethge et al, 2019), which is composed of 6 items that assess: justification of the review (item 1); objectives (item 2); search description (item 3); references (item 4); scientific argumentation (item 5), and presentation of the data (item 6). Each of these is scored on a three-value scale from 0 (low level) to 2 (high level); the sum of the scores (maximum 12) provides a measure of the quality of a narrative review article. Scores less than or equal to 6 were considered as low quality, scores between 7-9 as medium and between 10 - 12 as high quality.

In the third phase, a list was made of all the programs assessed as medium- and high-quality reviews; data from each

of these was extracted and grouped under three headings: I) Program results; II) Level of evidence and III) Characteristics. Under Heading I, in addition to results measured, information was collected on methodological aspects of the primary studies, such as sample characteristics, country of application and duration. The results were categorized into five groups, taking as a reference the studies by Goldberg et al. (2019) and Sklad et al. (2012). The primary results were: (a) Social adjustment (Adj-S) and (b) Emotional adjustment (Adj-E), these refer to general or specific measures of skills that had been improved or worked on during the intervention (e.g. social skills, communication, assertiveness; emotion recognition skills, expression and self-regulation). The secondary results were: (c) Behavioral adjustment (Adj-B), which includes positive social behaviors (e.g. cooperation, altruism), as well as behavior problems (disruptive behavior, aggressivity) and risk behaviors (e.g. tobacco and alcohol use); (d) Academic Outcomes (AO), which includes measures of scholastic achievement in different areas and specific skills (e.g. reading); and (e) Mental Health (MH), taking into account measurements of internalizing symptoms (e.g. anxiety, depression, stress) and well-being. In this phase, only two reviews presented clear quantitative results, so it was not possible to perform a quantitative analysis; instead, a qualitative and descriptive synthesis of the results and reported evidence was conducted.

Under Heading II, the programs' "level of evidence" was recorded: quality of the primary studies, instruments, and summary of evidence. Since the reviews used different tools to evaluate the quality of the primary studies and applied different criteria to classify the evidence, it was necessary to unify the criteria for interpretation. For this purpose, we used an ordinal categorization system based on previous work (Gálvez-Lara et al., 2018). In this manner, 5 levels of evidence were established and can be seen in Table 1: "no evidence," when no results from primary studies are reported; "weak," "moderate," or "strong" evidence, according to the quality reported for the primary studies collected; and "inconsistent evidence," when contradictory or unclear results are reported. As needed, when different primary results were found, the highest rating was taken as the reference, and cases of doubt were resolved by consensus.

Finally, taking as a reference the studies by Goldberg et al. (2019), Mackenzie and Williams (2018) and Siddiqui and Ventista (2018), the information collected under Heading III "program characteristics" contained: objectives; skills taught; methodology; theoretical frame of reference; agents who applied the program; components (families, educational community, others); and implementation (resources, reliability).

Qualitative and quantitative data were extracted and analyzed, the latter expressed in terms of frequency, percentages, means, ranges, minimums and maximums.

**Table 1**  
Ordinal scheme to classify the different levels of evidence.

	Goldberg et al. (2019)	Mackenzi & Williams (2018)	Pandey (2018)	Siddiqui & Ventista (2018)	Schlesier et al. (2019)	Weissberg & O'Brien (2004)
Tool	(a)	(b)	(a)	(c)	(d)	Not rated
No evidence	NR	NR	NR	NR	NR	NR
Weak evidence	Weak	0-50% (Poor)	Poor	1-2 (Low)	7-17	-
Moderate evidence	Moderate	51-60% (Reasonable) 61-74% (Good)	Moderate	3 - 4	18-23	“well designed studies”
Strong evidence	Strong	75-100% (Excellent)	High	5 (Strong)	24-28	-

Inconsistent evidence In case of contradiction or lack of clarity between studies

Notes: (a) *Quality assessment tool for quantitative studies*; (b) Downs & Black checklist, (c) Combination of *Maryland Scale of Scientific Method* and other criteria; (d) Own evaluation system; NR, Not reported.

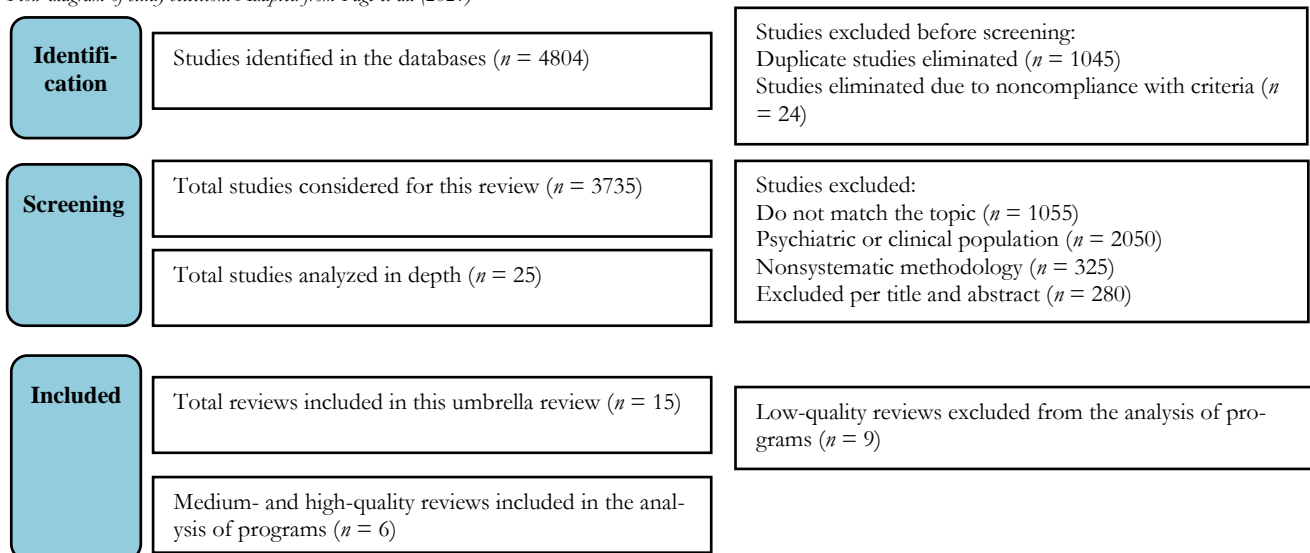
## Results

### Selection process

After searching the different databases, 4804 records were identified. After eliminating duplicates, the search was refined according to the eligibility criteria, resulting in 305

documents. After examining the titles and abstracts of all these documents, 280 were excluded. The full text of the 25 articles was evaluated, and 10 articles that did not meet the criteria were eliminated. Finally, for this investigation, 15 reviews were included in the overall analysis, and 6 were subsequently selected for program analysis on the basis of their quality, as can be seen in Figure 1.

**Figure 1**  
Flow diagram of study selection. Adapted from Page et al. (2021)



### Characteristics of the reviews

#### General characteristics

Table 2 shows the characteristics of the reviews analyzed. The main results show that 8 studies (53.3%) were systematic reviews, 4 (26.6%) were meta-analyses, and 3 (20%) were nonsystematic narrative reviews. Nine studies (60%) were published between the years 2015 and 2020, and 40% ( $n = 6$ ) were from the United States. Twelve studies (80%) included primary studies with experimental or quasi-experimental designs. In 60% of the reviews ( $n = 9$ ), the language of the primary studies was English only. The three narrative re-

views did not report data on the characteristics of the studies included (20%).

Regarding the programs analyzed, all 15 reviews (100%) included curricular programs, targeting a universal population, and were implemented at the level of primary education. Two of these (13.3%) were exclusive to primary education (Schlesier et al., 2019; Siddiqui & Ventista, 2018), and 13 (86.6%) additionally included students from other ages or stages of education. In the 2 reviews that addressed primary education exclusively, the participants' age in the primary studies ranged from 5 to 12 years; in the reviews with a mixed population the ages ranged from 3 to 21.

### Quality of the reviews

Table 2 presents results regarding the quality and critical elements of all the secondary studies evaluated using AMSTAR-2 and SANRA.

Of all the systematic reviews and meta-analyses, 41.6% ( $n = 5$ ) showed critically low quality; they did not meet the critical elements related to presenting an evaluation of biases or presenting a list or justification of excluded studies. Three

systematic reviews (25%) showed high quality, 2 studies (16.6%) had medium quality, with certain deficiencies in discussing the study biases, and 2 meta-analyses (16.6%) showed low quality.

On the other hand, 2 non-systematic reviews or narratives showed low quality, while one of them obtained medium quality. The most important limitations were inadequate presentation of data and a lack of description of search strategies.

**Table 2**

*Characteristics and quality of reviews included*

Secondary study	Primary Study	Target Population	Quality	Items
1° Author (Year); Country; Period covered	Total; Design ( $n =$ number); Language	Kindergarten/ Primary/ Secondary	AMSTAR SANRA	Not meet
<b>Systematic Review</b>				
Canet-Juric et al. (2020) Argentina; 2000 to 2018	Total ( $n = 18$ ); RCT/E ( $n = 9$ ); CE ( $n = 9$ ) Spanish and English	Kindergarten-Primary	Critically low	2*, 4*, 9, 13
Fernández et al. (2014) Spain; 1982 to 2013	RCT/E ( $n =$ NR) Spanish, Portuguese and English	Kindergarten-Primary	Critically low	7, 9, 13
**Mackenzie & Williams, (2018) USA; 2000 to 2016	Total ( $n = 12$ ) RCT/E ( $n = 12$ ) English	Primary-Secondary	High	2*
McCallops et al. (2019) USA; 2008 to 2018	Total ( $n = 51$ ); RCT/E ( $n = 22$ ); CE ( $n = 21$ ); Other ( $n = 8$ ) English	Primary-Secondary	Critically low	2*, 4*, 7, 9, 13
Rubiales et al. (2018) Argentina; 2005 to 2016	Total ( $n = 19$ ); RCT/E/CE Spanish and English	Primary-Secondary	Critically low	7*, 9, 13
Sancassiani et al. (2015) Italy; 2000 to 2014	Total ( $n = 22$ ); RCT/E English	Kindergarten to Secondary	Critically low	7*, 9, 13
**Schlesier et al. (2019) Germany; 2011 to 2016	Total ( $n = 14$ ); RCT/E ( $n = 8$ ); Other ( $n = 6$ ) English	Primary	Medium	13*
**Siddiqui et al. (2018) United Kingdom; 1995 to 2017	Total ( $n = 13$ ); RCT/E ( $n = 13$ ) English	Primary	High	-
<b>Meta-analysis</b>				
Durlak et al. (2011) USA; 1970 to 2007	Total ( $n = 213$ ); RCT/E/CE English	Primary-Secondary	Low	4*, 7
**Goldberg et al. (2019) Netherlands; 1998 to 2017	Total ( $n = 50$ ); RCT/E ( $n = 36$ ); CE ( $n = 14$ ) English	Primary-Secondary	Medium	2*
**Pandey et al. (2018) United Kingdom; 1977 to 2017	Total ( $n = 41$ ); RCT/E ( $n = 32$ ); Otros ( $n = 9$ ) English	Kindergarten to Secon- dary	High	-
Sklad et al. (2012) Netherlands; 1995 to 2008	Total ( $n = 75$ ); RCT/E/CE English	Primary-Secondary	Low	4*, 7, 11*, 15*
<b>Narrative Reviews</b>				
Opengart (2007) USA; Date NR	Total ( $n =$ NR) Language NR	School (Unspecified)	5 (Low)	1, 6
Pelco et al. (2007) USA; Date NR	Total. ( $n =$ NR) Language NR	School (Unspecified)	5 (Low)	3, 6
** Weissberg & O'Brien (2004) USA; Date NR	Total. ( $n =$ NR) Language NR	School (Unspecified)	8 (Medium)	2, 3

Note. (\*\*) Moderate or High quality reviews; (\*) Critical items partially met in AMSTAR; CE, Cuasi Experimental; E, Experimental; NR, Not Reported; RCT, Randomized Controlled Trial.

### Characteristics of the programs

For the program analysis, we considered only those programs that were reported in medium- or high-quality reviews or meta-analyses ( $n = 6$ ). A total of 39 social-emotional development programs, applied in primary education, were counted. Table 3 presents the data on level of evidence, results, and characteristics for all of these.

### Level of evidence

To establish the level of evidence for each program, the quality of the primary studies and the summary of the evidence reported in each review were taken into account; these elements were reported in 5 of the 6 reviews (83.3%). Three systematic reviews used different tools and criteria to summarize the evidence (ordinal, numeric, percentages), and two

meta-analyses used the same instrument, in different versions, and applied similar ordinal criteria. The narrative review did not evaluate the quality of the primary studies. After unifying the criteria for interpreting the evidence (see Table 1), we found that 6 programs (15.3%) were reported with strong evidence; 14 (35.8%) with moderate evidence and 12

(30.7%) with weak evidence. Six programs (15.3%) were considered as having “no evidence”; and one program was rated with “inconsistent evidence”, due to a contradiction between two reviews in the quality analysis of one of the primary studies.

**Table 3**  
Summary of program characteristics and evidence reported in high- and medium-quality reviews

Program name/ Level of evidence	Studies	Outcomes/Quality	Primary Study Data	Program Data
	Review/ Primary study (Design)	Variables (outcomes)/ Primary study quality	Sample $N =$ ; Age or educational level/ Country-application; Duration (Sessions)	Objectives; Skills taught; Methodology; Theoretical frame of reference; Who applied it; Components (family, community, other); Implementation.
<b>AI's Pals</b> Weak	Goldberg et al. (2019)/ 1 (CE) 2 (CE)	Adj-E and Adj-S/ (significant positive effects) /Weak Adj-B; Adj-E and Adj-S/ (significant positive effects) /Weak	$N = 333$ /USA $N = 399$ /USA	<b>C:</b> Involves the family
<b>aRRabbiadiario</b> No evidence	Schlesier et al. (2019)/ 1 (NR)	NR (descriptive analysis) / 10/28	$N =$ NR; 8-11 years.	<b>M:</b> based on a daily anger log
<b>BPY School</b> Weak	Goldberg et al. (2019)/ 1(CE)	Adj-B; Adj-E and Adj-S/ (significant positive effects) /Weak	$N = 3497$ /USA	<b>O:</b> Addresses bullying. <b>C:</b> Involves the family. <b>I:</b> reports reliability problems in the implementation.
<b>Caring school communities</b> (CSC) No evidence	Siddiqui & Ventista (2018)/ 1 (NR) Weissberg & O'Brien (2004)/ 2 (NR)	NR (effect size not calculated due to lack of data) / NR “participant reported less substance use at post test, engaged in more prosocial activities, and more problem-solving compared with control” / NR	$N =$ NR $N = 5000$	<b>O:</b> help primary schools become caring communities, strengthening connections among peers, teachers, students, families, and school. <b>M:</b> Uses four interactive approaches: a) discuss problems affecting classroom climate, seek solutions, plan and make decisions together, b) pair students of different ages in academic and recreational activities, c) organize family involvement activities, d) 15 community activities are carried out, along with optional modules on social and emotional skills. <b>TF:</b> based on the idea of the classroom as a shared community. <b>C:</b> family and community
<b>*Child Development</b> Inconsistent evidence	Goldberg et al.(2019)/ 1 (CE) Goldberg et al. (2019)/ 2** (CE) Siddiqui & Ventista (2018)/ 2**(CE)	Adj-B; Adj-E and Adj-S/ (significant positive effects) /Weak Adj-B; Adj-E; Adj-S; MH (internalizing symptoms) and AO/ (significant positive effects) / Moderate NR (“unreliable data reporting”) / NR	$N = 1334$ /USA $N = 13464$ /USA $N = 13464$ /USA	<b>O:</b> Develop social skills and peer relationships, motivation, self-regulation, and self-esteem. <b>M:</b> relies on three elements: development of discipline through decision making, cooperative learning, and instructions. <b>C:</b> Involves the family and the educational community <b>I:</b> reports reliability problems in the implementation.
<b>Dare to Care</b> Weak	Goldberg et al. (2019)/ 1 (CE)	Adj-B; Adj-E and Adj-S/ (significant positive effects) /Weak	$N = 102$ /Canada	<b>O:</b> Addresses bullying. <b>C:</b> Involves the family.
<b>*FRIENDS</b> Feelings, Relax, I can do it, Explore solutions, Now reward, Don't forget practice, Smile Strong	Mackenzie & Williams (2018)/ 1(NR) Mackenzie & Williams (2018)/ 2(NR)	MH (social and general anxiety)/ (significant improvements with moderate effect size)/ 43.4% / Weak MH (social and general anxiety)/ (significant improvements with small effect size / 75% (+) / Strong	$N = 106$ /England; 10 $N = 1448$ //England; 9 sessions (x60 min)	<b>O and S:</b> Identify feelings of anxiety and learn to relax; identify unhelpful thoughts and replace them with helpful ones; cope with problems and overcome challenges. <b>M:</b> lessons involving large and small group work, workbook exercises, role-plays and games. <b>M:</b> Based on Cognitive Behavioral approach. <b>WA:</b> School nurses, healthcare personnel or trained teachers <b>C:</b> Parents attend a prior informational session <b>I:</b> Those responsible for the intervention receive two days of training.

Program name/ Level of evidence	Studies	Outcomes/Quality	Primary Study Data	Program Data
	Review/ Primary study (Design)	Variables (outcomes)/ Primary study quality	Sample $N =$ ; Age or educative level/ Country-application; Duration (Sessions)	Objectives; Skills taught; Methodology; Theoretical frame of reference; Who applied it; Components (family, community, other); Implementation.
<b>*Health Promoting Schools</b> Moderate	Goldberg et al. (2019)/ 1(RCT)	Adj-E and Adj-S/(significant positive effects / Moderate	$N = 2758$ /Australia	<b>C:</b> Involves the family and the community
<b>I Can Problem Solve (ICPS)</b> No evidence	Pandey (2018)/ 1(NR)	Adj-E (Self-regulation)/(NR) / NR lack of data	$N = \text{NR}$ ; 1°/USA; 1 year	<b>O and S:</b> Self-regulation, problem solving. <b>M:</b> 28 sessions with games, stories, puppets and role-playing, 20 minutes each, integrated in the classroom study plan <b>WA:</b> Trained teachers
<b>*Incredible Years</b> Moderate	Pandey (2018)/ 1(NR)	Adj-B (behavioral problems) and Adj-E (Self-regulation)/ (significant improvements) / Moderate	$N = 1768$ /USA; 5 months	<b>O and S:</b> Self-regulation, conflict resolution. <b>M:</b> 30 sessions per year, with 300 activities in small groups and 100 videos with models of children, for working on social-emotional skills and conflict management strategies. <b>WA:</b> Trained teachers or certified research personnel.
<b>*INSIGHTS</b> Insight into Children's Temperament Moderate	Pandey (2018)/ 1(NR)	Adj-B (behavioral problems); Adj-E (self-regulation); Adj-S (SS); AO and MH/ (significant improvements) / Moderate	$N = 435$ /USA; 10 weeks	<b>O:</b> improve social skills and reduce behavior problems. <b>S:</b> Self-control, attention, behavioral repertoire. <b>M:</b> 2 hours per week for 10 weeks. <b>WA:</b> facilitators with a degree in psychology <b>C:</b> Involves the family.
	Siddiqui & Ventista (2018)/ 2(RCT)	Adj-B (behavior at home)/ (small effect size in favor of the intervention) / 3/5 Moderate	$N = 148$ ; 10 weeks	
<b>Lessons for living: Think Well, Do Well</b> Weak	Mackenzie y Williams (2018)/1(E)	Adj-E and MH (anxiety, avoidance)/ (significant improvements with medium effect size)/ 46.9% / Weak	$N = 317$ / Scotland; 10 sessions	<b>O:</b> Develop coping skills. <b>S:</b> coping, recognizing emotional symptoms, reducing avoidant coping strategies, proactive problem solving and support seeking. <b>M:</b> interactive, guided instruction. <b>TF:</b> Based on the Cognitive Behavioral approach. <b>WA:</b> psychologist or trained teacher <b>I:</b> Teachers receive one day of training and an intervention manual.
<b>MC (Making Choices) y MC+</b> Weak	Schlesier et al. (2019)/ 1(NR)	NR (No positive effects in emotional regulation) / 17/28 / Weak	$N = \text{NR}$ ; 3rd.	<b>TF:</b> Based on the theory of social information processing
<b>Mindfulness Education Programme</b> Weak	Siddiqui & Ventista (2018)/ 1(CE)	Adj-E and Adj-S (Socioemotional Competencies)/ (positive effect size, "weak evaluation") / 2 a 3/5 / Weak	$N = 246$ ; 4th to 6th/England; 10 weeks	<b>O:</b> improve social-emotional health and well-being. <b>M:</b> consists of quieting the mind by sitting in a comfortable position and listening to a single sound. Pupils focus on their breathing, their thoughts and their sensations.
<b>MOM Mindfulness-Orientated Meditation</b> Weak	Schlesier et al. (2019)/ 1(NR)	NR (No positive effects in emotional regulation) 14/28 / Weak	$N = \text{NR}$	<b>O:</b> promote psychological well-being. <b>M:</b> based on meditation.
<b>*Open circle</b> Moderate	Goldberg et al. (2019)/ 1(CE)	Adj-B; Adj-E; Adj-S and AO/ (significant positive effects / Moderate	$N = 154$ / USA	<b>C:</b> Involves the family and the community. <b>I:</b> adequate
<b>*PATHS</b> Promoting Alternative Thinking Strategies Moderate	Goldberg et al. (2019)/ 1(RCT)	Adj-E; Adj-S and AO/ (small effect size) / Moderate	$N = 5081$ /USA	<b>O:</b> provide opportunities to develop emotional awareness, self-control, interpersonal problem solving and peer relationships (Pandey et al, 2018; Weissberg & O'Brien, 2004). <b>S:</b> identify and describe feelings; relax with breathing techniques; take perspective from others; and solve problems with an 11-step model. Study skills and habits are also worked on (listening and
	Mackenzie & Williams (2018)/ 2 (RCT/E)	Adj-B (aggression); Adj-S (Social Competencies, social interaction) and MH (inattention)/ (significant improvements with small effect size) / 68.8% (+) / Moderate	$N = 5075$ /England 2 years (x44 sessions during 1st. year, and x 47 during 2nd)	

Program name/ Level of evidence	Studies	Outcomes/Quality	Primary Study Data	Program Data
	Review/ Primary study (Design)	Variables (outcomes)/ Primary study quality	Sample $N =$ ; Age or educative level/ Country-application; Duration (Sessions)	Objectives; Skills taught; Methodology; Theoretical frame of reference; Who applied it; Components (family, community, other); Implementation.
	Pandey (2018)/ 3(NR)	Adj-E (self-regulation) and AO (verbal fluency)/ (significant improvements) / Moderate	$N = 318$ ; 2nd y 3rd (7-9 years)/USA; 1 year (3x week)	paying attention in class, organization and planning, setting academic goals) (Weissberg & O'Brien, 2004). <b>M</b> : between 30 and 45 lessons per year, in sequence and focusing on techniques such as self-control and interpersonal problem solving. <b>WA</b> : Trained teachers. <b>C</b> : involves the family and the community. <b>I</b> : A two-day training is offered for teachers as well as follow-on and support services. A manual is provided with teacher scripts, images, activity sheets, photos, posters and homework activities. Mackenzie and Williams (2018) and Goldberg et al. (2019) report adequate reliability.
	Schlesier et al. (2019)/ 4(NR)	NR/ 16/28 /Weak	$N = \text{NR}$ ; 3rd to 6th (Mean 8.9years)	
	Siddiqui & Ventista (2018)/ 5(NR)	Adj-E (feelings, knowledge of self and others)/ (positive small effect size)/ 2 a 3 / Weak	$N = 192$ ; 2nd y 3rd.1 year aprox.	
	Weissberg & O'Brien (2004)/ 6(NR) 7(NR)	"Decrease in aggressive and disruptive behaviors, improvements in following rules, expression of emotions, concentration on classroom tasks" / NR ("well designed studies") Moderate	$N = \text{NR}$	
<b>*Philosophy for Children</b> Moderate	Siddiqui & Ventista (2018)/ 1(CE)	Adj-B (cooperation,teamwork) y Adj-S (SS, comunicacion)/ (low effect size in favor of the intervention)/ 3/5 /Moderate	$N = 2722$ ; 4th; England; 1 year.	<b>O and S</b> : Dialogue, reflect, use argumentation <b>M</b> : organized activities for classroom dialogue, thinking and reflecting, justifying beliefs and ideas, developing appropriate language for argumentation, and becoming aware of discussion skills.
<b>*Positive action</b> Moderate	Siddiqui & Ventista (2018)/ 1** (RCT)	NR (effect size not calculated due to lack of data) / NR	$N = 131$ / USA	<b>M</b> : 140 fifteen-minute lessons. A kit of materials is provided, including posters, games and worksheets. <b>C</b> : family and community <b>I</b> : adequate
	Goldberg et al. (2019)/ 1** (RCT)	Adj-B; Adj-E and Adj-S/ (NR) / Moderate	$N = 1170$ / USA	
<b>*Positive Behavioral Interventions and Supports (PBIS)</b> Moderate	Schlesier et al. (2019)/ 1(NR)	Adj-B (externalizing behavior)/ (positive effects, large effect size)/ 21/28/Moderate	$N = \text{NR}$ ; (Mean 9.8 years)	<b>TF</b> : Based on the theory of social learning, cognitive behavioral theory, and applied behavioral analysis.
<b>*Raising Healthy Children</b> Strong	Goldberg et al. (2019)/ 1(RCT)	Adj-B; Adj-E and Adj-S/ (significant positive effects) / Moderate	$N = 938$ / USA	<b>C</b> : Involves the family and the community. <b>I</b> : adequate
	Goldberg et al. (2019)/ 2 (RCT)	Adj-B and AO / (significant positive effects) / Strong	$N = 1040$ /USA	
<b>Rochester Social Problem Solving</b> Weak	Siddiqui & Ventista (2018)/ 1(E)	Adj-S (SS)/ (tlow effect size) / 2 a 3 /Weak	$N = 133$ ; 3rd-4th/Australia; 20 weeks	<b>M</b> : 34 sessions, teaching social skills and how these affect feelings and behavior; situations that occur in real life are discussed in the classroom.
<b>RULER</b> Weak	Siddiqui & Ventista (2018)/ 1(NR)	Adj-E (self-regulation) y Adj-S (SS)/ (significant differences with small to moderate effect size in favor of the intervention) / 2 a 3 / Weak	$N = 343$ ; 5th y 6th/ USA; 1 year	<b>O</b> : promote social, emotional and academic learning, <b>S</b> : recognize emotions in self and others, understand causes and consequences of emotions; label, express and regulate emotions appropriately. <b>M</b> : structured multi-year curriculum, can be applied from early childhood through secondary education.
<b>*School Community Intervention</b> Moderate	Goldberg et al. (2019)/ 1 (RCT)	Adj-B/ (significant positive effects) / Moderate	$N = 1153$ /USA	<b>C</b> : Involves the family and the community.



Program name/ Level of evidence	Studies	Outcomes/Quality	Primary Study Data	Program Data
	Review/ Primary study (Design)	Variables (outcomes)/ Primary study quality	Sample $N =$ ; Age or educative level/ Country-application; Duration (Sessions)	Objectives; Skills taught; Methodology; Theoretical frame of reference; Who applied it; Components (family, community, other); Implementation.
<b>Schooldog-Teacher-Team</b> Weak	Schlesier et al. (2019)/ 1 (NR)	NR (NR) / 10/28 / Weak	$N =$ NR; (Mean 8.5 years)	<b>M:</b> based on animals.
<b>*Seattle Social Development Project</b> Strong	Goldberg et al. (2019)/ 1(CE)	Adj-B; AO and MH (internalizing symptoms)/ (significant positive effects) / Strong	$N =$ 643/USA	<b>C:</b> Involves the family and the community <b>I:</b> adequate
<b>Second step program</b> No evidence	Siddiqui & Ventista (2018)/ 1(E)	NR (effect size not calculated due to lack of data) / NR	$N =$ 500; 2 years	<b>O:</b> social and emotional learning and prevention of bullying at school. <b>S:</b> self-regulation, behavioral skills. <b>M:</b> 25- to 40-minute sessions. Photo cards and videotaped stories are used to introduce and stimulate key issues. Application of self-regulation activities, role playing. <b>I:</b> material for application and training is available commercially.
<b>SOAR</b> Skills, Opportunity, and Recognition No evidence	Weissberg & O'Brien (2004)/ 1(NR)	"higher scores in reading, arts, NR language, math, SS; less anti-social interactions and less tobacco use in participants compared to controls" / (NR)		<b>O and S:</b> teach prosocial skills such as regulating emotions; listening, sharing and respecting others; cooperating in learning teams; problem solving; and developing strong bonds with peers, teachers and families. <b>M:</b> school-wide activities are presented to encourage teacher and family participation.
<b>*Social skills improvement system-classwide intervention programme (SSIS-CIP)</b> Moderate	Siddiqui & Ventista (2018)/ 1(RCT)	Adj-B (cooperation, comitment); Adj-E (self-regulation) and Adj-S (SS, assertiveness)/ (low to moderate effect sizes in favor of the intervention)/ 4 a 5/ Moderate	$N =$ 479; 1st y 2nd/USA; 12 weeks.	<b>O:</b> develop SS and reduce behavior problems in the classroom. <b>S and M:</b> 10 different SS in units of three 20-minute lessons: listening to others, following directions and classroom rules, ignoring distractions, asking for help, talking in conversation, cooperating with others, controlling anger, acting responsibly, and kindness.
<b>*Steps to Respect</b> Strong	Goldberg et al. (2019)/ 1(RCT)	Adj-B; Adj-E; Adj-S and AO (significant positive effects) / Moderate	$N =$ 3119/USA	<b>O:</b> Addresses bullying. <b>C:</b> Involves the family
	Goldberg et al. (2019)/ 2(RCT)	Adj-B; Adj-E; Adj-S and AO (significant positive effects) / Strong	$N =$ 1126/USA	
<b>*TFGA</b> Tools for getting along Moderate	Pandey (2018)/ 1**(NR)	Adj-E (self-regulation) (significant improvements) / Moderate	$N =$ 1296; 4th y 5th/ USA; 1 year	<b>O and S:</b> emotion regulation, social problem solving. <b>M:</b> 27 sessions which apply cognitive modeling, role-playing, small group activities, and strategies in real-life social situations. <b>TF:</b> Based on Cognitive Behavioral approach. <b>WA:</b> Trained teachers and school counselors.
	Schlesier et al. (2019)/ 1** (NR)	Adj-B (aggression)/(moderate influence)/ and Adj-E (very positive impact) / 23/28 / Moderate	NR	
	Schlesier et al. (2019)/ 2(NR)	Adj-B (aggression) and Adj-E/(positive impact, without maintenance at one year) / 21/28 / Moderate	NR	
<b>Think, Feel, Do</b> Weak	Mackenzie & Williams (2018)/ 1 (E)	MH (Anxiety) / (significant reduction with medium effect sizes) / 34% / Weak	$N =$ 13/England 6 interactive multimedia sessions (x45 min)	<b>O and S:</b> recognizing emotions, thoughts and behavior; identifying and changing negative thoughts; and problem-solving. <b>M:</b> carried out using an interactive multimedia CD-ROM; an animated character guides activities. Uses practical exercises, questionnaires, videos, music and animation. <b>TF:</b> Based on the Cognitive Behavioral approach, with psychoeducation. <b>WA:</b> Researcher (in the baseline study); within the school.

Program name/ Level of evidence	Studies	Outcomes/Quality	Primary Study Data	Program Data
	Review/ Primary study (Design)	Variables (outcomes)/ Primary study quality	Sample $N =$ ; Age or educative level/ Country-application; Duration (Sessions)	Objectives; Skills taught; Methodology; Theoretical frame of reference; Who applied it; Components (family, community, other); Implementation.
<b>Together at School</b> Weak	Goldberg et al. (2019)/ 1(RCT)	Adj-B; Adj-E and Adj-S/ (significant positive effects)/ Weak	$N = 3704$ /Finland	<b>C:</b> Involves the family.
<b>Tribes learning communities</b> No evidence	Siddiqui & Ventista (2018)/ 1(RCT)	NR (effect size not calculated due to lack of data) / NR	$N = 2309$	<b>O:</b> creates a sense of belonging where students feel supported and appreciated by their peers and teachers. <b>M:</b> community-building type of intervention, to resemble a family. <b>C:</b> family and community
<b>*WHO</b> Moderate	Siddiqui & Ventista (2018)/ 1(CE)	Adj-B (aggressive and disruptive behavior); Adj-S (social interaction) and MH (well-being and self-esteem)/ (small effect size)/ 3 / Moderate	$N = 91$ ;(8-10 years) /Italy; 1 year	<b>O:</b> improve interpersonal relations and self-esteem. <b>S:</b> self-image (2 meetings), self-esteem (6 meetings), corporeality (3 meetings), active listening (2 meetings). <b>M:</b> 15 two-hour meetings, with different activities for each area. <b>TF:</b> based on recommendations from the WHO.
<b>*WITS</b> Moderate	Goldberg et al. (2019)/ 1(CE)	Adj-B; Adj-E; Adj-S; AO and MH (internalizing symptoms) / (significant positive effects) / Moderate	$N = 432$ / Canada	<b>O:</b> Addresses bullying. <b>C:</b> Involves the family and the community <b>I:</b> reports reliability problems in the implementation.
	2(CE)	Adj-B / (significant positive effects) / Moderate	$N = 830$ / Canada	
<b>Zero Strong</b>	Goldberg et al. (2019)/ 1(CE)	Adj-B / (significant positive effects) / Strong	$N = 20446$ / Norway	<b>O:</b> Addresses bullying. <b>C:</b> Involves the family. <b>I:</b> adequate
<b>*Zippy's Friends</b> Moderate	Pandey (2018)/ 1(NR)	Adj-B (behavior problems), Adj-E (self-awareness, self-regulation, motivation) and Adj-S (SS)/ (significant improvements) / Moderate	$N = 730$ (Mean 7.3 years)/Ireland; 1 year	<b>O:</b> improve self-regulation and promote emotional well-being. <b>S:</b> Self-regulation <b>M:</b> 24 sessions divided across 6 modules, each containing 6 illustrated stories with an imaginary insect named Zippy and his friends. The teacher reads the stories and the pupils carry out related activities, discussions, or drawings. <b>WA:</b> Trained teachers or healthcare professionals.
	Siddiqui & Ventista (2018)/ 2(RCT)	Adj-E (emotional literacy, self-regulation, motivation) and Adj-S (SS)/ (low effect sizes) / 2 to 3 / Weak	$N = 766$ /Ireland; 1 year	
	3 (RCT)	Adj-E (regulation)/ (low effect size) / 3 to 4 / Moderate	$N = 1483$ ; (7-8 years)/Norway; 1 year	
	4 (E)	Adj-B (cooperation); Adj-E (self-regulation) and Adj-S (assertiveness)/ (low effect sizes) / 2 to 3 / Weak	$N = 432$ /Denmark; 1 year	
<b>*4R's</b> Strong	Goldberg et al. (2019)/ 1(RCT)	Adj-B; Adj-E; Adj-S; AO and MH (internalizing symptoms)/ (significant positive effects) / Strong	$N = 942$ / USA	<b>C:</b> Involves the family. <b>I:</b> adequate

Note. (\*) Programs with moderate or strong evidence; (\*\*) Using the same primary studies; C: Components; S: Skills; SS, Social skills; I: Implementation; M: Methodology; TF: Theoretical framework; NR: Not reported; O: Objectives; WA: Who applied the intervention; PS: Problem solving; MH: Mental health.

#### Effects of high- and medium-quality programs

The programs with strong evidence ( $n = 6$ ) showed significant positive effects in different variables in favor of the intervention, five (83.3%) showed results in Adj-B; 4 (66.6%) in AO; 3 of them (50%) also showed effects in Adj-E and Adj-S; and three in MH; in the latter group, only the FRIENDS program reported results in MH, with significant improvements in social and general anxiety. Effectiveness of these programs was tested through randomized experimental or quasi-experimental designs with control groups and samples of primary school subjects that contained from 643 to

20,446 subjects. Four of the programs (66.6%) were applied in the United States, one in England, and another in Norway.

Programs with moderate evidence ( $n = 14$ ) presented the following effects. Thirteen (92.8%) showed significant positive results in favor of the intervention in Adj-B; 10 (71.4%) in Adj-E and Adj-C; and 4 (28.5%) in AO and SM. Effects were tested with randomized experimental studies, experimental studies or quasi-experimental studies with samples ranging from 91 to 5081 subjects. Seven of these (50%) were applied in the United States and 2 (14.2%) were implemented in more than one country.

### *Descriptive characteristics*

Among the high-quality programs, one (16.6%) provided more detailed information about its characteristics, the other 5 (83.3%) provided little information; for this reason we decided to group together the programs with strong and moderate quality to present the descriptive characteristics. Together, these represent 51.2% ( $n = 20$ ) of all the interventions identified ( $n = 39$ ).

Regarding their components, all of them (100%) implemented activities with the students, 14 (70%) also involved the family and 10 (50%) additionally involved the educational community. Information on who implemented the programs was reported in 7 cases (35%): 6 were implemented by trained teachers, in some cases they were also carried out by professionals, researchers or trained school counselors. Implementation was considered in 9 programs (45%), 7 (35%) reported adequate reliability in implementation; in addition, 5 (25%) presented support materials such as workbooks, manuals, worksheets or videos to facilitate implementation and 2 (10%) mentioned training as an implementation strategy. Nine programs reported their duration (45%), 6 (30%) had a one-year duration, and the number of sessions ranged from 9 to 45.

The objectives were mentioned in 8 programs (40%), 3 (15%) reflected general objectives (to address bullying) and 5 (25%) reported more specific objectives, on 4 occasions relating to Adj-S (e.g. improved peer relations); in 3 cases related to Adj-B (e.g. reducing behavior problems); in 2 programs related to Adj-E (e.g. improving self-regulation), and in 2 programs related to MH (promoting well-being, self-esteem).

Nine programs (45%) described the skills worked on; in 4 of these, objectives and skills could not be clearly differentiated; we decided to include them as skills, considering the way in which they were expressed. Seven skills related to Adj-B were cited: coping, problem-solving, and conflict resolution were mentioned in 4 programs; 2 programs mentioned following instructions and rules; expanding the behavioral repertoire, cooperation, responsibility and kindness were each included in one program. Six skills were related to Adj-S: of these, listening skills were worked on in 3 programs; social skills, dialogue, argumentation, help-seeking and conversational skills each appeared in one program. Five skills were related to Adj-E: of these, cognitive self-regulation (regulating thoughts, perspective taking, reflection) and emotional self-regulation (e.g. control of anger) were mentioned in three; relaxation and identification of emotions and thoughts were each mentioned in two programs. Five skills were related to Academic Outcomes (AO): attention skills were named in 2 interventions; study skills, language skills, organization and planning were each mentioned in one program.

Work methodology was reported in 10 interventions (50%): 9 (45%) referred to an organized, sequenced structure. In addition, 7 (35%) mentioned more specific activities;

group work was mentioned in 5 programs, use of games in 2, and use of real-life situations or social stories in one case each. Use of classroom support materials (books, videos, work sheets) was reported in 5 programs (25%). Specific techniques were mentioned in 3 interventions (15%); modeling and role playing in 2 programs, and problem solving in one. The theoretical framework of reference was reported in 4 programs (20%); cognitive-behavioral approach was mentioned in 3 cases (15%), along with social learning and behavioral analysis in one case; another program was based on WHO recommendations.

## **Discussion**

The general aim of this study was to identify and analyze reviews and meta-analyses that included primary studies on intervention programs to develop social-emotional skills in primary education. For this purpose, the umbrella-type systematic review methodology was used. This procedure has seldom been used in this context; we found only one previous study that analyzed the effectiveness of programs published between 1997 and 2007, and incorporating all educational levels (Diekstra & Gravesteyn, 2008). The present research limited its study population to the sphere of primary education, and stipulated a publication period covering the past 20 years. In addition, study and program quality were evaluated, and we incorporated systematic reviews, meta-analyses and narrative reviews, overcoming some of the limitations of the previous study.

Our research identified 15 secondary studies that contain school-based, universal social-emotional development programs, confirming that there are numerous reviews in this field. However, only two reviews were identified that exclusively addressed the primary education population; the rest included two or more different stages of education, meaning that a large part of the data were not segregated. In line with previous findings, the scope of these reviews was found to be very broad and not very specific, as they include heterogeneous populations, with students of all ages, grade levels and stages of education, making it difficult to extract information, and limiting comparison between programs (Miñaca et al., 2013; Taylor et al., 2017; Vilorio, 2005).

The first review question was directed toward ascertaining the methodological characteristics and quality of the studies identified. We found that most of them were published in the past five years, come from the United States, and incorporate experimental or quasi-experimental primary studies, with control groups. Very few reviews included primary studies not reported in English; in the medium- and high-quality reviews, English was the exclusive language. A potential risk of bias is detected here, as limiting studies to a single language may exclude relevant research or leave out cultural considerations; this may have a negative impact on generalization and interpretation of results (Jackson & Kuriyama, 2019; Stern & Kleijnen, 2020).

Regarding quality, 60% of the secondary studies present-

ed low or critically low quality. The systematic reviews and meta-analyses, as a whole, comply with the basic methodological aspects; however, most of them do not comply with such fundamental aspects as evaluation of biases, explanation of reasons for excluding studies, or application of blind peer reviewing. On the other hand, the overall low quality of the non-systematic or narrative reviews was related to shortcomings in data presentation, reporting of search strategies, selection criteria, objectives and justification. Other umbrella reviews conducted in the educational field find similar results, regardless of whether they are systematic, meta-analytic or literature reviews, mainly mentioning the lack of transparency in the methods used as the central element affecting the reproducibility of research in this field (Barbosa et al., 2020; Gessler & Siemer, 2020).

For their part, the six reviews with moderate or high quality reflected an acceptable level of confidence in their results; for this reason they were selected for the next phase of answering the second review question, related to the level of program evidence. In this phase 39 programs were identified, and we determined the level of evidence for each of these. The general results showed that the reviews show different ways of studying and presenting the evidence, they apply different instruments and report values expressed in numerical ranges, percentages or ordinal categories that are not comparable with each other. For this reason, it was essential to unify the criteria for interpreting evidence by applying a nominal system used in previous studies (Gálvez-Lara et al., 2018). We found that half of these programs show moderate to strong levels of evidence. The remainder are split between those that show weak quality, a small proportion of programs with insufficient evidence, and a single program with inconsistent evidence. In this line, other previous studies underscore the high number of existing programs, as opposed to the low quantity and quality of evidence supporting them. By way of explanation, they suggest the characteristics of the primary studies, whose most frequent limitations are sample size, lack of randomization of the subjects, absence of control groups or the absence of follow-on measures (Goldberg et al., 2019; Schlesier et al., 2019; Siddiqui & Ventista, 2018).

In this phase, it is striking that most of the programs appear only in a single review, and they have only one source of support. In this regard, only the PATHS program was represented in all the reviews, and it presented seven primary sources of support; six other programs were included in two reviews only, and for the most part have two primary sources of support (Caring School Communities, Child Development Project, INSIGHT, Positive action, TFGA and Zippy's Friends). Another trend found was the coexistence of primary studies with weak and moderate quality as support for the same program. All this indicates that there is a low level of agreement between the reviews in terms of the programs identified and also in the primary studies analyzed, despite the fact that the time period and general objectives were similar. This variability may be associated with the use

of different strategies, search terms, eligibility criteria, and databases. These findings imply that there is still a need for more rigorous research showing robust effects of social-emotional interventions in the school context, and that the scope of secondary studies that synthesize this evidence also needs to be more clearly defined (Siddiqui & Ventista, 2018; Slavin, 2019).

The third review question aimed to establish the primary program effects that were reported. In general, and in line with Rubiales et al. (2018), we found that the variables evaluated were highly heterogeneous; for this reason we grouped the results using the categories defined by Goldberg et al. (2019) and Sklad et al. (2012). We found that the high- and moderate-quality programs presented significant positive effects mainly in one measure of secondary results, that is, behavioral adjustment (Adj-B), most notably a decrease in behavior problems or improvement in disruptive behaviors; other programs showed significant effects in positive behaviors, such as cooperation. Of the primary results, emotional and social adjustment were most often reported as significant, positive results. Significant improvements were seen primarily in emotional self-regulation, followed by emotion recognition and expression; significant improvements were also reported in SSs, communication and interpersonal problem solving. Most of the high-quality programs showed significant positive effects in another secondary variable, academic outcomes (AO), along with other medium-quality programs that also reflected these effects. Finally, a smaller proportion of medium- and high-quality programs reported effects on measures of mental health, mainly reflecting a significant decrease in internalizing symptoms, often mentioning anxiety.

In comparison with results from Diekstra and Gravesteyn (2008), one can observe coincidence in some of the reported effects, confirming that these types of programs significantly improve emotional and social skills, increase academic performance in certain specific areas, and reduce mental problems. As for behavior improvement, the present umbrella review reflects an important role attributed to behavior results, even though such important results were not reported in the studies by Diekstra and Gravesteyn (2008) or Goldberg et al. (2019). In the reviews by Sklad et al. (2012) and Fernandez et al. (2014), the programs reported significant effects on social-emotional and behavioral skills, in similar proportions. One possible explanation for this variability in findings could be the diversity of categories of analysis used in the different reviews, along with diversity in the initial frame of reference (Rubiales et al., 2018). Siddiqui and Ventista (2018) also suggest an explanation in the way results are studied and reported in the primary studies, where social-emotional skills are sometimes studied as a unit, and sometimes are reported as independent variables. This is undoubtedly a limitation for synthesizing and interpreting the findings.

On the other hand, it is difficult to establish which components affect each study variable, due to the lack of clarity

in reporting the data. In this regard, few reviews analyze the possible moderators of the effects with any specificity (Goldberg et al., 2019; Sklad et al., 2012). Neither could we analyze program effect size in this review, since this data, as Schlesier et al. (2019) also notes, is rarely reported. Even so, some secondary studies report significant effects in favor of programs with mainly small effect sizes, or moderate in some cases (Goldberg et al., 2019; Mackenzie & Williams, 2018; Pandey et al., 2018). All this may limit our interpretation of the effectiveness of the programs under discussion, as also occurs in previous studies (Canet-Juric et al., 2020; McCallops et al., 2019; Schlesier et al., 2019).

It should also be noted that medium- and high-quality programs have been applied mainly in an English-speaking population, in the United States and England, and to a lesser extent in countries like Norway. This coincides with the trend indicated by certain specialized agencies and in some previous studies (EEF, 2021; ESSA, 2021; Diekstra & Gravesteyn, 2008). Although some programs have been applied in different countries, such as the Zippy's friends implemented in Ireland, Norway and Denmark with rural and urban populations; or TFGA, in white and African-American populations; this is not the general trend. In this matter, we observe a lack of information on sample characteristics in many of the reviews, and there is also a possible risk of selection bias due to the limited representativeness across cultures, races and contexts.

The fourth question aimed to ascertain program characteristics of the programs with the most evidence. In this regard, one difficulty we encountered was the large variation in how descriptive information was reported, along with a high degree of unreported information. Information on objectives and skills worked on was reported in less than half of the quality interventions. In addition, the objectives were expressed in a very general way, or it was not possible to clearly differentiate them from the skills being worked on. Among the objectives, the most notable aims were to address bullying, improve social skills and interpersonal relationships, and reduce behavioral problems. The skills most frequently worked on in medium- and high-quality programs coincide with the results reported, and have to do with behavioral adjustment, especially coping and problem-solving skills; social adjustment, where developing communication skills is the main focus; and emotional adjustment, where cognitive and emotional regulation or relaxation are worked on, among

other skills. Some programs reinforce specific academic skills, like attention or planning.

Half of the programs offered data on the working methodology. In general terms, there is a tendency to use structured, active, guided and interactive methodologies in small groups, with the use of cognitive and behavioral modeling strategies, role playing, or problem-solving techniques, along with the use of specific material. The theoretical model of reference is seldom mentioned, though the cognitive-behavioral approach stands out. In the same way, there was little data reported for duration, but the length most often mentioned was about one year. Regarding components, most medium- and high-quality interventions involved the family, and half of them also involved the educational community. The programs were applied by trained teachers or professionals. Noteworthy among the implementation strategies was the use of manuals or specific materials, and training to support the implementation.

The high degree of unreported data precludes a more exhaustive analysis of the characteristics, effects, and possible moderators of program effects.

Authors such as Slavin (2019) highlight the value of secondary studies as resources that help to synthesize evidence for practical purposes, whether for selecting appropriate programs for classroom work, or for making political decisions on implementing broader strategies that are supported by scientific bases. Even though we have found an important number of secondary studies that contain universal school programs for social-emotional development, in future research it would be advisable that the population be segregated by age or stage of education; that primary studies not be limited to the English language; that more precise data be reported on sample characteristics and the context of program application. Primary studies should be extended to diverse contexts; primary and secondary results be reported in detail; possible moderators of these results should be described and analyzed; the quality of the primary studies should be assessed; and clear levels of evidence should be established to facilitate interpretation of the effects. Along the line of reasoning of Joyce (2019) and Slavin, (2019), all of this taken together could help make it possible to generalize findings and establish practical evidence-based recommendations.

**Conflict of interest.-** The authors declare no conflict of interest.

**Financial support.-** No funding.

## References

- Alzahrani, M., Alharbi, M., & Alodwani, A. (2019). The effect of social-emotional competence on children academic achievement and behavioral development. *International Education Studies*, 12(12), 141-149. <https://doi.org/10.5539/ies.v12n12p141>
- Aromataris E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132-140. <https://doi.org/10.1097/XEB.0000000000000055>
- Baethge, C., Goldbeck-Wood, S., & Mertens, S. (2019). SANRA a scale for the quality assessment of narrative review articles. *Research integrity and peer review*, 4(1), 1-7.
- Barbosa, A., Whiting, S., Simmonds, P., Scotini Moreno, R., Mendes, R., & Breda, J. (2020). Physical activity and academic achievement: an um-

- rella review. *International Journal of Environmental Research and Public Health*, 17(16), 5972, <https://doi.org/10.3390/ijerph17165972>
- Bear, G.G., y Watkins, J. M. (2006). Developing self-discipline. In Bear G., & K. Minke M. (Eds.). *Children's Social and Emotional Learning 421needs III: Development, prevention, and intervention* (pp. 29–44). Bethesda, MD: National Association of School Psychologists.
- Björklund, K., Liski, A., Samposalo, H., Lindblom, J., Hella, J., Huhtinen, H., & Santalahti, P. (2014). "Together at school" - a school-based intervention program to promote socio-emotional skills and mental health in children: study protocol for a cluster randomized controlled trial. *BMC Public Health* 14(1), 1-11. <http://www.biomedcentral.com/1471-2458/14/1042>
- \*Canet-Juric, L., García-Coni, A., Andrés, M. L., Vernucci, S., Aydmune, Y., Stelzer, F., Richard's, M. M. (2020). Intervención sobre autorregulación cognitiva, conductual y emocional en niños: Una revisión de enfoques basados en procesos y en el currículo escolar en Argentina [Intervention on cognitive, behavioral and emotional self-regulation in children: A review of process-based and school curriculum approaches in Argentina]. *Revista Argentina de Ciencias del Comportamiento*, 12(1), 1-25. <https://doi.org/10.32348/1852.4206.v12.n1.24999>
- Ciapponi, A. (2018). AMSTAR-2: herramienta de evaluación crítica de revisiones sistemáticas de estudios de intervenciones de salud [AMSTAR-2: Critical appraisal tool for systematic reviews of health intervention studies]. *Evidencia, Actualización en Práctica Ambulatoria*, 21(1). <https://doi.org/10.51987/evidencia.v21i1.6834>
- Cornell, C., Kiernan, N., Kaufman, D., Dobe, P., Frydenberg, E., & Deans, J. (2017). Developing social emotional competence in the early years. In Frydenberg, E., Martin, A., Collie, R. (eds) *Social and emotional learning in Australia and the Asia-Pacific* (p. 391–411). Springer Science Business Media. [https://psycnet.apa.org/doi/10.1007/978-981-10-3394-0\\_21](https://psycnet.apa.org/doi/10.1007/978-981-10-3394-0_21)
- Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333, 959-964. <https://doi.org/10.1126/science.1204529>
- Diekstra, R., & Gravesteyn, C. (2008). *A review of meta-analytic literature reviews in social and emotional education. Social and emotional education: An international analysis*. 258-284. Fundación Botín.
- \*Durlak, J., Weissberg, R. Dymnicki, A., Taylor, R., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, 45, 294–309. <https://link.springer.com/article/10.1007%2F10464-010-9300-6>
- Every Student Succeeds Act (ESSA) (2021). *Evidence-based Social-Emotional Learning Programs*. Center for Research and Reform in Education, Johns Hopkins University. <https://www.evidenceforessa.org/programs/social-emotional>
- Education Endowment Foundation (EEF) (2021). *Social and emotional learning*. Department for Education. What Works Network. <https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/social-and-emotional-learning/>
- \*Fernández, M., Díaz, M., Corredor A., & Arráez, A. (2014). Revisión sistemática de los programas de entrenamiento en competencia social [Systematic review of social competence training programs]. *International Journal of Developmental and Educational Psychology*, 4 (1), 229-237. <https://doi.org/10.17060/ijodaep.2014.n1.v4.607>
- Gálvez-Lara, M., Corpas, J., Moreno, E., Venceslá, J. F., Sanchez-Raya, A., & Moriana, J. A. (2018). Psychological treatments for mental disorders in children and adolescents: A review of the evidence of leading international organizations. *Clinical Child and Family Psychology Review*, 21(3), 366-387. <https://doi.org/10.1007/s10567-018-0257-6>
- Gessler, M., & Siemer, C. (2020). Umbrella review: Methodological review of reviews published in peer-reviewed journals with a substantial focus on vocational education and training research. *International Journal for Research in Vocational Education and Training*, 7(1), <https://doi.org/10.13152/IJRVET.7.1.5>
- Grajales, C. (2003). La escuela: un escenario de formación y socialización para la construcción de identidad moral [The school as an scenario of socialization for the development of moral self-identity]. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 1, 15-43. <http://www.scielo.org.co/pdf/rics/v1n2/v1n2a06.pdf>
- \*Goldberg, J. M., Sklad, M., Elfrink, T. R., Schreurs, K. M., Bohlmeijer, E. T., & Clarke, A. M. (2019). Effectiveness of interventions adopting a whole school approach to enhancing social and emotional development: a meta-analysis. *European Journal of Psychology of Education*, 34(4), 755-782. <https://doi.org/10.1007/s10212-018-0406-9>
- Jackson, J. L., & Kuriyama, A. (2019). How often do systematic reviews exclude articles not published in English? *Journal of General Internal Medicine*, 34(8), 1388-1389. <https://doi.org/10.1007/s11606-019-04976-x>
- Joyce, K. E. (2019). The key role of representativeness in evidence-based education. *Educational Research and Evaluation*, 25(1-2), 43-62. <https://doi.org/10.1080/13803611.2019.1617989>
- Low, S., Cook, C. R., Smolkowski, K., & Buntain Ricklefs, J. (2015). Promoting social-emotional competence: An evaluation of the elementary version of Second Step. *Journal of School Psychology*, 53, 463-477. <https://doi.org/10.1016/j.jsp.2015.09.002>
- \*Mackenzie, K., & Williams, C. (2018). Universal, school-based interventions to promote mental and emotional well-being: what is being done in the UK and does it work? A systematic review. *BMJ Open*, 8(9). <https://doi.org/10.1136/bmjopen-2018-022560>
- \*McCallops, K., Barnes, T. N., Berte, I., Fenniman, J., Jones, I., Navon, R., & Nelson, M. (2019). Incorporating culturally responsive pedagogy within social-emotional learning interventions in urban schools: An international systematic review. *International Journal of Educational Research*, 94, 11-28. <https://doi.org/10.1016/j.ijer.2019.02.007>
- Miñaca, M. I., Hervás, M., & Laprida, I. (2013). Análisis de programas relacionados con la Educación Emocional desde el modelo propuesto por Salovey & Mayer [Analysis of programs related to Emotional Education from the model proposed by Salovey & Mayer.]. *Revista de Educación Social*, 17, 1-17. <http://hdl.handle.net/10272/10243>
- \*Opengart R. (2007). Integrative literature review: Emotional intelligence in the k-12 curriculum and its relationship to american workplace needs: A literature review. *Human Resource Development Review*, 6(4), 442-458. <https://doi.org/10.1177%2F1534484307307556>
- Oros, L. B., Manucci, V., & Richaud, M. C. (2011). Desarrollo de emociones positivas en la niñez. Lineamientos para la intervención escolar [Development of positive emotions in children: guidelines for school intervention]. *Educación y Educadores*, 14(3), 493-509. <http://www.redalyc.org/articulo.oa?id=83422605004>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E.A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. J., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S.,... & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/bmj.n71>
- \*Pandey, A., Hale, D., Das, S., Goddings, A. L., Blakemore, S. J., & Viner, R. M., (2018). Effectiveness of universal self-regulation-Based interventions in children and adolescents. *JAMA Pediatrics*, 172 (6), 566. <https://doi.org/10.1001/jamapediatrics.2018.0232>
- \*Pelco, L. E., & Reed, V. E. (2007). Self-regulation and learning related social skills: Intervention ideas for elementary school students. *Preventing School Failure*, 51(3), 36–43. <https://doi.org/10.3200/PSFL.51.3.36-42>
- Pérez-Escoda, N., Filella, G., Alegre, A., & Bisquerra, R. (2012). Developing the emotional competence of teachers and pupils in school contexts. *Electronic Journal of Research in Educational Psychology*, 10(3), 1183-1208. [http://repositorio.ual.es/bitstream/handle/10835/1888/Art\\_28\\_756\\_eng.pdf?sequence=1](http://repositorio.ual.es/bitstream/handle/10835/1888/Art_28_756_eng.pdf?sequence=1)
- \*Rubiales, J., Russo, D., Paneiva, P., & González, R. (2018). Revisión sistemática sobre los programas de entrenamiento socioemocional para niños y adolescentes de 6 a 18 años publicados entre 2011 y 2015 [Systematic review of social-emotional training programs for children and adolescents]. *Revista Costarricense de Psicología*, 37(2),163-186. <https://doi.org/10.22544/rcps.v37i2.05>
- \*Sancassiani, F., Pintus, E., Holte, A., Paulus, P., Moro, M. F., Cossu, G., & Lindert, J. (2015). Enhancing the emotional and social skills of the

- youth to promote their wellbeing and positive development: A systematic review of universal school-based randomized controlled trials. *Clinical Practice and Epidemiology in Mental Health*, 11(1), 21-40. <https://doi.org/10.2174/1745017901511010021>
- \*Schlesier, J., Roden, I., & Moschner, B. (2019). Emotion regulation in primary school children: A systematic review. *Children and Youth Services Review*, 100, 239–257. <https://doi.org/10.1016/j.childyouth.2019.02.044>
- \*Siddiqui, N., & Ventista, O. M. (2018). A review of school-based interventions for the improvement of social emotional skills and wider outcomes of education. *International Journal of Educational Research*, 90, 117-132. <https://doi.org/10.1016/j.ijer.2018.06.003>
- Slavin, R. E. (2019). How evidence-based reform will transform research and practice in education. *Educational Psychologist*, 55(1), 21-31. <https://doi.org/10.1080/00461520.2019.1611432>
- Smith, V., Devane, D., Begley, C. M., & Clarke, M. (2011). Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology*, 11(1), 1-6. <https://doi.org/10.1186/1471-2288-11-15>
- Shea, B.J, Reeves, B.C, Wells, G., Thuku, M., Hamel, C., Moran J., & Henry D. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358. <https://doi.org/10.1136/bmj.j4008>
- \*Sklad, M., Diekstra, R., Ritter, M. Ben, J., & Gravezteijn, C. (2012). Effectiveness of school-based universal social,emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment?. *Psychology in the Schools*, 49(9), 892-909. <https://doi.org/10.1002/pits.21641>
- Stern, C., & Kleijnen, J. (2020). Language bias in systematic reviews: you only get out what you put in. *JBIE Evidence Synthesis*, 18(9), 1818-1819. <https://doi.org/10.11124/JBIES-20-00361>
- Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: a meta-analysis of follow-up effects. *Child Development*, 88(4), 1156-1171. <https://doi.org/10.1111/cdev.12864>
- Viloria, C. (2005). La educación emocional en edades tempranas y el interés de su aplicación en la escuela: programas de educación emocional, nuevo reto en la formación de los profesores [Emotional education at an early age and interest in its application at school: emotional education programs, a new challenge in teacher training]. *Tendencias pedagógicas*, 10, 107-124. [https://repositorio.uam.es/bitstream/handle/10486/4739/31241\\_2005\\_10\\_05.pdf](https://repositorio.uam.es/bitstream/handle/10486/4739/31241_2005_10_05.pdf)
- \*Weissberg, R. P., & O'Brien, M. U. (2004). What works in school-based social and emotional learning programs for positive youth development. *Annals AAPSS*, 591, 86-97. <https://doi.org/10.1177%2F0002716203260093>
- What Works Clearinghouse (WWC) (2021). *A searchable online database of available school and classroom programs and interventions: The What Works Clearinghouse*. Institute of Education Sciences. U.S Department of Education. <https://ies.ed.gov/ncee/wwc/>
- Wu, Z., Hu, B. Y., Fan, X., Zhang, X., & Zhang, J. (2018). The associations between social skills and teacher-child relationships: A longitudinal study among Chinese preschool children. *Children and Youth Services Review*, 88, 582-590. <https://doi.org/10.1016/j.childyouth.2018.03.052>
- Zins, J. E., Elias, M. J., Greenberg, M. T., & Weissberg, R. P. (2000). Promoting social and emotional competence in children. *Preventing school problems promoting that work*, 71-100. Bethesda: National Association of School Psychologists.